Optical Fiber Communications

Gerd Keiser
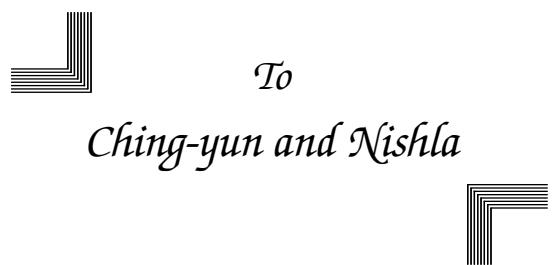
Optical Fiber Communications

——Fourth Edition——

Optical Fiber Communications

Fourth Edition

Gerd Keiser

National Taiwan University of Science and Technology

To
Ching-yun and Nishla

About the Author

Gerd Keiser is a National Science Council Chair Professor in the Department of Electronic Engineering at the National Taiwan University of Science and Technology. His teaching and research interests include photonic component development, telecom optical transmission systems, fiber-to-the-premises (FTTP) networks, intelligent-building and smart-home networks, and biomedical photonics. In addition, he is the founder and principal consultant at PhotonicsComm Solutions, a firm specializing in consulting and education for the optical communications industry. Previously he worked at Honeywell, GTE, and General Dynamics in the development and application of optical network and digital switch technologies. His technical achievements at GTE earned him the prestigious Leslie Warner Award. As an outside activity, he was an adjunct professor at Northeastern University, Tufts University, and Boston University. Dr. Keiser is a Fellow of the IEEE, a member of OSA and SPIE, an associate editor of the journal *Optical Fiber Technology*, and the author of four graduate-level books.

Brief Contents

Contents

Preface

Objective

Optical fiber communications has undergone a fascinating history since the first edition of this book in 1983. Especially exciting was the 2009 Nobel Prize in Physics received by Charles K. C. Kao for his pioneering insight into using glass fibers as a medium for data transmission and for his enthusiastic international follow-ups in promoting the further development of low-loss optical fibers. The first ultrapure fiber was fabricated in 1970, only four years after Kao's prediction. This breakthrough led to a series of technology developments related to optical fibers. Initially the technology focused on simple transmission links but then quickly moved to increasingly sophisticated networks. Along the way many new components and communication techniques were tried. Some of these were highly successful, some faded away perhaps because of their implementation complexity, and others, which were ahead of their time, are reappearing after being in hibernation for a while. Modern high-capacity telecommunication networks based on optical fiber technology now have become an integral and indispensable part of society. Applications for these sophisticated networks range from simple web browsing and e-mail exchanges to critical health care diagnosis, grid and cloud computing, and complex business transactions. Due to the importance of these networks to everyday life, users have come to expect the communication services to always be available and to function properly. Meeting such a stringent demand requires careful engineering in all technological aspects ranging from component development to system design and installation to network operation and maintenance.

To address the attainment and implementation of these skills, this expanded fourth edition presents the fundamental principles for understanding and applying a wide range of optical fiber technologies to modern communication networks. The sequence of topics takes the reader systematically from the underlying principles of components and their interactions with other devices in an optical fiber link, through descriptions of the architectures and performance characteristics of complex optical links and networks, to essential measurement and test procedures required during network installation and maintenance. By mastering these fundamental topics the reader will be prepared not only to contribute to disciplines such as current device, communication link, or equipment designs, but also to understand quickly any further technology developments for future enhanced networks.

Contents

To accomplish these objectives, Chapter 1 gives a basic overview of how optical communications blend into telecommunication systems. The discussions include the motivations and advantages for using optical fibers, the spectral bands being used, how wavelength division multiplexing can boost the transmission capacity of an optical fiber, what standards are being applied, and the use of simulation tools.

Chapters 2 through 11 describe the purpose and performance characteristics of the major elements in an optical link. These elements include optical fibers, light sources, photodetectors, passive optical devices, optical amplifiers, and active optoelectronic devices used in multiple-wavelength networks. Despite its apparent simplicity, an optical fiber is one of the most important elements in a fiber link. Chapters 2 and 3

give details on the physical structures, constituent materials, attenuation behavior, lightwave propagation mechanisms, and signal distortion characteristics of the wide variety of optical fibers that exist. In addition, Chapter 3 introduces optical fiber fabrication methods and illustrates several generic fiber cable configurations. The new topics in these chapters include a discussion of photonic crystal fibers and a streamlining of the discussions on modal effects and pulse broadening.

Chapter 4 addresses the structures, light-emitting principles, and operating characteristics of light sources used in optical communications. In addition, the discussion includes direct and external modulation techniques, temperature effects, device lifetime considerations, and line coding methods for transporting signals over a fiber. How to couple the light source to a fiber is described in Chapter 5, as well as how to join two fibers in order to ensure a low optical power loss at the joints.

The lightwave receiver has the task of detecting the arriving optical signal and converting it into an electrical signal that can be processed by the receiver electronics. Chapter 6 covers the structures and responses of photodetectors, and Chapter 7 describes the principles and functions of lightwave receivers. The new features in Chapter 7 include a simplified presentation of the operational characteristics of lightwave receivers, the concepts of signal-detection statistics, eye-diagram measurement schemes, and a description of burst-mode receivers used for passive optical networks.

Chapters 8 and 9 examine the design methods for digital and analog links, respectively. For Chapter 8 this includes discussions of link power budgets and bandwidth limitations. The new features are an expanded coverage of power penalties, basic coherent detection schemes, and details of error-control methods for digital signals. Additions to Chapter 9 include the concepts of sending radio-frequency (RF) analog signals at microwave frequencies over optical fibers. An expanding application of these RF-over-fiber techniques is for broadband radio-over-fiber networks for mobile subscribers outdoors and in buildings.

Chapter 10 addresses the principles of wavelength division multiplexing (WDM), examines the functions of a generic WDM link, and discusses international standards for different WDM schemes. New features in this chapter include expanded descriptions and application examples of passive and active WDM devices, such as fiber Bragg gratings, thin-film filters, arrayed waveguide gratings, diffraction gratings, and variable optical attenuators.

Chapter 11 describes different concepts for creating optical amplification. Among the topics are semiconductor optical amplifiers, doped-fiber amplifiers, and a new section on Raman amplification schemes. In addition to discussions of the traditional erbium-doped fiber amplifier (EDFA), new structures such as a thulium-doped fiber amplifier (TDFA) for use in the S-band and a gain-shifted EDFA for the L-band are described.

Next, Chapters 12 through 14 show how the elements are put together to form links and networks, and explain measurement methodologies used to evaluate the performance of lightwave components and links. A new Chapter 12 is devoted to the origins and effects of nonlinear processes in optical fibers. Some of these nonlinear effects degrade system performance and need to be controlled, whereas others, such as stimulated Raman scattering, can have beneficial uses.

Optical networking concepts for long-haul, metro, local area, and access networks are presented in an extensively expanded Chapter 13. Among the new topics are high-speed optical links operating up to 160 Gb/s, the concepts of optical add/drop multiplexing and optical crossconnects, wavelength routing, optical packet switching, optical burst switching, passive optical networks, IP over WDM, optical Ethernet, and mitigation techniques for transmission impairments in high-speed networks.

The final chapter discusses performance measurement and monitoring. The topics include a discussion of internationally recognized measurement standards, basic test instruments for optical fiber link characterization, methods for characterizing optical fibers, and evaluation of link performance through

eye-pattern measurements. Particular emphasis is placed on evaluating WDM links. New features in Chapter 14 concerning eye diagrams include the concepts of eye masks, stressed-eye tests, and bit-error rate eye contours. Another new feature is a discussion of optical performance monitoring. This has become an essential function in optical communication networks, particularly in relation to error monitoring, network maintenance, and fault management.

New to This Edition

The following material is new to this edition of the book:

- Designation of spectral bands used in optical fiber communications
- Descriptions of photonic crystal fibers, which have an internal microstructure that adds another dimension of light control within a fiber
- An overview of optical fiber cable-installation methods used in environments ranging from indoor ducts to undersea links
- Descriptions of specialty fibers that are designed to interact with light in order to control and manipulate optical signals
- A discussion of international standards used to designate the characteristics of various types of optical fibers in order to have industry-wide compatibility between fibers
- An illustrative section on the characteristics and form factors of popular commercial transceiver packages
- An illustrative section on the characteristics and form factors of popular commercial optical fiber connectors
- A discussion on the characteristics of burst-mode optical receivers needed for passive optical networks
- Expanded coverage of power penalties incurred in transmission links
- Expanded coverage of single-mode links operating at 10 Gb/s and higher
- A new section on coherent detection methods that offer higher spectral purity and greater resistance to dispersion effects compared to conventional direct-detection methods
- A new section on digital quadrature phase-shift-keying (DQPSK) methods used for transmission links operating at rates greater than 10 Gb/s
- A new section on digital error detection and correction methods including the use of polynomial codes and forward-error-correction (FEC) techniques
- A new section on radio-over-fiber technologies for wireless access networks, wireless services for indoor environments, and connections to personal area networks in homes
- Expanded coverage of photonic devices used in wavelength division multiplexing (WDM)
- A new section on Raman optical amplification techniques and an expanded discussion on erbium-doped fiber amplifiers (EDFA)
- A new chapter on the impact of nonlinear effects in optical fibers
- A significantly expanded chapter on optical networks that now includes high-speed lightwave links, optical add/drop multiplexing, optical switching, WDM network examples, IP over WDM, optical Ethernet, and passive optical networks for fiber-to-the-premises (FTTP) applications
- A revised chapter on performance measurement and monitoring that includes eye diagram tests, optical performance monitoring (OPM) functions, and performance testing methods; the measurement tests include bit-error rate (BER), optical signal-to-noise ratio (OSNR), the Q factor, optical modulation amplitude (OMA), and timing jitter

Use of the Book

This fourth edition provides the basic material for a senior-level or graduate course in the theory and application of optical fiber communication technology. It also will serve well as a working reference for practicing engineers dealing with the design and development of components, transmission equipment, test instruments, and cable plants for optical fiber communication systems. The background required to study the book is that of typical senior-level engineering students. This includes introductory electromagnetic theory, calculus and elementary differential equations, and basic concepts of optics as presented in a freshman physics course. Concise reviews of several background topics, such as optics concepts, electromagnetic theory, and basic semiconductor physics, are included in the main body of the text. Various sections dealing with advanced material (e.g., the applications of Maxwell's equations to cylindrical dielectric waveguides) are designated by a star and can be skipped over without loss of continuity. To assist readers in learning the material and applying it to practical designs, 143 examples are given throughout the book. A collection of 277 homework problems is included to help test the reader's comprehension of the material covered, and to extend and elucidate the text. Instructors can obtain the problem solutions from the publisher.

Numerous references are provided at the end of each chapter as a start for delving deeper into any given topic. Since optical fiber communications brings together research and development efforts from many different scientific and engineering disciplines, there are hundreds of articles in the literature relating to the material covered in each chapter. Even though not all these articles can be cited in the references, the selections represent some of the major contributions to the fiber optics field and can be considered as a good introduction to the literature. Supplementary material and references for up-to-date developments can be found in specialized textbooks and various conference proceedings.

To help the reader understand and use the material in the book, a table in the inside front cover provides a quick reference for various physical constants and units. Appendices A through D give an overview of the international system of units, listings of mathematical formulas needed for homework problems, and discussions on decibels. Appendices E through G provide listings of acronyms, Roman symbols, and Greek symbols, respectively, that are used in the book.

Computer-based modeling and simulation tools offer a powerful method to assist in analyzing the design of an optical component, circuit, or network before costly prototypes are built. The book website (www.mhhe.com/keiserOFC) describes abbreviated interactive demonstration versions of simulation modules. These modules may be downloaded from the websites of three companies that produce simulation tools. The simplified versions contain over 100 predefined component and link configurations that allow interactive concept demonstrations. Although the configurations are fixed, the user can vary certain parameters or turn them on and off to see the effect on system performance.

Book Website

In addition to providing web links for learning about and downloading interactive simulation demonstrations, the book website (www.mhhe.com/keiserOFC) has further information on new technology developments and updated reference material related to the book. The problem solutions are available to instructors, as well as a selection of lecture slides and figures from the text. These are available for download at the password-protected section of the website for instructors only. Permission for instructors to access the solutions manual can be obtained from the publisher.

COSMOS

(available to Instructors only)
McGraw-Hill's COSMOS (Complete Online Solutions Manual Organization System) allows instructors to streamline the creation of assignments, quizzes, and tests by using problems and solutions from the textbook, as well as their own custom material. COSMOS is now available online at http://cosmos.mhhe.com/

Gerd Keiser

Acknowledgments

edition were all the positive comments on the previous editions received from users and adapters at numerous academic institutions worldwide.

CHAPTER 1

Overview of Optical Fiber Communications

Ever since ancient times, people had a principal need to communicate with one another. This need created interests in devising communication systems for sending messages from one distant place to another. Optical communication methods were of special interest among the many systems that people tried to use. One of the earliest known optical transmission links was a *fire-signal method* used by the Greeks in the eighth century BC for sending alarms, calls for help, or announcements of certain events. Improvements of these optical transmission systems were not pursued very actively due to technological limitations at the time. For example, the speed of sending information over the communication link was limited because the transmission rate depended on how fast the senders could move their hands, the optical signal receiver was the error-prone human eye, line-of-sight transmission paths were required, and atmospheric effects such as fog and rain made the transmission path unreliable. Thus it turned out to be faster, more efficient, and more dependable to send messages by a courier over the road network.

Subsequently, no significant advances for optical communications appeared until the invention of the laser in the early 1960s. Optical frequencies generated by such a coherent optical source are on the order of 5×10^{14} Hz, so the laser has a theoretical information capacity exceeding that of microwave systems by a factor of 10^5. With the potential of such wideband transmission capacities in mind, experiments using atmospheric optical channels were carried out in the early 1960s. These experiments showed the feasibility of modulating a coherent optical carrier wave at very high frequencies. However, the high cost of developing and implementing such systems, together with the limitations imposed on the atmospheric optical channels by rain, fog, snow, and dust, make such extremely high-speed links economically unattractive.

At the same time it was recognized that an optical fiber can provide a more reliable transmission channel because it is not subject to adverse environmental conditions.[1,2] Initially, the extremely large losses of more than 1000 dB/km made optical fibers appear impractical. This changed in 1966 when Kao and Hockman[3] speculated that the high losses were a result of impurities in the fiber material, and that the losses potentially could be reduced significantly in order to make optical fibers a viable transmission medium. In 2009 Charles K. C. Kao was awarded the Nobel Prize in Physics for his pioneering insight and his enthusiastic international follow-ups in promoting the further development of low-loss optical fibers. These efforts led to the first ultrapure fiber being fabricated in 1970, only four years after his

prediction. This breakthrough led to a series of technology developments related to optical fibers. These events finally allowed practical *lightwave communication systems* to start being fielded worldwide in 1978. These systems operate in the near-infrared region of the electromagnetic spectrum (nominally 770 to 1675 nm) and use optical fibers as the transmission medium.

The goal of this book is to describe the various technologies, implementation methodologies, and performance measurement techniques that make optical fiber communication systems possible. The reader can find additional information on the theory of light propagation in fibers, the design of links and networks, and the evolution of optical fibers, photonic devices, and optical fiber communication systems in a variety of books and conference proceedings.[4–23]

This chapter gives an overview of fundamental communications concepts and illustrates how optical fiber transmission systems operate. First, Sec. 1.1 gives the motivations behind the development of optical fiber transmission systems. Next, Sec. 1.2 defines the different spectral bands that describe various operational wavelength regions used in optical communications. Section 1.3 explains decibel notation for expressing optical power levels. Section 1.4 gives the basic hierarchy for multiplexing digitized information streams used on optical links, and Sec. 1.5 describes how wavelength division multiplexing can boost the transmission capacity of an optical fiber significantly. Next, Sec. 1.6 introduces the functions and implementation considerations of the key elements used in optical fiber systems.

An important aspect of realizing a smoothly interacting worldwide lightwave network is to have well-established international standards for all aspects of components and networks. Section 1.7 introduces the organizations involved with this standardization activity and lists the main classes of standards related to optical communication components, system operations, and installation procedures. Finally, Sec. 1.8 gives an introduction to modeling and simulation tools that have been developed to aid in the design of optical fibers, passive and active devices, links, and networks.

Chapters 2 through 10 describe the purpose and performance characteristics of the major elements in an optical link. These elements include optical fibers, light sources, photodetectors, passive optical devices, optical amplifiers, and active optoelectronic devices used in multiple-wavelength networks. Chapters 11 through 14 show how the elements are put together to form links and networks and explain measurement methodologies used to evaluate the performance of lightwave components and links.

1.1 | Motivations for Lightwave Communications

1.1.1 The Path to Optical Networks

Prior to about 1980 most communication technologies involved some type of electrical transmission mechanism. The era of electrical communications started in 1837 with the invention of the telegraph by Samuel F. B. Morse. The telegraph system used the Morse code, which represents letters and numbers by a coded series of dots and dashes. The encoded symbols were conveyed by sending short and long pulses of electricity over a copper wire at a rate of tens of pulses per second. More advanced telegraph schemes, such as the Baudot system invented in 1874, enabled the information speeds to increase to about 120 bits per second (b/s) but required the use of skilled operators. Shortly thereafter in 1876 Alexander Graham Bell developed a fundamentally different, easily useable device that could transmit the entire voice signal in an analog form.

Both the telegraph and the analog voice signals were sent using a baseband transmission mode. *Baseband* refers to the technology in which a signal is transmitted directly over a channel. For example,

this method is used on standard twisted-pair wire links running from an analog telephone to the nearest switching interface equipment. The same baseband method is used widely in optical communications; that is, the optical output from a light source is turned on and off in response to the variations in voltage levels of an information-bearing electrical signal.

In the ensuing years an increasingly larger portion of the electromagnetic spectrum was utilized to develop and deploy progressively more sophisticated and reliable electrical communication systems with larger capacities for conveying information from one place to another. The basic motivations behind each new system application were (a) to improve the transmission fidelity so that fewer distortions or errors occur in the received message, (b) to increase the data rate or capacity of a communication link so that more information can be sent, or (c) to increase the transmission distance between in-line repeater or amplification stations so that messages can be sent farther without the need to restore the signal amplitude or fidelity periodically along its path. These activities led to the birth of a wide variety of communication systems that are based on using high-capacity long-distance terrestrial and undersea copper-based wire lines and wireless radio-frequency (RF), microwave, and satellite links.

In these developments the basic trend for advancing the link capacity was to use increasingly higher channel frequencies. The reason for this trend is that a time-varying baseband information-bearing signal may be transferred over a communication channel by superimposing it onto a sinusoidal electromagnetic wave, which is known as the *carrier wave* or simply *carrier*. At the destination the baseband information signal is removed from the carrier wave and processed as desired. Since the amount of information that can be transmitted is directly related to the frequency range over which the carrier operates, increasing the carrier frequency theoretically increases the available transmission bandwidth and, consequently, provides a larger information capacity.[24–28] For example, Fig. 1.1 shows the electromagnetic spectral bands used for radio transmission. As the diverse radio technologies move from high frequency (HF) to very high frequency (VHF) to ultra high frequency (UHF) bands with nominal carrier frequencies of 10^7, 10^8, and 10^9 Hz, respectively, increasingly higher information transmission speeds can be employed to provide a higher link capacity. Thus the trend in electrical communication system developments was to use progressively higher frequencies, which offer corresponding increases in bandwidth or information capacity.

As Fig. 1.1 also shows, optical frequencies are several orders of magnitude higher than those used by electrical communication systems. Thus the invention of the laser in the early 1960s aroused a curiosity about the possibility of using the optical region of the electromagnetic spectrum for transmitting information. Of particular interest is the near-infrared spectral band ranging from about 770 to 1675 nm, since this is a low-loss region in silica glass fibers. The technical breakthrough for optical fiber communications started in 1970 when researchers at Corning demonstrated the feasibility of producing a glass fiber having an optical power loss that was low enough for a practical transmission link.[29]

As research progressed, it became clear that many complex problems made it extremely difficult to extend the carrier concept for achieving a super broadband optical communication link. Nevertheless, the unique properties of optical fibers gave them a number of performance advantages compared to copper wires, so that optical links operating in a simple on-off keyed baseband mode were attractive applications.

The first installed optical fiber links appeared in the late 1970s and were used for transmitting telephony signals at about 6 Mb/s over distances of around 10 km. As research and development progressed, the sophistication and capabilities of these systems increased rapidly during the 1980s to create links carrying aggregated data rates beyond terabits per second over distances of hundreds of kilometers without the need to restore signal fidelity along the path length.
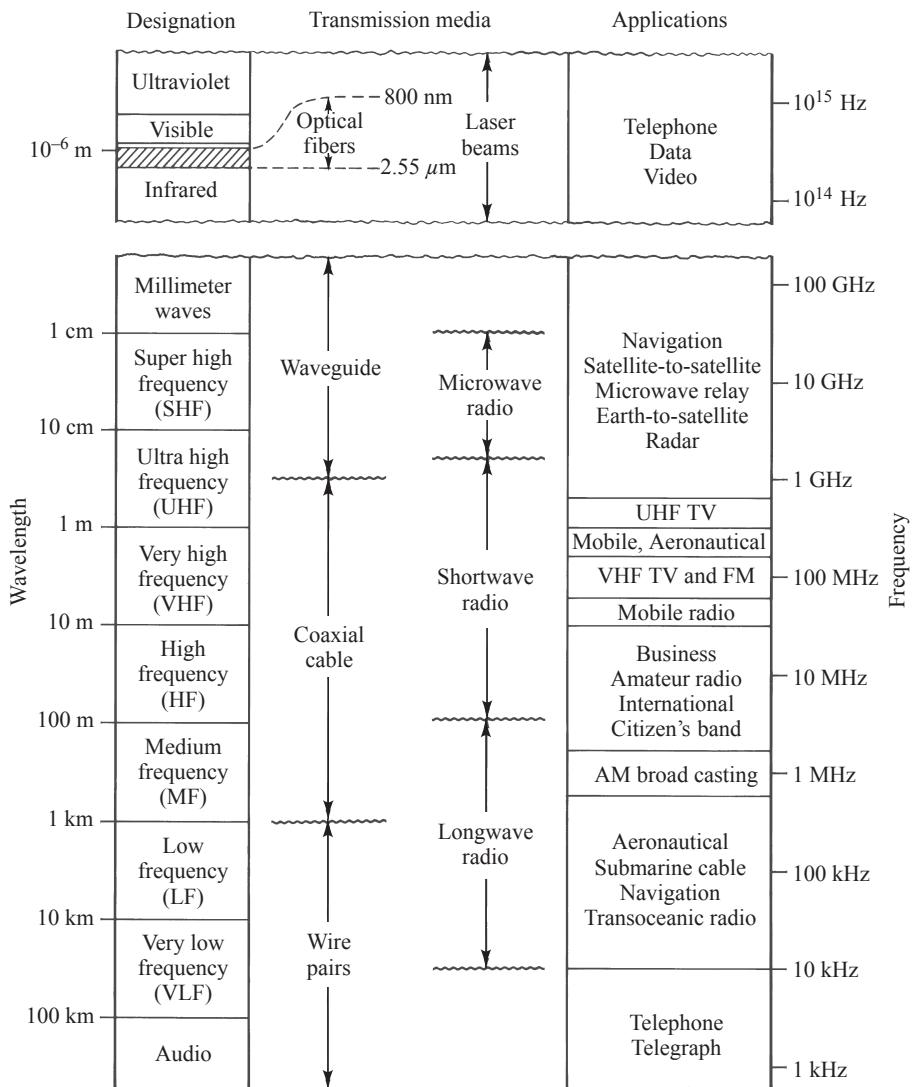
Fig. 1.1 *The regions of the electromagnetic spectrum used for radio and optical fiber communications. (Used with permission from A. B. Carlson,* Communication Systems, © *1986, McGraw-Hill Book Company.)*

Starting in the 1990s there was a burgeoning demand on communication-network assets for bandwidth-hungry services such as database queries, home shopping, high-definition interactive video, remote education, telemedicine and e-health, high-resolution editing of home videos, blogging, and large-scale high-capacity e-science and Grid computing.[30–33] This demand was fueled by the rapid proliferation of personal computers (PCs) coupled with a phenomenal increase in their storage capacity and processing capabilities. Furthermore, the widespread availability and continuous expansion of the Internet, and an extensive choice of remotely accessible programs and information databases, resulted in a dramatic rise

in PC usage. To handle the ever-increasing demand for high-bandwidth services ranging from home-based PC users to large businesses and research organizations, telecommunication companies worldwide greatly enhanced the capacity of fiber lines. This was accomplished by adding more independent signal-carrying wavelengths on individual fibers and increasing the transmission speed of information being carried by each wavelength.

1.1.2 Advantages of Optical Fibers

The advantages of optical fibers compared to copper wires include the following:

Long Distance Transmission   Optical fibers have lower transmission losses compared to copper wires. Consequently data can be sent over longer distances, thereby reducing the number of intermediate repeaters needed to boost and restore signals in long spans. This reduction in equipment and components decreases system cost and complexity.

Large Information Capacity   Optical fibers have wider bandwidths than copper wires, so that more information can be sent over a single physical line. This property decreases the number of physical lines needed for sending a given amount of information.

Small Size and Low Weight   The low weight and the small dimensions of fibers offer a distinct advantage over heavy, bulky wire cables in crowded underground city ducts or in ceiling-mounted cable trays. This feature also is of importance in aircraft, satellites, and ships where small, low-weight cables are advantageous, and in tactical military applications where large amounts of cable must be unreeled and retrieved rapidly.[34]

Immunity to Electrical Interference   An especially important feature of an optical fiber relates to the fact that it is a dielectric material, which means it does not conduct electricity. This makes optical fibers immune to the electromagnetic interference effects seen in copper wires, such as inductive pickup from other adjacent signal-carrying wires or coupling of electrical noise into the line from any type of nearby equipment.

Enhanced Safety   Optical fibers offer a high degree of operational safety because they do not have the problems of ground loops, sparks, and potentially high voltages inherent in copper lines. However, precautions with respect to laser light emissions need to be observed to prevent possible eye damage.

Increased Signal Security   An optical fiber offers a high degree of data security because the optical signal is well-confined within the fiber and an opaque coating around the fiber absorbs any signal emissions. This feature is in contrast to copper wires where electrical signals potentially could be tapped off easily. Thus optical fibers are attractive in applications where information security is important, such as financial, legal, government, and military systems.

1.2 Optical Spectral Bands

1.2.1 Electromagnetic Energy

All telecommunication systems use some form of electromagnetic energy to transmit signals. The *spectrum* of electromagnetic (EM) radiation is shown in Fig. 1.2. *Electromagnetic energy* is a combination of electrical and magnetic fields and includes power, radio waves, microwaves, infrared light, visible
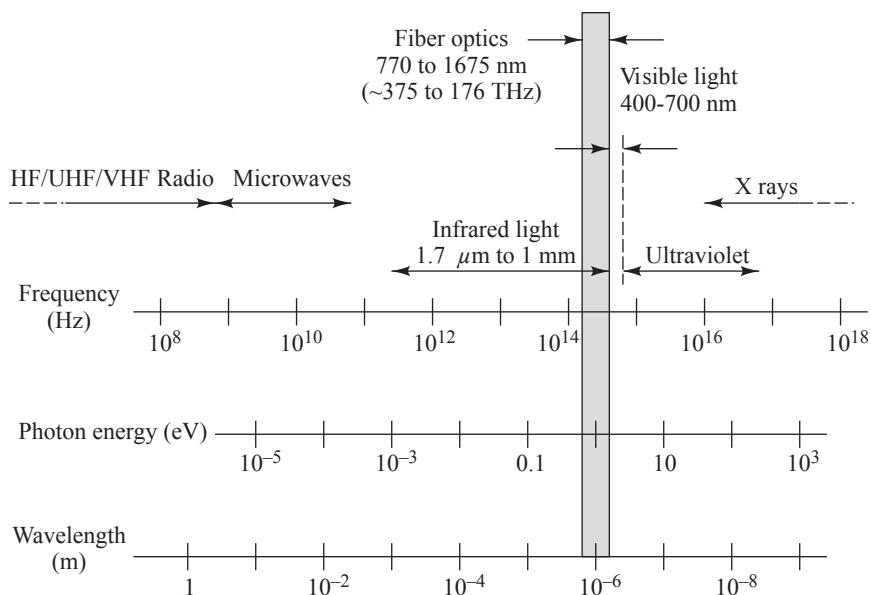
Fig. 1.2 *The spectrum of electromagnetic radiation*

light, ultraviolet light, X rays, and gamma rays. Each discipline takes up a portion (or band) of the electromagnetic spectrum. The fundamental nature of all radiation within this spectrum is that it can be viewed as electromagnetic waves that travel at the speed of light, which is about $c = 3 \times 10^8$ m/s in a vacuum. Note that the speed of light s in a material is smaller by the refractive-index factor n than the speed c in a vacuum, as described in Chapter 2. For example, $n \approx 1.45$ for silica glass, so that the speed of light in this material is about $s = 2 \times 10^8$ m/s.

The physical properties of the waves in different parts of the spectrum can be measured in several interrelated ways. These are the length of one period of the wave, the energy contained in the wave, or the oscillating frequency of the wave. Whereas electrical signal transmission tends to use frequency to designate the signal operating bands, optical communication generally uses *wavelength* to designate the spectral operating region and *photon energy* or *optical power* when discussing topics such as signal strength or electro-optical component performance.

As can be seen from Fig. 1.2, there are three different ways to measure the physical properties of a wave in various regions in the EM spectrum. These measurement units are related by some simple equations. First of all, in a vacuum the speed of light c is equal to the wavelength λ times the frequency v, so that

$$c = \lambda v \tag{1.1}$$

where the frequency v is measured in cycles per second or *hertz* (Hz).

The relationship between the energy of a photon and its frequency (or wavelength) is determined by the equation known as *Planck's Law*

$$E = hv \tag{1.2}$$

where the parameter $h = 6.63 \times 10^{-34}$ J-s $= 4.14 \times 10^{-15}$ eV-s is *Planck's constant*. The unit J means *joules* and the unit eV stands for *electron volts*. In terms of wavelength (measured in units of μm), the energy

| O-Band | E-Band | S-Band | C-Band | L-Band | U-Band |

1260   1360   1460   1530 1565   1625 1675

Wavelength (nm)
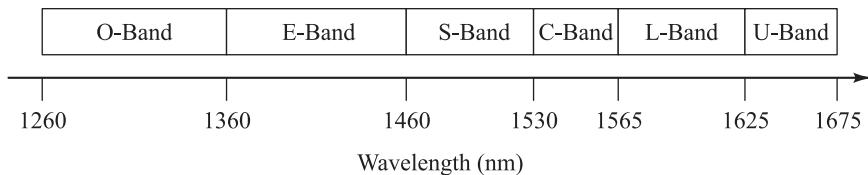
Fig. 1.3 *Designations of spectral bands used for optical fiber communications*

in electron volts is given by

$$E(\text{eV}) = \frac{1.2406}{\lambda(\mu\text{m})} \tag{1.3}$$

Figure 1.2 shows the optical spectrum ranges from about 5 nm in the ultraviolet region to 1 mm for far-infrared radiation. In between these limits is the 400-to-700-nm *visible band*. Optical fiber communication uses the *near-infrared spectral band* ranging from nominally 770 to 1675 nm.

The International Telecommunications Union (ITU) has designated six spectral bands for use in optical fiber communications within the 1260-to-1675-nm region.[35] These *long-wavelength band* designations arose from the attenuation characteristics of optical fibers and the performance behavior of an erbium-doped fiber amplifier (EDFA), as described in Chapters 3 and 10, respectively. Figure 1.3 shows and Table 1.1 defines the regions, which are known by the letters O, E, S, C, L, and U.

The 770-to-910 nm band is used for shorter-wavelength multimode fiber systems. Thus this region is designated as the *short-wavelength* or *multimode fiber band*. Later chapters describe the operational performance characteristics and applications of optical fibers, electro-optic components, and other passive optical devices for use in the short- and long-wavelength bands.

Example 1.1 Show that photon energies decrease with increasing wavelength. Use wavelengths at 850, 1310, and 1550 nm.

Solution: From Eq. (1.3) we find $E(850\,nm) = 1.46$ eV, $E(1310\,nm) = 0.95$ eV, and $E(1550\,nm) = 0.80$ eV.

Table 1.1 *Spectral band designations used in optical fiber communications*

Name	Designation	Spectrum (nm)	Origin of name
Original band	O-band	1260 to 1360	Original (first) region used for single-mode fiber links
Extended band	E-band	1360 to 1460	Link use can extend into this region for fibers with low water content
Short band	S-band	1460 to 1530	Wavelengths are shorter than the C-band but higher than the E-band
Conventional band	C-band	1530 to 1565	Wavelength region used by a conventional EDFA
Long band	L-band	1565 to 1625	Gain of an EDFA decreases steadily to 1 at 1625 nm in this longer wavelength band
Ultra-long band	U-band	1625 to 1675	Region beyond the response capability of an EDFA

1.2.2   Windows and Spectral Bands

Figure 1.4 shows the operating range of optical fiber systems and the characteristics of the four key components of a link: the optical fiber, light sources, photodetectors, and optical amplifiers. Here the dashed vertical lines indicate the centers of the three main traditional operating wavelength bands of optical fiber systems, which are the short-wavelength region, the O-band, and the C-band. One of the principal characteristics of an optical fiber is its attenuation as a function of wavelength, as shown at the top in Fig. 1.4. Early applications in the late 1970s made exclusive use of the 770-to-910-nm wavelength band where there was a low-loss window and GaAlAs optical sources and silicon photodetectors operating at these wavelengths were available. Originally this region was referred to as the *first window* because around 1000 nm there was a large attenuation spike due to absorption by water molecules. As a result of this spike, early fibers exhibited a local minimum in the attenuation curve around 850 nm.
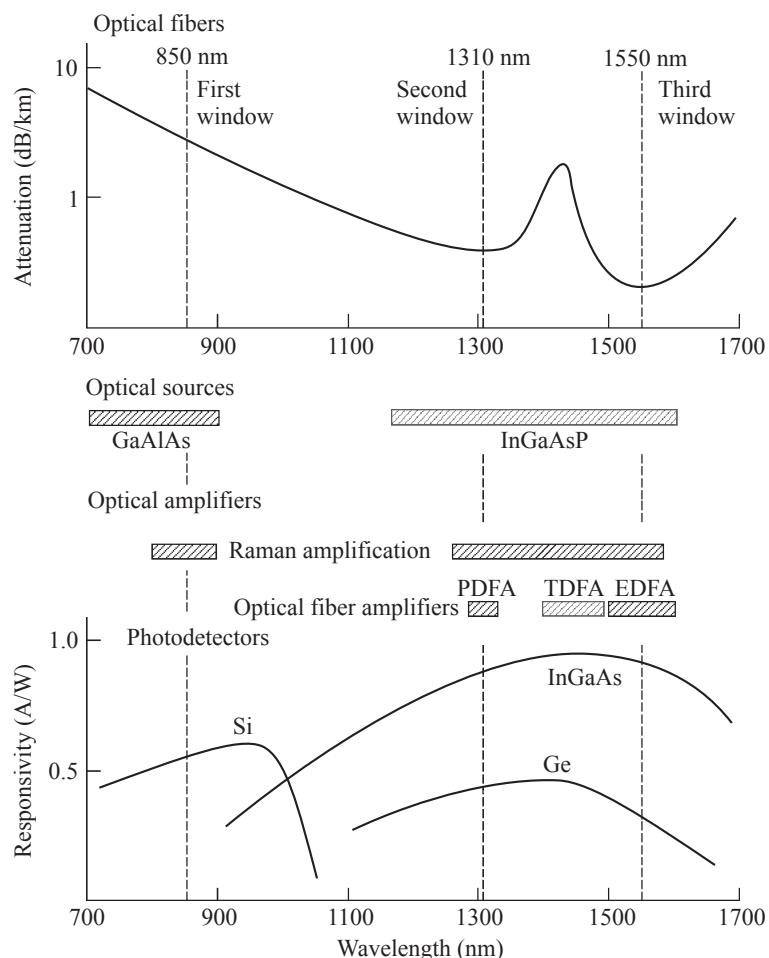
Fig. 1.4   *Characteristics and operating ranges of the four key optical fiber link components*

By reducing the concentration of hydroxyl ions and metallic impurities in the fiber material, in the 1980s manufacturers could fabricate optical fibers with very low losses in the 1260-to-1675-nm region. This spectral band is called the *long-wavelength region*. Because the glass still contained some water molecules, a third-order absorption spike remained around 1400 nm. This spike defined two low-loss windows, these being the *second window* centered at 1310 nm and the *third window* centered at 1550 nm. These two windows now are called the O-band and C-band, respectively. The desire to use the low-loss long-wavelength regions prompted the development of InGaAsP-based light sources and InGaAs photodetectors that can operate at 1310 and 1550 nm. In addition, doping optical fibers with rare-earth elements such as Pr, Th, and Er creates optical fiber amplifiers (called PDFA, TDFA, and EDFA devices, respectively). These devices and the use of Raman amplification gave a further capacity boost to high-capacity long-wavelength systems.

Special material-purification processes can eliminate almost all water molecules from the glass fiber material, thereby dramatically reducing the water-attenuation peak around 1400 nm. This process opens the E-band (1360-to-1460 nm) transmission region to provide around 100 nm more spectral bandwidth in these specially fabricated fibers than in conventional single-mode fibers.

Systems operating at 1550 nm provide the lowest attenuation, but the signal dispersion in a standard silica fiber is larger at 1550 nm than at 1310 nm. Fiber manufacturers overcame this limitation first by creating dispersion-shifted fibers for single-wavelength operation and then by devising non-zero dispersion-shifted fiber (NZDSF) for use with multiple-wavelength implementations. The latter fiber type has led to the widespread use of multiple-wavelength S-band and C-band systems for high-capacity, long-span terrestrial and undersea transmission links. These links routinely carry traffic at 10 Gb/s over nominally 90-km distances between amplifiers or repeaters. By 2005 links operating at 40 Gb/s were being installed and field trials of 160-Gb/s long-distance transmission systems were tested successfully.[36–39]

1.3 Decibel Units

As described in all of the following chapters of this book, a critical consideration when designing and implementing an optical fiber link is to establish, measure, and/or interrelate the optical signal levels at each of the elements of a transmission link. Thus it is necessary to know parameter values such as the optical output power from a light source, the power level needed at the receiver to properly detect a signal, and the amount of optical power lost at the constituent elements of the transmission link.

Reduction or attenuation of signal strength arises from various loss mechanisms in a transmission medium. For example, electric power is lost through heat generation as an electric signal flows along a wire, and optical power is attenuated through scattering and absorption processes in a glass or plastic fiber or in an atmospheric channel. To compensate for these energy losses, amplifiers are used periodically along a channel path to boost the signal level, as shown in Fig. 1.5.

A standard and convenient method for measuring attenuation through a link or a device is to reference the output signal level to the input level. For guided media such as an optical fiber, the signal strength normally decays exponentially. Thus for convenience one can designate it in terms of a logarithmic power ratio measured in *decibels* (dB). The dB unit is defined by

$$\text{Power ratio in dB} = 10 \log \frac{P_2}{P_1} \tag{1.4}$$

where P_1 and P_2 are the electrical or optical power levels of a signal at points 1 and 2 in Fig. 1.6, and *log* is the base-10 logarithm. The logarithmic nature of the decibel allows a large ratio to be expressed
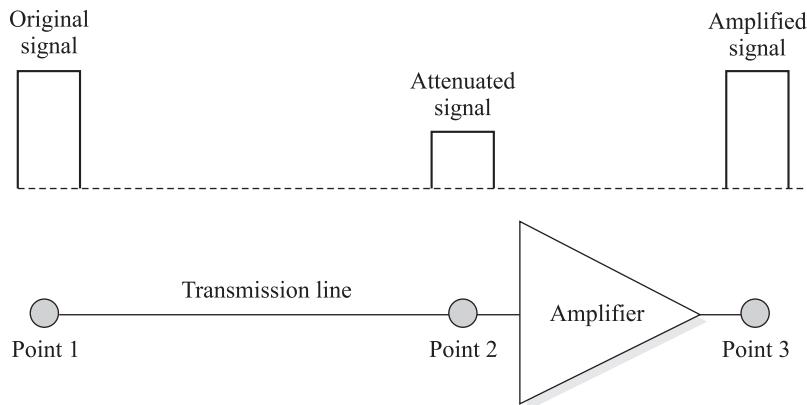
Fig. 1.5   *Periodically placed amplifiers compensate for energy losses along a link*
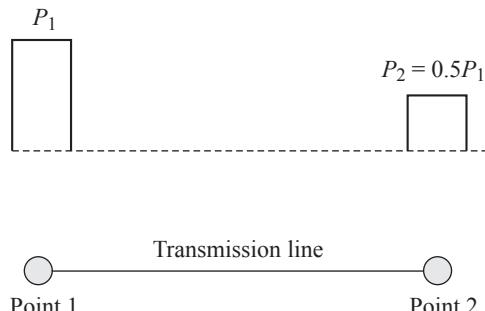
Fig. 1.6   *Example of pulse attenuation in a link. P_1 and P_2 are the power levels of a signal at points 1 and 2*

in a fairly simple manner. Power levels differing by many orders of magnitude can be compared easily when they are in decibel form. Another attractive feature of the decibel is that to measure the changes in the strength of a signal, one merely adds or subtracts the decibel numbers between two different points.

Table 1.2 shows some sample values of power loss given in decibels and the percent of power remaining after this loss. These types of numbers are important when considering factors such as the effects of tapping off a small part of an optical signal for monitoring purposes, for examining the power loss through some optical element, or when calculating the signal attenuation in a specific length of optical fiber.

Table 1.2   *Representative values of decibel power loss and the remaining percentages*

Power loss (in dB)	Percent of power left
0.1	98
0.5	89
1	79
2	63
3	50
6	25
10	10
20	1

Example 1.2   Assume that after traveling a certain distance in some transmission medium, the power of a signal is reduced to half, that is, $P_2 = 0.5 \, P_1$ in Fig. 1.6. At this point, using Eq. (1.4) the attenuation or loss of power is

$$10 \log \frac{P_2}{P_1} = 10 \log \frac{0.5 \, P_1}{P_1} = 10 \log 0.5 = 10(-0.3) = -3 \text{ dB}$$

Thus, –3 dB (or a 3-dB attenuation or loss) means that the signal has lost half its power. If an amplifier is inserted into the link at this point to boost the signal back to its original level, then that amplifier has a 3-dB gain. If the amplifier has a 6-dB gain, then it boosts the signal power level to twice the original value.

Example 1.3   Consider the transmission path from point 1 to point 4 shown in Fig. 1.7. Here the signal is attenuated by 9 dB between points 1 and 2. After getting a 14-dB boost from an amplifier at point 3, it is again attenuated by 3 dB between points 3 and 4. Relative to point 1, the signal level in dB at point 4 is

dB level at point 4 = (loss in line 1) + (amplifier gain)
+ (loss in line 2)
= (–9 dB) + (14 dB) + (–3 dB) = +2 dB

Thus the signal has a 2-dB (a factor of $10^{0.2} = 1.58$) gain in power in going from point 1 to point 4.
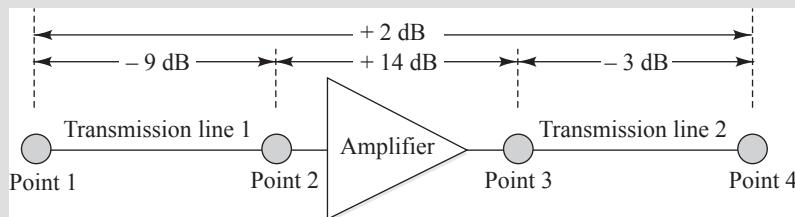
Fig. 1.7   *Example of signal attenuation and amplification in a transmission path*

Since the decibel is used to refer to ratios or relative units, it gives no indication of the absolute power level. However, a derived unit can be used for this purpose. Such a unit that is particularly common in optical fiber communications is the dBm (simply pronounced *dee bee em*). This unit expresses the power level P as a logarithmic ratio of P referred to 1 mW. In this case, the power in dBm is an absolute value defined by

$$\text{Power level (in dBm)} = 10 \log \frac{P \, (\text{in mW})}{1 \text{ mW}} \qquad (1.5)$$

An important rule-of-thumb relationship to remember for optical fiber communications is 0 dBm = 1 mW. Therefore, positive values of dBm are greater than 1 mW and negative values are less than 1 mW.

Table 1.3 lists some examples of optical power levels and their dBm equivalents.

Table 1.3   *Examples of optical power levels and their dBm equivalents*

Power	dBm equivalent
200 mW	23
100 mW	20
10 mW	10
1 mW	0
100 μW	–10
10 μW	–20
1 μW	–30
100 nW	–40
10 nW	–50
1 nW	–60
100 pW	–70
10 pW	–80
1 pW	–90

Example 1.4 Consider three different light sources having the following optical output powers: 50 μW, 1 mW, and 50 mW. What are the power levels in dBm units?

Solution: Using Eq. (1.5) to express the light levels in dBm units, we find that the output powers of these sources are –13 dBm, 0 dBm, and +17 dBm, respectively.

Example 1.5 Consider a product data sheet for a photodetector that states that an optical power level of –32 dBm is needed at the photodetector to satisfy a specific performance requirement. What is the power level in nW (nanowatt) units?

Solution: From Eq. (1.5) we find that a –32-dBm level corresponds to a power in nW of

$$P = 10^{-32/10} \text{ mW} = 0.631 \ \mu\text{W} = 631 \text{ nW}$$

1.4 Network Information Rates

To handle the continuously rising demand for high-bandwidth services from users ranging from individuals to large businesses and research organizations, telecommunication companies worldwide are implementing increasingly sophisticated digital multiplexing techniques that allow a larger number of independent information streams to share the same physical transmission channel simultaneously. This section describes some common digital signal multiplexing techniques.

1.4.1 Telecom Signal Multiplexing

Table 1.4 gives examples of information rates for some typical telecom services. To send these services from one user to another, network providers combine the signals from many different users and send the aggregate signal over a single transmission line. This scheme is known as *time-division-multiplexing* (TDM) wherein N independent information streams, each running at a data rate of R b/s, are interleaved electrically into a single information stream operating at a higher rate of $N \times R$ b/s. To get a detailed perspective of this methodology, let us look at the multiplexing schemes used in telecommunications.

Early applications of fiber optic transmission links were mainly for large capacity telephone lines. These digital links consisted of time-division-multiplexed 64-kb/s voice channels. The multiplexing was developed in the 1960s and is based on what is known as the *plesiochronous digital hierarchy* (PDH). Figure 1.8 shows the digital transmission hierarchy used in the North American telephone network.

Table 1.4 *Examples of information rates for some typical services*

Type of service	Data rate
Video on demand/interactive TV	1.5 to 6 Mb/s
Video games	1 to 2 Mb/s
Remote education	1.5 to 3 Mb/s
Electronic shopping	1.5 to 6 Mb/s
Data transfer or telecommuting	1 to 3 Mb/s
Video conferencing	0.384 to 2 Mb/s
Voice (single telephone channel)	33.6 to 56 kb/s
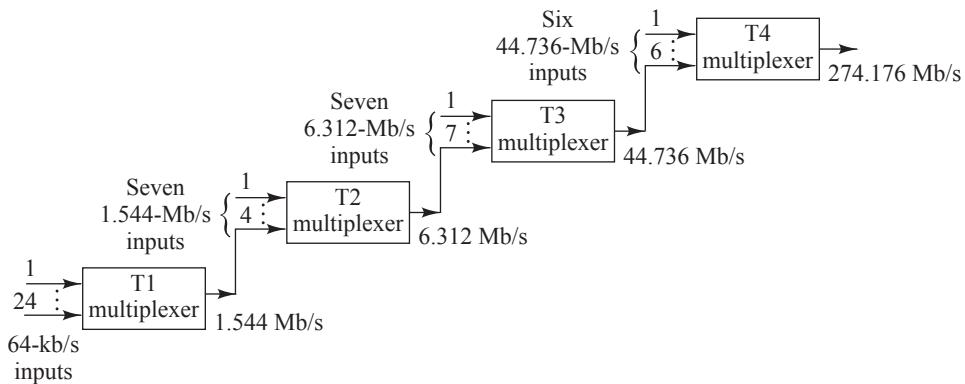
Fig. 1.8 *Digital transmission hierarchy used in the North American telephone network*

Example 1.6 As can be seen from Fig. 1.8, at each multiplexing level some overhead bits are added for synchronization purposes. What is the overhead for $T1$?

Solution: At $T1$ the overhead is

$$1544 \text{ kb/s} - 24 \times 64 \text{ kb/s} = 8 \text{ kb/s}$$

The fundamental building block is a 1.544-Mb/s transmission rate known as a *DS*1 rate, where *DS* stands for *digital system*. It is formed by time-division-multiplexing twenty-four voice channels, each digitized at a 64-kb/s rate (which is referred to as *DS*0). *Framing bits*, which indicate where an information unit starts and ends, are added along with these voice channels to yield the 1.544-Mb/s bit stream. Framing and other control bits that may get added to an information unit in a digital stream are called *overhead* bits. At any multiplexing level a signal at the designated input rate is combined with other input signals at the same rate.

DSx* versus *Tx In describing telephone network data rates, one sees terms such as *T*1, *T*3, and so on. Often the terms *Tx* and *DSx* (e.g., *T*1 and *DS*1 or *T*3 and *DS*3) are used interchangeably. However, there is a subtle difference in their meaning. Designations such as *DS*1, *DS*2, and *DS*3 refer to a *service type*; for example, a user who wants to send information at a 1.544 Mb/s rate would subscribe to a *DS*1 service. Abbreviations such as *T*1, *T*2, and *T*3 refer to the data rate the transmission-line technology uses to deliver that service over a physical link. For example, the *DS*1 service is transported over a physical wire or optical fiber using electrical or optical pulses sent at a *T*1 = 1.544 Mb/s rate.

The TDM scheme is not restricted to multiplexing voice signals. For example, at the *DS*1 level, any 64-kb/s digital signal of the appropriate format could be transmitted as one of the 24 input channels shown in Fig. 1.8. As noted there and in Table 1.5, the main multiplexed rates for North American applications are designated as *DS*1 (1.544 Mb/s), *DS*2 (6.312 Mb/s), and *DS*3 (44.736 Mb/s). European and Japanese networks define similar hierarchies using different bit-rate levels as shown in Table 1.5. In Europe the multiplexing hierarchy is labeled *E*1, *E*2, *E*3, and so on.

1.4.2 SONET/SDH Multiplexing Hierarchy

With the advent of high-capacity fiber optic transmission lines in the 1980s, service providers established a standard signal format called *synchronous optical network* (SONET) in North America and *synchronous*

Table 1.5 *Digital multiplexing levels used in North America, Europe, and Japan*

Digital multiplexing level	Number of 64-kb/s channels	Bit rate (Mb/s)		
		North America	Europe	Japan
DS0	1	0.064	0.064	0.064
DS1	24	1.544		1.544
	30		2.048	
	48	3.152		3.152
DS2	96	6.312		6.312
	120		8.448	
DS3	480		34.368	32.064
	672	44.736		
	1344	91.053		
	1440			97.728
DS4	1920		139.264	
	4032	274.176		
	5760			397.200

digital hierarchy (SDH) in other parts of the world.[40–42] These standards define a synchronous frame structure for sending multiplexed digital traffic over optical fiber trunk lines. The basic building block and first level of the SONET signal hierarchy is called the *Synchronous Transport Signal - Level 1* (STS-1), which has a bit rate of 51.84 Mb/s. Higher-rate SONET signals are obtained by byte-interleaving N of these STS-1 frames, which then are scrambled and converted to an *Optical Carrier - Level N* (OC-N) signal. Thus the OC-N signal will have a line rate exactly N times that of an OC-1 signal. For SDH systems the fundamental building block is the 155.52-Mb/s *Synchronous Transport Module - Level 1* (STM-1). Again, higher-rate information streams are generated by synchronously multiplexing N different STM-1 signals to form the STM-N signal. Table 1.6 shows commonly used SDH and SONET signal levels, the line rate, and the popular numerical name for that rate.

Table 1.6 *Common SDH and SONET line rates and their popular numerical name*

SONET level	Electrical level	SDH level	Line rate (Mb/s)	Popular rate name
OC-1	STS-1	—	51.84	—
OC-3	STS-3	STM-1	155.52	155 Mb/s
OC-12	STS-12	STM-4	622.08	622 Mb/s
OC-48	STS-48	STM-16	2488.32	2.5 Gb/s
OC-192	STS-192	STM-64	9953.28	10 Gb/s
OC-768	STS-768	STM-256	39813.12	40 Gb/s

1.5 WDM Concepts

The use of *wavelength division multiplexing* (WDM) offers a further boost in fiber transmission capacity. The basis of WDM is to use multiple sources operating at slightly different wavelengths to transmit several independent information streams simultaneously over the same fiber. Figure 1.9 shows the basic WDM concept. Here N independent optically formatted information streams, each transmitted at a different wavelength, are combined by means of an optical multiplexer

Fig. 1.9 *Basic concept of wavelength division multiplexing*

and sent over the same fiber. Note that each of these streams could be at a different data rate. Each information stream maintains its individual data rate after being multiplexed with the other traffic streams, and still operates at its unique wavelength. Conceptually, the WDM scheme is the same as frequency division multiplexing (FDM) used in microwave radio and satellite systems.

Although researchers started looking at WDM techniques in the 1970s, during the ensuing years it generally turned out to be easier to transmit only a single wavelength on a fiber using high-speed electronic and optical devices, than to invoke the greater system complexity called for in WDM. However, a dramatic surge in WDM popularity started in the early 1990s owing to several factors. These include new fiber types that provide better performance of multiple-wavelength operation at 1550 nm, advances in producing WDM devices that can separate closely spaced wavelengths, and the development of optical amplifiers that can boost C-band optical signal levels completely in the optical domain.

1.6 Key Elements of Optical Fiber Systems

Similar to electrical communication systems, the basic function of an optical fiber link is to transport a signal from communication equipment (e.g., a computer, telephone, or video device) at one location to corresponding equipment at another location with a high degree of reliability and accuracy. Figure 1.10 shows the main constituents of an optical fiber communications link. The key sections are
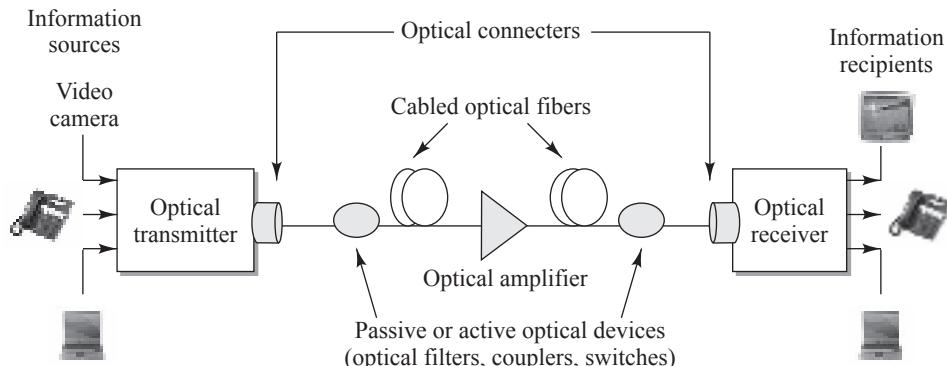
Fig. 1.10 *Main constituents of an optical fiber communications link*

a transmitter consisting of a light source and its associated drive circuitry, a cable offering mechanical and environmental protection to the optical fibers contained inside, and a receiver consisting of a photodetector plus amplification and signal-restoring circuitry. Additional components include optical amplifiers, connectors, splices, couplers, regenerators (for restoring the signal-shape characteristics), and other passive components and active photonic devices.
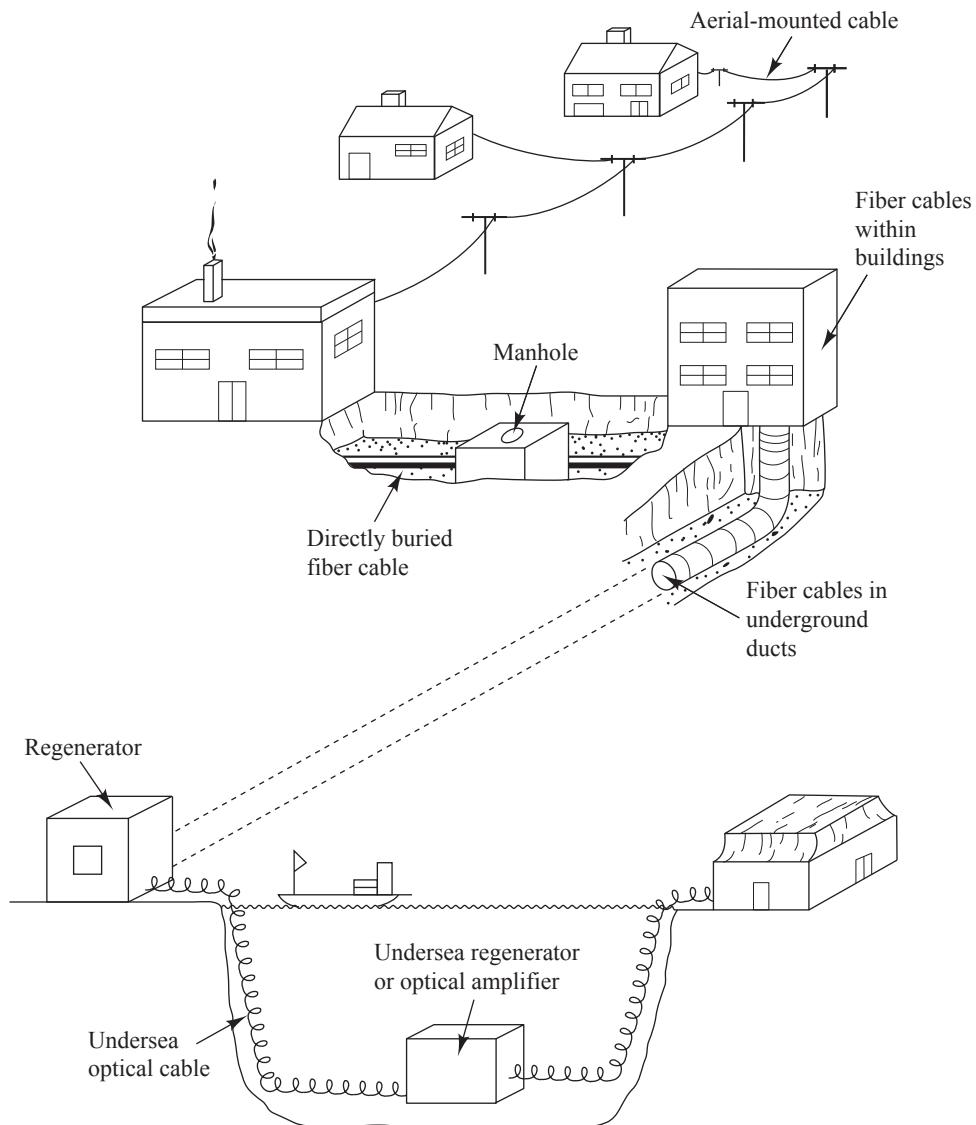
Fig. 1.11   *Optical fiber cables can be installed on poles, in ducts, and underwater, or they can be buried directly in the ground.*

The cabled fiber is one of the most important elements in an optical fiber link as is described in Chapters 2 and 3. In addition to protecting the glass fibers during installation and service, the cable may contain copper wires for powering optical amplifiers or signal regenerators, which are needed periodically in long-distance links for amplifying and reshaping the signal. A variety of fiber types with different performance characteristics exist for a wide range of applications. To protect the glass fibers during installation and service, there are many different cable configurations depending on whether the cable is to be installed inside a building, underground in ducts or through direct-burial methods, outside on poles, or under water. Very low-loss optical connectors and splices are needed in all categories of optical fiber networks for joining cables and for attaching one fiber to another.

Analogous to copper cables, the installation of optical fiber cables can be either aerial, in ducts, undersea, or buried directly in the ground, as illustrated in Fig. 1.11. As described in Chapter 2, the cable structure will vary greatly depending on the specific application and the environment in which it will be installed. Owing to installation and/or manufacturing limitations, individual cable lengths for in-building or terrestrial applications will range from several hundred meters to several kilometers. Practical considerations such as reel size and cable weight determine the actual length of a single cable section. The shorter segments tend to be used when the cables are pulled through ducts. Longer lengths are used in aerial, direct-burial, or underwater applications.

Workers can install optical fiber cables by pulling or blowing them through ducts (both indoor and outdoor) or other spaces, laying them in a trench outside, plowing them directly into the ground, suspending them on poles, or laying or plowing them underwater. Although each method has its own special handling procedures, they all need to adhere to a common set of precautions. These include avoiding sharp bends of the cable, minimizing stresses on the installed cable, periodically allowing extra cable slack along the cable route for unexpected repairs, and avoiding excessive pulling or hard yanks on the cable. For *direct-burial installations* a fiber optic cable can be plowed directly underground or placed in a trench that is filled in later. Figure 1.12 illustrates a plowing operation that may be carried out in nonurban areas. The cables are mounted on large reels on the plowing vehicle and are fed directly into the ground by means of the plow mechanism.

Fig. 1.12 *Plowing operation for direct burial of optical fiber cables. (Photo © VERMEER CORPORATION. All Rights Reserved; www.vermeer.com.)*

Transoceanic cable lengths can be many thousands of kilometers long and include periodically spaced (on the order of 80–120 km) optical repeaters to boost the signal level. The cables are assembled in onshore factories and then are loaded into special cable-laying ships, as illustrated in Fig. 1.13. Splicing together individual cable sections forms continuous transmission lines for these long-distance links.

Once the cable is installed, a transmitter can be used to launch a light signal into the fiber. The transmitter consists of a light source that is dimensionally compatible with the fiber core and associated electronic control and modulation circuitry. Semiconductor light-emitting diodes (LEDs) and laser diodes are suitable for this purpose. For these devices the light output can be modulated rapidly by simply varying the input current at the desired transmission rate, thereby producing an optical signal. The electric input signals to the transmitter circuitry for driving the optical source can be either of an analog or digital form. The

Fig. 1.13 *Ship used to lay optical fiber cables across a sea or an ocean. (Photo courtesy of TE SubCom: www.SubCom.com.)*

functions of the associated transmitter electronics are to set and stabilize the source operating point and output power level. For high-rate systems (usually greater than about 2.5 Gb/s), direct modulation of the source can lead to unacceptable optical signal distortion. In this case, an external modulator is used to vary the amplitude of a continuous light output from a laser diode source. In the 770-to-910-nm region the light sources are generally alloys of GaAlAs. At longer wavelengths (1260 to 1675 nm) an InGaAsP alloy is the principal optical source material.

After an optical signal is launched into a fiber, it will become progressively attenuated and distorted with increasing distance because of scattering, absorption, and dispersion mechanisms in the glass material. At the destination of an optical fiber transmission line, there is a receiving device that interprets the information contained in the optical signal. Inside the receiver is a photodiode that detects the weakened and distorted optical signal emerging from the end of an optical fiber and converts it to an electrical signal (referred to as a *photocurrent*). The receiver also contains electronic amplification devices and circuitry to restore signal fidelity. Silicon photodiodes are used in the 770-to-910-nm region. The primary material in the 1260-to-1675-nm region is an InGaAs alloy.

The design of an optical receiver involves the requirement to correctly interpret the content of the weakened and degraded signal received by the photodetector. The principal figure of merit for a receiver is the minimum optical power necessary at the desired data rate to attain either a given error probability for digital systems or a specified signal-to-noise ratio for an analog system. The ability of a receiver to achieve a certain performance level depends on the photodetector type, the effects of noise in the system, and the characteristics of the successive amplification stages in the receiver.

Included in any optical fiber link are various passive optical devices that assist in controlling and guiding the light signals. Passive devices are optical components that require no electronic control for their operation. Among these are optical filters that select only a narrow spectrum of desired light, optical splitters that divide the power in an optical signal into a number of different branches, optical multiplexers that combine signals from two or more distinct wavelengths onto the same fiber (or that separate the wavelengths at the receiving end) in multiple-wavelength optical fiber networks, and couplers used to tap off a certain percentage of light, usually for performance monitoring purposes.

In addition, modern sophisticated optical fiber networks contain a wide range of active optical components, which require an electronic control for their operation. These include light signal modulators, tunable (wavelength-selectable) optical filters, reconfigurable elements for adding and dropping wavelengths at intermediate nodes, variable optical attenuators, and optical switches.

After an optical signal has traveled a certain distance along a fiber, it becomes greatly weakened due to power loss along the fiber. Therefore, when setting up an optical link, engineers formulate a power budget and add amplifiers or repeaters when the path loss exceeds the available power margin. The periodically placed amplifiers merely give the optical signal a power boost, whereas a repeater also will attempt to restore the signal to its original shape. Prior to 1990, only repeaters were available for signal amplification. For an incoming optical signal, a repeater performs photon-to-electron conversion, electrical amplification, retiming, pulse shaping, and then electron-to-photon conversion. This process can be fairly complex for high-speed multiwavelength systems. Thus researchers expended a great deal of effort to develop all-optical amplifiers, which boost the light power level completely in the optical domain. Optical amplification mechanisms for WDM links include the use of devices based on rare-earth-doped lengths of fiber and distributed amplification by means of a stimulated Raman scattering effect.

The installation and operation of an optical fiber communication system require measurement techniques for verifying that the specified performance characteristics of the constituent components are satisfied. In addition to measuring optical fiber parameters, system engineers are interested in knowing the characteristics of passive splitters, connectors, and couplers, and electro-optic components, such as sources, photodetectors, and optical amplifiers. Furthermore, when a link is being installed and tested, operational parameters that should be measured include bit error rate, timing jitter, and signal-to-noise ratio as indicated by the eye pattern. During actual operation, measurements are needed for maintenance and monitoring functions to determine factors such as fault locations in fibers and the status of remotely located optical amplifiers.

1.7 | Standards for Optical Fiber Communications

To allow components and equipment from different vendors to interface with one another, numerous international standards have been developed.[43–45] The three basic classes for fiber optics are primary standards, component testing standards, and system standards.

Primary standards refer to measuring and characterizing fundamental physical parameters such as attenuation, bandwidth, operational characteristics of fibers, and optical power levels and spectral widths. In the USA the main organization involved in primary standards is the National Institute of Standards and Technology (NIST). This organization carries out fiber optic and laser standardization work, and it sponsors an annual conference on optical fiber measurements. Other national organizations include the National Physical Laboratory (NPL) in the United Kingdom and the Physikalisch-Technische Bundesanstalt (PTB) in Germany.

Component testing standards define tests for fiber-optic component performance and establish equipment-calibration procedures. Several different organizations are involved in formulating testing standards, some very active ones being the Telecommunications Industry Association (TIA) in association with the Electronics Industries Alliance (EIA), the Telecommunication Sector of the International Telecommunication Union (ITU-T), and the International Electrotechnical Commission (IEC). The TIA has a list of over 120 fiber optic test standards and specifications under the general designation TIA/EIA-455-XX-YY, where XX refers to a specific measurement technique and YY refers to the publication year. These standards are also called *Fiber Optic Test Procedures* (FOTP), so that TIA/EIA-455-XX becomes FOTP-XX. These include a wide variety of recommended methods for testing the response of fibers, cables, passive devices, and electro-optic components to environmental factors and operational conditions. For example, TIA/EIA-455-60-1997, or FOTP-60, is a method published in 1997 for measuring fiber or cable length.

System standards refer to measurement methods for links and networks. The major organizations are the American National Standards Institute (ANSI), the Institute for Electrical and Electronic Engineers (IEEE), the ITU-T, and Telcordia Technologies. Of particular interest for fiber optics system are test standards and recommendations from the ITU-T. Within the G series (in the number range G.650 and higher) the recommendations relate to fiber cables, optical amplifiers, wavelength multiplexing, optical transport networks (OTN), system reliability and availability, and management and control for passive optical networks (PON). The L and O series of the ITU-T address methods and equipment for the construction, installation, maintenance support, monitoring, and testing of cable and other elements in the optical fiber outside plant, that is, the fielded cable system. Telcordia Technologies provides a wide range of generic requirements for telecommunication network components and systems. For example, the GR-3120 *Generic Requirements for Hardened Fiber Optic Connectors* document describes the necessary specifications for optical connectors that are environmentally hardened and can be mated in the field.

1.8   Modeling and Simulation Tools

Computer-based simulation and modeling tools that integrate component, link, and network functions can make the design process of complex optical links and networks more efficient, less expensive, and faster.[46–51] The rapid proliferation and increase in capabilities of personal computers led to the development of many sophisticated simulation programs for these machines for predicting photonic component, link, and network performance behavior. These software tools are based on well-established numerical models and can simulate factors such as connector losses due to geometric or position mismatches of fibers, efficiencies of coupling optical power from light sources into fibers, behaviors of passive and active optical components, and the performance of complex optical networks. They also can model many types of passive and active devices, such as waveguide couplers, optical filters, waveguide grating arrays, and optical sources, to a high degree of sophistication.

1.8.1   Simulation Tool Characteristics

Computer-aided design (CAD) tools can offer a powerful method to assist in analyzing the design of an optical component, circuit, or network before costly prototypes are built. Important points to consider, however, are the approximations and modeling assumptions made in the software design. Since most telecommunication systems are designed with several decibels of safety margin, approximations for calculating operating behavior that are reasonably accurate are not only acceptable but in general necessary to allow tractable computation times.

The theoretical models used in computer simulations nominally include the following characteristics:
- Enough detail so that all factors that could influence the performance of the component, circuit, or network can be appropriately evaluated.
- A common set of parameters so that simulated devices can be interconnected with each other to form circuits or networks.
- Interfaces that pass sufficient information between the constituent components so that all possible interactions are identified.
- Computational efficiency that allows a trade-off between accuracy and speed, so that quick estimates of system performance can be made in the early stages of a design.
- The capability to simulate devices over the desired spectral bandwidth.
- The ability to simulate factors such as nonlinear effects, crosstalk between optical channels, distortion in lasers, and dispersion in optical fibers.

To enable a user to visualize and simulate a system quickly, the simulation programs normally have the following features:
- The ability to create a system schematic based on a library of graphical icons and a graphical user interface (GUI). The icons represent various system components (such as optical fibers, filters, amplifiers) and instrumentation (e.g., data sources, power meters, and spectrum analyzers).
- The ability for the user to interact with the program during a simulation. For example, the user may want to modify a parameter or some operating condition in order to evaluate its effect. This is especially important in the early stages of a design when the operating range of interest is being established.
- A wide range of statistical-analysis, signal-processing, and display tools.
- Common display formats, including time waveforms, electrical and optical spectra, eye diagrams, and error-rate curves.

1.8.2 Graphical Programming

Commercially available simulation tools for lightwave applications are based on well-established graphical-programming languages. In these languages, system components (e.g., lasers, modulators, optical amplifiers, optical fibers), measurement instruments, and plotting tools are represented by a module library of programmed icons that have bidirectional optical and electrical interfaces. For example, as Fig. 1.14 illustrates, the left-hand side of a typical modeling tool window lists classes of items such as lasers, optical receivers, optical fibers, and measurement equipment. Using a mouse to click on a specific class will open a window with a selection of icons for that class. Associated with each icon is a menu window where the user specifies the values of the operational parameters of the component and its interface characteristics. In addition to using preprogrammed modules, users can create their own custom devices with either the underlying software code or the graphical-programming language.

Using such a set of graphical icons, one can put together a simulation of a complex component, a simple link, or a sophisticated multichannel transmission path in several minutes. One simply selects the icons representing the desired components and measurement instruments, and connects them together with a wiring tool to create a model of the optical transmission system. When the design is completed, the diagram compiles rapidly and the simulation can be run using control buttons on a tool bar. Figure 1.14 also shows an example of a simple link consisting of a 1-mW optical transmitter, a 20-km fiber, an optical receiver, and an optical spectrum analyzer (OSA). Figure 1.15 illustrates the performance simulation result of this link in the form of an eye diagram for the bit error rate (see Chapter 14) at a 10 Gb/s data transmission rate.
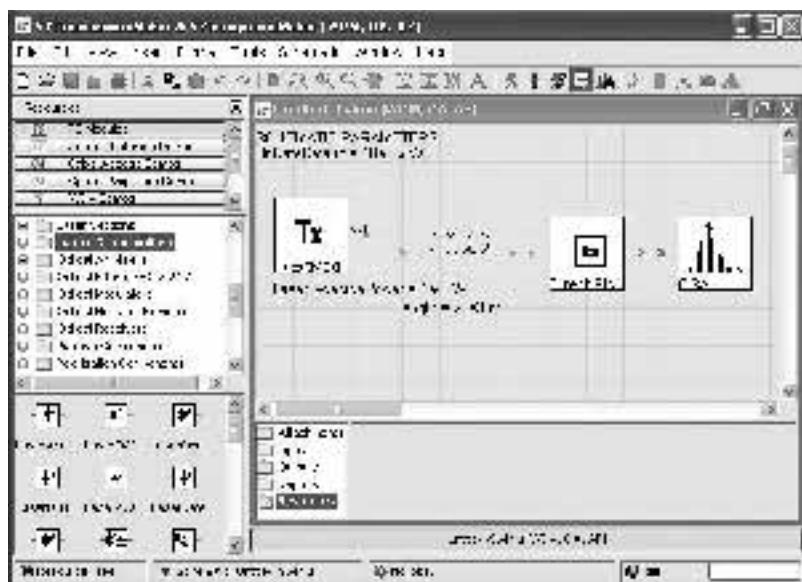
Fig. 1.14 *A screen view showing the tools used to create a simulation schematic using a modeling program from VPIphotonics.*
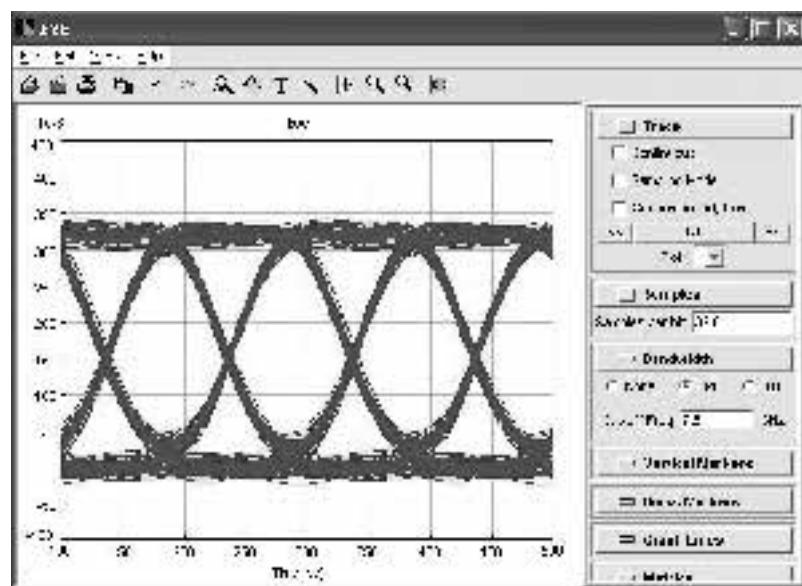
Fig. 1.15 *A screen view showing an eye-diagram result from the simulation schematic shown in Fig. 1.14.*

Once the icons have been selected and connected together, the complex and challenging part starts for the user. This involves choosing realistic ranges of the parameters of the electrical and optical components and submodules. It is important that the parameter values make sense in an actual application. In some cases this may entail examining the specification sheets of vendors.

1.8.3 Example Programs for Student Use

Several commercial vendors offer various suites of software-based modeling-tool modules for optical fiber communication systems.[52–54] These design and planning tools are intended for use across all levels of lightwave network analyses, performance evaluations, and technology comparisons ranging from passive and active components and modules, to complex transmission links, to entire optical networks. Familiar measurement instruments are built into the software to offer a wide range of settable options when displaying data from multiple simulation runs and multidimensional sweeps across a parameter range. Signal-processing modules allow data to be manipulated to mimic any laboratory setup.

Such tools are in use by component and system manufacturers, system integrators, network operators, and access service providers for functions such as capacity planning, comparative assessments of various technologies, optimization of transport and service networks, syntheses and analyses of WDM system and link designs, and component designs. In addition, many universities are using these simulation tools for both research and teaching purposes.

Abbreviated versions of several simulation modules may be downloaded for noncommercial educational use from the websites of the tool vendors. These simplified versions contain predefined component and link configurations that allow interactive concept demonstrations. Among the numerous demonstration setups are optical amplifier structures, simple single-wavelength links, and WDM links. The configurations are fixed, but the reader has the ability to interactively change the operational parameter values of components such as optical fibers, light sources, optical filters, and optical amplifiers. As part of the results that are possible, images of what would appear on the display screen of a standard instrument, such as a spectrum analyzer or an oscilloscope, enable the user to see the effects on link performance when various component values change.

The book website (http://www.mhhe.com/keiserOFC) describes current offerings available on vendor websites of the demonstration modules that the reader may download and learn from.[55]

▌ PROBLEMS ▌

1.1 (*a*) What are the energies in electron volts (eV) of light at wavelengths 850, 1310, 1490, and 1550 nm?

(*b*) Consider a 1-ns pulse with a 100-nW amplitude at each of these wavelengths. How many photons are in such a pulse at each wavelength?

1.2 A WDM optical transmission system is designed so that each channel has a spectral width of 0.8 nm. How many wavelength channels can be used in the C-band?

1.3 Three sine waves have the following periods: 25 μs, 250 ns, 125 ps. What are their frequencies?

1.4 A sine wave is offset 1/6 of a cycle with respect to time zero. What is its phase in degrees and in radians?

1.5 Consider two signals that have the same frequency. When the amplitude of the first signal is at its maximum, the amplitude of the second signal is at half its maximum from the zero level. What is the phase shift between the two signals?

1.6 What is the duration of a bit for each of the following three signals which have bit rates of 64 kb/s, 5 Mb/s, and 10 Gb/s?

1.7 (*a*) Convert the following absolute power gains to decibel power gains: 10^{-3}, 0.3, 1, 4, 10, 100, 500, 2^n.

(*b*) Convert the following decibel power gains to absolute power gains: –30 dB, 0 dB, 13 dB, 30 dB, 10*n* dB.

1.8	(*a*) Convert the following absolute power levels to dBm values: 1 pW, 1 nW, 1 mW, 10 mW, 50 mW.

(*b*) Convert the following dBm values to power levels in units of mW: –13 dBm, –6 dBm, 6 dBm, 17 dBm.

1.9	A signal travels from point *A* to point *B*.

(*a*) If the signal power is 1.0 mW at point *A* and 0.125 mW at point *B*, what is the attenuation in dB?

(*b*) What is the signal power at point *B* if the attenuation is 15 dB?

1.10	A signal passes through three cascaded amplifiers, each of which has a 5-dB gain. What is the total gain in dB? By what numerical factor is the signal amplified?

1.11	A 50-km long optical fiber has a total attenuation of 24 dB. If 500 μW of optical power get launched into the fiber, what is the output optical power level in dBm and in μW?

1.12	Based on the Shannon theorem, the maximum data rate R of a channel with a bandwidth B is R = B $\log_2$(1+S/N), where S/N is the signal-to-noise ratio. Suppose a transmission line has a bandwidth of 2 MHz. If the signal-to-noise ratio at the receiving end is 20 dB, what is the maximum data rate that this line can support?

1.13	(*a*) At the lowest TDM level of the digital service scheme, 24 channels of 64 kb/s each are multiplexed into a 1.544-Mb/s *DS*1 channel. How much is the overhead that is added?

(*b*) The next higher multiplexed level, the *DS*2 rate, is 6.312 Mb/s. How many *DS*1 channels can be accommodated in the *DS*2 rate, and what is the overhead?

(*c*) If the *DS*3 rate that is sent over a *T*3 line is 44.376 Mb/s, how many *DS*2 channels can be accommodated on a *T*3 line, and what is the overhead?

(*d*) Using the above results, find how many *DS*0 channels can be sent over a *T*3 line. What is the total added overhead?

REFERENCES

1. J. Hecht, *City of Light*, Oxford University Press, New York, revised expanded ed., 2004. This book gives a comprehensive account of the history behind the development of optical fiber communication systems.

2. E. Snitzer, "Cylindrical dielectric waveguide modes," *J. Opt. Soc. Amer.*, vol. 51, pp. 491–498, May 1961.

3. K. C. Kao and G. A. Hockman, "Dielectric-fibre surface waveguides for optical frequencies," *Proceedings IEE*, vol. 113, pp. 1151–1158, July 1966.

4. This series of books contains dozens of topics in all areas of optical fiber technology presented by researchers from AT&T Bell Laboratories over a period of thirty years. (*a*) S. E. Miller and A. G. Chynoweth, eds., *Optical Fiber Telecommunications*, Academic, New York, 1979; (*b*) S. E. Miller and I. P. Kaminow, eds., *Optical Fiber Telecommunications–II*, Academic, New York, 1988; (*c*) I. P. Kaminow and T. L. Koch, eds., *Optical Fiber Telecommunications–III*, vols. A and B, Academic, New York, 1997; (*d*) I. P. Kaminow and T. Li, eds., *Optical Fiber Telecommunications–IV*, vols. A and B, Academic, New York, 2002; (*e*) I. P. Kaminow, T. Li, and A. E. Willner, eds., *Optical Fiber Telecommunications–V*, vols. A and B, Academic, New York, 2008.

5. C. L. Chen, *Foundations of Guided-Wave Optics*, Wiley, New York, 2007.

6. L. N. Binh, *Digital Optical Communications*, CRC Press, Boca Raton, FL, 2009.

7. S. L. Chuang, *Physics of Photonic Devices*, Wiley, New York, 2nd ed., 2009.

8. J. M. Senior, *Optical Fiber Communications*, Prentice-Hall, Englewood Cliffs, NJ, 3rd ed., 2009.

9. R. Ramaswami, K. Sivarajan, and G. Sasaki, *Optical Networks*, Morgan Kaufmann, San Francisco, 3rd ed., 2009.

10. R. L. Freeman, *Fiber-Optic Systems for Telecommunications*, Wiley, Hoboken, NJ, 2002.

11. G. P. Agrawal, *Fiber Optic Communication Systems*, Wiley, Hoboken, NJ, 3rd ed., 2002.

12. E. Desurvire, *Erbium-Doped Fiber Amplifiers*, Wiley, Hoboken, NJ, 2002.

13. E. Desurvire, D. Bayart, B. Desthieux, and S. Bigo, *Erbium-Doped Fiber Amplifiers: Devices and System Developments*, Wiley, New York, 2002.

14. B. Razavi, *Design of Integrated Circuits for Optical Communications*, McGraw-Hill, New York, 2003.

15. J. A. Buck, *Fundamentals of Optical Fibers*, Wiley, New York, 2004.

16. K. P. Ho, *Phase-Modulated Optical Communication Systems*, Springer, New York, 2005.

17. G. Keiser, *Optical Communications Essentials*, McGraw-Hill, New York, 2003.

18. G. Keiser, *FTTX Concepts and Applications*, Wiley, Hoboken, NJ, 2006.

19. A. Yariv and P. Yeh, *Photonics: Optical Electronics in Modern Communications*, Oxford University Press USA, New York, 6th ed., 2006.

20. B.E.A. Saleh and M. C. Teich, *Fundamentals of Photonics*, Wiley, Hoboken, NJ, 2nd ed., 2007.

21. *Optical Fiber Communications (OFC) Conf.* is cosponsored annually by the Optical Society of America (OSA), Washington, DC, and the Institute of Electrical and Electronic Engineers (IEEE), New York, NY.

22. *European Conference on Optical Fibre Communications (ECOC)* is held annually in Europe; sponsored by various European engineering organizations.

23. *Photonics West*, held in the United States, and *Asia-Pacific Optical Communications (APOC)*, held in Asia, are two of a number of annual conferences sponsored by SPIE, Bellingham, WA, USA.

24. S. Haykin, *Communication Systems*, Wiley, New York, 5th ed., 2009.

25. A. Leon-Garcia and I. Widjaja, *Communication Networks*, McGraw-Hill, New York, 2nd ed., 2004.

26. J. Proakis and M. Salehi, *Digital Communications*, McGraw-Hill, Burr Ridge, IL, 5th ed., 2008.

27. A. B. Carlson and P. B. Crilly, *Communication Systems*, McGraw-Hill, Burr Ridge, IL, 5th ed., 2009.

28. L. W. Couch II, *Digital and Analog Communication Systems*, Prentice Hall, Upper Saddle River, NJ, 7th ed., 2007.

29. F. P. Kapron, D. B. Keck, and R. D. Maurer, "Radiation losses in glass optical waveguides," *Appl. Phys. Lett.*, vol. 17, pp. 423–425, Nov. 1970.

30. W. A. Gambling, "The rise and rise of optical fibers," *IEEE J. Sel. Topics Quantum Electron.*, vol. 6, no. 6, pp. 1084–1093 Nov./Dec. 2000.

31. B. St. Arnaud, J. Wu, and B. Kalali, "Customer-controlled and -managed optical networks," *J. Lightwave Technology*, vol. 21, pp. 2804–2810, Nov. 2003.

32. D. Simeonidou, R. Nejabati, G. Zervas, D. Klonidis, A. Tzanakaki, M. J. O'Mahony, "Dynamic optical-network architectures and technologies for existing and emerging grid services," *J. Lightwave Technology*, vol. 23, pp. 3347–3357, Oct. 2005.

33. N. Taesombut, F. Uyeda, A. A. Chien, L. Smarr, T. A. DeFanti, P. Papadopoulos, J. Leigh, M. Ellisman, and J. Orcutt, "The OptIPuter: High-performance, QoS-guaranteed network service for emerging e-science applications," *IEEE Commun. Mag.*, vol. 44, pp. 38–45, May 2006.

34. D. H. Rice and G. Keiser, "Applications of fiber optics to tactical communication systems," *IEEE Commun. Mag.*, vol. 23, pp. 46–57, May 1985.

35. ITU-T Recommendation G.Sup39, *Optical System Design and Engineering Considerations*, Feb. 2006.

36. A. Scavennec and O. Leclerc, "Toward high-speed 40-Gb/s transponders," *Proc. IEEE*, vol. 94, pp. 986–996, May 2006.

37. J. X. Cai, C. R. Davidson, M. Nissov, H. Li, W. T. Anderson, Y. Cai, L. Liu, A. N. Pilipetskii, D. G. Foursa, W. W. Patterson, P. C. Corbett, A. J. Lucero, and N. S. Bergano, "Transmission of 40-Gb/s WDM signals over transoceanic distance using conventional NZ-DSF with

receiver dispersion slope compensation," *J. Lightwave Technology*, vol. 24, pp. 191–200, Jan. 2006.

38. E. Le Rouzic and S. Gosselin, "160-Gb/s optical networking: A prospective techno-economical analysis," *J. Lightwave Technology*, vol. 23, pp. 3024–3033, Oct. 2005.

39. T.-R. Gong, F.-P. Yan, D. Lu, M. Chen, P. Liu, P.-L. Tao, M. G. Wang, T. J. Li, and S. S. Jian, "Demonstration of single channel 160-Gb/s OTDM 100-km transmission system," *Optics Commun.*, vol. 282, pp. 3460–3463, Sept. 2009.

40. H. van Helvoort, *The ComSoc Guide to Next Generation Optical Transport: SDH/SONET/ODN*, Wiley/IEEE Press, Hoboken, NJ, 2009.

41. H. G. Perros, *Connection-Oriented Networks: SONET/SDH, ATM, MPLS and Optical Networks*, Wiley, Hoboken, NJ, 2005.

42. W. Goralski, *SONET/SDH*, McGraw-Hill, New York, 3rd ed., 2003.

43. M. L. Jones, "Optical networking standards," *J. Lightwave Technology*, vol. 22, pp. 275–280, Jan. 2004.

44. Telecommunications Sector—International Telecommunication Union (ITU-T), various G.600 Series Recommendations for all aspects of optical fiber communications.

45. Special issue on "Standards activities: Addressing the challenges of next-generation optical networks," K. Kazi, guest editor, *Optical Networks Magazine*, vol. 4, issue 1, Jan./Feb. 2003.

46. J. Piprek, ed., *Optoelectronic Devices: Advanced Simulation and Analysis*, Springer, New York, 2005.

47. K. Kawano and T. Kitoh, *Introduction to Optical Waveguide Analysis: Solving Maxwell's Equation and the Schrödinger Equation*, Wiley, Hoboken, NJ, 2002.

48. A. J. Lowery, "Photonic simulation tools," in I. Kaminow and T. Li, eds., *Optical Fiber Telecommunications IV-B: Systems and Impairments*, Academic, New York, 2002.

49. A. J. Lowery, "WDM systems simulations," in A. Gumaste and T. Anthony, eds., *DWDM Network Designs and Engineering Solutions*, Cisco Press, Indianapolis, IN, 2003.

50. R. Scarmozzino, "Simulation tools for devices, systems, and networks," in I. P. Kaminow, T. Koch, and A. E. Willner, eds., *Optical Fiber Telecommunications–V*, vol. B, Chap. 20, Academic, New York, 2008.

51. N. Antoniades, I. Roudas, G. Ellinas, and V. Grigoryan, *Modeling and Computer-Aided Design of Optical Communications Systems and Networks,* Springer, New York, 2007.

52. VPIsystems, Inc., Holmdel, New Jersey, USA, www.vpisystems.com.

53. RSoft Design Group, Inc., Ossining, New York, USA, www.rsoftdesign.com.

54. Optiwave Systems, Inc., Toronto, Ontario, Canada, www.optiwave.com.

55. The book website can be found at http://www.mhhe.com/KeiserOFC.

CHAPTER 2

Optical Fibers: Structures, Waveguiding, and Fabrication

The operational characteristics of an optical fiber largely determine the overall performance of a lightwave transmission system. Some of the questions that arise concerning optical fibers are

1. What is the structure of an optical fiber?
2. How does light propagate along a fiber?
3. Of what materials are fibers made?
4. How is the fiber fabricated?
5. How are fibers incorporated into cable structures?
6. What is the signal loss or attenuation mechanism in a fiber?
7. Why and to what degree does a signal get distorted as it travels along a fiber?

The purpose of this chapter is to present some of the fundamental answers to the first five questions in order to attain a good understanding of the physical structure and waveguiding properties of optical fibers. Questions 6 and 7 are answered in Chapter 3. The discussions address both conventional silica and photonic crystal fibers.

Fiber optics technology involves the emission, transmission, and detection of light, so we begin our discussion by first considering the nature of light, and then we shall review a few basic laws and definitions of optics.[1] Following a description of the structure of optical fibers, two methods are used to describe how an optical fiber guides light. The first approach uses the geometrical or ray optics concept of light reflection and refraction to provide an intuitive picture of the propagation mechanisms. In the second approach, light is treated as an electromagnetic wave that propagates along the optical fiber waveguide. This involves solving Maxwell's equations subject to the cylindrical boundary conditions of the fiber.

2.1 | The Nature of Light

The concepts concerning the nature of light have undergone several variations during the history of physics. Until the early seventeenth century, it was generally believed that light consisted of a stream of minute particles that were emitted by luminous sources. These particles were pictured as traveling in straight lines, and it was assumed that they could penetrate transparent materials but were reflected from
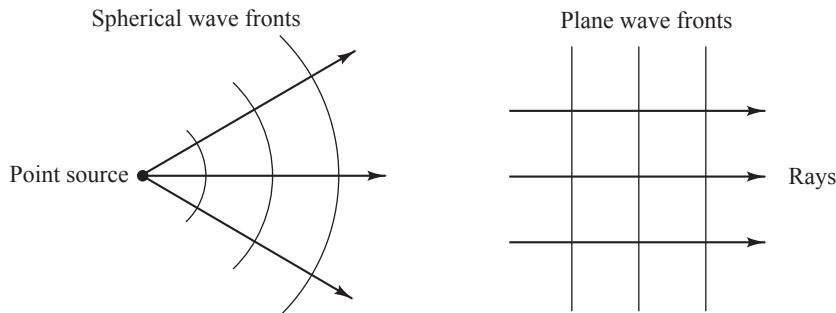
Fig. 2.1 *Representations of spherical and plane wave fronts and their associated rays*

opaque ones. This theory adequately described certain large-scale optical effects, such as reflection and refraction, but failed to explain finer-scale phenomena, such as interference and diffraction.

The correct explanation of diffraction was given by Fresnel in 1815. Fresnel showed that the approximately rectilinear propagation character of light could be interpreted on the assumption that light is a wave motion, and that the diffraction fringes could thus be accounted for in detail. Later, the work of Maxwell in 1864 theorized that light waves must be electromagnetic in nature. Furthermore, observation of polarization effects indicated that light waves are transverse (i.e., the wave motion is perpendicular to the direction in which the wave travels). In this *wave* or *physical optics viewpoint,* the electromagnetic waves radiated by a small optical source can be represented by a train of spherical wave fronts with the source at the center as shown in Fig. 2.1. A *wave front* is defined as the locus of all points in the wave train that have the same phase. Generally, one draws wave fronts passing through either the maxima or the minima of the wave, such as the peak or trough of a sine wave, for example. Thus the wave fronts (also called *phase fronts*) are separated by one wavelength.

When the wavelength of the light is much smaller than the object (or opening) that it encounters, the wave fronts appear as straight lines to this object or opening. In this case, the light wave can be represented as a *plane wave,* and its direction of travel can be indicated by a *light ray*, which is drawn perpendicular to the phase front. The light-ray concept allows large-scale optical effects such as reflection and refraction to be analyzed by the simple geometrical process of *ray tracing.* This view of optics is referred to as *ray* or *geometrical optics.* The concept of light rays is very useful because the rays show the direction of energy flow in the light beam.

2.1.1 Linear Polarization

The electric or magnetic field of a train of *plane linearly polarized waves* traveling in a direction **k** can be represented in the general form

$$\mathbf{A}(\mathbf{r}, t) = \mathbf{e}_i A_0 \exp[\, j \, (\omega t - \mathbf{k} \cdot \mathbf{r})] \tag{2.1}$$

with $\mathbf{r} = x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z$ representing a general position vector and $\mathbf{k} = k_x\mathbf{e}_x + k_y\mathbf{e}_y + k_z\mathbf{e}_z$ representing the wave propagation vector.

Here, A_0 is the maximum amplitude of the wave, $\omega = 2\pi v$, where v is the frequency of the light; the magnitude of the wavevector **k** is $k = 2\pi/\lambda$, which is known as the *wave propagation constant,* with λ being the wavelength of the light; and $\mathbf{e}_i$ is a unit vector lying parallel to an axis designated by i.

Note that the components of the actual (measurable) electromagnetic field represented by Eq. (2.1) are obtained by taking the real part of this equation. For example, if $\mathbf{k} = k\mathbf{e}_z$, and if $\mathbf{A}$ denotes the electric field $\mathbf{E}$ with the coordinate axes chosen such that $\mathbf{e}_i = \mathbf{e}_x$, then the real measurable electric field is given by

$$\mathbf{E}_x(z, t) = \text{Re}(\mathbf{E}) = \mathbf{e}_x E_{0x} \cos(\omega t - kz) = \mathbf{e}_x E_x \tag{2.2}$$

which represents a plane wave that varies harmonically as it travels in the z direction. Here E_{0x} is the maximum amplitude of the wave along the x axis and E_x is the amplitude at a given value of z. The reason for using the exponential form is that it is more easily handled mathematically than equivalent expressions given in terms of sine and cosine. In addition, the rationale for using harmonic functions is that any waveform can be expressed in terms of sinusoidal waves using Fourier techniques.

The electric and magnetic field distributions in a train of plane electromagnetic waves at a given instant in time are shown in Fig. 2.2. The waves are moving in the direction indicated by the vector $\mathbf{k}$. Based on Maxwell's equations, it can be shown[2] that $\mathbf{E}$ and $\mathbf{H}$ are both perpendicular to the direction of propagation. This condition defines a *plane wave*; that is, the vibrations in the electric field are parallel to each other at all points in the wave. Thus, the electric field forms a plane called the *plane of vibration*. Likewise all points in the magnetic field component of the wave lie in another plane of vibration. Furthermore, $\mathbf{E}$ and $\mathbf{H}$ are mutually perpendicular, so that $\mathbf{E}$, $\mathbf{H}$, and $\mathbf{k}$ form a set of orthogonal vectors.

The plane wave example given by Eq. (2.2) has its electric field vector always pointing in the $\mathbf{e}_x$ direction. Such a wave is *linearly polarized* with polarization vector $\mathbf{e}_x$. A general state of polarization
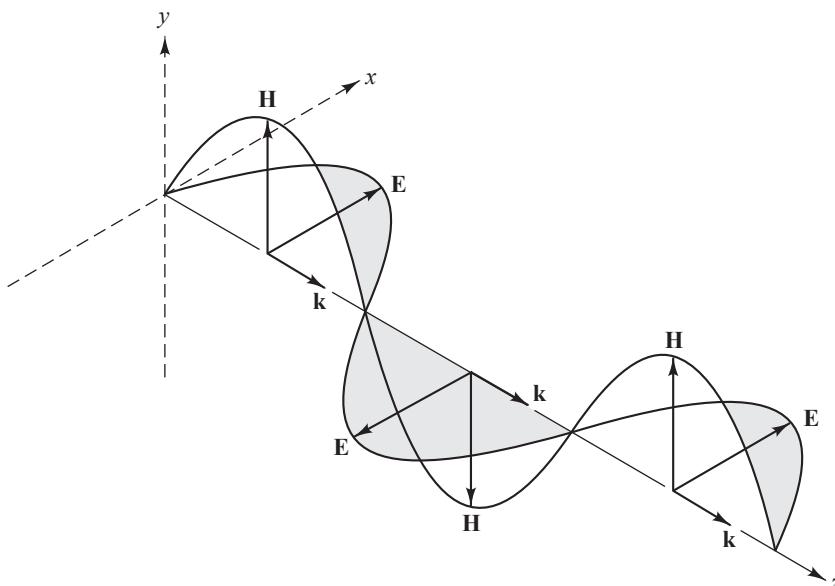
Fig. 2.2   *Electric and magnetic field distributions in a train of plane electromagnetic waves at a given instant in time*

is described by considering another linearly polarized wave that is independent of the first wave and orthogonal to it. Let this wave be

$$\mathbf{E}_y(z, t) = \mathbf{e}_y E_{0y}\cos(\omega t - kz + \delta) = \mathbf{e}_y E_y \tag{2.3}$$

where δ is the relative phase difference between the waves. Similar to Eq. (2.2), E_{0y} is the maximum amplitude of the wave along the y axis and E_y is the amplitude at a given value of z. The resultant wave is

$$\mathbf{E}(z, t) = \mathbf{E}_x(z, t) + \mathbf{E}_y(z, t) \tag{2.4}$$

If δ is zero or an integer multiple of 2π, the waves are in phase. Equation (2.4) is then also a linearly polarized wave with a polarization vector making an angle

$$\theta = \arctan \frac{E_{0y}}{E_{0x}} \tag{2.5}$$

with respect to $\mathbf{e}_x$ and having a magnitude

$$E = \left(E_{0x}^2 + E_{0y}^2 \right)^{1/2} \tag{2.6}$$

This case is shown schematically in Fig. 2.3. Conversely, just as any two orthogonal plane waves can be combined into a linearly polarized wave, an arbitrary linearly polarized wave can be resolved into two independent orthogonal plane waves that are in phase.
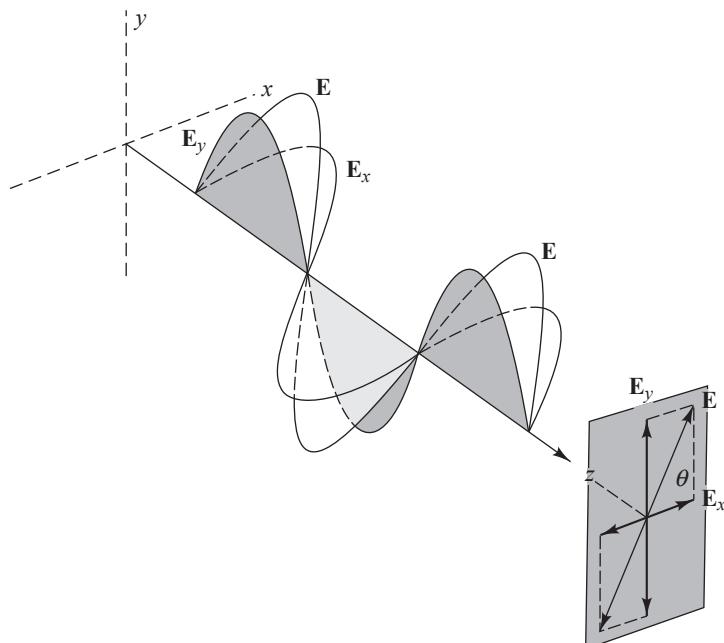
Fig. 2.3 *Addition of two linearly polarized waves having a zero relative phase between them*

Example 2.1   The general form of an electromagnetic wave is

$$y = (\text{amplitude in } \mu\text{m}) \times \cos(\omega t - kz)$$
$$= A \cos\left[2\pi(\nu t - z/\lambda)\right]$$

Find the (a) amplitude, (b) the wavelength, (c) the angular frequency, and (d) the displacement at time $t = 0$ and $z = 4\ \mu$m of a given plane electromagnetic wave specified by the equation $y = 12 \cos\left[2\pi(3t - 1.2z)\right]$.

Solution: From the above general wave equation we find that

(*a*) Amplitude = 12 μm
(*b*) Wavelength: $1/\lambda = 1.2\ \mu\text{m}^{-1}$ so that $\lambda = 833$ nm
(*c*) The angular frequency is $\omega = 2\pi\nu = 2\pi(3) = 6\pi$
(*d*) At time $t = 0$ and $z = 4\ \mu$m we have that the displacement is

$$y = 12 \cos\left[2\pi(-1.2\ \mu\text{m}^{-1})(4\ \mu\text{m})\right]$$
$$= 12 \cos\left[2\pi(-4.8)\right] = 10.38\ \mu\text{m}$$

2.1.2   Elliptical and Circular Polarization

For general values of δ the wave given by Eq. (2.4) is *elliptically polarized*. The resultant field vector **E** will both rotate and change its magnitude as a function of the angular frequency ω. From Eqs. (2.2) and (2.3) we can show that (see Prob. 2.4) for a general value of δ,

$$\left(\frac{E_x}{E_{0x}}\right)^2 + \left(\frac{E_y}{E_{0y}}\right)^2 - 2\left(\frac{E_x}{E_{0x}}\right)\left(\frac{E_y}{E_{0y}}\right)\cos\delta = \sin^2\delta \tag{2.7}$$

which is the general equation of an ellipse. Thus as Fig. 2.4 shows, the endpoint of **E** will trace out an ellipse at a given point in space. The axis of the ellipse makes an angle α relative to the x axis given by

$$\tan 2\alpha = \frac{2E_{0x}E_{0y}\cos\delta}{E_{0x}^2 - E_{0y}^2} \tag{2.8}$$
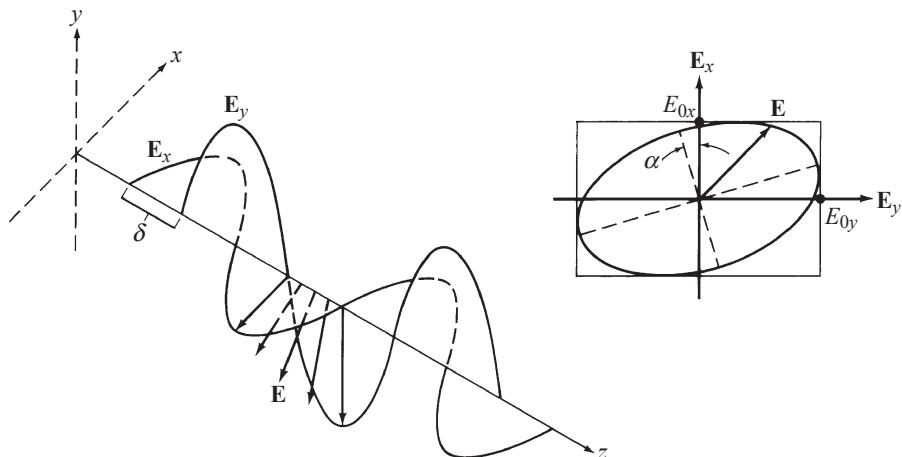
Fig. 2.4   *Elliptically polarized light results from the addition of two linearly polarized waves of unequal amplitude having a nonzero phase difference δ between them*

To get a better picture of Eq. (2.7), let us align the principal axis of the ellipse with the x axis. Then $\alpha = 0$, or, equivalently, $\delta = \pm\,\pi/2,\,\pm\,3\pi/2, \,...$, so that Eq. (2.7) becomes

$$\left(\frac{E_x}{E_{0x}}\right)^2 + \left(\frac{E_y}{E_{0y}}\right)^2 = 1 \tag{2.9}$$

This is the familiar equation of an ellipse with the origin at the center and semi-axes E_{0x} and E_{0y}.

When $E_{0x} = E_{0y} = E_0$ and the relative phase difference $\delta = \pm\pi/2 + 2m\pi$, where $m = 0, \pm1, \pm2, \,...$, then we have *circularly polarized* light. In this case, Eq. (2.9) reduces to

$$E_x^2 + E_y^2 = E_0^2 \tag{2.10}$$

which defines a circle. Choosing the positive sign for δ, Eqs (2.2) and (2.3) become

$$\mathbf{E}_x(z, t) = \mathbf{e}_x E_0 \cos(\omega t - kz) \tag{2.11}$$

$$\mathbf{E}_y(z, t) = -\mathbf{e}_y E_0 \sin(\omega t - kz) \tag{2.12}$$

In this case, the endpoint of $\mathbf{E}$ will trace out a circle at a given point in space, as Fig. 2.5 illustrates. To see this, consider an observer located at some arbitrary point z_{ref} toward whom the wave
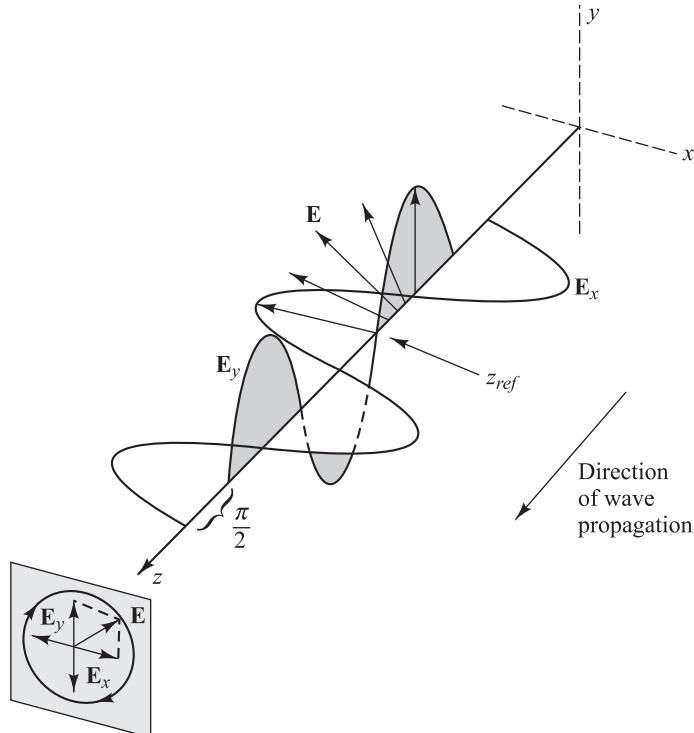
Fig. 2.5   *Addition of two equal-amplitude linearly polarized waves with a relative phase difference δ = π/2 + 2mπ results in a right circularly polarized wave*

is moving. For convenience, we will pick this point at $z = \pi/k$ at $t = 0$. Then, from Eqs (2.11) and (2.12) we have

$$\mathbf{E}_x(z, t) = -\mathbf{e}_x E_0 \quad \text{and} \quad \mathbf{E}_y(z, t) = 0$$

so that $\mathbf{E}$ lies along the negative x axis as Fig. 2.5 shows. At a later time, say $t = \pi/2\omega$, the electric field vector has rotated through 90° and now lies along the positive y axis at z_{ref}. Thus as the wave moves toward the observer with increasing time, the resultant electric field vector $\mathbf{E}$ rotates *clockwise* at an angular frequency ω. It makes one complete rotation as the wave advances through one wavelength. Such a light wave is *right circularly polarized*.

If we choose the negative sign for δ, then the electric field vector is given by

$$\mathbf{E} = E_0[\mathbf{e}_x \cos(\omega t - kz) + \mathbf{e}_y \sin(\omega t - kz)] \tag{2.13}$$

Now $\mathbf{E}$ rotates *counterclockwise* and the wave is *left circularly polarized*.

2.1.3   The Quantum Nature of Light

The wave theory of light adequately accounts for all phenomena involving the transmission of light. However, in dealing with the interaction of light and matter, such as occurs in dispersion and in the emission and absorption of light, neither the particle theory nor the wave theory of light is appropriate. Instead, we must turn to quantum theory, which indicates that optical radiation has particle as well as wave properties. The particle nature arises from the observation that light energy is always emitted or absorbed in discrete units called *quanta* or *photons*. In all experiments used to show the existence of photons, the photon energy is found to depend only on the frequency v. This frequency, in turn, must be measured by observing a wave property of light.

The relationship between the energy E and the frequency v of a photon is given by

$$E = hv \tag{2.14}$$

where $h = 6.625 \times 10^{-34}$ J·s is Planck's constant. When light is incident on an atom, a photon can transfer its energy to an electron within this atom, thereby exciting it to a higher energy level. In this process either all or none of the photon energy is imparted to the electron. The energy absorbed by the electron must be exactly equal to that required to excite the electron to a higher energy level. Conversely, an electron in an excited state can drop to a lower state separated from it by an energy hv by emitting a photon of exactly this energy.

2.2   Basic Optical Laws and Definitions

This section reviews some of the basic optics laws and definitions relevant to optical fiber transmission technology. These include Snell's law, the definition of the refractive index of a material, and the concepts of reflection, refraction, and polarization.

2.2.1   Refractive Index

A fundamental optical parameter of a material is the *refractive index* (or *index of refraction*). In free space a light wave travels at a speed $c = 3 \times 10^8$ m/s. The speed of light is related to the frequency v and the wavelength λ by $c = v\lambda$. Upon entering a dielectric or nonconducting medium the wave now travels

Table 2.1 *Indices of refraction for various substances*

Material	Refractive index
Acetone	1.356
Air	1.000
Diamond	2.419
Ethyl alcohol	1.361
Fused quartz (SiO_2): varies with wavelength	1.453 @ 850 nm
Gallium arsenide (GaAs)	3.299 (infrared region)
Glass, crown	1.52–1.62
Glycerin	1.473
Polymethylmethacrylate (PMMA)	1.489
Silicon (varies with wavelength)	3.650 @ 850 nm
Water	1.333

at a speed *s*, which is characteristic of the material and is less than *c*. The ratio of the speed of light in a vacuum to that in matter is the index of refraction *n* of the material and is given by

$$n = \frac{c}{s} \tag{2.15}$$

Table 2.1 lists the indices of refraction for various substances.

2.2.2 Reflection and Refraction

The concepts of reflection and refraction can be interpreted most easily by considering the behavior of light rays associated with plane waves traveling in a dielectric material. When a light ray encounters a boundary separating two different media, part of the ray is reflected back into the first medium and the remainder is bent (or refracted) as it enters the second material. This is shown in Fig. 2.6 where $n_2 < n_1$. The bending or refraction of the light ray at the interface is a result of the difference in the speed of
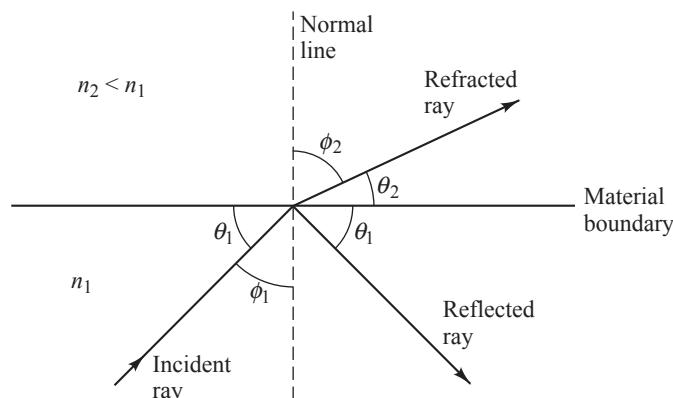
Fig. 2.6 *Refraction and reflection of a light ray at a material boundary*

light in two materials that have different refractive indices. The relationship at the interface is known as Snell's law and is given by

$$n_1 \sin \phi_1 = n_2 \sin \phi_2 \tag{2.16}$$

or, equivalently, as

$$n_1 \cos \theta_1 = n_2 \cos \theta_2 \tag{2.17}$$

where the angles are defined in Fig. 2.6. The angle ϕ_1 between the incident ray and the normal to the surface is known as the *angle of incidence.*

According to the law of reflection, the angle θ_1 at which the incident ray strikes the interface is exactly equal to the angle that the reflected ray makes with the same interface. In addition, the incident ray, the normal to the interface, and the reflected ray all lie in the same plane, which is perpendicular to the interface plane between the two materials. This plane is called the *plane of incidence.* When light traveling in a certain medium is reflected off an optically denser material (one with a higher refractive index), the process is referred to as *external reflection.* Conversely, the reflection of light off of less optically dense material (such as light traveling in glass being reflected at a glass–air interface) is called *internal reflection.*

As the angle of incidence ϕ_1 in an optically denser material becomes larger, the refracted angle ϕ_2 approaches $\pi/2$. Beyond this point no refraction is possible and the light rays become *totally internally reflected.* The conditions required for total internal reflection can be determined by using Snell's law [Eq. (2.16)]. Consider Fig. 2.7, which shows a glass surface in air. A light ray gets bent toward the glass surface as it leaves the glass in accordance with Snell's law. If the angle of incidence ϕ_1 is increased, a point will eventually be reached where the light ray in air is parallel to the glass surface. This point is known as the *critical angle of incidence ϕ_c.* When the incidence angle ϕ_1 is greater than the critical angle, the condition for total internal reflection is satisfied; that is, the light is totally reflected back into the glass with no light escaping from the glass surface. (This is an idealized situation. In practice, there is always some tunneling of optical energy through the interface. This can be explained in terms of the electromagnetic wave theory of light, which is presented in Sec. 2.4.)

As an example, consider the glass–air interface shown in Fig. 2.7. When the light ray in air is parallel to the glass surface, then $\phi_2 = 90°$ so that $\sin \phi_2 = 1$. The critical angle in the glass is thus

$$\sin \phi_c = \frac{n_2}{n_1} \tag{2.18}$$
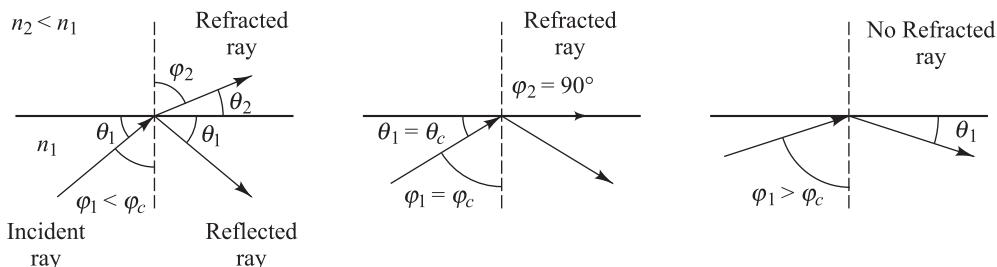
Fig. 2.7 *Representation of the critical angle and total internal reflection at a glass–air interface, where n_1 is the refractive index of glass*

Example 2.2   Consider the interface between a glass slab with $n_1 = 1.48$ and air for which $n_2 = 1.00$. What is the critical angle for light traveling in the glass?

Solution: From Eq. (2.18), for light traveling in the glass the critical angle is

$$\varphi_c = \sin^{-1}\frac{n_2}{n_1} = \sin^{-1} 0.676 = 42.5°$$

Thus any light ray traveling in the glass that is incident on the glass–air interface at a normal angle ϕ_1 (as shown in Fig. 2.7) greater than 42.5° is totally reflected back into the glass.

Example 2.3   A light ray traveling in air ($n_1 = 1.00$) is incident on a smooth, flat slab of crown glass, which has a refractive index $n_2 = 1.52$. If the incoming ray makes an angle of $\varphi_1 = 30.0°$ with respect to the normal, what is the angle of refraction φ_2 in the glass?

Solution: From Snell's law given by Eq. (2.16), we find

$$\sin \varphi_2 = \frac{n_1}{n_2}\sin\varphi_1 = \frac{1.00}{1.52}\sin 30°$$

$$= 0.658 \times 0.5 = 0.329$$

Solving for φ_2 then yields $\varphi_2 = \sin^{-1}(0.329) = 19.2°$.

In addition, when light is totally internally reflected, a phase change δ occurs in the reflected wave. This phase change depends on the angle $\theta_1 < \pi/2 - \phi_c$ according to the relationships[1]

$$\tan\frac{\delta_N}{2} = \frac{\sqrt{n^2\cos^2\theta_1 - 1}}{n\sin\theta_1} \qquad (2.19a)$$

$$\tan\frac{\delta_p}{2} = n\frac{\sqrt{n^2\cos^2\theta_1 - 1}}{\sin\theta_1} \qquad (2.19b)$$

Here, δ_N and δ_p are the phase shifts of the electric field wave components normal and parallel to the plane of incidence, respectively, and $n = n_1/n_2$. These phase shifts are shown in Fig. 2.8 for a glass–air interface ($n = 1.5$ and $\phi_c = 42.5°$). The values range from zero immediately at the critical angle when $\theta_1 = \pi/2 - \phi_c$ to 180° when $\theta_1 = 0°$.
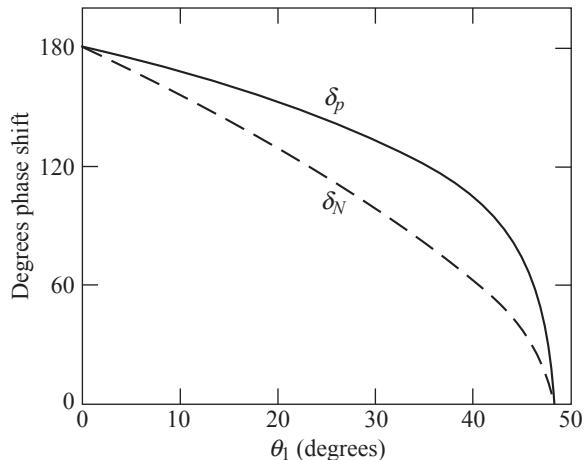
Fig. 2.8   *Phase shifts occurring from the reflection of wave components normal (δ_N) and parallel (δ_p) to the plane of incidence*

2.2.3   Polarization Components of Light

An ordinary lightwave consists of many transverse electromagnetic waves that vibrate in a variety of directions (i.e., in more than one plane) and is called *unpolarized light.* However, one can represent any arbitrary direction of vibration as a combination of a parallel vibration and a perpendicular vibration, as shown in Fig. 2.9. Therefore, one can consider unpolarized light as consisting of two orthogonal plane polarization components, one that lies in the plane of incidence (the plane containing the incident and reflected rays) and the other of which lies in a plane perpendicular to the plane of incidence. These are the *parallel polarization* and the *perpendicular polarization* components, respectively. In the case when
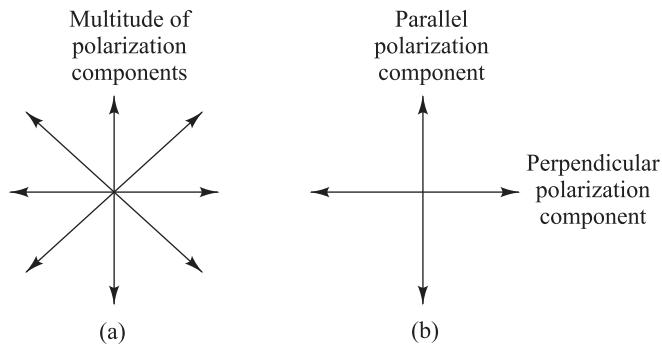
Fig. 2.9 *Polarization represented as a combination of a parallel vibration and a perpendicular vibration*

all the electric field planes of the different transverse waves are aligned parallel to each other, then the lightwave is linearly polarized. This is the simplest type of polarization, as Sec. 2.1.1 describes.

Unpolarized light can be split into separate polarization components either by reflection off of a nonmetallic surface or by refraction when the light passes from one material to another. As noted in Fig. 2.10, when an unpolarized light beam traveling in air impinges on a nonmetallic surface such as glass, part of the beam is reflected and part is refracted into the glass. A circled dot and an arrow designate the parallel and perpendicular polarization components, respectively, in Fig. 2.10. The reflected beam is partially polarized and at a specific angle (known as *Brewster's angle*) the reflected light is
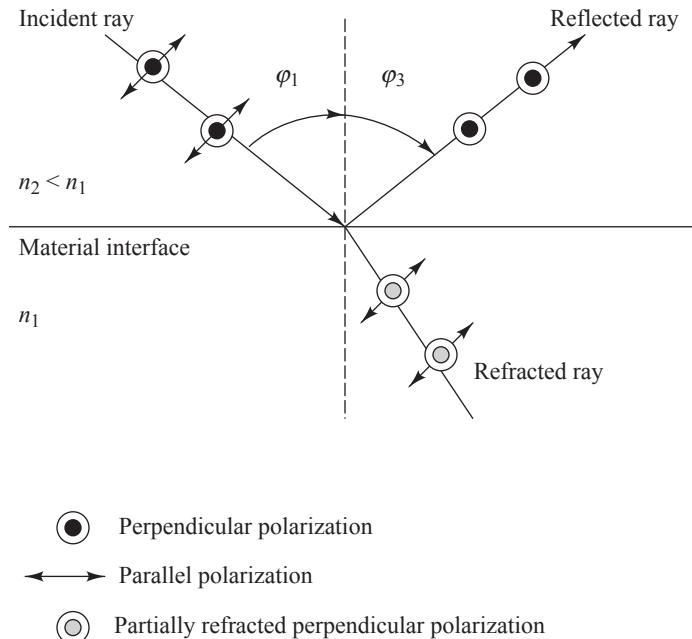
Fig. 2.10 *Behavior of an unpolarized light beam at the interface between air and a nonmetallic surface*
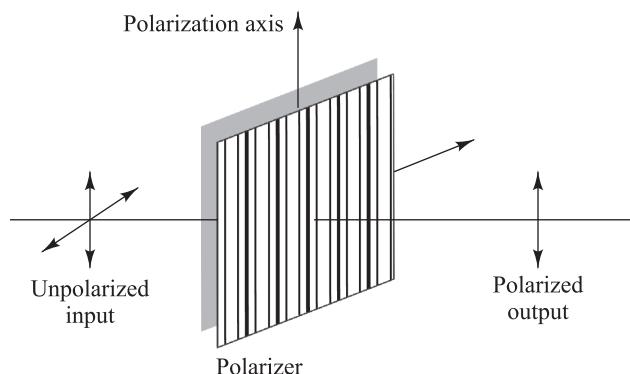
Fig. 2.11   *Only the vertical polarization component passes through a vertically oriented polarizer.*

completely perpendicularly polarized. The parallel component of the refracted beam is transmitted entirely into the glass, whereas the perpendicular component is only partially refracted. How much of the refracted light is polarized depends on the angle at which the light approaches the surface and on the material composition.

2.2.4 Polarization-Sensitive Materials

The polarization characteristics of light are important when examining the behavior of components such as optical isolators and light filters. Here we look at three polarization-sensitive materials or devices that are used in such components. These are polarizers, Faraday rotators, and birefringent crystals.

A *polarizer* is a material or device that transmits only one polarization component and blocks the other. For example, in the case when unpolarized light enters a polarizer that has a vertical transmission axis as shown in Fig. 2.11, only the vertical polarization component passes through the device. A familiar example of this concept
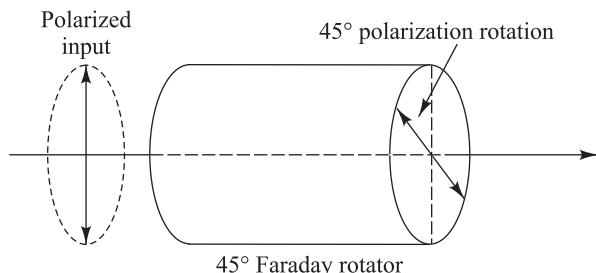
Fig. 2.12   *A Faraday rotator is a device that rotates the state of polarization, for example, clockwise by 45° or a quarter of a wavelength.*
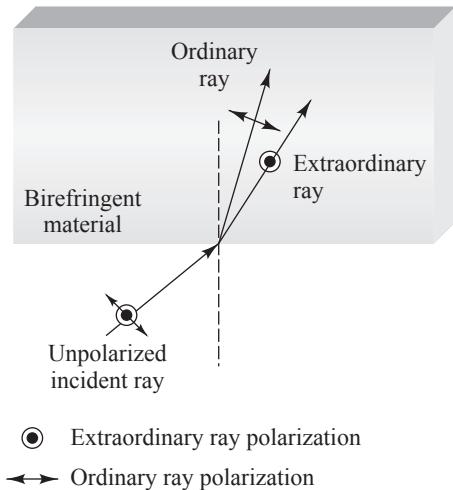
Fig. 2.13   *A birefringent crystal splits the light signal entering it into two perpendicularly polarized beams.*

Table 2.2   *Common birefringent crystals and some applications*

Crystal name	Symbol	n_o	n_e	Applications
Calcite	$CaCO_3$	1.658	1.486	Polarization controllers and beamsplitters
Lithium niobate	$LiNbO_3$	2.286	2.200	Light signal modulators
Rutile	TiO_2	2.616	2.903	Optical isolators and circulators
Yttrium vanadate	YVO_4	1.945	2.149	Optical isolators, circulators, and beam displacers

is the use of polarizing sunglasses to reduce the glare of partially polarized sunlight reflections from road or water surfaces. To see the polarization property of the sunglasses, a number of glare spots will appear when users tilt their head sideways. The polarization filters in the sunglasses block out the polarized light coming from these glare spots when the head is held normally.

A *Faraday rotator* is a device that rotates the *state of polarization* (SOP) of light passing through it by a specific amount. For example, a popular device rotates the SOP clockwise by 45° or a quarter of a wavelength, as shown in Fig. 2.12.

This rotation is independent of the SOP of input light, but the rotation angle is different depending on the direction in which the light passes through the device. That is, the rotation process is not reciprocal. In this process, the SOP of the input light is maintained after the rotation. For example, if the input light to a 45° Faraday rotator is linearly polarized in a vertical direction, then the rotated light exiting the crystal also is linearly polarized at a 45° angle. The Faraday rotator material usually is an asymmetric crystal such as yttrium iron garnet (YIG), and the degree of angular rotation is proportional to the thickness of the device.

Birefringent or *double-refractive crystals* have a property called *double refraction*. This means that the indices of refraction are slightly different along two perpendicular axes of the crystal as shown in Fig. 2.13. A device made from such materials is known as a *spatial walk-off polarizer* (SWP). The SWP splits the light signal entering it into two orthogonally (perpendicularly) polarized beams. One of the beams is called an *ordinary ray* or o-ray because it obeys Snell's law of refraction at the crystal surface. The second beam is called the *extraordinary ray* or e-ray because it refracts at an angle that deviates from the prediction of the standard form of Snell's law. Each of the two orthogonal polarization components thus is refracted at a different angle as shown in Fig. 2.13. For example, if the incident unpolarized light arrives at an angle perpendicular to the surface of the device, the o-ray can pass straight through the device whereas the e-ray component is deflected at a slight angle so it follows a different path through the material.

Table 2.2 lists the ordinary index n_o and the extraordinary index n_e of some common birefringent crystals that are used in optical communication components and gives some of their applications.

2.3 | Optical Fiber Modes and Configurations

Before going into details on optical fiber characteristics, this section first presents a brief overview of the underlying concepts of optical fiber modes and optical fiber configurations. The discussions in Sec. 2.3 through 2.7 address conventional optical fibers, which consist of solid dielectric structures. Section 2.8 describes the structure of photonic crystal fibers, which can be created to have a variety of internal microstructures. Chapter 3 describes the operational characteristics of both categories of fibers.

2.3.1 Fiber Types

An optical fiber is a dielectric waveguide that operates at optical frequencies. This fiber waveguide is normally cylindrical in form. It confines electromagnetic energy in the form of light to within its surfaces and guides the light in a direction parallel to its axis. The transmission properties of an optical waveguide are dictated by its structural characteristics, which have a major effect in determining how an optical signal is affected as it propagates along the fiber. The structure basically establishes the information-carrying capacity of the fiber and also influences the response of the waveguide to environmental perturbations.

The propagation of light along a waveguide can be described in terms of a set of guided electromagnetic waves called the *modes* of the waveguide. These guided modes are referred to as the *bound* or *trapped* modes of the waveguide. Each guided mode is a pattern of electric and magnetic field distributions that is repeated along the fiber at equal intervals. Only a certain discrete number of modes are capable of propagating along the guide. As will be seen in Sec. 2.4, these modes are those electromagnetic waves that satisfy the homogeneous wave equation in the fiber and the boundary condition at the waveguide surfaces.

Although many different configurations of the optical waveguide have been discussed in the literature,[3] the most widely accepted structure is the single solid dielectric cylinder of radius a and index of refraction n_1 shown in Fig. 2.14. This cylinder is known as the *core* of the fiber. The core is surrounded by a solid dielectric *cladding*, which has a refractive index n_2 that is less than n_1. Although, in principle, a cladding is not necessary for light to propagate along the core of the fiber, it serves several purposes. The cladding reduces scattering loss that results from dielectric discontinuities at the core surface, it adds mechanical strength to the fiber, and it protects the core from absorbing surface contaminants with which it could come in contact.

In standard optical fibers the core material is highly pure silica glass (SiO_2) compound and is surrounded by a glass cladding. Higher-loss plastic-core fibers with plastic claddings are also widely in use. In addition, most fibers are encapsulated in an elastic, abrasion-resistant plastic material. This material adds further strength to the fiber and mechanically isolates or buffers the fibers from small geometrical irregularities, distortions, or roughnesses of adjacent surfaces. These perturbations could otherwise cause scattering losses induced by random microscopic bends that can arise when the fibers are incorporated into cables or supported by other structures.

Variations in the material composition of the core give rise to the two commonly used fiber types shown in Fig. 2.15. In the first case, the refractive index of the core is uniform throughout and undergoes an abrupt change (or step) at the cladding boundary. This is called a *step-index fiber*. In the second case,
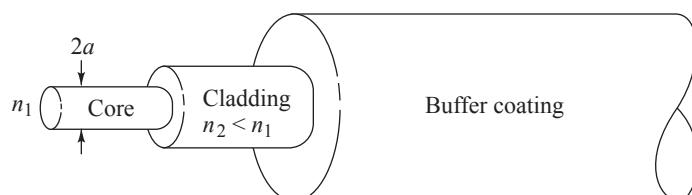
Fig. 2.14 *Schematic of a conventional silica fiber structure. A circular solid core of refractive index n_1 is surrounded by a cladding having a refractive index $n_2 < n_1$. An elastic plastic buffer encapsulates the fiber.*
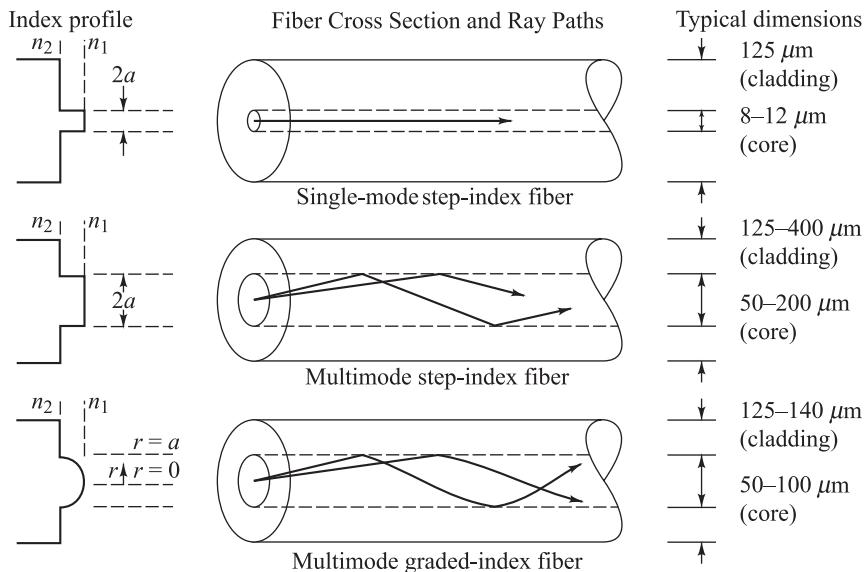
Fig. 2.15   *Comparison of conventional single-mode and multimode step-index and graded-index optical fibers*

the core refractive index is made to vary as a function of the radial distance from the center of the fiber. This type is a *graded-index fiber.*

Both the step- and the graded-index fibers can be further divided into single-mode and multimode classes. As the name implies, a single-mode fiber sustains only one mode of propagation, whereas multimode fibers contain many hundreds of modes. A few typical sizes of single- and multimode fibers are given in Fig. 2.15 to provide an idea of the dimensional scale. Multimode fibers offer several advantages compared with single-mode fibers. As we shall see in Chapter 5, the larger core radii of multimode fibers make it easier to launch optical power into the fiber and facilitate the connecting together of similar fibers. Another advantage is that light can be launched into a multimode fiber using a light-emitting-diode (LED) source, whereas single-mode fibers must generally be excited with laser diodes. Although LEDs have less optical output power than laser diodes (as we shall discuss in Chapter 4), they are easier to make, are less expensive, require less complex circuitry, and have longer lifetimes than laser diodes, thus making them more desirable in certain applications.

A disadvantage of multimode fibers is that they suffer from intermodal dispersion. We shall describe this effect in detail in Chapter 3. Briefly, intermodal dispersion can be described as follows. When an optical pulse is launched into a fiber, the optical power in the pulse is distributed over all (or most) of the modes of the fiber. Each of the modes that can propagate in a multimode fiber travels at a slightly different velocity. This means that the modes in a given optical pulse arrive at the fiber end at slightly different times, thus causing the pulse to spread out in time as it travels along the fiber. This effect, which is known as *intermodal dispersion* or *intermodal distortion,* can be reduced by using a graded-index profile in a fiber core. This allows graded-index fibers to have much larger bandwidths (data rate transmission capabilities) then step-index fibers. Even higher bandwidths are possible in single-mode fibers, where intermodal dispersion effects are not present.

2.3.2 Rays and Modes

The electromagnetic light field that is guided along an optical fiber can be represented by a superposition of bound or trapped modes. Each of these guided modes consists of a set of simple electromagnetic field configurations. For monochromatic light fields of radian frequency ω, a mode traveling in the positive z direction (i.e., along the fiber axis) has a time and z dependence given by

$$e^{j(\omega t - \beta z)}$$

The factor β is the z component of the wave propagation constant $k = 2\pi/\lambda$ and is the main parameter of interest in describing fiber modes. For guided modes, β can assume only certain discrete values, which are determined from the requirement that the mode field must satisfy Maxwell's equations and the electric and magnetic field boundary conditions at the core-cladding interface. This is described in detail in Sec. 2.4.

Another method for theoretically studying the propagation characteristics of light in an optical fiber is the geometrical optics or ray-tracing approach. This method provides a good approximation to the light acceptance and guiding properties of optical fibers when the ratio of the fiber radius to the wavelength is large. This is known as the *small-wavelength limit.* Although the ray approach is strictly valid only in the zero-wavelength limit, it is still relatively accurate and extremely valuable for nonzero wavelengths when the number of guided modes is large; that is, for multimode fibers. The advantage of the ray approach is that, compared with the exact electromagnetic wave (modal) analysis, it gives a more direct physical interpretation of the light propagation characteristics in an optical fiber.

Since the concept of a light ray is very different from that of a mode, let us see qualitatively what the relationship is between them. (The mathematical details of this relationship are beyond the scope of this book but can be found in the literature.[4–6]) A guided mode traveling in the z direction (along the fiber axis) can be decomposed into a family of superimposed plane waves that collectively form a standing-wave pattern in the direction transverse to the fiber axis. That is, the phases of the plane waves are such that the envelope of the collective set of waves remains stationary. Since with any plane wave we can associate a light ray that is perpendicular to the phase front of the wave, the family of plane waves corresponding to a particular mode forms a set of rays called a *ray congruence.* Each ray of this particular set travels in the fiber at the same angle relative to the fiber axis. We note here that, since only a certain number M of discrete guided modes exist in a fiber, the possible angles of the ray congruences corresponding to these modes are also limited to the same number M. Although a simple ray picture appears to allow rays at any angle greater than the critical angle to propagate in a fiber, the allowable quantized propagation angles result when the phase condition for standing waves is introduced into the ray picture. This is discussed further in Sec. 2.3.5.

Despite the usefulness of the approximate geometrical optics method, a number of limitations and discrepancies exist between it and the exact modal analysis. An important case is the analysis of single-mode or few-mode fibers, which must be dealt with by using electromagnetic theory. Problems involving coherence or interference phenomena must also be solved with an electromagnetic approach. In addition, a modal analysis is necessary when a knowledge of the field distribution of individual modes is required. This arises, for example, when analyzing the excitation of an individual mode or when analyzing the coupling of power between modes at waveguide imperfections (which we shall discuss in Chapter 3).

Another discrepancy between the ray optics approach and the modal analysis occurs when an optical fiber is uniformly bent with a constant radius of curvature. As shown in Chapter 3, wave optics correctly

predicts that every mode of the curved fiber experiences some radiation loss. Ray optics, on the other hand, erroneously predicts that some ray congruences can undergo total internal reflection at the curve and, consequently, can remain guided without loss.

2.3.3 Step-Index Fiber Structure

We begin our discussion of light propagation in an optical waveguide by considering the step-index fiber. In practical step-index fibers the core of radius a has a refractive index n_1, which is typically equal to 1.48. This is surrounded by a cladding of slightly lower index n_2, where

$$n_2 = n_1(1 - \Delta) \tag{2.20}$$

The parameter Δ is called the *core-cladding index difference* or simply the *index difference*. Values of n_2 are chosen such that Δ is nominally 0.01. Typical values range from 1 to 3 percent for multimode fibers and from 0.2 to 1.0 percent for single-mode fibers. Since the core refractive index is larger than the cladding index, electromagnetic energy at optical frequencies is made to propagate along the fiber waveguide through internal reflection at the core-cladding interface.

2.3.4 Ray Optics Representation

Since the core size of multimode fibers is much larger than the wavelength of the light we are interested in (which is approximately $1\,\mu$m), an intuitive picture of the propagation mechanism in an ideal multimode step-index optical waveguide is most easily seen by a simple ray (geometrical) optics representation.[6–11] For simplicity, in this analysis we shall consider only a particular ray belonging to a ray congruence that represents a fiber mode. The two types of rays that can propagate in a fiber are meridional rays and skew rays. *Meridional rays* are confined to the meridian planes of the fiber, which are the planes that contain the axis of symmetry of the fiber (the core axis). Since a given meridional ray lies in a single plane, its path is easy to track as it travels along the fiber. Meridional rays can be divided into two general classes: bound rays that are trapped in the core and propagate along the fiber axis according to the laws of geometrical optics, and unbound rays that are refracted out of the fiber core.

Skew rays are not confined to a single plane, but instead tend to follow a helical-type path along the fiber as shown in Fig. 2.16. These rays are more difficult to track as they travel along the fiber because they do not lie in a single plane. Although skew rays constitute a major portion of the total number of guided rays, their analysis is not necessary to obtain a general picture of rays propagating in a fiber. The examination of meridional rays will suffice for this purpose. However, a detailed inclusion of skew
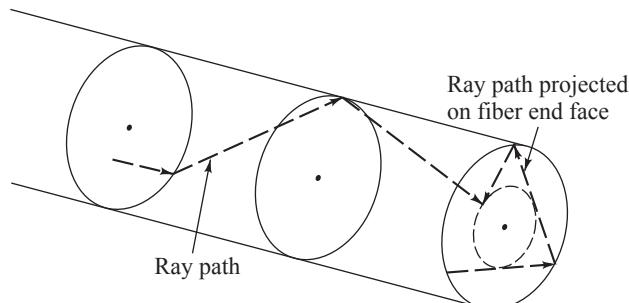
Fig. 2.16 *Ray optics representation of skew rays traveling in a step-index optical fiber core*

rays will change such expressions as the light-acceptance ability of the fiber and power losses of light traveling along a waveguide.[6,10]

A greater power loss arises when skew rays are included in the analyses because many of the skew rays that geometric optics predicts to be trapped in the fiber are actually leaky rays.[5,12,13] These leaky rays are only partially confined to the core of the circular optical fiber and attenuate as the light travels along the optical waveguide. This partial reflection of leaky rays cannot be described by pure ray theory alone. Instead, the analysis of radiation loss arising from these types of rays must be described by mode theory. This is explained further in Sec. 2.4.

Fig. 2.17 *Meridional ray optics representation of the propagation mechanism in an ideal step-index optical waveguide*

The meridional ray is shown in Fig. 2.17 for a step-index fiber. The light ray enters the fiber core from a medium of refractive index n at an angle θ_0 with respect to the fiber axis and strikes the core-cladding interface at a normal angle ϕ. If it strikes this interface at such an angle that it is totally internally reflected, then the meridional ray follows a zigzag path along the fiber core, passing through the axis of the guide after each reflection.

From Snell's law, the minimum or critcal angle ϕ_c that supports total internal reflection for the meridional ray is given by

$$\sin \phi_c = \frac{n_2}{n_1} \tag{2.21}$$

Rays striking the core-cladding interface at angles less than ϕ_c will refract out of the core and be lost in the cladding, as the dashed line shows. By applying Snell's law to the air–fiber face boundary, the condition of Eq. (2.21) can be related to the maximum entrance angle $\theta_{0,\,max}$, which is called the *acceptance angle* θ_A, through the relationship

$$n \sin \theta_{0,\,max} = n \sin \theta_A = n_1 \sin \theta_c = \left(n_1^2 - n_2^2 \right)^{1/2} \tag{2.22}$$

where $\theta_c = \pi/2 - \phi_c$. Thus those rays having entrance angles θ_0 less than θ_A will be totally internally reflected at the core–cladding interface. Thus θ_A defines an acceptance cone for an optical fiber.

Equation (2.22) also defines the *numerical aperture* (NA) of a step-index fiber for meridional rays:

$$\text{NA} = n \sin \theta_A = \left(n_1^2 - n_2^2 \right)^{1/2} \approx n_1 \sqrt{2\Delta} \tag{2.23}$$

The approximation on the right-hand side is valid for the typical case where Δ, as defined by Eq. (2.20), is much less than 1. Since the numerical aperture is related to the acceptance angle, it is commonly used to describe the light acceptance or gathering capability of a fiber and to calculate source-to-fiber optical power coupling efficiencies. This is detailed in Chapter 5. The numerical aperture is a dimensionless quantity which is less than unity, with values normally ranging from 0.14 to 0.50.
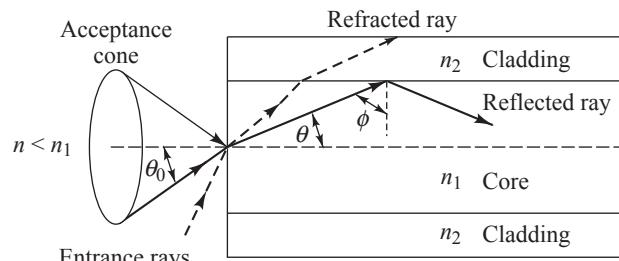
Example 2.4 Consider a multimode silica fiber that has a core refractive index $n_1 = 1.480$ and a cladding index $n_2 = 1.460$. Find (a) the critical angle, (b) the numerical aperture, and (c) the acceptance angle.

Solution: (a) From Eq. (2.21), the critical angle is given by

$$\varphi_c = \sin^{-1} \frac{n_2}{n_1} = \sin^{-1} \frac{1.460}{1.480} = 80.5°$$

(b) From Eq. (2.23) the numerical aperture is

$$NA = \left(n_1^2 - n_2^2 \right)^{1/2} = 0.242$$

(c) From Eq. (2.22) the acceptance angle in air ($n = 1.00$) is

$$\theta_A = \sin^{-1} NA = \sin^{-1} 0.242 = 14°$$

Example 2.5 Consider a multimode fiber that has a core refractive index of 1.480 and a core-cladding index difference 2.0 percent ($\Delta = 0.020$). Find the (a) numerical aperture, (b) the acceptance angle, and (c) the critical angle.

Solution: From Eq. (2.20), the cladding index is $n_2 = n_1(1 - \Delta) = 1.480(0.980) = 1.450$.

(a) From Eq. (2.23) we find that the numerical aperture is

$$NA = n_1 \sqrt{2\Delta} = 1.480(0.04)^{1/2} = 0.296$$

(b) Using Eq. (2.22) the acceptance angle in air ($n = 1.00$) is

$$\theta_A = \sin^{-1} NA = \sin^{-1} 0.296 = 17.2°$$

(c) From Eq. (2.21) the critical angle at the core–cladding interface is

$$\varphi_c = \sin^{-1} \frac{n_2}{n_1} = \sin^{-1} 0.980 = 78.5°$$

2.3.5 Wave Representation in a Dielectric Slab Waveguide

Referring to Fig. 2.17, the ray theory appears to allow rays at any angle ϕ greater than the critical angle ϕ_c to propagate along the fiber. However, when the interference effect due to the phase of the plane wave associated with the ray is taken into account, it is seen that only waves at certain discrete angles greater than or equal to ϕ_c are capable of propagating along the fiber.

To see this, let us consider wave propagation in an infinite dielectric slab waveguide of thickness d. Its refractive index n_1 is greater than the index n_2 of the material above and below the slab. A wave will thus propagate in this guide through multiple reflections, provided that the angle of incidence with respect to the upper and lower surfaces satisfies the condition given in Eq. (2.22).

Figure 2.18 shows the geometry of the waves reflecting at the material interfaces. Here, we consider two rays, designated ray 1 and ray 2, associated with the same wave. The rays are incident on the material interface at an angle $\theta < \theta_c = \pi/2 - \phi_c$. The ray paths in Fig. 2.18 are denoted by solid lines and their associated constant-phase fronts by dashed lines.

The condition required for wave propagation in the dielectric slab is that all points on the same phase front of a plane wave must be in phase. This means that the phase change occurring in ray 1 when traveling from point A to point B minus the phase change in ray 2 between points C and D must differ by an integer multiple of 2π. As the wave travels through the material, it undergoes a phase shift Δ given by

$$\Delta = k_1 s = n_1 k s = n_1 2\pi s/\lambda$$

where k_1 = the propagation constant in the medium of refractive index n_1
$k = k_1/n_1$ is the free-space propagation constant
s = the distance the wave has traveled in the material
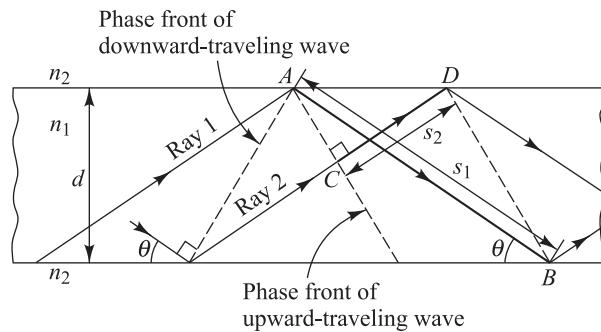
Fig. 2.18 *Light wave propagating along a fiber waveguide. Phase changes occur both as the wave travels through the fiber medium and at the reflection points.*

The phase of the wave changes not only as the wave travels but also upon reflection from a dielectric interface, as described in Sec. 2.2.

In going from point A to point B, ray 1 travels a distance $s_1 = d/\sin\theta$ in the material, and undergoes two phase changes δ at the reflection points. Ray 2 does not incur any reflections in going from point C to point D. To determine its phase change, first note that the distance from point A to point D is $\overline{AD} = (d/\tan\theta) - d\tan\theta$. Thus, the distance between points C and D is

$$s_2 = \overline{AD}\cos\theta = (\cos^2\theta - \sin^2\theta)\, d/\sin\theta$$

The requirement for wave propagation can then be written as

$$\frac{2\pi n_1}{\lambda}(s_1 - s_2) + 2\delta = 2\pi m \tag{2.24a}$$

where $m = 0, 1, 2, 3, \ldots$. Substituting the expressions for s_1 and s_2 into Eq. (2.24a) then yields

$$\frac{2\pi n_1}{\lambda}\left\{\frac{d}{\sin\theta} - \left[\frac{(\cos^2\theta - \sin^2\theta)\, d}{\sin\theta}\right]\right\} + 2\delta = 2\pi m \tag{2.24b}$$

which can be reduced to

$$\frac{2\pi n_1 d \sin\theta}{\lambda} + \delta = \pi m \tag{2.24c}$$

Considering only electric waves with components normal to the plane of incidence, we have from Eq. (2.19a) that the phase shift upon reflection is

$$\delta = -2\arctan\left[\frac{\sqrt{\cos^2\theta - \left(n_2^2/n_1^2\right)}}{\sin\theta}\right] \tag{2.25}$$

The negative sign is needed here since the wave in the medium must be a decaying and not a growing wave. Substituting this expression into Eq. (2.24c) yields

$$\frac{2\pi n_1 d \sin\theta}{\lambda} - \pi m = 2\arctan\left[\frac{\sqrt{\cos^2\theta - \left(n_2^2/n_1^2\right)}}{\sin\theta}\right] \tag{2.26a}$$

or

$$\tan\left(\frac{\pi n_1 d \sin\theta}{\lambda} - \frac{\pi m}{2}\right) = \left[\frac{\sqrt{n_1^2 \cos^2\theta - n_2^2}}{n_1 \sin\theta}\right] \tag{2.26b}$$

Thus only waves that have those angles θ that satisfy the condition in Eq. (2.26) will propagate in the dielectric slab waveguide (see Prob. 2.13).

2.4 Mode Theory for Circular Waveguides

To attain a more detailed understanding of the optical power propagation mechanism in a fiber, it is necessary to solve Maxwell's equations subject to the cylindrical boundary conditions at the interface between the core and the cladding of the fiber. This has been done in extensive detail in a number of works.[7,10,14–18] Since a complete treatment is beyond the scope of this book, only a general outline of a simplified (but still complex) analysis will be given here.

Before presenting the basic mode theory in circular optical fibers, in Sec. 2.4.1 we first give a qualitative overview of the concepts of modes in a waveguide. Next, Sec. 2.4.2 presents a brief summary of the fundamental results obtained from the analyses in Secs. 2.4.3 through 2.4.9, so that those who are not familiar with Maxwell's equations can skip over those sections designated by a star (*) without loss of continuity.

When solving Maxwell's equations for hollow metallic waveguides, only transverse electric (TE) modes and transverse magnetic (TM) modes are found. However, in optical fibers the core-cladding boundary conditions lead to a coupling between the electric and magnetic field components. This gives rise to hybrid modes, which makes optical waveguide analysis more complex than metallic waveguide analysis. The hybrid modes are designated as HE or EH modes, depending on whether the transverse electric field (the E field) or the transverse magnetic field (the H field) is larger for that mode. The two lowest-order modes are designated by HE_{11} and TE_{01}, where the subscripts refer to possible modes of propagation of the optical field.

Although the theory of light propagation in optical fibers is well understood, a complete description of the guided and radiation modes requires the use of six-component hybrid electromagnetic fields that have very involved mathematical expressions. A simplification[19–23] of these expressions can be carried out, in practice, since fibers usually are constructed so that the difference in the core and cladding indices of refraction is very small; that is, $n_1 - n_2 \ll 1$. With this assumption, only four field components need to be considered and their expressions become significantly simpler. The field components are called *linearly polarized* (LP) modes and are labeled LP_{jm} where j and m are integers designating mode solutions. In this scheme for the lowest-order modes, each LP_{0m} mode is derived from an HE_{1m} mode and each LP_{1m} mode comes from TE_{0m}, TM_{0m}, and HE_{0m} modes. Thus the fundamental LP_{01} mode corresponds to an HE_{11} mode.

Although the analysis required for even these simplifications is still fairly involved, this material is key to understanding the principles of optical fiber operation. In Secs. 2.4.3 through 2.4.9 we first solve Maxwell's equations for a circular step-index waveguide and then describe the resulting solutions for some of the lower-order modes.

2.4.1 Overview of Modes

Before we progress with a discussion of mode theory in circular optical fibers, let us qualitatively examine the appearance of modal fields in the planar dielectric slab waveguide shown in Fig. 2.19. The core of this waveguide is a dielectric slab of index n_1 that is sandwiched between two dielectric layers that have refractive indices $n_2 < n_1$. These surrounding layers are called the *cladding*. This represents the simplest form of an optical waveguide and can serve as a model to gain an understanding of wave propagation in optical fibers. In fact, a cross-sectional view of the slab waveguide looks the same as the cross-sectional view of an optical fiber cut along its axis. Figure 2.19 shows the field patterns of several of the lower-order transverse electric (TE) modes (which are solutions of Maxwell's equations for the slab waveguide[7–10]). The *order* of a mode is equal to the number of field zeros across the guide. The order of the mode is also related to the angle that the ray congruence corresponding to this mode makes with the plane of the waveguide (or the axis of a fiber); that is, the steeper the angle, the higher the order of the mode. The plots show that the electric fields of the guided modes are not completely confined to the central dielectric slab (i.e., they do not go to zero at the guide-cladding interface), but, instead, they extend partially into the cladding. The fields vary harmonically in the guiding region of refractive index n_1 and decay exponentially outside of this region. For low-order modes the fields are tightly concentrated near the center of the slab (or the axis of an optical fiber), with little penetration into the cladding region. On the other hand, for higher-order modes the fields are distributed more toward the edges of the guide and penetrate farther into the cladding region.

Solving Maxwell's equations shows that, in addition to supporting a finite number of guided modes, the optical fiber waveguide has an infinite continuum of *radiation modes* that are not trapped in the core
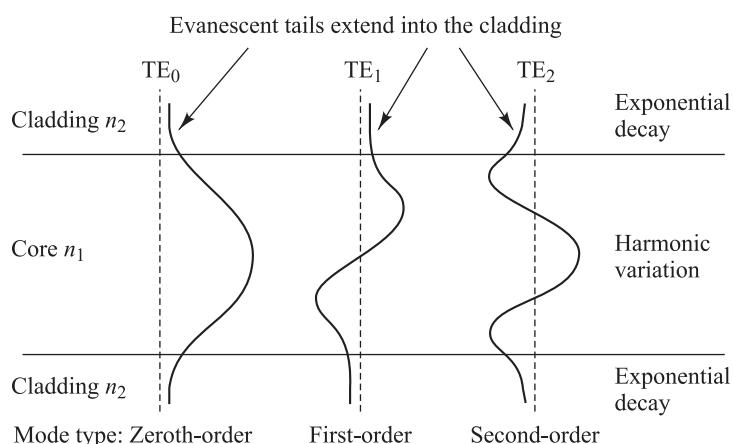
Fig. 2.19 *Electric field distributions for several of the lower-order guided modes in a symmetrical-slab waveguide*

and guided by the fiber but are still solutions of the same boundary-value problem. The radiation field basically results from the optical power that is outside the fiber acceptance angle being refracted out of the core. Because of the finite radius of the cladding, some of this radiation gets trapped in the cladding, thereby causing cladding modes to appear. As the core and cladding modes propagate along the fiber, mode coupling occurs between the cladding modes and the higher-order core modes. This coupling occurs because the electric fields of the guided core modes are not completely confined to the core but extend partially into the cladding (see Fig. 2.19) and likewise for the cladding modes. A diffusion of power back and forth between the core and cladding modes thus occurs; this generally results in a loss of power from the core modes.

Guided modes in the fiber occur when the values for β satisfy the condition $n_2 k < \beta < n_1 k$. At the limit of propagation when $\beta = n_2 k$, a mode is no longer properly guided and is called being *cut off*. Thus unguided or radiation modes appear for frequencies below the cutoff point where $\beta < n_2 k$. However, wave propagation can still occur below cutoff for those modes where some of the energy loss due to radiation is blocked by an angular momentum barrier that exists near the core-cladding interface.[17] These propagation states behave as partially confined guided modes rather than radiation modes and are called *leaky modes*.[5,6,12,13] These leaky modes can travel considerable distances along a fiber but lose power through leakage or tunneling into the cladding as they propagate.

2.4.2 Summary of Key Modal Concepts

An important parameter connected with the cutoff condition is the *V number* defined by

$$V = \frac{2\pi a}{\lambda}\left(n_1^2 - n_2^2\right)^{1/2} = \frac{2\pi a}{\lambda}\,\text{NA} \tag{2.27}$$

This parameter is a dimensionless number that determines how many modes a fiber can support. Except for the lowest-order HE_{11} mode, each mode can exist only for values of V that exceed a certain limiting value (with each mode having a different V limit). The modes are cut off when $\beta = n_2 k$. This occurs when $V \leq 2.405$. The HE_{11} mode has no cutoff and ceases to exist only when the core diameter is zero. This is the principle on which single-mode fibers are based. The details for these and other modes are given in Sec. 2.4.7.

Example 2.6 A step-index fiber has a normalized frequency $V = 26.6$ at a 1300-nm wavelength. If the core radius is 25 μm, what is the numerical aperture?

Solution: From Eq. (2.27) the NA is

$$NA = V\frac{\lambda}{2\pi a} = 26.6\,\frac{1.30\,\mu\text{m}}{2\pi \times 25\,\mu\text{m}} = 0.22$$

The *V* number also can be used to express the number of modes M in a multimode step-index fiber when V is large (see Sec. 2.6 for modes in a graded-index multimode fiber). For this case, an estimate of the total number of modes supported in such a fiber is

$$M \approx \frac{1}{2}\left(\frac{2\pi a}{\lambda}\right)^2\left(n_1^2 - n_2^2\right) = \frac{V^2}{2} \tag{2.28}$$

Example 2.7   Consider a multimode step-index fiber with a 62.5-μm core diameter and a core-cladding index difference of 1.5 percent. If the core refractive index is 1.480, estimate the normalized frequency of the fiber and the total number of modes supported in the fiber at a wavelength of 850 nm.

Solution: From Eq. (2.27) the normalized frequency is

$$V \approx \frac{2\pi a}{\lambda} n_1 \sqrt{2\Delta} = \frac{2\pi \times 31.25\,\mu\text{m} \times 1.48}{0.85\,\mu\text{m}} \sqrt{2 \times 0.015}$$
$$= 59.2$$

Using Eq. (2.28), the total number of modes is

$$M \approx \frac{V^2}{2} = 1752$$

Example 2.8   Suppose we have a multimode step-index optical fiber that has a core radius of 25 μm, a core index of 1.48, and an index difference $\Delta = 0.01$. What are the number of modes in the fiber at wavelengths 860, 1310, and 1550 nm?

Solution:

(a) First, from Eqs. (2.23) and (2.28), at an operating wavelength of 860 nm the value of V is

$$V \approx \frac{2\pi a}{\lambda} n_1 \sqrt{2\Delta} = \frac{2\pi \times 25\,\mu\text{m} \times 1.48}{0.86\,\mu\text{m}} \sqrt{2 \times 0.01}$$
$$= 38.2$$

Using Eq. (2.28), the total number of modes at 860 nm is

$$M \approx \frac{V^2}{2} = 729$$

(b) Similarly, at 1310 nm we have $V = 25.1$ and $M = 315$.

(c) Finally at 1550 nm we have $V = 21.2$ and $M = 224$.

Example 2.9   Suppose we have three multimode step-index optical fibers each of which has a core index of 1.48 and an index difference $\Delta = 0.01$. Assume the three fibers have core diameters of 50, 62.5, and 100 μm. What are the number of modes in these fibers at a wavelength of 1550 nm?

Solution: (a) First, from Eqs. (2.23) and (2.28), at 50-μm diameter the value of V is

$$V \approx \frac{2\pi a}{\lambda} n_1 \sqrt{2\Delta} = \frac{2\pi \times 25\,\mu\text{m} \times 1.48}{1.55\,\mu\text{m}} \sqrt{2 \times 0.01}$$
$$= 21.2$$

Using Eq. (2.28), the total number of modes in the 50-μm core diameter fiber is

$$M \approx \frac{V^2}{2} = 224$$

(b) Similarly, at 62.5-μm we have $V = 26.5$ and $M = 351$.

(c) Finally at 100-μm we have $V = 42.4$ and $M = 898$.

Since the field of a guided mode extends partly into the cladding, as shown in Fig. 2.19, a final quantity of interest for a step-index fiber is the fractional power flow in the core and cladding for a given mode. As the V number approaches cutoff for any particular mode, more of the power of that mode is in the cladding. At the cutoff point, the mode becomes radiative with all the optical power of the mode residing in the cladding. Far from cutoff—that is, for large values of V—the fraction of the average optical power residing in the cladding can be estimated by

$$\frac{P_{\text{clad}}}{P} \approx \frac{4}{3\sqrt{M}} \tag{2.29}$$

where P is the total optical power in the fiber. The details for the power distribution between the core and the cladding of various LP_{jm} modes are given in Sec. 2.4.9. Note that since M is proportional to V^2, the power flow in the cladding decreases as V increases. However, this increases the number of modes in the fiber, which is not desirable for a high-bandwidth capability.

Example 2.10 Consider a multimode step-index optical fiber that has a core radius of 25 μm, a core index of 1.48, and an index difference $\Delta = 0.01$. Find the percentage of optical power that propagates in the cladding at 840 nm.

Solution: From Eqs. (2.23) and (2.28), at an operating wavelength of 840 nm the value of V is

$$V \approx \frac{2\pi a}{\lambda} n_1 \sqrt{2\Delta} = \frac{2\pi \times 25\,\mu\text{m} \times 1.48}{0.84\,\mu\text{m}} \sqrt{2 \times 0.01}$$
$$= 39$$

Using Eq. (2.28), the total number of modes is

$$M \approx \frac{V^2}{2} = 760$$

From Eq. (2.29) we have

$$\frac{P_{\text{clad}}}{P} \approx \frac{4}{3\sqrt{M}} = 0.05$$

Thus approximately 5 percent of the optical power propagates in the cladding. If Δ is decreased to 0.03 in order to lower the signal dispersion (see Chapter 3), then there are 242 modes in the fiber and about 9 percent of the power propagates in the cladding.

2.4.3   Maxwell's Equations*

To analyze the optical waveguide we need to consider Maxwell's equations that give the relationships between the electric and magnetic fields. Assuming a linear, isotropic dielectric material having no currents and free charges, these equations take the form[2]

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{2.30a}$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} \tag{2.30b}$$

$$\nabla \cdot \mathbf{D} = 0 \tag{2.30c}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2.30d}$$

where $\mathbf{D} = \epsilon\mathbf{E}$ and $\mathbf{B} = \mu\mathbf{H}$. The parameter ϵ is the permittivity (or dielectric constant) and μ is the permeability of the medium.

A relationship defining the wave phenomena of the electromagnetic fields can be derived from Maxwell's equations. Taking the curl of Eq. (2.30a) and making use of Eq. (2.30b) yields

$$\nabla \times (\nabla \times \mathbf{E}) = -\mu\frac{\partial}{\partial t}(\nabla \times \mathbf{H}) = -\epsilon\mu\frac{\partial^2 \mathbf{E}}{\partial t^2} \tag{2.31a}$$

Using the vector identity (see App. B),

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}$$

and using Eq. (2.30*c*) (i.e., $\nabla \cdot \mathbf{E} = 0$), Eq. (2.31*a*) becomes

$$\nabla^2 \mathbf{E} = \epsilon\mu \frac{\partial^2 \mathbf{E}}{\partial t^2} \qquad\qquad (2.31b)$$

Similarly, by taking the curl of Eq. (2.30*b*), it can be shown that

$$\nabla^2 \mathbf{H} = \epsilon\mu \frac{\partial^2 \mathbf{H}}{\partial t^2} \qquad\qquad (2.31c)$$

Equations (2.31*b*) and (2.31*c*) are the standard *wave equations*.

2.4.4   Waveguide Equations*

Consider electromagnetic waves propagating along the cylindrical fiber shown in Fig. 2.20. For this fiber, a cylindrical coordinate system $\{r, \phi, z\}$ is defined with the *z* axis lying along the axis of the waveguide. If the electromagnetic waves are to propagate along the *z* axis, they will have a functional dependence of the form

$$\mathbf{E} = \mathbf{E}_0(r, \phi)\, e^{j(\omega t - \beta z)} \qquad\qquad (2.32a)$$

$$\mathbf{H} = \mathbf{H}_0(r, \phi)\, e^{j(\omega t - \beta z)} \qquad\qquad (2.32b)$$

which are harmonic in time *t* and coordinate *z*. The parameter β is the *z* component of the propagation vector and will be determined by the boundary conditions on the electromagnetic fields at the core-cladding interface described in Sec. 2.4.6. When Eqs (2.32*a*) and (2.32*b*) are substituted into Maxwell's curl equations, we have, from Eq. (2.30*a*)

$$\frac{1}{r}\left(\frac{\partial E_z}{\partial \phi} + jr\beta E_\phi \right) = -j\omega\mu H_r \qquad\qquad (2.33a)$$
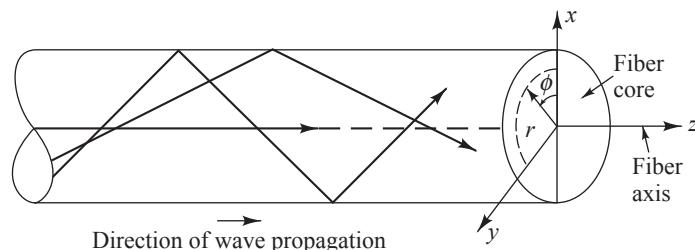
Fig. 2.20   *Cylindrical coordinate system used for analyzing electromagnetic wave propagation in an optical fiber*

$$j\beta E_r + \frac{\partial E_z}{\partial r} = j\omega\mu H_\phi \tag{2.33b}$$

$$\frac{1}{r}\left(\frac{\partial}{\partial r}(rE_\phi) - \frac{\partial E_r}{\partial \phi} \right) = -j\mu\omega H_z \tag{2.33c}$$

and, from Eq. (2.30b),

$$\frac{1}{r}\left(\frac{\partial H_z}{\partial \phi} + jr\beta H_\phi \right) = j\epsilon\omega E_r \tag{2.34a}$$

$$j\beta H_r + \frac{\partial H_z}{\partial r} = -j\epsilon\omega E_\phi \tag{2.34b}$$

$$\frac{1}{r}\left[\frac{\partial}{\partial r}(rH_\phi) - \frac{\partial H_r}{\partial \phi} \right] = j\epsilon\omega E_z \tag{2.34c}$$

By eliminating variables these equations can be rewritten such that, when E_z and H_z are known, the remaining transverse components E_r, E_ϕ, H_r, and H_ϕ can be determined. For example, E_ϕ or H_r can be eliminated from Eqs (2.33a) and (2.34b) so that the component H_ϕ or E_r, respectively, can be found in terms of E_z or H_z. Doing so yields

$$E_r = -\frac{j}{q^2}\left(\beta\frac{\partial E_z}{\partial r} + \frac{\mu\omega}{r}\frac{\partial H_z}{\partial \phi} \right) \tag{2.35a}$$

$$E_\phi = -\frac{j}{q^2}\left(\frac{\beta}{r}\frac{\partial E_z}{\partial \phi} - \mu\omega\frac{\partial H_z}{\partial r} \right) \tag{2.35b}$$

$$H_r = \frac{-j}{q^2}\left(\beta\frac{\partial H_z}{\partial r} - \frac{\omega\epsilon}{r}\frac{\partial E_z}{\partial \phi} \right) \tag{2.35c}$$

$$H_\phi = \frac{-j}{q^2}\left(\frac{\beta}{r}\frac{\partial H_z}{\partial \phi} + \omega\epsilon\frac{\partial E_z}{\partial r} \right) \tag{2.35d}$$

where $q^2 = \omega^2\epsilon\mu - \beta^2 = k^2 - \beta^2$.

Substitution of Eqs (2.35c) and (2.35d) into Eq. (2.34c) results in the wave equation in cylindrical coordinates,

$$\frac{\partial^2 E_z}{\partial r^2} + \frac{1}{r}\frac{\partial E_z}{\partial r} + \frac{1}{r^2}\frac{\partial^2 E_z}{\partial \phi^2} + q^2 E_z = 0 \tag{2.36}$$

and substitution of Eqs (2.35*a*) and (2.35*b*) into Eq. (2.33*c*) leads to

$$\frac{\partial^2 H_z}{\partial r^2} + \frac{1}{r}\frac{\partial H_z}{\partial r} + \frac{1}{r^2}\frac{\partial^2 H_z}{\partial \phi^2} + q^2 H_z = 0 \tag{2.37}$$

It is interesting to note that Eqs. (2.36) and (2.37) each contain either only E_z or only H_z. This appears to imply that the longitudinal components of **E** and **H** are uncoupled and can be chosen arbitrarily provided that they satisfy Eqs (2.36) and (2.37). However, in general, coupling of E_z and H_z is required by the boundary conditions of the electromagnetic field components described in Sec. 2.4.6. If the boundary conditions do not lead to coupling between the field components, mode solutions can be obtained in which either $E_z = 0$ or $H_z = 0$. When $E_z = 0$ the modes are called *transverse electric* or TE modes, and when $H_z = 0$ they are called *transverse magnetic* or TM modes. *Hybrid* modes exist if both E_z and H_z are nonzero. These are designated as HE or EH modes, depending on whether H_z or E_z, respectively, makes a larger contribution to the transverse field. The fact that the hybrid modes are present in optical waveguides makes their analysis more complex than in the simpler case of hollow metallic waveguides where only TE and TM modes are found.

2.4.5 Wave Equations for Step-Index Fibers*

We now use the above results to find the guided modes in a step-index fiber. A standard mathematical procedure for solving equations such as Eq. (2.36) is to use the separation-of-variables method, which assumes a solution of the form

$$E_z = AF_1(r)F_2(\phi)F_3(z)F_4(t) \tag{2.38}$$

As was already assumed, the time- and *z*-dependent factors are given by

$$F_3(z)F_4(t) = e^{j(\omega t - \beta z)} \tag{2.39}$$

since the wave is sinusoidal in time and propagates in the *z* direction. In addition, because of the circular symmetry of the waveguide, each field component must not change when the coordinate ϕ is increased by 2π. We thus assume a periodic function of the form

$$F_2(\phi) = e^{j\nu\phi} \tag{2.40}$$

The constant ν can be positive or negative, but it must be an integer since the fields must be periodic in ϕ with a period of 2π.

Substituting Eq. (2.40) into Eq. (2.38), the wave equation for E_z [Eq. 2.36)] becomes

$$\frac{\partial^2 F_1}{\partial r^2} + \frac{1}{r}\frac{\partial F_1}{\partial r} + \left(q^2 - \frac{\nu^2}{r^2}\right)F_1 = 0 \tag{2.41}$$

which is a well-known differential equation for Bessel functions.[24–26] An exactly identical equation can be derived for H_z.

For the configuration of the step-index fiber we consider a homogeneous core of refractive index n_1 and radius *a*, which is surrounded by an infinite cladding of index n_2. The reason for assuming an infinitely thick cladding is that the guided modes in the core have exponentially decaying fields outside the

core and these must have insignificant values at the outer boundary of the cladding. In practice, optical fibers are designed with claddings that are sufficiently thick so that the guided-mode field does not reach the outer boundary of the cladding. To get an idea of the field patterns, the electric field distributions for several of the lower-order guided modes in a symmetrical-slab waveguide were shown in Fig. 2.14. The fields vary harmonically in the guiding region of refractive index n_1 and decay exponentially outside of this region.

Equation (2.41) must now be solved for the regions inside and outside the core. For the inside region the solutions for the guided modes must remain finite as $r \to 0$, whereas on the outside the solutions must decay to zero as $r \to \infty$. Thus, for $r < a$ the solutions are Bessel functions of the first kind of order v. For these functions we use the common designation $J_v(ur)$. Here, $u^2 = k_1^2 - \beta^2$ with $k_1 = 2\pi n_1/\lambda$. The expressions for E_z and H_z inside the core are thus

$$E_z(r < a) = A J_v(ur)\, e^{jv\phi} e^{j(\omega t - \beta z)} \tag{2.42}$$

$$H_z(r < a) = B J_v(ur)\, e^{jv\phi} e^{j(\omega t - \beta z)} \tag{2.43}$$

where A and B are arbitrary constants.

Outside of the core the solutions to Eq. (2.41) are given by modified Bessel functions of the second kind, $K_v(wr)$, where $w^2 = \beta^2 - k_2^2$ with $k_2 = 2\pi n_2/\lambda$. The expressions for E_z and H_z outside the core are therefore

$$E_z(r > a) = C K_v(wr)\, e^{jv\phi} e^{j(\omega t - \beta z)} \tag{2.44}$$

$$H_z(r > a) = D K_v(wr)\, e^{jv\phi} e^{j(\omega t - \beta z)} \tag{2.45}$$

with C and D being arbitrary constants.

The definitions of $J_v(ur)$ and $K_v(wr)$ and various recursion relations are given in App. C. From the definition of the modified Bessel function, it is seen that $K_v(wr) \to e^{-wr}$ as $wr \to \infty$. Since $K_v(wr)$ must go to zero as $r \to \infty$, it follows that $w > 0$. This, in turn, implies that $\beta \geq k_2$, which represents a cutoff condition. The *cutoff condition* is the point at which a mode is no longer bound to the core region. A second condition on β can be deduced from the behavior of $J_v(ur)$. Inside the core the parameter u must be real for F_1 to be real, from which it follows that $k_1 \geq \beta$. The permissible range of β for bound solutions is therefore

$$n_2 k = k_2 \leq \beta \leq k_1 = n_1 k \tag{2.46}$$

where $k = 2\pi/\lambda$ is the free-space propagation constant.

2.4.6 Modal Equation*

The solutions for β must be determined from the boundary conditions. The boundary conditions require that the tangential components E_ϕ and E_z of **E** inside and outside of the dielectric interface at $r = a$ must be the same, and similarly for the tangential components H_ϕ and H_z. Consider first the tangential components of **E**. For the z component we have, from Eq. (2.42) at the inner core-cladding boundary ($E_z = E_{z1}$) and from Eq. (2.44) at the outside of the boundary ($E_z = E_{z2}$), that

$$E_{z1} - E_{z2} = A J_v(ua) - C K_v(wa) = 0 \tag{2.47}$$

The ϕ component is found from Eq. (2.35*b*). Inside the core the factor q^2 is given by

$$q^2 = u^2 = k_1^2 - \beta^2 \tag{2.48}$$

where $k_1 = 2\pi n_1/\lambda = \omega\sqrt{\epsilon_1\mu}$, while outside the core

$$w^2 = \beta^2 - k_2^2 \tag{2.49}$$

with $k_2 = 2\pi n_2/\lambda = \omega\sqrt{\epsilon_2\mu}$. Substituting Eqs (2.42) and (2.43) into Eq. (2.35*b*) to find $E_{\phi 1}$, and, similarly, using Eqs. (2.44) and (2.45) to determine $E_{\phi 2}$, yields, at $r = a$,

$$
\begin{aligned}
E_{\phi 1} - E_{\phi 2} &= -\frac{j}{u^2}\left[A\frac{jv\beta}{a}J_v(ua) - B\omega\mu u J_v'(ua) \right] \\
&\quad -\frac{j}{w^2}\left[C\frac{jv\beta}{a}K_v(wa) - D\omega\mu w K_v'(wa) \right] = 0
\end{aligned}
\tag{2.50}
$$

where the prime indicates differentiation with respect to the argument.

Similarly, for the tangential components of **H** it is readily shown that, at $r = a$,

$$H_{z1} - H_{z2} = BJ_v(ua) - DK_v(wa) = 0 \tag{2.51}$$

and

$$
\begin{aligned}
H_{\phi 1} - H_{\phi 2} &= -\frac{j}{u^2}\left[B\frac{jv\beta}{a}J_v(ua) + A\omega\epsilon_1 u J_v'(ua) \right] \\
&\quad -\frac{j}{w^2}\left[D\frac{jv\beta}{a}K_v(wa) + C\omega\epsilon_2\omega K_v'(wa) \right] = 0
\end{aligned}
\tag{2.52}
$$

Equations (2.47), (2.50), (2.51), and (2.52) are a set of four equations with four unknown coefficients, *A, B, C,* and *D*. A solution to these equations exists only if the determinant of these coefficients is zero:

$$
\begin{vmatrix}
J_v(ua) & 0 & -K_v(wa) & 0 \\
\dfrac{\beta v}{au^2}J_v(ua) & \dfrac{j\omega\mu}{u}J_v'(ua) & \dfrac{\beta v}{aw^2}K_v(wa) & \dfrac{j\omega\mu}{w}K_v'(wa) \\
0 & J_v(ua) & 0 & -K_v(wa) \\
-\dfrac{j\omega\epsilon_1}{u}J_v'(ua) & \dfrac{\beta v}{au^2}J_v(ua) & -\dfrac{j\omega\epsilon_2}{w}K_v'(wa) & \dfrac{\beta v}{aw^2}K_v(wa)
\end{vmatrix} = 0
\tag{2.53}
$$

Evaluation of this determinant yields the following eigenvalue equation for β:

$$(\mathscr{J}_v + \mathscr{K}_v)(k_1^2\mathscr{J}_v + k_2^2\mathscr{K}_v) = \left(\frac{\beta v}{a}\right)^2\left(\frac{1}{u^2} + \frac{1}{w^2}\right)^2 \tag{2.54}$$

where

$$\mathscr{J}_v = \frac{J'_v(ua)}{uJ_v(ua)} \quad \text{and} \quad \mathscr{K}_v = \frac{K'_v(wa)}{wK_v(wa)}$$

Upon solving Eq. (2.54) for β, it will be found that only discrete values restricted to the range given by Eq. (2.46) will be allowed. Although Eq. (2.54) is a complicated transcendental equation that is generally solved by numerical techniques, its solution for any particular mode will provide all the characteristics of that mode. We shall now consider this equation for some of the lowest-order modes of a step-index waveguide.
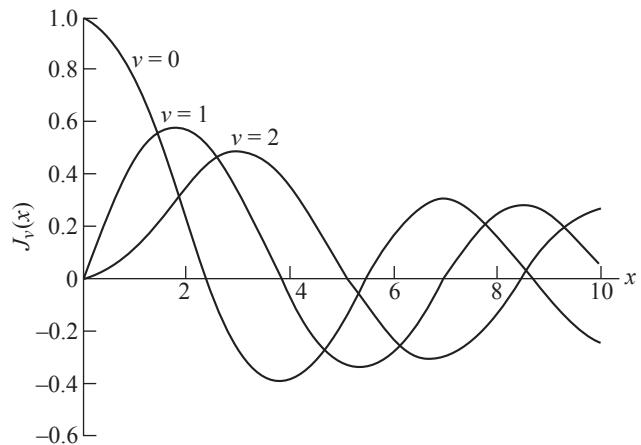
Fig. 2.21 *Variation of the Bessel function $J_v(x)$ for the first three orders ($v = 0, 1, 2$) plotted as a function of x*

2.4.7   Modes in Step-Index Fibers*

To help describe the modes, we shall first examine the behavior of the *J*-type Bessel functions. These are plotted in Fig. 2.21 for the first three orders. The *J*-type Bessel functions are similar to harmonic functions since they exhibit oscillatory behavior for real k, as is the case for sinusoidal functions. Because of the oscillatory behavior of J_v, there will be m roots of Eq. (2.54) for a given v value. These roots will be designated by β_{vm}, and the corresponding modes are either TE_{vm}, TM_{vm}, EH_{vm}, or HE_{vm}. Schematics of the transverse electric field patterns for the four lowest-order modes over the cross section of a step-index fiber are shown in Fig. 2.22.
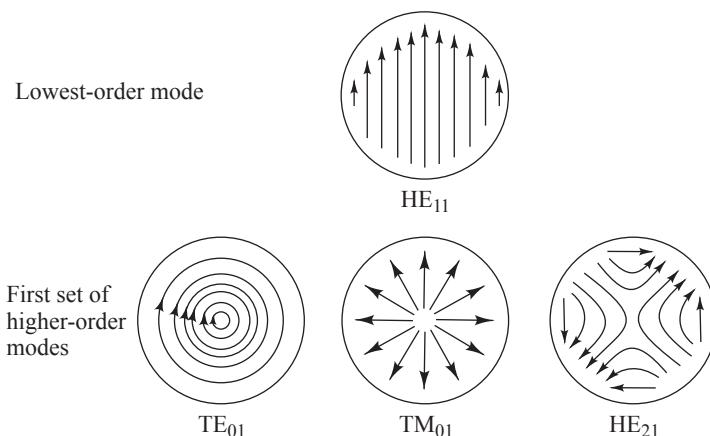
Fig. 2.22 *Fiber end views of the transverse electric field vectors for the four lowest-order modes in a step-index fiber*

For the dielectric fiber waveguide, all modes are hybrid modes except those for which $v = 0$. When $v = 0$, the right-hand side of Eq. (2.54) vanishes and two different eigenvalue equations result. These are

$$\mathcal{J}_0 + \mathcal{K}_0 = 0 \tag{2.55a}$$

or, using the relations for J_v' and K_v' in App. C,

$$\frac{J_1(ua)}{uJ_0(ua)} + \frac{K_1(wa)}{wK_0(wa)} = 0 \tag{2.55b}$$

which corresponds to TE_{0m} modes ($E_z = 0$), and

$$k_1^2 \mathcal{J}_0 + k_2^2 \mathcal{K}_0 = 0 \tag{2.56a}$$

or

$$\frac{k_1^2 J_1(ua)}{uJ_0(ua)} + \frac{k_2^2 K_1(wa)}{wK_0(wa)} = 0 \tag{2.56b}$$

which corresponds to TM_{0m} modes ($H_z = 0$). The proof of this is left as an exercise (see Prob. 2.16).

When $v \neq 0$ the situation is more complex and numerical methods are needed to solve Eq. (2.54) exactly. However, simplified and highly accurate approximations based on the principle that the core and cladding refractive indices are nearly the same have been derived.[19,20,27] The condition that $n_1 - n_2 \ll 1$ is referred to as giving rise to *weakly guided* modes. A treatment of these derivations is given in Sec. 2.4.8.

Let us examine the cutoff conditions for fiber modes. As was mentioned in relation to Eq. (2.46), a mode is referred to as being cut off when it is no longer bound to the core of the fiber, so that its field no longer decays on the outside of the core. The cutoffs for the various modes are found by solving Eq. (2.54) in the limit $w^2 \rightarrow 0$. This is, in general, fairly complex, so that only the results,[14,16] which are listed in Table 2.3, will be given here.

Table 2.3   *Cutoff conditions for some lower-order modes*

v	Mode	Cutoff condition
0	TE_{0m}, TM_{0m}	$J_0(ua) = 0$
1	HE_{1m}, EH_{1m}	$J_1(ua) = 0$
≥ 2	EH_{vm}	$J_v(ua) = 0$
	HE_{vm}	$\left(\dfrac{n_1^2}{n_2^2} + 1\right) J_{v-1}(ua) = \dfrac{ua}{v - 1} J_v(ua)$

An important parameter connected with the cutoff condition is the *normalized frequency V* (also called the *V number* or *V parameter*) defined by

$$V^2 = (u^2 + w^2)a^2 = \left(\frac{2\pi a}{\lambda}\right)^2 (n_1^2 - n_2^2) = \left(\frac{2\pi a}{\lambda}\right)^2 NA^2 \tag{2.57}$$

which is a dimensionless number that determines how many modes a fiber can support. The number of modes that can exist in a waveguide as a function of V may be conveniently represented in terms of a *normalized propagation constant b* defined by[20]

$$b = \frac{a^2 w^2}{V^2} = \frac{(\beta/k)^2 - n_2^2}{n_1^2 - n_2^2}$$

A plot of b (in terms of β/k) as a function of V is shown in Fig. 2.23 for a few of the low-order modes. This figure shows that each mode can exist only for values of V that exceed a certain limiting value. The modes are cut off when $\beta/k = n_2$. The HE_{11} mode has no cutoff and ceases to exist only when the core diameter is zero. This is the principle on which the single-mode fiber is based. By appropriately choosing a, n_1, and n_2 so that

$$V = \frac{2\pi a}{\lambda}(n_1^2 - n_2^2)^{1/2} \le 2.405 \tag{2.58}$$

which is the value at which the lowest-order Bessel function $J_0 = 0$ (see Fig. 2.21), all modes except the HE_{11} mode are cut off.
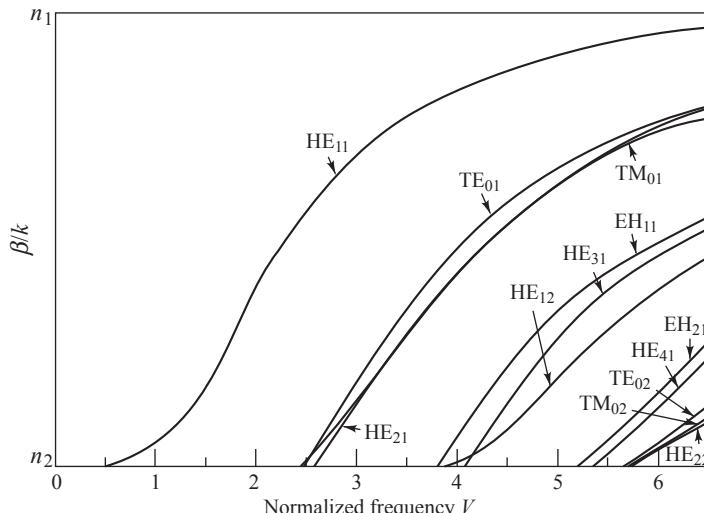
Fig. 2.23 *Plots of the propagation constant (in terms of β/k) as a function of V for a few of the lowest-order modes. (Modified with permission from Gloge.[20])*

The parameter V can also be related to the number of modes M in a multimode fiber when M is large. An approximate relationship for step-index fibers can be derived from ray theory. A ray congruence will be accepted by the fiber if it lies within an angle θ defined by the numerical aperture as given in Eq. (2.23):

$$NA = \sin \theta = (n_1^2 - n_2^2)^{1/2} \tag{2.59}$$

For practical numerical apertures, $\sin \theta$ is small so that $\sin \theta \simeq \theta$. The solid acceptance angle for the fiber is therefore

$$\Omega = \pi\theta^2 = \pi(n_1^2 - n_2^2) \tag{2.60}$$

For electromagnetic radiation of wavelength λ emanating from a laser or a waveguide, the number of modes per unit solid angle is given by $2A/\lambda^2$, where A is the area the mode is leaving or entering.[28] The area A in this case is the core cross section πa^2. The factor 2 comes from the fact that the plane wave can have two polarization orientations. The total number of modes M entering the fiber is thus given by

$$M \simeq \frac{2A}{\lambda^2}\Omega = \frac{2\pi^2 a^2}{\lambda^2}(n_1^2 - n_2^2) = \frac{V^2}{2} \tag{2.61}$$

2.4.8   Linearly Polarized Modes*

In place of a lengthy exact analysis for the modes of a fiber, a simpler but highly accurate approximation can be used, based on the principle that in a typical step-index fiber the difference between the indices of refraction of the core and cladding is very small; that is, $\Delta \ll 1$. This is the basis of the *weakly guiding fiber approximation*.[7,19,20,27] In this approximation the electromagnetic field patterns and the propagation constants of the mode pairs $HE_{v+1, m}$ and $EH_{v-1, m}$ are very similar. This holds likewise for the three modes TE_{0m}, TM_{0m}, and HE_{2m}. This can be seen from Fig. 2.23 with $(v, m) = (0, 1)$ and $(2, 1)$ for the mode groupings $\{HE_{11}\}$, $\{TE_{01}, TM_{01}, HE_{21}\}$, $\{HE_{31}, EH_{11}\}$, $\{HE_{12}\}$, $\{HE_{41}, EH_{21}\}$, and $\{TE_{02}, TM_{02}, HE_{22}\}$. The result is that only four field components need to be considered instead of six, and the field description is further simplified by the use of caretesian instead of cylindrical coordinates.

When $\Delta \ll 1$ we have that $k_1^2 \simeq k_2^2 \simeq \beta^2$. Using these approximations, Eq. (2.54) becomes

$$\mathscr{J}_v + \mathscr{K}_v = \pm\frac{v}{a}\left(\frac{1}{u^2} + \frac{1}{w^2}\right) \tag{2.62}$$

Thus Eq. (2.55b) for TE_{0m} modes is the same as Eq. (2.56b) for TM_{0m} modes. Using the recurrence relations for J'_v and K'_v given in App. C, we get two sets of equations for Eq. (2.62) for the positive and negative signs. The positive sign yields

$$\frac{J_{v+1}(ua)}{uJ_v(ua)} + \frac{K_{v+1}(wa)}{wK_v(wa)} = 0 \tag{2.63}$$

The solution of this equation gives a set of modes called the EH modes. For the negative sign in Eq. (2.62) we get

$$\frac{J_{v-1}(ua)}{uJ_v(ua)} - \frac{K_{v-1}(wa)}{wK_v(wa)} = 0 \tag{2.64a}$$

or, alternatively, taking the inverse of Eq. (2.64a) and using the first expressions for $J_v(ua)$ and $K_v(wa)$ from Sec. C.1.2 and Sec. C.2.2,

$$-\frac{uJ_{v-2}(ua)}{J_{v-1}(ua)} = \frac{wK_{v-2}(wa)}{wK_{v-1}(wa)} \tag{2.64b}$$

This results in a set of modes called the HE modes.

If we define a new parameter

$$j = \begin{cases} 1 & \text{for TE and TM modes} \\ v+1 & \text{for EH modes} \\ v-1 & \text{for HE modes} \end{cases} \tag{2.65}$$

then Eqs (2.55b) (2.63), and (2.64b) can be written in the unified form

$$\frac{uJ_{j-1}(ua)}{J_j(ua)} = -\frac{wK_{j-1}(wa)}{K_j(wa)} \tag{2.66}$$

Equations (2.65) and (2.66) show that within the weakly guiding approximation all modes characterized by a common set of j and m satisfy the same characteristic equation. This means that these modes are degenerate. Thus if an $\text{HE}_{v+1, m}$ mode is degenerate with an $\text{EH}_{v-1, m}$ mode (i.e., if HE and EH modes of corresponding radial order m and equal circumferential order v form degenerate pairs), then any combination of an $\text{HE}_{v+1, m}$ mode with an $\text{EH}_{v-1, m}$ mode will likewise constitute a guided mode of the fiber.

Such degenerate modes are called *linearly polarized* (LP) modes, and are designated LP_{jm} modes regardless of their TM, TE, EH, or HE field configuration.[20] The normalized propagation constant b as a function of V is given for various LP_{jm} modes in Fig. 2.24. In general, we have the following:

1. Each LP_{0m} mode is derived from an HE_{1m} mode.
2. Each LP_{1m} mode comes from TE_{0m}, TM_{0m}, and HE_{2m} modes.
3. Each LP_{vm} mode ($v \geq 2$) is from an $\text{HE}_{v+1, m}$ and an $\text{EH}_{v-1, m}$ mode.

The correspondence between the ten lowest LP modes (i.e., those having the lowest cutoff frequencies) and the traditional TM, TE, EH, and HE modes is given in Table 2.4. This table also shows the number of degenerate modes.

A very useful feature of the LP-mode designation is the ability to readily visualize a mode. In a complete set of modes only one electric and one magnetic field component are significant. The electric field vector **E** can be chosen to lie along an arbitrary axis, with the magnetic field vector **H** being perpendicular to it. In addition, there are equivalent solutions with the field polarities reversed. Since each of the two possible polarization directions can be coupled with either a $\cos j\phi$ or a $\sin j\phi$ azimuthal dependence, four discrete mode patterns can be obtained from a single LP_{jm} label. As an example, the four possible
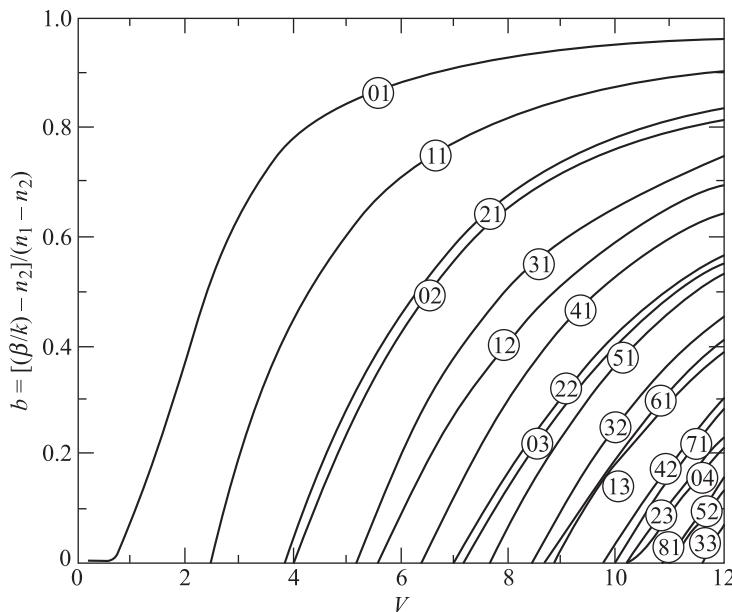
Fig. 2.24   *Plots of the propagation constant b as a function of V for various LP$_{jm}$ modes. (Reproduced with permission from Gloge.[20])*

Table 2.4   *Composition of the lower-order linearly polarized modes*

LP-mode designation	Traditional-mode designation and number of modes	Number of degenerate modes
LP$_{01}$	HE$_{11}\times 2$	2
LP$_{11}$	TE$_{01}$, TM$_{01}$, HE$_{21}\times 2$	4
LP$_{21}$	EH$_{11}\times 2$, HE$_{31}\times 2$	4
LP$_{02}$	HE$_{12}\times 2$	2
LP$_{31}$	EH$_{21}\times 2$, HE$_{41}\times 2$	4
LP$_{12}$	TE$_{02}$, TM$_{02}$, HE$_{22}\times 2$	4
LP$_{41}$	EH$_{31}\times 2$, HE$_{51}\times 2$	4
LP$_{22}$	EH$_{12}\times 2$, HE$_{32}\times 2$	4
LP$_{03}$	HE$_{13}\times 2$	2
LP$_{51}$	EH$_{41}\times 2$, HE$_{61}\times 2$	4
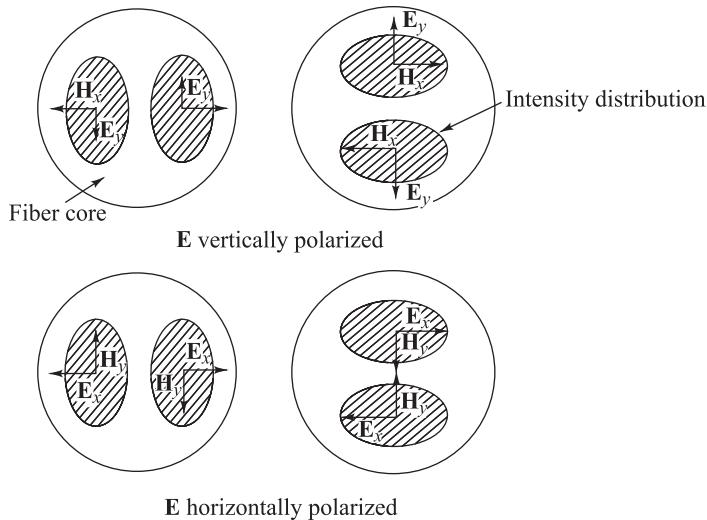
Fig. 2.25 *The four possible transverse electric field and magnetic field directions and the corresponding intensity distributions for the LP$_{11}$ mode*

electric and magnetic field directions and the corresponding intensity distributions for the LP$_{11}$ mode are shown in Fig. 2.25. Figure 2.26*a* and 2.26*b* illustrate how two LP$_{11}$ modes are composed from the exact HE$_{21}$ plus TE$_{01}$ and the exact HE$_{21}$ plus TM$_{01}$ modes, respectively.

2.4.9 Power Flow in Step-Index Fibers*

A final quantity of interest for step-index fibers is the fractional power flow in the core and cladding for a given mode. As illustrated in Fig. 2.19, the electromagnetic field for a given mode does not go to zero at the core-cladding interface, but changes from an oscillating form in the core to an exponential decay in the cladding. Thus the electromagnetic energy of a guided mode is carried partly in the core and partly in the cladding. The farther away a mode is from its cutoff frequency, the more concentrated its energy is in the core. As cutoff is approached, the field penetrates farther into the cladding region and a greater percentage of the energy travels in the cladding. At cutoff the field no longer decays outside the core and the mode now becomes a fully radiating mode.

The relative amounts of power flowing in the core and the cladding can be obtained by integrating the Poynting vector in the axial direction,

$$S_z = \frac{1}{2} \operatorname{Re}(\mathbf{E} \times \mathbf{H}^*) \cdot \mathbf{e}_z \tag{2.67}$$

over the fiber cross section. Thus the powers in the core and cladding, respectively, are given by

$$P_{\text{core}} = \frac{1}{2} \int_0^a \int_0^{2\pi} r(E_x H_y^* - E_y H_x^*)\, d\phi\, dr \tag{2.68}$$

$$P_{\text{clad}} = \frac{1}{2} \int_a^\infty \int_0^{2\pi} r(E_x H_y^* - E_y H_x^*)\, d\phi\, dr \tag{2.69}$$

where the asterisk denotes the complex conjugate. Based on the weakly guided mode approximation, which has an accuracy on the order of the index difference Δ between the core and cladding, the relative core and cladding powers for a particular mode v are given by[20,29]

$$\frac{P_{\text{core}}}{P} = \left(1 - \frac{u^2}{V^2}\right)\left[1 - \frac{J_v^2(ua)}{J_{v+1}(ua)\,J_{v-1}(ua)}\right]$$

$$(2.70)$$

and

$$\frac{P_{\text{clad}}}{P} = 1 - \frac{P_{\text{core}}}{P} \qquad (2.71)$$

where P is the total power in the mode v. The relationships between P_{core} and P_{clad} are plotted in Fig. 2.27 in terms of the fractional powers P_{core}/P and P_{clad}/P for various LP_{jm} modes. In addition, far from cutoff the average total power in the cladding has been derived for fibers in which many modes can propagate. Because of this large number of modes, those few modes that are appreciably close to cutoff can be ignored to a reasonable approximation. The derivation assumes an incoherent source, such as a tungsten filament lamp or a light-emitting diode, which, in general, excites every fiber mode with the same amount of power. The total average cladding power is thus approximated by[20]

$$\left(\frac{P_{\text{clad}}}{P}\right)_{\text{total}} = \frac{4}{3}M^{-1/2} \qquad (2.72)$$

(a)

(b)

Fig. 2.26 *Composition of two LP_{11} modes from exact modes and their transverse electric field and intensity distributions*

<u>***Example 2.11***</u> As an example, consider a fiber having a core radius of 25 μm, a core index of 1.48, and $\Delta = 0.01$. At an operating wavelength of 0.84 μm the value of V is 39 and there are 760 modes in the fiber. From Eq. (2.72), approximately 5 percent of the power propagates in the cladding. If Δ is decreased to, say, 0.003 in order to decrease signal dispersion (see Chapter 3), then 242 modes propagate in the fiber and about 9 percent of the power resides in the cladding. For the case of the single-mode fiber, considering the LP_{01} mode (the HE_{11} mode) in Fig. 2.27, it is seen that for $V = 1$ about 70 percent of the power propagates in the cladding, whereas for $V = 2.405$, which is where the LP_{11} mode (the TE_{01} mode) begins, approximately 84 percent of the power is now within the core.
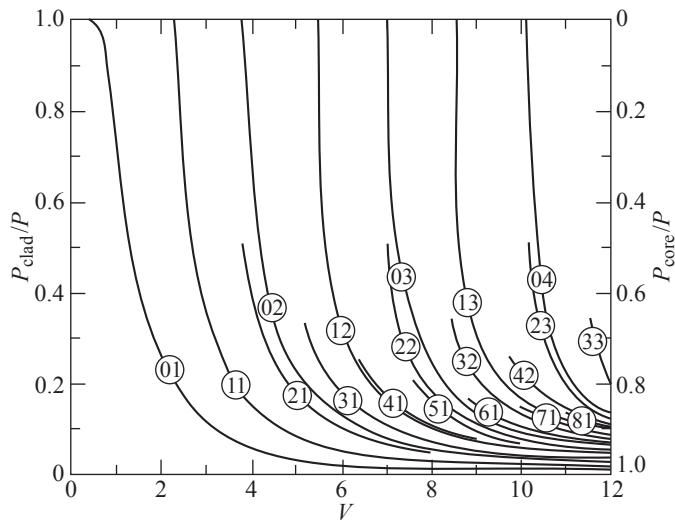
Fig. 2.27 *Fractional power flow in the cladding of a step-index optical fiber as a function of V. When v ≠ 1, the curve numbers vm designate the $HE_{v+1,m}$ and $EH_{v-1,m}$ modes. For v = 1, the curve numbers vm give the HE_{2m}, TE_{0m}, and TM_{0m} modes. (Reproduced with permission from Gloge.[20])*

where, from Eq. (2.61), *M* is the total number of modes entering the fiber. From Fig. 2.27 and Eq. (2.72) it can be seen that, since *M* is proportional to V^2, the power flow in the cladding decreases as *V* increases.

2.5 | Single-Mode Fibers

2.5.1 Construction

Single-mode fibers are constructed by letting the dimensions of the core diameter be a few wavelengths (usually 8–12) and by having small index differences between the core and the cladding. From Eq. (2.27) or (2.58) with *V* = 2.4, it can be seen that single-mode propagation is possible for fairly large variations in values of the physical core size *a* and the core-cladding index differences Δ. However, in practical designs of single-mode fibers,[27] the core-cladding index difference varies between 0.2 and 1.0 percent, and the core diameter should be chosen to be just below the cutoff of the first higher-order mode; that is, for *V* slightly less than 2.4.

2.5.2 Mode-Field Diameter

For multimode fibers the core diameter and numerical aperture are key parameters for describing the signal transmission properties. In single-mode fibers the geometric distribution of light in the propagating mode is what is needed when predicting the performance characteristics of these fibers. Thus a fundamental parameter of a single-mode fiber is the *mode-field diameter* (MFD). This parameter can be determined from the mode-field distribution of the fundamental fiber mode and is a function of the optical source wavelength,

Example 2.12 A manufacturing engineer wants to make an optical fiber that has a core index of 1.480 and a cladding index of 1.478. What should the core size be for single-mode operation at 1550 nm?

Solution: Using the condition that $V \leq 2.405$ must be satisfied for single-mode operation, then from Eq. (2.27) we have

$$a = \frac{V\lambda}{2\pi} \frac{1}{\sqrt{n_1^2 - n_2^2}}$$

$$\leq \frac{2.405 \times 1.55\mu m}{2\pi} \frac{1}{\sqrt{(1.480)^2 - (1.478)^2}} = 7.7\mu m$$

If this fiber also should be single-mode at 1310 nm, then the core radius must be less than 6.50 μm.

Example 2.13 An applications engineer has an optical fiber that has a 3.0-μm core radius and a numerical aperture of 0.1. Will this fiber exhibit single-mode operation at 800 nm?

Solution: From Eq. (2.27)

$$V \approx \frac{2\pi a}{\lambda} NA = \frac{2\pi \times 3\mu m}{0.80\mu m} 0.10 = 2.356$$

Since $V < 2.405$, this fiber will exhibit single-mode operation at 800 nm.

the core radius, and the refractive index profile of the fiber. The mode-field diameter is analogous to the core diameter in multimode fibers, except that in single-mode fibers not all the light that propagates through the fiber is carried in the core (see Sec. 2.4). Figure 2.28 illustrates this effect. For example, at $V = 2$ only 75 percent of the optical power is confined to the core. This percentage increases for larger values of V and is less for smaller V values.

The MFD is an important parameter for single-mode fiber because it is used to predict fiber properties such as splice loss, bending loss, cutoff wavelength, and waveguide dispersion. Chapters 3 and 5 describe these parameters and their effects on fiber performance. A variety of models have been proposed for characterizing and measuring the MFD.[30–35] These include far-field scanning, near-field scanning, transverse offset, variable aperture in the far field, knife-edge, and mask methods.[30] The main consideration of all these methods is how to approximate the optical power distribution.

A standard technique to find the MFD is to measure the far-field intensity distribution $E^2(r)$ and then calculate the MFD using the Petermann II equation[32]

$$\text{MFD} = 2w_0 = 2 \left[\frac{2\int_0^\infty E^2(r)r^3 \, dr}{\int_0^\infty E^2(r)r \, dr} \right]^{1/2} \qquad (2.73)$$
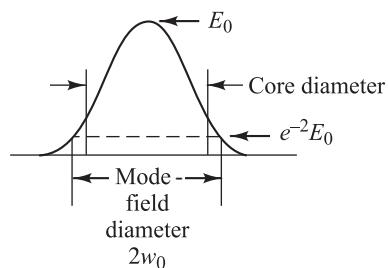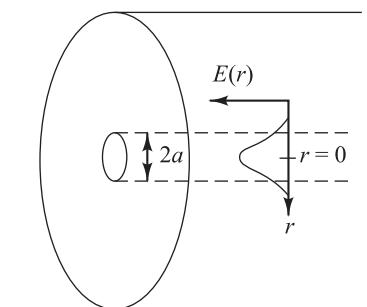
Fig. 2.28 *Distribution of light in a single-mode fiber above its cutoff wavelength. For a Gaussian distribution the MFD is given by the 1/e² width of the optical power*

where $2w_0$ (called the *spot size*) is the full width of the far-field distribution. For calculation simplicity the exact field distribution can be fitted to a Gaussian function[21]

$$E(r) = E_0 \exp\left(-r^2 \big/ w_0^2\right) \tag{2.74}$$

where r is the radius and E_0 is the field at zero radius, as shown in Fig. 2.28. Then the MFD is given by the $1/e^2$ width of the optical power.

2.5.3 Propagation Modes in Single-Mode Fibers

As described in Sec. 2.4.8, in any ordinary single-mode fiber there are actually two independent, degenerate propagation modes.[36–39] These modes are very similar, but their polarization planes are orthogonal. These may be chosen arbitrarily as the horizontal (H) and the vertical (V) polarizations as shown in Fig. 2.29. Either one of these two polarization modes constitutes the fundamental HE_{11} mode. In general, the electric field of the light propagating along the fiber is a linear superposition of these two polarization modes and depends on the polarization of the light at the launching point into the fiber.

Suppose we arbitrarily choose one of the modes to have its transverse electric field polarized along the x direction and the other independent, orthogonal mode to be polarized in the y direction as shown in Fig. 2.29. In ideal fibers with perfect rotational symmetry, the two modes are degenerate with equal propagation constants ($k_x = k_y$), and any polarization state injected into the fiber will propagate unchanged. In actual fibers there are imperfections, such as asymmetrical lateral stresses, noncircular cores, and variations in refractive-index profiles. These imperfections break the circular symmetry of the ideal fiber and lift the degeneracy of the two modes. The modes propagate with different phase velocities, and the difference between their effective refractive indices is called the fiber *birefringence,*

$$B_f = n_y - n_x \tag{2.75}$$

Equivalently, we may define the birefringence as

$$\beta = k_0(n_y - n_x) \tag{2.76}$$

where $k_0 = 2\pi/\lambda$ is the free-space propagation constant.

If light is injected into the fiber so that both modes are excited, then one will be delayed in phase relative to the other as they propagate. When this phase difference is an integral multiple of 2π, the two

Horizontal mode   Vertical mode
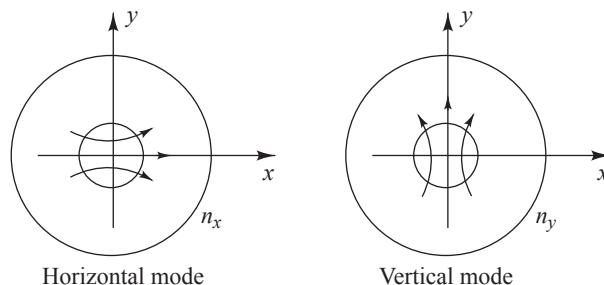
Fig. 2.29 *Two polarizations of the fundamental HE_{11} mode in a single-mode fiber*

modes will beat at this point and the input polarization state will be reproduced. The length over which this beating occurs is the *fiber beat length,*

$$L_p = 2\pi/\beta \tag{2.77}$$

Example 2.14 A single mode optical fiber has a beat length of 8 cm at 1300 nm. What is the birefringence?

Solution: From Eqs. (2.75) to (2.77) we have that the modal birefringence is

$$B_f = n_y - n_x = \frac{\lambda}{L_p} = \frac{1.3 \times 10^{-6}\,\text{m}}{8 \times 10^{-2}\,\text{m}} = 1.63 \times 10^{-5}$$

or, alternatively,

$$\beta = \frac{2\pi}{L_p} = \frac{2\pi}{0.08\,\text{m}} = 78.5\,\text{m}^{-1}$$

This indicates an intermediate-type fiber because birefringences can vary from $B_f = 1 \times 10^{-3}$ (a typical high-birefringence fiber) to $B_f = 1 \times 10^{-8}$ (a typical low-birefringence fiber).

2.6 Graded-Index Fiber Structure

In the graded-index fiber design the core refractive index decreases continuously with increasing radial distance *r* from the center of the fiber but is generally constant in the cladding. The most commonly used construction for the refractive-index variation in the core is the power law relationship

$$n(r) = \begin{cases} n_1\left[1 - 2\Delta\left(\dfrac{r}{a}\right)^{\alpha}\right]^{1/2} & \text{for} \quad 0 \le r \le a \\ n_1(1 - 2\Delta)^{1/2} \simeq n_1(1 - \Delta) = n_2 & \text{for} \quad r \ge a \end{cases} \tag{2.78}$$

Here, *r* is the radial distance from the fiber axis, *a* is the core radius, n_1 is the refractive index at the core axis, n_2 is the refractive index of the cladding, and the dimensionless parameter α defines the shape of the index profile. The index difference Δ for the graded-index fiber is given by

$$\Delta = \frac{n_1^2 - n_2^2}{2n_1^2} \simeq \frac{n_1 - n_2}{n_1} \tag{2.79}$$

The approximation on the right-hand side of this equation reduces the expression for Δ to that of the step-index fiber given by Eq. (2.20). Thus, the same symbol is used in both cases. For $\alpha = \infty$, inside the core Eq. (2.78) reduces to the step-index profile $n(r) = n_1$.

Determining the NA for graded-index fibers is more complex than for step-index fibers because it is a function of position across the core end face. This is in contrast to the step-index fiber, where the NA is constant across the core. Geometrical optics considerations show that light incident on the fiber core at position *r* will propagate as a guided mode only if it is within the local numerical aperture NA(*r*) at that point. The local numerical aperture is defined as[40]

$$\text{NA}(r) = \begin{cases} [n^2(r) - n_2^2]^{1/2} \simeq \text{NA}(0)\sqrt{1 - (r/a)^{\alpha}} & \text{for} \quad r \le a \\ 0 & \text{for} \quad r > a \end{cases} \tag{2.80a}$$
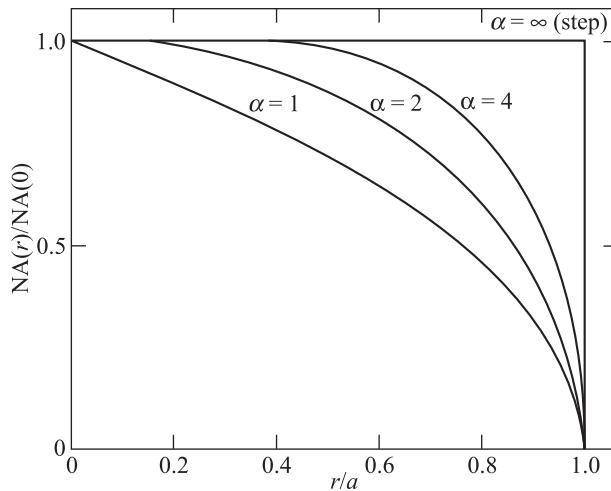
Fig. 2.30 *A comparison of the numerical apertures for fibers having various core index profiles*

where the axial numerical aperture is defined as

$$\text{NA}(0) = \left[n^2(0) - n_2^2 \right]^{1/2} = \left(n_1^2 - n_2^2 \right)^{1/2} \simeq n_1 \sqrt{2\Delta} \tag{2.80b}$$

Thus, the NA of a graded-index fiber decreases from NA(0) to zero as r moves from the fiber axis to the core-cladding boundary. A comparison of the numerical apertures for fibers having various α profiles is shown in Fig. 2.30. The number of bound modes in a graded-index fiber is[17–18,40]

$$M_g = \frac{\alpha}{\alpha + 2} a^2 k^2 n_1^2 \Delta \simeq \frac{\alpha}{\alpha + 2} \frac{V^2}{2} \tag{2.81}$$

where $k = 2\pi/\lambda$ and the right-hand approximation is derived using Eqs (2.23) and (2.27). Fiber manufacturers typically choose a *parabolic refractive index profile* given by $\alpha = 2.0$. In this case, $M_g = V^2/4$, which is half the number of modes supported by a step-index fiber (for which $\alpha = \infty$) that has the same V value, as Eq. (2.61) shows.

Example 2.15   Suppose we have a 50-μm diameter graded-index fiber that has a parabolic refractive index profile ($\alpha = 2$). If the fiber has a numerical aperture NA = 0.22, what is the total number of guided modes at a wavelength of 1310 nm?

Solution: First, from Eq. (2.27)

$$V = \frac{2\pi a}{\lambda} NA = \frac{2\pi \times 25\mu\text{m}}{1.31\mu\text{m}} \times 0.22 = 26.4$$

Then from Eq. (2.81) the total number of modes for $\alpha = 2$ is

$$M \approx \frac{a}{a + 2} \frac{V^2}{2} = \frac{V^2}{4} = 174$$

2.7 Fiber Materials

In selecting materials for optical fibers, a number of requirements must be satisfied. For example:

1. It must be possible to make long, thin, flexible fibers from the material.
2. The material must be transparent at a particular optical wavelength in order for the fiber to guide light efficiently.
3. Physically compatible materials that have slightly different refractive indices for the core and cladding must be available.

Materials that satisfy these requirements are glasses and plastics.

The majority of fibers are made of glass consisting of either silica (SiO_2) or a silicate. The variety of available glass fibers ranges from moderate-loss fibers with large cores used for short-transmission distances to very transparent (low-loss) fibers employed in long-haul applications. Plastic fibers are less widely used because of their substantially higher attenuation than glass fibers. The main use of plastic fibers is in short-distance applications (several hundred meters) and in abusive environments, where the greater mechanical strength of plastic fibers offers an advantage over the use of glass fibers.

2.7.1 Glass Fibers

Glass is made by fusing mixtures of metal oxides, sulfides, or selenides.[42–44] The resulting material is a randomly connected molecular network rather than a well-defined ordered structure as found in crystalline materials. A consequence of this random order is that glasses do not have well-defined melting points. When glass is heated up from room temperature, it remains a hard solid up to several hundred degrees centigrade. As the temperature increases further, the glass gradually begins to soften until at very high temperatures it becomes a viscous liquid. The expression "melting temperature" is commonly used in glass manufacture. This term refers only to an extended temperature range in which the glass becomes fluid enough to free itself fairly quickly of gas bubbles.

The largest category of optically transparent glasses from which optical fibers are made consists of the oxide glasses. Of these, the most common is silica (SiO_2), which has a refractive index ranging from 1.458 at 850 to 1.444 at 1550 nm. To produce two similar materials that have slightly different indices of refraction for the core and cladding, either fluorine or various oxides (referred to as *dopants*), such as B_2O_3, GeO_2, or P_2O_5, are added to the silica. As shown in Fig. 2.31, the addition of GeO_2 or P_2O_5 increases the refractive index, whereas doping the silica with fluorine or B_2O_3 decreases it. Since the cladding must have a lower index than the core, examples of fiber compositions are

1. GeO_2–SiO_2 core; SiO_2 cladding
2. P_2O_5–SiO_2 core; SiO_2 cladding
3. SiO_2 core; B_2O_3–SiO_2 cladding
4. GeO_2–B_2O_3–SiO_2 core; B_2O_3–SiO_2 cladding

Here, the notation GeO_2–SiO_2, for example, denotes a GeO_2-doped silica glass.
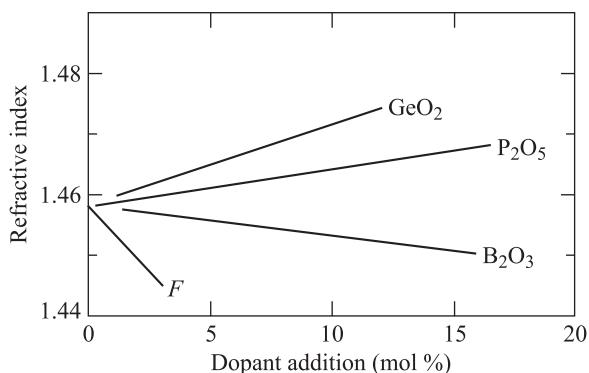
Fig. 2.31 *Variation in refractive index as a function of doping concentration in silica glass*

The principal raw material for silica is high-purity sand. Glass composed of pure silica is referred to as either *silica glass, fused silica,* or *vitreous silica.* Some of its desirable properties are a resistance to deformation at temperatures as high as 1000°C, a high resistance to breakage from thermal shock because of its low thermal expansion, good chemical durability, and high transparency in both the visible and infrared regions of interest to fiber optic communication systems. Its high melting temperature is a disadvantage if the glass is prepared from a molten state. However, this problem is partially avoided when using vapor deposition techniques.

2.7.2 Active Glass Fibers

Incorporating rare-earth elements (atomic numbers 57–71) into a normally passive glass gives the resulting material new optical and magnetic properties. These new properties allow the material to perform amplification, attenuation, and phase retardation on the light passing through it.[45–47] Doping (i.e., adding impurities) can be carried out for silica, tellurite, and halide glasses.

Two commonly used doping materials for fiber lasers are erbium and neodymium. The ionic concentrations of the rare-earth elements are low (on the order of 0.005–0.05 mole percent) to avoid clustering effects. To make use of the absorption and fluorescence spectra of these materials, one can use an optical source that emits at an absorption wavelength of the doping material to excite electrons to higher energy levels in the rare-earth dopants. When these excited electrons are stimulated by a signal photon to drop to lower energy levels, the transition process results in the emission of light in a narrow optical spectrum at the fluorescence wavelength. Chapter 11 discusses the applications of fibers doped with rare-earth elements to create optical amplifiers.

2.7.3 Plastic Optical Fibers

The growing demand for delivering high-speed services directly to the workstation has led fiber developers to create high-bandwidth graded-index polymer (plastic) optical fibers (POF) for use in a customer premises.[48,49] The core of these fibers is either polymethylmethacrylate or a perfluorinated polymer. These fibers are hence referred to as PMMA POF and PF POF, respectively. Although they exhibit considerably greater optical signal attenuations than glass fibers, they are tough and durable. For example, since the modulus of these polymers is nearly two orders of magnitude lower than that of silica, even a 1-mm-diameter graded-index POF is sufficiently flexible to be installed in conventional fiber cable routes. Standard optical connectors can be used on plastic fibers having core sizes that are compatible with the core diameters of standard multimode glass telecom fibers. Thus coupling between similar sized plastic and glass fibers is straightforward. In addition, for the plastic fibers inexpensive plastic injection-molding technologies can be used to fabricate connectors, splices, and transceivers.

Table 2.5 gives sample characteristics of PMMA and PF polymer optical fibers.

Table 2.5   *Sample characteristics of PMMA and PF polymer optical fibers*

Characteristic	*PMMA POF*	*PF POF*
Core diameter	0.4 mm	0.050–0.30 mm
Cladding diameter	1.0 mm	0.25–0.60 mm
Numerical aperture	0.25	0.20
Attenuation	150 dB/km at 650 nm	< 40 dB/km at 650–1300 nm
Bandwidth	2.5 Gb/s over 200 m	2.5 Gb/s over 550 m

2.8 │ **Photonic Crystal Fibers**

In the early 1990s researchers envisioned and demonstrated a new optical fiber structure. Initially this was called a *holey fiber* and later became known as a *photonic crystal fiber* (PCF) or a *microstructured* fiber.[50–54] The difference between this new structure and that of a conventional fiber is that the cladding and, in some cases, the core regions of a PCF contain air holes, which run along the entire length of the fiber. Whereas the material properties of the core and cladding define the light transmission characteristics of conventional fibers, the structural arrangement in a PCF creates an internal microstructure, which offers extra dimensions in controlling the optical properties of light, such as the dispersion, nonlinearity, and birefringence effects in optical fibers.

The sizes of the holes and the hole-to-hole spacing (known as the *pitch*) in the microstructure and the refractive index of its constituent material determine the light-guiding characteristics of photonic crystal fibers. The two basic PCF categories are index-guiding fibers and photonic bandgap fibers. The light transmission mechanism in an *index-guiding fiber* is similar to that in a conventional fiber as it has a high-index core surrounded by a lower-index cladding. However, for a PCF the effective refractive index of the cladding depends on the wavelength and the size and pitch of the holes. In contrast, in a *photonic bandgap fiber* light is guided by means of a photonic bandgap effect in either a hollow or microstructured core, which is surrounded by a microstructured cladding.

2.8.1 Index-Guiding PCF

Figure 2.32 shows the two-dimensional cross-sectional end views of two basic structures of an index-guiding PCF. The fibers have a solid core that is surrounded by a cladding region, which contains air holes that run along the length of the fiber and have a variety of different shapes, sizes, and distribution patterns. As an illustration, in Fig. 2.32a the air holes are arranged in a uniform hexagonal array. The holes all have a diameter d and a hole-to-hole spacing or pitch Λ. In comparison, in Fig. 2.32b the holes in the hexagonal array have different sizes. For example, a very low-loss PCF has been made using a configuration with 60 holes, where the hole diameter and hole pitch were 4 and 8 μm, respectively.[52]
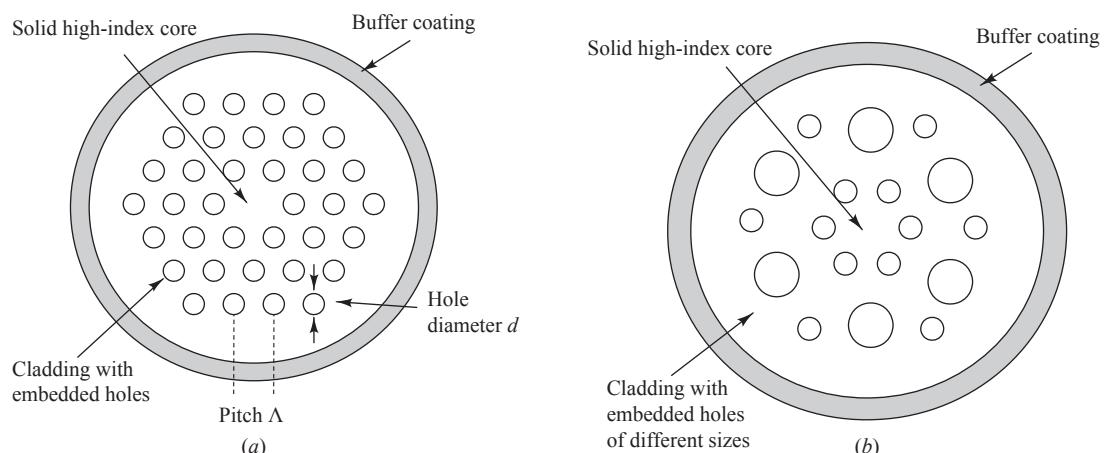
Fig. 2.32 *Cross-sectional end views of two basic structures of index-guided photonic crystal fibers: (a) air holes of uniform size; (b) air holes of different sizes*

The values of the hole diameter and the pitch are important for determining the operational characteristics of an index-guiding PCF. For a diameter-to-pitch ratio $d/\Lambda < 0.4$ the fiber exhibits single-mode properties over a wide range of wavelengths (from about 300 to 2000 nm).[50] This characteristic is not possible to achieve in standard optical fibers and is useful for the simultaneous transmission of multiple wavelengths over the same fiber. For example, ultra-wideband WDM transmission has been demonstrated using the visible to infrared wavelength region to send simultaneous signals at wavelengths of 658, 780, 853, 1064, 1309, and 1556 nm over a 1 km PCF. The 658-nm signal operated at 1 Gb/s whereas all the other wavelengths carried 10-Gb/s signals. All signals experienced a power penalty of less than 0.4 dB at a 10^{-9} bit error rate.[52]

Although the core and the cladding in a PCF are made of the same material (for example, pure silica), the air holes lower the effective index of refraction in the cladding region, since $n = 1.00$ for air and 1.45 for silica. The large difference in refractive indices together with the small dimensions of the microstructures causes the effective index of the cladding to depend strongly on wavelength. The fact that the core can be made of pure silica gives the PCF a number of operational advantages over conventional fibers, which typically have a germanium-doped silica core. These include very low losses, the ability to transmit high optical power levels, and a high resistance to darkening effects from nuclear radiation. The fibers can support single-mode operation over wavelengths ranging from 300 nm to more than 2000 nm. The mode-field area of a PCF can be greater than 300 μm^2 compared to the 80 μm^2 area of conventional single-mode fibers. This allows the PCF to transmit high optical power levels without encountering the nonlinear effects exhibited by standard fibers (see Chapter 12).

Optical filters are important components in lightwave networks for reducing noise or unwanted signals, or for flattening (equalizing) the gain of multiple signals at different wavelengths after optical amplification. For such applications, optical fiber-based filters are desirable because they can be spliced easily to the input and output fibers. Fiber-based optical filters can be created with a PCF structure by filling the air holes with a polymer or a high-index fluid. With such a structure, light is guided in the solid core only over certain wavelength ranges. By thermally tuning the refractive index of the fluid in the holes in the range of 30–90ºC along a few centimeters-long section of the fiber, one can create versatile fiber-based spectral filters that have the ability to be tuned dynamically over a wavelength range of greater than 150 nm at specific locations along the fiber.[53]

For most fluids the temperature-dependent index variation is around 0.1 to 0.4 K^{-1}, but the variation can be much larger if liquid crystals are used. In addition, liquid crystals have the further benefit of being tunable using external electric fields.[53a] Depending on the fluid and the fiber material, inserting fluids into photonic crystal fiber holes can be achieved by simply submerging one PCF end-facet into a liquid reservoir, while leaving the other end-facet open to atmospheric pressure.[53b]

2.8.2 Photonic Bandgap Fiber

Photonic bandgap (PBG) fibers have a different light-guiding mechanism, which is based on a two-dimensional photonic bandgap in the transverse plane of the cladding region. This photonic bandgap results from a periodic arrangement of the air holes in the cladding. Wavelengths within this bandgap are prevented from traveling in the cladding and thus are confined to travel in a region where the index is lower than the surrounding material. The functional principle of a photonic bandgap fiber is analogous to the role of a periodic crystalline lattice in a semiconductor, which blocks electrons from occupying a bandgap region. In a traditional PBG fiber the hollow core acts as a defect in the photonic bandgap structure, which creates a region in which the light can propagate. Whereas modes at all wavelengths can propagate along index-guiding fibers, the guided light in a PBG fiber is allowed in only a relatively narrow wavelength region with a width of approximately 100 to 200 nm.

Figure 2.33 shows a two-dimensional cross-sectional end view of the structure of an example PBG fiber. Here a large hollow core is formed by removing material in the center of the fiber from an area that can be occupied by seven air holes. Such a structure is called an *air-guiding* or a *hollow-core* PBG fiber, and allows about 98 percent of the power in the guided modes to propagate in the air hole regions. Analogous to an index-guiding fiber, the holes in the cladding region have a diameter *d* and a pitch Λ. Such hollow-core fibers can have a very low nonlinearity and a high damage threshold. Thus PBG fibers can be used for dispersive pulse compression with high optical intensities. In addition, fiber optic sensors or variable power attenuators can be constructed by filling PBG fibers that have larger core holes with gases or liquids.[53]

Fig. 2.33 *Cross-sectional end view of one type of photonic bandgap fiber*

In addition to using an array of air holes in the core and cladding regions to create a PBG fiber, it also is possible to use an arrangement of solid silica components.[54] One such PBG structure employs a two-dimensional periodic array of glass rods in the core region. This particular array consists of high-index 2.5 percent germanium-doped silica rods (n_2) immersed in a low-index fused silica background material ($n_1 = 1.45$). The high-index rods have an average diameter *d* and a rod-to-rod spacing (pitch) Λ. Since the fiber rods have subwavelength dimensions, the fiber architecture is called a *nanostructure core fiber* (NCF). As a particular example, a fiber was fabricated and tested with 37 rods arranged in four ring layers. The diameters of the high-index fiber rods were 800 nm and the pitch was 1 μm, thus giving the ratio $d/\Lambda = 0.8$. This type of fiber has potential applications in areas such as optical sensing, fiber lasers, and gyroscopes.

2.9 Fiber Fabrication

Two basic techniques[55–58] are used in the fabrication of all-glass optical waveguides. These are the vapor-phase oxidation process and the direct-melt methods. The direct-melt method follows traditional glass-making procedures in that optical fibers are made directly from the molten state of purified components of silicate glasses. In the vapor-phase oxidation process, highly pure vapors of metal halides (e.g., $SiCl_4$ and $GeCl_4$) react with oxygen to form a white powder of SiO_2 particles. The particles are then collected on the surface of a bulk glass by one of four different commonly used processes and are sintered (transformed to a homogeneous glass mass by heating without melting) by one of a variety of techniques to form a clear glass rod or tube (depending on the process). This rod or tube is called a *preform*. It is typically around 10–25 mm in diameter and 60–120 cm long. Fibers are made from the preform[57–59] by using the equipment shown in Fig. 2.34. The preform is precision-fed into a circular heater called the *drawing furnace*. Here the preform end is softened to the point where it can be drawn into a very thin filament, which becomes the optical fiber. The turning speed of the takeup drum at the bottom of the draw tower determines how fast the fiber is drawn. This, in turn, will determine the thickness of the fiber, so that a precise rotation rate must be maintained. An optical fiber thickness monitor is used in a feedback loop
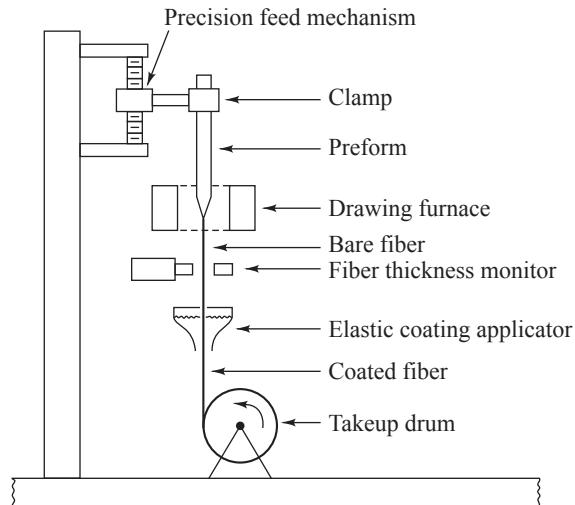
Fig. 2.34 *Schematic of a fiber-drawing apparatus*

for this speed regulation. To protect the bare glass fiber from external contaminants, such as dust and water vapor, an elastic coating is applied to the fiber immediately after it is drawn.

2.9.1 Outside Vapor-Phase Oxidation

The first fiber to have a loss of less than 20 dB/km was made at the Corning Glass Works[60–62] by the *outside vapor-phase oxidation* (OVPO) process. This method is illustrated in Fig. 2.35. First, a layer of SiO_2 particles called a *soot* is deposited from a burner onto a rotating graphite or ceramic mandrel. The glass soot adheres to this bait rod and, layer by layer, a cylindrical, porous glass preform is built up. By properly controlling the constituents of the metal halide vapor stream during the deposition process, the glass compositions and dimensions desired for the core and cladding can be incorporated into the preform. Either step- or graded-index preforms can thus be made.

When the deposition process is completed, the mandrel is removed and the porous tube is then vitrified in a dry atmosphere at a high temperature (above 1400°) to a clear glass preform. This clear preform is subsequently mounted in a fiber-drawing tower and made into a fiber, as shown in Fig. 2.34. The central hole in the tube preform collapses during this drawing process.

2.9.2 Vapor-Phase Axial Deposition

The OVPO process described in Sec. 2.9.1 is a lateral deposition method. Another OVPO-type process is the *vapor-phase axial deposition* (VAD) method,[63,64] illustrated in Fig. 2.36. In this method, the SiO_2 particles are formed in the same way as described in the OVPO process. As these particles emerge from the torches, they are deposited onto the end surface of a silica glass rod, which acts as a seed. A porous preform is grown in the axial direction by moving the rod upward. The rod is also continuously rotated to maintain cylindrical symmetry of the particle deposition. As the porous preform moves upward, it is transformed into a solid, transparent rod preform by zone melting (heating in a narrow localized zone) with the carbon ring heater shown in Fig. 2.36. The resultant preform can then be drawn into a fiber by heating it in another furnace, as shown in Fig. 2.34.
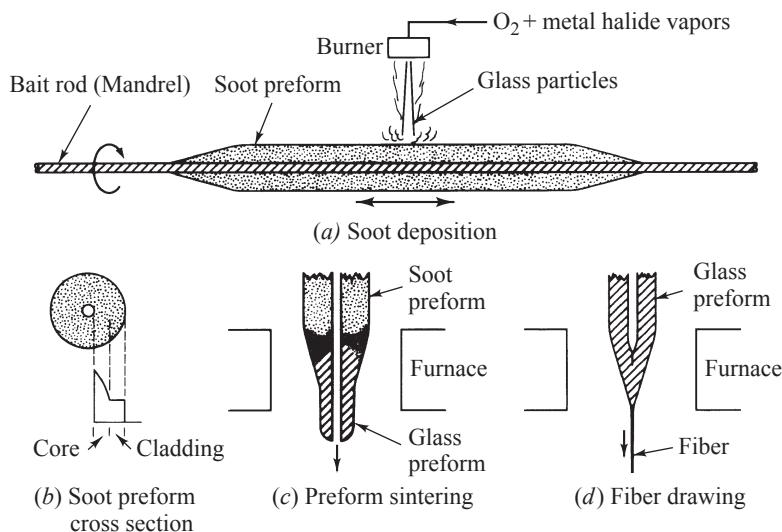
Fig. 2.35 *Basic steps in preparing a preform by the OVPO process. (a) Bait rod rotates and moves back and forth under the burner to produce a uniform deposition of glass soot particles along the rod; (b) profiles can be step or graded index; (c) following deposition, the soot preform is sintered into a clear glass preform; (d) fiber is drawn from the glass preform. (Reproduced with permission from Schultz.[55])*

Both step- and graded-index fibers in either multimode or single-mode varieties can be made by the VAD method. The advantages of the VAD method are (1) the preform has no central hole as occurs with the OVPO process, (2) the preform can be fabricated in continuous lengths, which can affect process costs and product yields, and (3) the fact that the deposition chamber and the zone-melting ring heater are tightly connected to each other in the same enclosure allows the achievement of a clean environment.

2.9.3 Modified Chemical Vapor Deposition

The *modified chemical vapor deposition* (MCVD) process shown in Fig. 2.37 was pioneered at Bell Laboratories[43,65–67] and widely adopted elsewhere to produce very low-loss graded-index fibers. The glass vapor particles, arising from the reaction of the constituent metal halide gases and oxygen, flow through the inside of a revolving silica tube. As the SiO_2 particles are deposited, they are sintered to a clear glass layer by an oxyhydrogen torch, which travels back and forth along the tube. When the desired thickness of glass has been deposited, the vapor flow is shut off and the tube is heated strongly to cause it to collapse into a solid rod preform. The fiber that is subsequently drawn from this preform rod will have a core that consists of the vapor-deposited material and a cladding that consists of the original silica tube.

2.9.4 Plasma-Activated Chemical Vapor Deposition

Scientists at Philips Research invented the *plasma-activated chemical vapor deposition* (PCVD) process.[68–70] As shown in Fig. 2.38, the PCVD method is similar to the MCVD process in that deposition occurs within a silica tube. However, a nonisothermal microwave plasma operating at low pressure initiates the chemical reaction. With the silica tube held at temperatures in the range of 1000–1200°C to reduce mechanical stresses in the growing glass films, a moving microwave resonator operating at 2.45 GHz

generates a plasma inside the tube to activate the chemical reaction. This process deposits clear glass material directly on the tube wall; there is no soot formation. Thus, no sintering is required. When one has deposited the desired glass thickness, the tube is collapsed into a preform just as in the MCVD case.

2.9.5 Photonic Crystal Fiber Fabrication

The fabrication of a photonic crystal fiber also is based on first creating a preform.[71–74] The preform is made by means of an array of hollow capillary silica tubes. To make a preform for an index-guiding fiber, the capillary tubes first are bundled in an array around a solid silica rod. For a photonic bandgap fiber, the hollow core is established by leaving an empty space at the center of the array. Following the array stacking processes, these configurations are fused together to create a preform and then made into a fiber using a conventional optical fiber drawing tower.

In the drawing process, the holes keep their original arrangement. This allows the creation of any type of array pattern (for example, close-packed arrays, a single circle surrounding a large solid core, or a star-shaped hollow core) and hole shape (for example, circular, hexagonal, or oval openings) in the final fiber.

In addition to using silica, polymer and soft glass materials, such as PMMA and SF6 or SF57 Schott (lead silicate) glass, also have been used to make photonic crystal fibers.[53,71–76] For these fiber types the preform is made by an extrusion process. In an example extrusion process a glass disk is heated until it is soft and then is pressed through a die. The structure of the die determines the configuration of the fiber
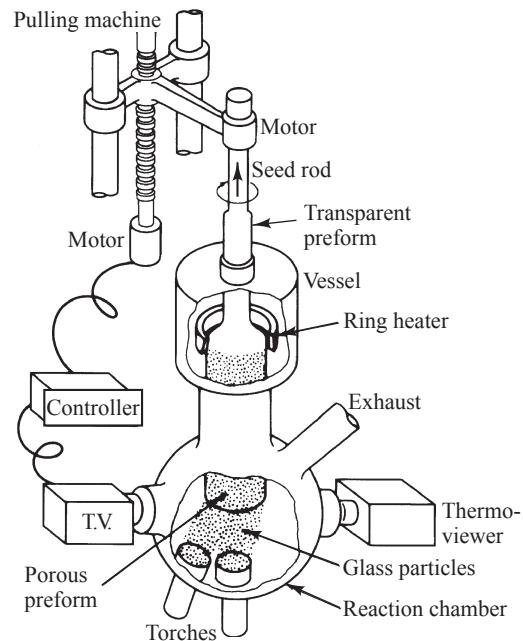
Fig. 2.36 *Apparatus for the VAD (vapor-phase axial deposition) process. (Reproduced with permission from Izawa and Inagaki,[63] © 1980, IEEE.)*
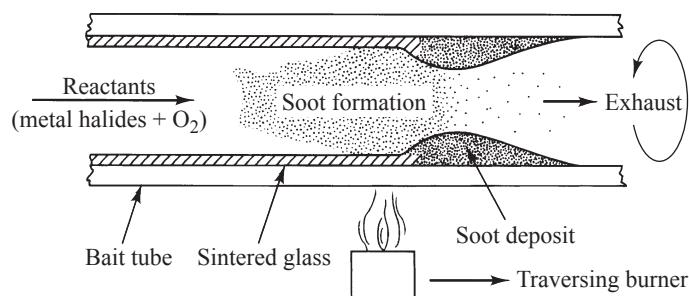
Fig. 2.37 *Schematic of MCVD (modified chemical vapor deposition) process. (Reproduced with permission from Schultz.[55])*
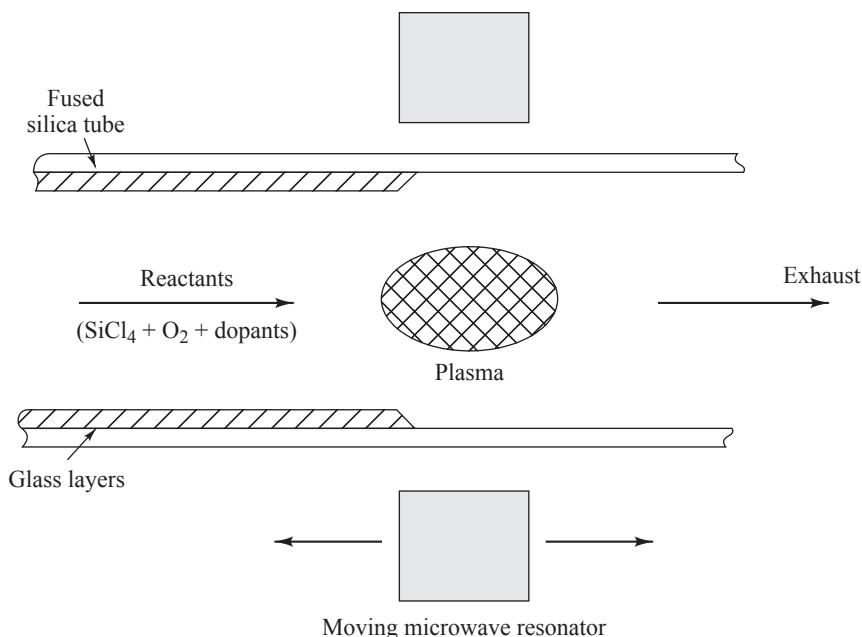
Fig. 2.38 *Schematic of PCVD (plasma-activated chemical vapor deposition) process*

cross section. The extrusion time typically is around two hours. After the preform is fabricated, fibers are made using standard optical fiber drawing equipment.

2.10 Mechanical Properties of Fibers

In addition to the transmission properties of optical waveguides, mechanical characteristics play a very important role when fibers are used as the transmission medium in optical communication systems.[77–80] Fibers must be able to withstand the stresses and strains that occur during the cabling process and the loads induced during the installation and service of the cable. During cable manufacture and installation, the loads applied to the fiber can be either impulsive or gradually varying. Once the cable is in place, the service loads are usually slowly varying loads, which can arise from temperature variations or a general settling of the cable following installation.

Strength and *static fatigue* are the two basic mechanical characteristics of glass optical fibers. Since the sight and sound of shattering glass are quite familiar, one intuitively suspects that glass is not a very strong material. However, the longitudinal breaking stress of pristine glass fibers is comparable to that of metal wires. The cohesive bond strength of the constituent atoms of a glass fiber governs its theoretical intrinsic strength. Maximum tensile strengths of 14 GPa (2×10^6 lb/in.^{2}) have been observed in short-gauge-length glass fibers. This is close to the 20-GPa tensile strength of steel wire. The difference between glass and metal is that, under an applied stress, glass will extend elastically up to its breaking strength, whereas metals can be stretched plastically well beyond their true elastic range. Copper wires, for example, can be elongated plastically by more than 20 percent before they fracture. For glass fibers, elongations of only about 1 percent are possible before fracture occurs.

In practice, the existence of stress concentrations at surface flaws or microcracks limits the median strength of long glass fibers to the 700-to-3500-MPa (1–5×10^5 lb/in.^{2}) range. The fracture strength of a given length of glass fiber is determined by the size and geometry of the severest flaw (the one that produces the largest stress concentration) in the fiber. A hypothetical, physical flaw model is shown in Fig. 2.39. This elliptically shaped crack is generally referred to as a *Griffith*

Fig. 2.39 *A hypothetical model of a microcrack in an optical fiber*

microcrack.[81] It has a width w, a depth χ, and a tip radius ρ. The strength of the crack for silica fibers follows the relation

$$K = Y \chi^{1/2} \sigma \tag{2.82}$$

where the stress intensity factor K is given in terms of the stress σ in megapascals applied to the fiber, the crack depth χ in millimeters, and a dimensionless constant Y that depends on flaw geometry. For elliptically shaped surface flaws, which are the most critical in glass fibers, $Y = \sqrt{\pi}$. From this equation the maximum crack size allowable for a given applied stress level can be calculated. The maximum values of K depend upon the glass composition but tend to be in the range 0.6–0.9 MN/m$^{3/2}$.

Since an optical fiber generally contains many flaws that have a random distribution of size, the fracture strength of a fiber must be viewed statistically. If $F(\sigma, L)$ is defined as the cumulative probability that a fiber of length L will fail below a stress level σ, then, under the assumption that the flaws are independent and randomly distributed in the fiber and that the fracture will occur at the most severe flaw, we have

$$F(\sigma, L) = 1 - e^{-LN(\sigma)} \tag{2.83}$$

where $N(\sigma)$ is the cumulative number of flaws per unit length with a strength less than σ. A widely used form for $N(\sigma)$ is the empirical expression

$$N(\sigma) = \frac{1}{L_0} \left(\frac{\sigma}{\sigma_0} \right)^m \tag{2.84}$$

where m, σ_0, and L_0 are constants related to the initial inert strength distribution. This leads to the so-called *Weibull expression*[82]

$$F(\sigma, L) = 1 - \exp\left[-\left(\frac{\sigma}{\sigma_0} \right)^m \frac{L}{L_0} \right] \tag{2.85}$$

A plot of the Weibull expression is shown in Fig. 2.40 for measurements performed on long fiber samples.[78,83] These data points were obtained by testing to destruction a large number of fiber samples. The fact that a single curve can be drawn through the data indicates that the failures arise from a single type of flaw. By careful environmental control of the fiber-drawing furnace, lengths of silica fiber greater than 50 km with a single failure distribution can be fabricated.
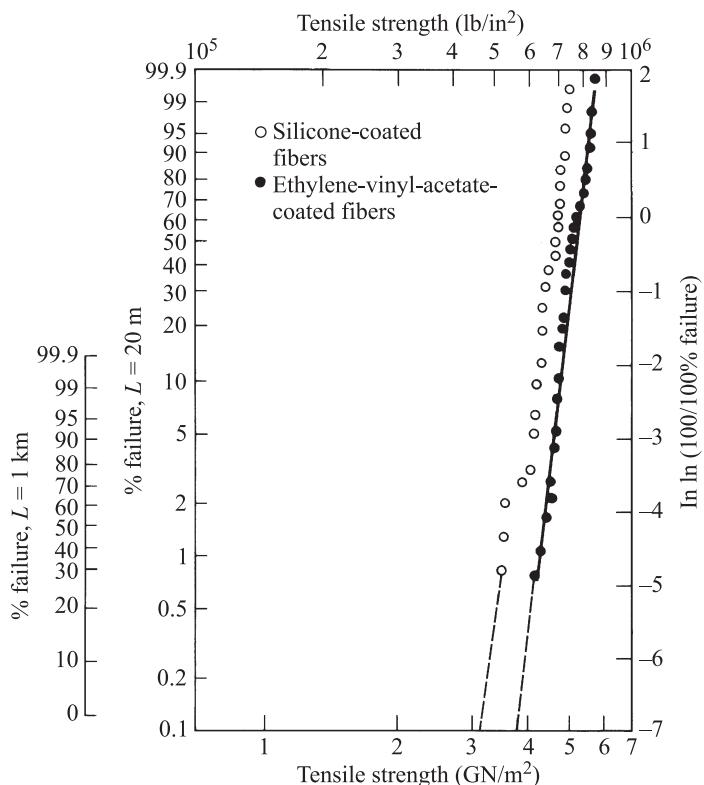
Fig. 2.40 *A Weibull-type plot showing the cumulative probability that fibers of 20-m and 1-km lengths will fracture at the indicated applied stress. (Reproduced with permission from Miller, Hart, Vroom, and Bowden.[83])*

In contrast to strength, which relates to instantaneous failure under an applied load, *static fatigue* relates to the slow growth of preexisting flaws in the glass fiber under humid conditions and tensile stress.[78] This gradual flaw growth causes the fiber to fail at a lower stress level than that which could be reached under a strength test. A flaw such as the one shown in Fig. 2.39 propagates through the fiber because of chemical erosion of the fiber material at the flaw tip. The primary cause of this erosion is the presence of water in the environment, which reduces the strength of the SiO_2 bonds in the glass. The speed of the growth reaction is increased when the fiber is put under stress. Certain fiber materials are more resistant to static fatigue than others, with fused silica being the most resistant of the glasses in water. In general, coatings that are applied to the fiber immediately during the manufacturing process afford a good degree of protection against environmental corrosion.[84]

Another important factor to consider is *dynamic fatigue.* When an optical cable is being installed in a duct, it experiences repeated stress owing to surging effects. The surging is caused by varying degrees of friction between the optical cable and the duct or guiding tool in a manhole on a curved route. Varying stresses also arise in aerial cables that are set into transverse vibration by the wind. Theoretical and experimental investigations[85,86] have shown that the time to failure under these conditions is related to

the maximum allowable stress by the same lifetime parameters that are found from the cases of static stress and stress that increases at a constant rate.

A high assurance of fiber reliability can be provided by proof testing.[33,87,88] In this method, an optical fiber is subjected to a tensile load greater than that expected at any time during the cable manufacturing, installation, and service. Any fibers that do not pass the proof test are rejected. Empirical studies of slow crack growth show that the growth rate $d\chi/dt$ is approximately proportional to a power of the stress intensity factor; that is,

$$\frac{d\chi}{dt} = AK^b \tag{2.86}$$

Here, A and b are material constants and the stress intensity factor is given by Eq. (2.82). For most glasses, b ranges between 15 and 50.

If a proof test stress σ_p is applied for a time t_p, then from Eq. (2.86) we have

$$B\left(\sigma_i^{b-2} - \sigma_p^{b-2}\right) = \sigma_p^b \, t_p \tag{2.87}$$

where σ_i is the initial inert strength and

$$B = \frac{2}{b-2}\left(\frac{K}{Y}\right)^{2-b}\frac{1}{AY^b} \tag{2.88}$$

When this fiber is subjected to a static stress σ_s after proof testing, the time to failure t_s is found from Eq. (2.86) to be

$$B\left(\sigma_p^{b-2} - \sigma_s^{b-2}\right) = \sigma_s^b \, t_s \tag{2.89}$$

Combining Eqs (2.87) and (2.89) yields

$$B\left(\sigma_i^{b-2} - \sigma_s^{b-2}\right) = \sigma_p^b \, t_p + \sigma_s^b \, t_s \tag{2.90}$$

To find the failure probability F_s of a fiber after a time t_s after proof testing, we first define $N(t, \sigma)$ to be the number of flaws per unit length that will fail in a time t under an applied stress σ. Assuming that $N(\sigma_i) \gg N(\sigma_s)$, then

$$N(t_s, \sigma_s) \simeq N(\sigma_i) \tag{2.91}$$

Solving Eq. (2.90) for σ_i and substituting into Eq. (2.84), we have, from Eq. (2.91),

$$N(t_s, \sigma_s) = \frac{1}{L_0}\left\{\frac{\left[\left(\sigma_p^b \, t_p + \sigma_s^b \, t_s\right)\big/ B + \sigma_s^{b-2}\right]^{1/(b-2)}}{\sigma_0}\right\}^m \tag{2.92}$$

The failure number $N(t_p, \sigma_p)$ per unit length during proof testing is found from Eq. (2.92) by setting $\sigma_s = \sigma_p$ and letting $t_s = 0$, so that

$$N(t_p, \sigma_p) = \frac{1}{L_0} \left[\frac{\left(\sigma_p^b t_p / B + \sigma_p^{b-2} \right)^{1/(b-2)}}{\sigma_0} \right]^m \tag{2.93}$$

Letting $N(t_x, \sigma_x) = N_x$, the failure probability F_s for a fiber after it has been proof tested is given by

$$F_s = 1 - \exp[-L(N_s - N_p)] \tag{2.94}$$

Substituting Eqs (2.92) and (2.93) into Eq. (2.94), we have

$$F_s = 1 - \exp\left(-N_p L \left\{ \left[\left(1 + \frac{\sigma_s^b t_s}{\sigma_p^b t_p} \right) \frac{1}{1 + C} \right]^{m/(b-2)} - 1 \right\} \right) \tag{2.95}$$

where $C = B / \sigma_p^2 t_p$, and where we have ignored the term

$$\left(\frac{\sigma_s}{\sigma_p} \right)^b \frac{B}{\sigma_s^2 t_p} \ll 1 \tag{2.96}$$

This holds because typical values of the parameters in this term are $\sigma_s / \sigma_p \simeq 0.3 - 0.4$, $t_p \simeq 10\,s$, $b > 15$, $\sigma_p = 350$ MN/m^2, and $B \simeq 0.05 - 0.5$ (MN/m^2)^{2}s.

The expression for F_s given by Eq. (2.95) is valid only when the proof stress is unloaded immediately, which is not the case in actual proof testing of optical fibers. When the proof stress is released within a finite duration, the C value should be rewritten as

$$C = \gamma \frac{B}{\sigma_p^2 t_p} \tag{2.97}$$

where γ is a coefficient of slow-crack-growth effect arising during the unloading period.

2.11 | Fiber Optic Cables

In any practical application of optical waveguide technology, the fibers need to be incorporated in some type of cable structure.[89-95] The cable structure will vary greatly, depending on whether the cable is to be pulled into underground or intrabuilding ducts, buried directly in the ground, installed on outdoor poles, or submerged underwater. Different cable designs are required for each type of application, but certain fundamental cable design principles will apply in every case. The objectives of cable manufacturers have been that the optical fiber cables should be installable with the same equipment, installation techniques, and precautions as those used for conventional wire cables. This requires special cable designs because of the mechanical properties of glass fibers.

2.11.1 Cable Structures

One important mechanical property is the maximum allowable axial load on the cable because this factor determines the length of cable that can be reliably installed. In copper cables the wires themselves are generally the principal load-bearing members of the cable, and elongations of more than 20 percent are possible without fracture. On the other hand, extremely strong optical fibers tend to break at 4-percent elongation, whereas typical good-quality fibers exhibit long-length breaking elongations of about 0.5–1.0 percent. Since static fatigue occurs very quickly at stress levels above 40 percent of the permissible elongation and very slowly below 20 percent, fiber elongations during cable manufacture and installation should be limited to 0.1–0.2 percent.

Steel wire has been used extensively for reinforcing electric cables and also can be used as a *strength member* for optical fiber cables. For some applications it is desirable to have a completely nonmetallic construction, either to avoid the effects of electromagnetic induction or to reduce cable weight. In such cases, plastic strength members and high-tensile-strength synthetic yarns are used. A popular yarn is Kevlar®, which is a soft but tough yellow synthetic nylon material belonging to a generic yarn family known as *aramids*. Good fabrication practices will isolate the fibers from other cable components, keep them close to the neutral axis of the cable, and allow the fibers to move freely when the cable is flexed or stretched.

The generic cable configuration shown in Fig. 2.41 illustrates some common materials that are used in the optical fiber cabling process. Individual fibers or modules of bundled fiber groupings and optional copper wires for powering in-line equipment are wound loosely around the central buffered strength member. A cable wrapping tape and other strength members such as Kevlar then encapsulate and bind these fiber groupings together. Surrounding all these components is a tough polymer *jacket* that provides crush resistance and handles any tensile stresses applied to the cable so that the fibers inside are not damaged. The jacket also protects the fibers inside against abrasion, moisture, oil, solvents, and other contaminants. The jacket type largely defines the application characteristics; for example, heavy-duty outside-plant cables for direct-burial and aerial applications have much thicker and tougher jackets than light-duty indoor cables.

The two basic fiber optic cable structures are the *tight-buffered fiber cable* design and the *loose-tube cable* configuration. Cables with tight-buffered fibers nominally are used indoors whereas the loose-tube structure is intended for long-haul outdoor applications. A *ribbon cable* is an extension of the tight-buffered cable. In all cases the fibers themselves consist of the normally manufactured glass core and cladding, which is surrounded by a protective 250-μm diameter coating.

As shown in Fig. 2.42, in the *tight-buffered* design each fiber is individually encapsulated within its own 900-μm diameter plastic buffer structure, hence the designation *tight-buffered design*. The 900-μm buffer is nearly four times the diameter and five times the thickness of the 250-μm protective coating material. This construction feature contributes to the excellent moisture and temperature

Outer sheath
Yarn strength member
Buffered strength member
Paper/plastic binding tape
Basic fiber building block
Insulated copper conductors
Polyurethane/PVC jacket
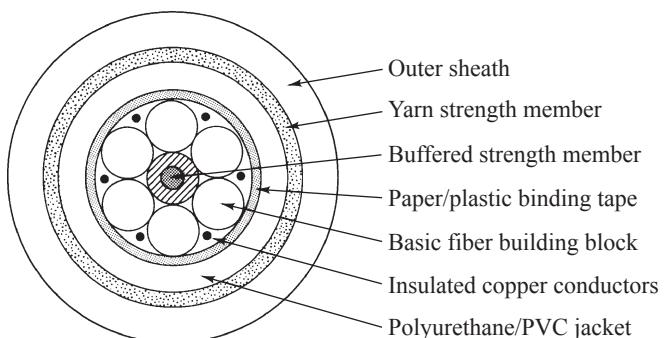
Fig. 2.41 *A typical six-fiber cable created by stranding six basic fiber building blocks around a central strength member*

Glass fiber with cladding

Fiber coating (250 μm)

Fiber buffer (900 μm)

Yarn strength member

Outer cable jacket
(e.g., 2.4 mm diameter)
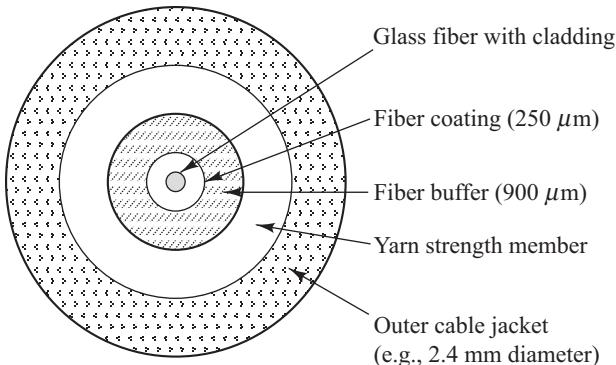
Fig. 2.42   *Construction of a simplex tight-buffered fiber cable module*

Fig. 2.43   *A loose-tube fiber optic cable that has a protective steel armoring layer inside the jacket. (Photo courtesy of OFS Fitel LLC; www .ofsoptics.com.)*

performance of tight-buffered cables and also permits their direct termination with connectors. In a single-fiber module, a layer of amarid strength material surrounds the 900-μm fiber structure. This configuration then is encapsulated within a PVC outer jacket.

In the *loose-tube* cable configuration one or more standard-coated fibers are enclosed in a thermoplastic tube that has an inner diameter much larger than the fiber diameter. The fibers in the tube are slightly longer than the cable itself. The purpose of this construction is to isolate the fiber from any stretching of the surrounding cable structure caused by factors such as temperature changes, wind forces, or ice loading. The tube is filled with either a *gel* or a dry *water-blocking material* that acts as a buffer, enables the fibers to move freely within the tube, and prevents moisture from entering the tube. Figure 2.43 gives an example of a loose-tube cable that also has an armoring layer just inside the jacket to offer crush resistance and protection against gnawing rodents. Such a cable can be used for direct-burial or aerial-based outside plant applications.

To facilitate the field operation of splicing cables that contain a large number of fibers, cable designers devised the fiber-ribbon structure. As shown in Fig. 2.44, the *ribbon cable* is an arrangement of fibers that are aligned precisely next to each other and then are encapsulated in a plastic buffer or jacket to form a long continuous ribbon. The number of fibers in a ribbon typically ranges from four to twelve. These ribbons can be stacked on top of each other to form a densely packed arrangement of many fibers (e.g., 144 fibers) within a cable structure.

There are many different ways in which to arrange fibers inside a cable. The particular arrangement of fibers and the cable design itself need to take into account issues such as the physical environment, the services that the optical link will provide, and any anticipated maintenance and repair that may be needed. Two examples of ribbon cables for duct and aerial applications are shown in Figs. 2.44b and 2.44c.

2.11.2   Indoor Cable Designs

Indoor cables can be used for interconnecting instruments, for distributing signals among office users, for connections to printers or servers, and for short patch cords in telecommunication equipment racks. The three main types are described here.

Interconnect cable serves light-duty low-fiber count indoor applications such as fiber-to-the-desk links, patch cords, and point-to-point runs in conduits and trays. The cable is flexible, compact, and
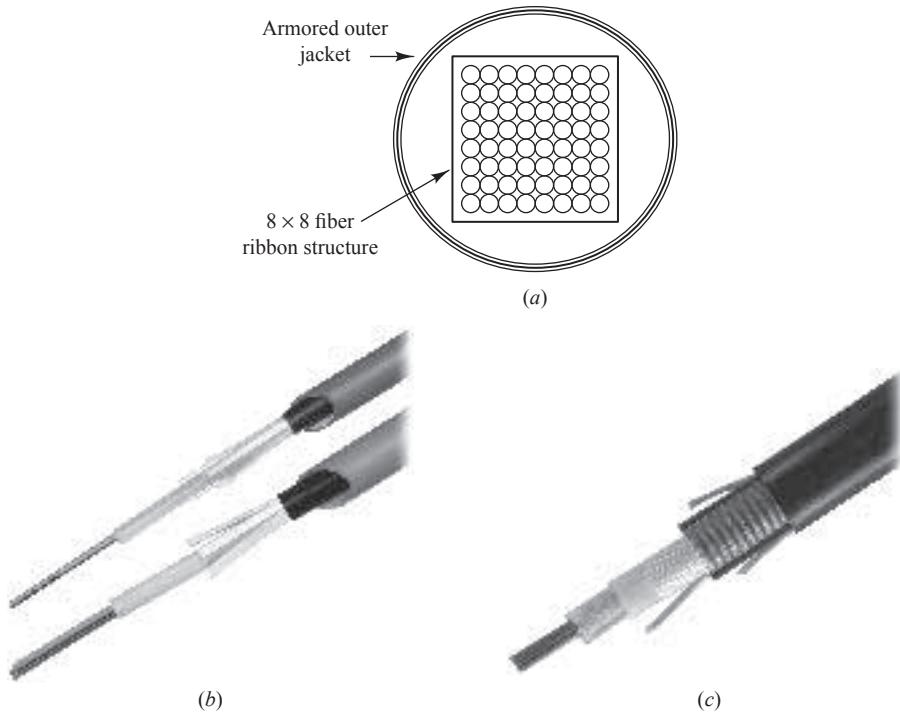
Fig. 2.44 *Fiber ribbon cable: (a) layered ribbon structure of a 64-fiber design; (b) small cable design for air-blown installation into ducts; (c) lightly armored cable for aerial, direct-burial, or duct installations. (Photos courtesy of OFS Fitel LLC: AccuRibbon Ductsaver FX® and AccuRibbon DC® metallic cables; www.ofsoptics.com.)*

lightweight with a tight-buffered construction. A popular indoor cable type is the *duplex cable,* which consists of two fibers that are encapsulated in an outer PVC jacket. Fiber optic *patch cords,* also known as *jumper cables,* are short lengths (usually less than 2 m) of simplex or duplex cable with connectors on both ends. They are used to connect lightwave test equipment to a fiber patch panel or to interconnect optical transmission modules within an equipment rack.

Breakout or fanout cable consists of up to 12 tight-buffered fibers stranded around a central strength member. Such cables serve low- to medium-fiber-count applications where it is necessary to protect individual jacketed fibers. The breakout cable allows easy installation of connectors on individual fibers in the cable. With such a cable configuration, routing the individually terminated fibers to separate pieces of equipment can be achieved easily.

Distribution cable consists of individual or small groupings of tight-buffered fibers stranded around a central strength member. This cable serves a wide range of network applications for sending data, voice, and video signals. Distribution cables are designed for use in intrabuilding cable trays, conduits, and loose placement in dropped-ceiling structures. A main feature is that they enable groupings of fibers within the cable to be branched (distributed) to various locations.

2.11.3 Outdoor Cables

Outdoor cable installations include aerial, duct, direct-burial, and underwater applications. Invariably these cables consist of a loose-tube structure. Many different designs and sizes of outdoor cables are available depending on the physical environment in which the cable will be used and the particular application.

Aerial cable is intended for mounting outside between buildings or on poles or towers. The two popular designs are the self-supporting and the facility-supporting cable structures. The *self-supporting cable*

Fig. 2.45 *Example configuration of an armored outdoor fiber optic cable*

contains an internal strength member that permits the cable to be strung between poles without using any additional support mechanisms for the cable. For the *facility-supporting cable,* first a separate wire or strength member is strung between the poles and then the cable is lashed or clipped to this member.

Armored cable for direct-burial or underground-duct applications has one or more layers of steel-wire or steel-sheath protective armoring below a layer of polyethylene jacketing as shown in Fig. 2.45. This not only provides additional strength to the cable but also protects it from gnawing animals such as squirrels or burrowing rodents, which often cause damage to underground cables. For example, in the United States the plains pocket gopher (*Geomys busarius*) can destroy unprotected cable that is buried less than 2 m (6 ft) deep. Other cable components include a central strength member, wrapping and binding tapes, and water-blocking materials.

Fig. 2.46 *Example configuration of an underwater fiber optic cable*

Underwater cable, also known as *submarine cable,* is used in rivers, lakes, and ocean environments. Since such cables normally are exposed to high water pressures, they have much more stringent requirements than underground cables. For example, as shown in Fig. 2.46, cables that can be used in rivers and lakes have various water-blocking layers, one or more protective inner polyethylene sheaths, and a heavy outer armor jacket. Cables that run under the ocean have further layers of armoring and contain copper wires to provide electrical power for submersed optical amplifiers or regenerators.

2.12 | Cable Installation Methods

Optical fiber cables can be installed by plowing them directly into the ground, laying them in an outside trench, pulling or blowing them through ducts (both indoor and outdoor) or other spaces, suspending them on poles, or placing them underwater.[89–101] Although each method has its own special handling procedures, each one needs to adhere to a common set of precautions. These precautions include avoiding sharp bends of the cable, minimizing stresses on the installed cable, periodically allowing extra cable slack along the cable route for unexpected repairs, and avoiding excessive pulling forces on the cable.

2.12.1 Direct-Burial Installations

For *direct-burial installations* a fiber optic cable can be plowed directly underground or placed in a trench that is filled in later. Figure 2.47 illustrates a plowing operation that may be carried out in nonurban areas. The cables are mounted on large reels on the plowing vehicle and are fed directly into the ground by means of the plow mechanism. A plowing operation normally is not feasible in an urban environment, so a *trenching* method must be used. Trenching is more time-consuming than direct plowing as it requires a trench to be dug by hand or by machine to some specified depth. However, trenching allows the

Fig. 2.47 *A cable plowing machine that is used in a nonurban environment. (Photo courtesy of © VERMEER CORPORATION. All Rights Reserved; www.vermeer.com.)*

installation to be more controlled than in plowing. For example, in direct plowing it is not known if a sharp rock is left pressing against the installed cable or if the cable was damaged in some way that may cause it to fail later.

Usually a combination of the two methods is used, with plowing being done in isolated open areas and trenching being done where plowing is not possible, such as in urban areas. In addition, *directional boring* or *horizontal drilling* may be needed in areas where the surface cannot be disturbed. For example, if the cable needs to go under a busy road, a stream, or a landscaped area, it is better to drill an underground duct-encapsulated hole and then run the cable through that duct. Drilling machines come in various sizes depending on the depth and distance that holes need to be bored. For example, the 48-in-wide horizontal drilling machine for urban use illustrated in Fig. 2.48 can bore a hole below the surface for a 1.9-in-diameter (5-cm) pipe over distances of up to 400 ft (122 m).

During direct-burial installations a bright (usually orange) *warning tape* normally is placed a short distance (typically 18 inches) above the cable to alert future digging operators to the presence of a cable. The tape may contain metallic strips so that it can be located from above ground with a metal detector. In addition, a warning post or a cable marker that is flush with the ground may be used to indicate where a cable is buried. Figure 2.49 illustrates some typical signs and markers for indicating underground cable locations. Besides indicating to repair crews where

Fig. 2.48 *Compact horizontal drilling machine for urban use. (Photo © VERMEER CORPORATION. All Rights Reserved; www.vermeer.com.)*

a cable is located, these precautions also are intended to minimize the occurrence of what is known popularly in the telecommunications world as "backhoe fade" (the rupture of a cable by errant digging of a backhoe).

2.12.2 Pulling Cable into Ducts

Most *ducts* are constructed of high-density polyurethane, PVC, or an epoxy fiberglass compound. To reduce pulling tensions during cable installation, the inside walls may have longitudinal or corrugated ribs, or they may have been lubricated at the factory. Alternatively a variety of pulling lubricants are available that can be applied to the cable itself as it is pulled into a long duct or one that has numerous bends. Ducts also can contain a pulling tape running along its length that was installed by the duct manufacturer. This is a flat tape similar to a measuring tape that is marked every meter for easy identification of distance. If the duct does not already contain a pulling tape, it can be fished through or blown into

Fig. 2.49 *Buried cable location marking methods: (a) underground tapes, (b) metal sign posts, (c) plastic sign posts, (d) ground-flush markers. (Photos courtesy of William Frick and Company; www.fricknet.com.)*

a duct length. After the fiber optic cable is installed in a duct, end plugs can be added to prevent water and debris from entering the duct. Similar to direct-burial installations, a warning tape may be placed underground above the duct, or warning posts or markers may be placed aboveground to alert future digging operators to its presence.

2.12.3 Cable Jetting Installation

An alternative method to a pulling procedure is to employ a high-pressure air flow to blow a fiber cable into a duct. The installation scheme of utilizing the friction of forced air pushing on the cable jacket is referred to as an *air-assisted* or a *cable jetting* method. Cable jetting must overcome the same frictional forces to move a cable as in a pulling operation, but it does this differently and with much less mechanical stress on the cable.

As Fig. 2.50 shows, the two driving forces in the cable jetting method come from a mechanical roller mechanism that feeds the cable into the duct and from the force of air pushing on the cable jacket. The advantage of cable jetting is that the cable moves freely around bends, whereas the pulling method puts a high lateral stress on the cable when it passes through bends in a duct. This is illustrated by the diagram in

High-pressure air flow into duct

Ductwall

Fiber cable or microduct

Flowing air pressure

Mechanical feeder
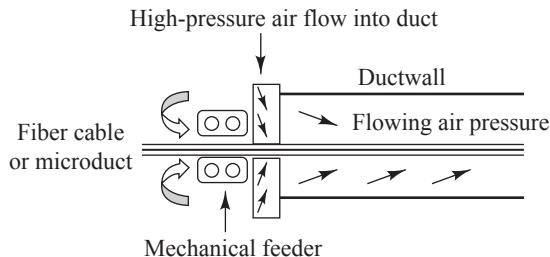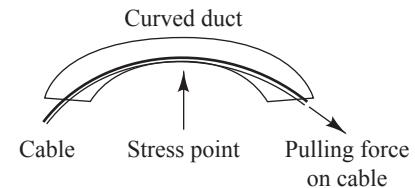
Fig. 2.50 *Mechanical rollers and air pressure are the driving forces in the cable jetting method.*

Curved duct

Cable   Stress point   Pulling force on cable

(*a*) Cable pulling method (localized stresses)

Curved duct

Cable   Stress-free flow around corner   Air-pushing force on cable

(*b*) Cable jetting method (distributed stresses)

Fig. 2.51 *(a) The pulling method puts a high lateral stress on the cable at duct bends. (b) With cable jetting the cable moves freely around bends in a duct.*

Fig. 2.51. Figure 2.51a shows that during a pulling operation a high lateral stress is imposed on a cable at bends in the duct. With a cable jetting procedure such lateral forces are mitigated because the forced air pushes on the cable from all sides. This tends to keep the cable in the middle of the duct as it is blown through, as shown in Fig. 2.51b.
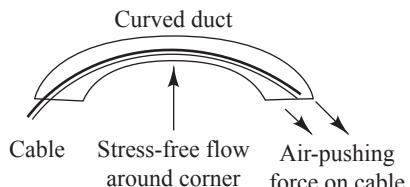
Figure 2.52 shows an example of a cable jetting machine. The cable or duct is fed in from the top right. The central part is the mechanical feeder, and the air-pressure connection pipe is seen in the leftmost section. Cables or microducts can be installed at rates of 150–300 FPM (feet per minute) with up to 20 corners in a duct. For long installation lengths the jetting machines can be cascaded, as Fig. 2.53 illustrates. This figure shows how a cable is fed from a truck-mounted reel into a cable jetting machine at the beginning of a large duct. After a distance of 3000–6000 ft (1–1.8 km), a second cable jetting machine and its associated air compressor and other peripheral equipment continues the cable installation.

2.12.4 Aerial Installation

Cable crews can install an aerial cable either by lashing it onto an existing steel *messenger wire* that runs from pole to pole or by directly suspending the cable between poles if it is a self-supporting design. Several different methods can be used to install the fiber optic cables. The primary method for installing self-supporting cable is a *stationary reel technique*. This method stations the payoff reel at one end of the cable route and the takeup reel at the other end. A pull rope is attached to the cable and is threaded through pulleys on each pole. The takeup reel gradually pulls the cable from the payoff reel, the pulleys guide it into position along the route, and the cable is then attached to the poles.
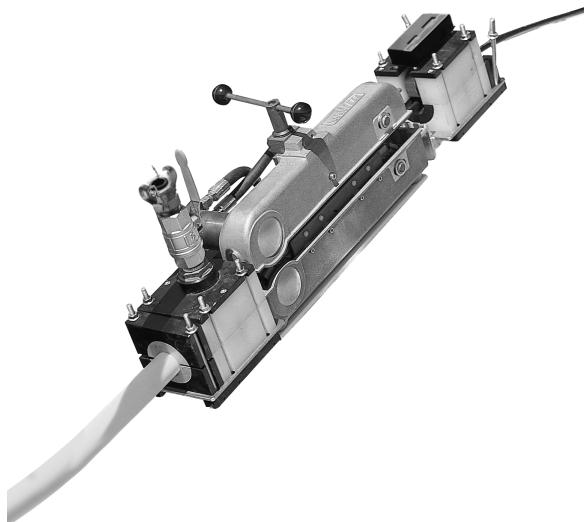
Fig. 2.52 *Cable jetting machine. (Photo courtesy of Sherman & Reilly, Inc.; www.sherman-reilly.com.)*
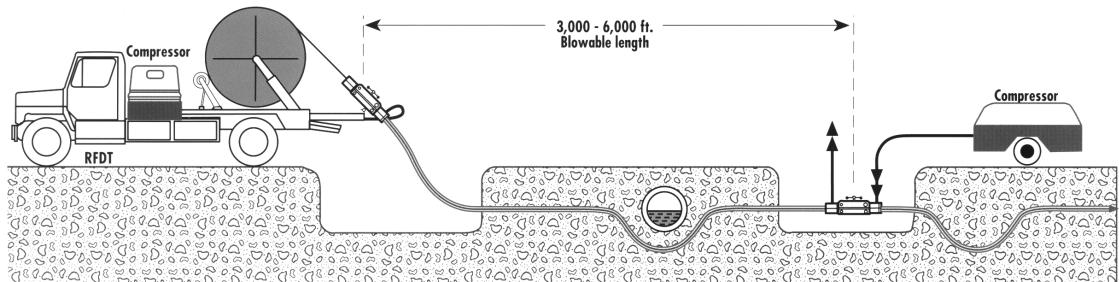
Fig. 2.53 *Jetting machines can be cascaded for long installation lengths. (Photo courtesy of Sherman & Reilly, Inc.; www.sherman-reilly.com.)*

If a messenger wire is used, first this wire is installed between poles with an appropriate tension and sag calculated to support the fiber optic cable. The messenger wire must be grounded properly and should be kept on one pole side along the route whenever possible. One of at least three techniques then can be used to attach the fiber optic cable to the messenger wire. Each of these methods uses a special lashing machine that hangs on the messenger wire and attaches the cable to the messenger as it moves along the wire length.

2.12.5 Underwater Installation

Over a million kilometers of submarine cable already is submerged in oceans around the world, which is enough to circle the globe thirty times.[96–98,102] Specially designed cable-laying ships are used to install an undersea cable (see Fig. 1.15). As illustrated in Fig. 2.54, such ships have several large circular containers inside of them called *cable tanks*. In modern cable ships the tanks can hold around 5000 tons of underwater cable, which is enough to cross the Atlantic Ocean. Such a length of cable is assembled onshore in a factory environment along with underwater signal amplifiers that need to be located every 80–120 km. The amplifiers are housed in beryllium-copper cylinders that are about a meter long and 50 cm in diameter. After being assembled, this cable unit is coiled by hand into the cable holding tanks on the ship at a rate

Cable handling machinery includes two 30 ton cable drums and a 20 wheel pair linear cable engine.

The state-of-the art bridge houses the latest technology in navigation, DP (dynamic positioning) and communication.

The ROV (remotely operated vehicle) is a versatile underwater robot used for various cable installation and repair tasks to 2500 meters of water depth.

The 65 ton universal A-frame is used to maneuver various sea plows that are able to bury cable up to 3 meters safely below the ocean floor.

The 5 main diesel engine generators provide enough electrical power to the motors to propel the ship at service speeds of 14 knots.

The 3 main cable tanks can hold in excess of 5000 tons of cable-enough to cross the Atlantic Ocean.

Fig. 2.54 *Internal view of an optical fiber cable-laying ship showing three large cable tanks. The cable is dropped into the water from the aft part of the ship. (Diagram courtesy of TE SubCom: drawn by Nick Rotondo; www.SubCom.com)*

Table 2.6 *Some ITU-T recommendations for optical fiber cables and cable components*

Recommendation name	Description
L.10, *Optical fiber cables for duct and tunnel application,* Dec. 2002	Describes characteristics, construction, and test methods of optical fiber cables for duct and tunnel application.
L.26, *Optical fiber cables for aerial application,* Dec. 2002	Describes characteristics, construction, and test methods of optical fiber cables for aerial application but does not apply to Optical Fiber Ground Wire (OPGW) cables (described in L.34).
L.36, *Single-mode fiber optic connectors,* Jan. 2008	Describes the main features of fiber optic connectors in terms of types, fields of application, configurations, and technical aspects; examines the optical, mechanical, and environmental characteristics of fiber optic connectors.
L.37, *Optical branching components (non-wavelength selective)*, Feb. 2007	Describes the types, applications, configurations, and technical aspects of fiber optic branching devices; describes mechanical and environmental requirements, physical performance, reliability, and test methods with regard to the optical performance of passive optical networks (PONs).
L.43, *Optical fiber cables for buried application,* Dec. 2002	Describes characteristics, construction, and test methods of optical fiber cables for buried application.
L.50, *Requirements for passive optical nodes: Optical distribution frames for central office environments,* Nov. 2003	Deals with general requirements, performance evaluation, and mechanical and optical evaluation testing for optical distribution frames (ODF) in a central office environment.
L.59, *Optical fiber cables for indoor applications,* Jan. 2008	Describes optical and mechanical characteristics of optical fibers and cables, mechanical and environmental test methods, and test methods for fire safety.
L.78, *Optical fiber cable construction for sewer duct applications,* May 2008	Describes the characteristics, construction, and test methods for optical fiber cables to be installed in sewer ducts and drainpipes.

of around 80 km per day. During installation, near the shore a sea plow buries the cable to a depth of about one meter under the ocean floor to protect it from fishing nets and other factors that might damage the cable. In the middle of the ocean, the cable simply lays exposed on the ocean floor.

2.12.6   Industry Installation Standards

The Telecommunications Section of the International Telecommunications Union (ITU-T) has published a number of L-series recommendations for the construction, installation, and protection of fiber optic cables. Table 2.6 lists some ITU-T recommendations for the characteristics, features, constituent components, and other requirements to which various types of cables must adhere. For example, indoor cables must meet more stringent smoke and toxic fume emission criteria than outdoor cables. Table 2.7 lists some ITU-T recommendations and guidelines for various optical cable installation techniques used in locations such as underground, in ducts and air plenums, on poles, in trenches, along railway lines, and inside sewer pipes.

Table 2.7   *Some ITU-T recommendations for optical fiber cable installations*

Recommendation name	Description
L.35, *Installation of optical fiber cables in the access network,* Oct. 1998; Amendment 1, Nov. 2007	Gives guidance for installing optical cables in ducts, on poles, and through direct burial.
L.38, *Use of trenchless techniques for the construction of underground infrastructures for telecommunication cable installation,* Sept. 1999	Describes underground drilling techniques for installing telecommunication cables without the need for disruptive excavation or plowing.
L.39, *Investigation of the soil before using trenchless techniques,* May 2000	Describes soil-investigation techniques for obtaining information about the position of buried objects and the nature of the ground.
L.48, *Mini-trench installation techniques,* March 2003	Describes a technique for putting duct-based underground optical cables in small trenches. This allows quicker installations, lower cost, and limited surface disruption.
L.49, *Micro-trench installation techniques,* March 2003	Describes installing underground cables at a shallow depth in small grooves.
L.56, *Installation of optical fiber cables along railways*, May 2003	Describes duct installation, direct-buried cable installation, and installation of optical fiber cables suspended on poles.
L.57, *Air-assisted installation of optical fiber cables,* May 2003	Describes air-assisted methods for installation of optical fiber cables in ducts.
L.77, *Installation of optical fiber cables inside sewer ducts,* May 2008	Discusses installation of infrastructure and/or optical cables inside man-accessible and non-man-accessible sewer ducts and gives safety guidelines.

▌ PROBLEMS ▐

2.1 Consider an electric field represented by the expression

$$\mathbf{E} = \left[100e^{j30°}\mathbf{e}_x + 20e^{-j50°}\mathbf{e}_y + 40e^{j210°}\mathbf{e}_z \right] e^{j\omega t}$$

Express this as a measurable electric field as described by Eq. (2.2) at a frequency of 100 MHz.

2.2 A wave is specified by $y = 8 \cos 2\pi(2t - 0.8z)$, where y is expressed in micrometers and the propagation constant is given in μm^{-1}. Find (*a*) the amplitude, (*b*) the wavelength, (*c*) the angular frequency, and (*d*) the displacement at time $t = 0$ and $z = 4\ \mu m$.

2.3 Consider two plane waves X_1 and X_2 traveling in the same direction. If they have the same frequency ω but different amplitudes a_i and phases δ_i, then we can represent them by

$$X_1 = a_1 \cos (\omega t - \delta_1)$$
$$X_1 = a_2 \cos (\omega t - \delta_2)$$

According to the principle of superposition, the resultant wave X is simply the sum of X_1 and X_2. Show that X can be written in the form

$$X = A \cos (\omega t - \phi)$$

where

$$A^2 = a_1^2 + a_2^2 + 2a_1 a_2 \cos(\delta_1 - \delta_2)$$

and

$$\tan \phi = \frac{a_1 \sin \delta_1 + a_2 \sin \delta_2}{a_1 \cos \delta_1 + a_2 \cos \delta_2}$$

2.4 Elliptically polarized light can be represented by the two orthogonal waves given in Eqs (2.2) and (2.3). Show that elimination of the $(\omega t - kz)$ dependence between them yields

$$\left(\frac{E_x}{E_{0x}}\right)^2 + \left(\frac{E_y}{E_{0y}}\right)^2 - 2\frac{E_x}{E_{0x}}\frac{E_y}{E_{0y}}\cos\delta = \sin^2\delta$$

which is the equation of an ellipse making an angle α with the x axis, where α is given by Eq. (2.8).

2.5 Let $E_{0x} = E_{0y} = 1$ in Eq. (2.7). Using a computer or a graphical calculator, write a program to plot this equation for values of $\delta = (n\pi)/8$, where $n = 0, 1, 2, ... 16$. What does this show about the state of polarization as the angle δ changes?

2.6 Show that any linearly polarized wave may be considered as the superposition of left and right circularly polarized waves that are in phase and have equal amplitudes and frequencies.

2.7 Light traveling in air strikes a glass plate at an angle $\theta_1 = 33°$, where θ_1 is measured between the incoming ray and the glass surface. Upon striking the glass, part of the beam is reflected and part is refracted. If the refracted and reflected beams make an angle of 90° with each other, what is the refractive index of the glass? What is the critical angle for this glass?

2.8 A point source of light is 12 cm below the surface of a large body of water ($n = 1.33$ for water). What is the radius of the largest circle on the water surface through which the light can emerge?

2.9 A 45°–45°–90° prism is immersed in alcohol ($n = 1.45$). What is the minimum refractive index the prism must have if a ray incident normally on one of the short faces is to be totally reflected at the long face of the prism?

2.10 Show that the critical angle at an interface between doped silica with $n_1 = 1.460$ and pure silica with $n_2 = 1.450$ is 83.3°.

2.11 The reflection coefficient R_p for parallel polarized light is given by

$$R_p = \frac{n_1 \cos \varphi_2 - n_2 \cos \varphi_1}{n_1 \cos \varphi_2 + n_2 \cos \varphi_1}$$

Show that the condition $R_p = 0$ for the Brewster angle occurs when $\tan \varphi_1 = n_2/n_1$.

2.12 Calculate the numerical aperture of a step-index fiber having $n_1 = 1.48$ and $n_2 = 1.46$. What is the acceptance angle θ_A for this fiber if the outer medium is air with $n = 1.00$?

2.13 Consider a dielectric slab having a thickness $d = 10$ mm and index of refraction $n_1 = 1.50$. Let the medium above and below the slab be air, in which $n_2 = 1$. Let the wavelength be $\lambda = 10$ mm (equal to the thickness of the waveguide).
   (a) What is the critical angle for the slab waveguide?
   (b) Solve Eq. (2.26b) graphically to show that there are three angles of incidence that satisfy this equation.
   (c) What happens to the number of angles as the wavelength is decreased?

2.14 Derive the approximation of the right-hand side of Eq. (2.23) for $\Delta \ll 1$. What is the difference in the approximate and exact expressions for the value of NA if $n_1 = 1.49$ and $n_2 = 1.48$?

2.15 Using the expressions in Eqs. (2.33) and (2.34) derived from Maxwell's curl equations, derive the radial and transverse electric and magnetic field components given in Eqs. (2.35a) to (2.35d). Show that these expressions lead to Eqs. (2.36) and (2.37).

2.16 Show that for $\nu = 0$, Eq. (2.55b) corresponds to TE_{0m} modes ($E_z = 0$) and that Eq. (2.56b) corresponds to TM_{om} modes ($H_z = 0$).

2.17 Verify that $k_1^2 \approx k_2^2 \approx \beta^2$ when $\Delta \ll 1$, where k_1 and k_2 are the core and cladding propagation constants, respectively, as defined in Eq. (2.46).

2.18 A step-index multimode fiber with a numerical aperture of 0.20 supports approximately 1000 modes at an 850-nm wavelength.

(*a*) What is the diameter of its core?

(*b*) How many modes does the fiber support at 1320 nm?

(*c*) How many modes does the fiber support at 1550 nm?

2.19 (*a*) Determine the normalized frequency at 820 nm for a step-index fiber having a 25-μm core radius, $n_1 = 1.48$, and $n_2 = 1.46$.

(*b*) How many modes propagate in this fiber at 820 nm?

(*c*) How many modes propagate in this fiber at 1320 nm?

(*d*) How many modes propagate in this fiber at 1550 nm?

(*e*) What percent of the optical power flows in the cladding in each case?

2.20 Consider a fiber with a 25-μm core radius, a core index $n_1 = 1.48$, and $\Delta = 0.01$.

(*a*) If $\lambda = 1320$ nm, what is the value of V and how many modes propagate in the fiber?

(*b*) What percent of the optical power flows in the cladding?

(*c*) If the core-cladding difference is reduced to $\Delta = 0.003$, how many modes does the fiber support and what fraction of the optical power flows in the cladding?

2.21 Find the core radius necessary for single-mode operation at 1320 nm of a step-index fiber with $n_1 = 1.480$ and $n_2 = 1.478$. What are the numerical aperture and acceptance angle of this fiber?

2.22 A manufacturer wishes to make a silica-core, step-index fiber with $V = 75$ and a numerical aperture NA = 0.30 to be used at 820 nm. If $n_1 = 1.458$, what should the core size and cladding index be?

2.23 Draw a design curve of the fractional refractive-index difference Δ versus the core radius a for a silica-core ($n_1 = 1.458$), single-mode fiber to operate at 1300 nm. Suppose the fiber we select from this curve has a 5-μm core radius. Is this fiber still single-mode at 820 nm? Which modes exist in the fiber at 820 nm?

2.24 Using the following approximation for w_0 given by[103]

$w_0 = a(0.65 + 1.619V^{-3/2} + 2.879\ V^{-6})$

evaluate and plot $E(r)/E_0$ with r ranging from 0 to 3 for values of $V = 1.0, 1.4, 1.8, 2.2, 2.6$, and 3.0. Here a is the fiber radius.

2.25 Commonly available single-mode fibers have beat lengths in the range 10 cm $< L_p <$ 2 m. What range of birefringent refractive-index differences does this correspond to for $\lambda = 1300$ nm?

2.26 Plot the refractive-index profiles from n_1 to n_2 as a function of radial distance $r \leq a$ for graded-index fibers that have α values of 1, 2, 4, 8, and ∞ (step index). Assume the fibers have a 25-μm core radius, $n_1 = 1.48$, and $\Delta = 0.01$.

2.27 Calculate the number of modes at 820 nm and 1.3 μm in a graded-index fiber having a parabolic-index profile ($\alpha = 2$), a 25-μm core radius, $n_1 = 1.48$, and $n_2 = 1.46$. How does this compare to a step-index fiber?

2.28 Calculate the numerical apertures of (*a*) a plastic step-index fiber having a core refractive index of $n_1 = 1.60$ and a cladding index of $n_2 = 1.49$, (*b*) a step-index fiber having a silica core ($n_1 = 1.458$) and a silicone resin cladding ($n_2 = 1.405$).

2.29 When a preform is drawn into a fiber, the principle of conservation of mass must be satisfied under steady-state drawing conditions. Show that for a solid rod preform this is represented by the expression

$$s = S\left(\frac{D}{d}\right)^2$$

where D and d are the preform and fiber diameters, and S and s are the preform feed and fiber-draw speeds, respectively. A typical drawing speed is 1.2 m/s for a 125-μm outer-diameter fiber. What is the preform feed rate in cm/min for a 9-mm-diameter preform?

2.30 A silica tube with inside and outside radii of 3 and 4 mm, respectively, is to have a certain thickness of glass deposited on the inner surface. What should the thickness of this glass deposition be if a fiber having a core diameter

of 50 μm and an outer cladding diameter of 125 μm is to be drawn from this preform using the MCVD fabrication process.

2.31 (*a*) The density of fused silica is 2.6 g/cm^3. How many grams are needed for a 1-km-long 50-μm-diameter fiber core?

(*b*) If the core material is to be deposited inside of a glass tube at a 0.5-g/min deposition rate, how long does it take to make the preform for this fiber?

2.32 During fabrication of optical fibers, dust particles incorporated into the fiber surface are prime examples of surface flaws that can lead to reduced fiber strength. What size dust particles are tolerable if a glass fiber having a 20-N/mm$^{3/2}$ stress intensity factor is to withstand a 700-MN/m^2 stress?

2.33 Static fatigue in a glass fiber refers to the condition where a fiber is stressed to a level σ_a, which is much less than the fracture stress associated with the weakest flaw. Initially, the fiber will not fail but, with time, cracks in the fiber will grow as a result of chemical erosion at the crack tip. One model for the growth rate of a crack of depth χ assumes a relation of the form given in Eq. (2.86).

(*a*) Using this equation, show that the time required for a crack of initial depth χ_i to grow to its failure size χ_f is given by

$$t = \frac{2}{(b-2)\,A(Y\sigma)^b}\left[\chi_i^{(2-b)/2} - \chi_f^{(2-b)/2}\right]$$

(*b*) For long, static fatigue times (on the order of 20 years), $K_i^{2-b} \ll K_f^{2-b}$ for large values of b. Show that under this condition the failure time is

$$t = \frac{2K_i^{2-b}}{(b-2)\,A\sigma^2 Y^2}$$

REFERENCES

1. See any general physics book or introductory optics book; for example: (*a*) D. Halliday, R. Resnick, and J. Walker, *Fundamentals of Physics*, Wiley, Hoboken, NJ, 9th ed., 2010; (*b*) E. Hecht, *Optics*, Addison-Wesley, Boston, 4th ed., 2002; (*c*) F. A. Jenkins and H. E. White, *Fundamentals of Optics*, McGraw-Hill, Burr Ridge, IL, 4th ed., 2001; (*d*) C. A. Bennett, *Principles of Physical Optics*, Wiley, New York, 2008; (*e*) F. L. Pedrotti, L. M. Pedrotti, and L. S. Pedrotti, *Introduction to Optics*, Addison-Wesley, Reading, MA, 3rd ed., 2007.

2. See any introductory electromagnetics book; for example: (*a*) J. Franklin, *Classical Electromagnetism*, Prentice Hall, Upper Saddle River, NJ, 2005; (*b*) W. H. Hayt Jr. and J. A. Buck, *Engineering Electromagnetics*, McGraw-Hill, Burr Ridge, IL, 7th ed., 2006; (*c*) N. Ida, *Engineering Electromagnetics*, Springer, New York, 2nd ed., 2007; (*d*) F. T. Ulaby, *Fundamentals of Applied Electromagnetics*, Prentice Hall, Upper Saddle River, NJ, 6th ed., 2010.

3. E.A.J. Marcatili, "Objectives of early fibers: Evolution of fiber types," in S. E. Miller and A. G. Chynoweth, eds., *Optical Fiber Telecommunications*, Academic, New York, 1979.

4. L. B. Felsen, "Rays and modes in optical fibers," *Electron. Lett.*, vol. 10, pp. 95–96, Apr. 1974.

5. A. W. Snyder and D. J. Mitchell, "Leaky rays on circular optical fibers," *J. Opt. Soc. Amer.*, vol. 64, pp. 599–607, May 1974.

6. A. W. Snyder and J. D. Love, *Optical Waveguide Theory*, Chapman & Hall, New York, 1983.

7. D. Marcuse, *Theory of Dielectric Optical Waveguides*, Academic, New York, 2nd ed., 1991.

8. K. Okamoto, *Fundamentals of Optical Waveguides*, Academic, New York, 2nd ed., 2006.

9. C. L. Chen, *Foundations of Guided-Wave Optics*, Wiley, New York, 2007.

10. K. Kawano and T. Kitoh, *Introduction to Optical Waveguide Analysis: Solving Maxwell's Equation and the Schrödinger Equation*, Wiley, Hoboken, NJ, 2002.

11. J. A. Buck, *Fundamentals of Optical Fibers*, Wiley, Hoboken, NJ, 2nd ed., 2004.

12. R. Olshansky, "Leaky modes in graded index optical fibers," *Appl. Opt.*, vol. 15, pp. 2773–2777, Nov. 1976.

13. A. Tomita and L. G. Cohen, "Leaky-mode loss of the second propagation mode in single-mode fibers with index well profiles," *Appl. Opt.*, vol. 24, pp. 1704–1707, 1985.

14. E. Snitzer, "Cylindrical dielectric waveguide modes," *J. Opt. Soc. Amer.*, vol. 51, pp. 491–498, May 1961.

15. M. Koshiba, *Optical Waveguide Analysis*, McGraw-Hill, New York, 1992.

16. D. Marcuse, *Light Transmission Optics*, Van Nostrand-Reinhold, New York, 2nd ed., 1982.

17. R. Olshansky, "Propagation in glass optical waveguides," *Rev. Mod. Phys.*, vol. 51, pp. 341–367, Apr. 1979.

18. D. Gloge, "The optical fiber as a transmission medium," *Rep. Progr. Phys.*, vol. 42, pp. 1777–1824, Nov. 1979.

19. A. W. Snyder, "Asymptotic expressions for eigenfunctions and eigenvalues of a dielectric or optical waveguide," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, pp. 1130–1138, Dec. 1969.

20. D. Gloge, "Weakly guiding fibers," *Appl. Opt.*, vol. 10, pp. 2252–2258, Oct. 1971.

21. D. Marcuse, "Gaussian approximation of the fundamental modes of graded index fibers," *J. Opt. Soc. Amer*, vol. 68, pp. 103–109, Jan. 1978.

22. H. M. DeRuiter, "Integral equation approach to the computation of modes in an optical waveguide," *J. Opt. Soc. Amer*, vol. 70, pp. 1519–1524, Dec. 1980.

23. A. W. Snyder, "Understanding monomode optical fibers," *Proc. IEEE*, vol. 69, pp. 6–13, Jan. 1981.

24. A. Jeffery and H. H. Dai, *Handbook of Mathematical Formulas and Integrals*, Academic, New York, 4th ed., 2008.

25. M. Kurtz, *Handbook of Applied Mathematics for Engineers and Scientists*, McGraw-Hill, New York, 1991.

26. D. Zwillinger, ed., *Standard Mathematical Tables and Formulae*, CRC Press, Boca Raton, FL, 31st ed., 2003.

27. D. Marcuse, D. Gloge, and E.A.J. Marcatili, "Guiding properties of fibers," in S. E. Miller and A. G. Chynoweth, eds., *Optical Fiber Telecommunications*, Academic, New York, 1979.

28. R. M. Gagliardi and S. Karp, *Optical Communications*, Wiley, New York, 2nd ed., 1995.

29. D. Gloge, "Propagation effects in optical fibers," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-23, pp. 106–120, Jan. 1975.

30. M. Artiglia, G. Coppa, P. DiVita, M. Potenza, and A. Sharma, "Mode field diameter measurements in single-mode optical fibers," *J. Lightwave Tech.*, vol. 7, pp. 1139–1152, Aug. 1989.

31. T. J. Drapela, D. L. Franzen, A. H. Cherin, and R. J. Smith, "A comparison of far-field methods for determining mode field diameter of single-mode fibers using both gaussian and Petermann definitions," *J. Lightwave Tech.*, vol. 7, pp. 1153–1157, Aug. 1989.

32. K. Petermann, "Constraints for fundamental mode spot size for broadband dispersion-compensated single-mode fibers," *Electron. Lett.*, vol. 19, pp. 712–714, Sept. 1983.

33. (*a*) ITU-T Recommendation G.650.1, *Definitions and test methods for linear, deterministic attributes of single-mode fibre and cable*, June 2004; (*b*) ITU-T Recommendation G.650.2, *Definitions and test methods for statistical and nonlinear related attributes of single-mode fibre and cable*, June 2007; (*c*) ITU-T Recommendation G.650.3, *Test methods for installed single-mode optical fibre cable links*, March 2008.

34. TIA/EIA FOTP-191, *Measurement of Mode Field Diameter of Single-Mode Fiber*, Sept. 1998.

35. F. Kapron, "Fiber-optic test methods," in F. Allard, ed., *Fiber Optics Handbook for Engineers and Scientists*, McGraw-Hill, New York, 1990.

36. D. K. Mynbaev and L. L. Scheiner, *Fiber-Optic Communications Technology*, Prentice Hall, Upper Saddle River, NJ, 2001.

37. I. P. Kaminow, "Polarization in optical fibers," *IEEE J. Quantum Electron.*, vol. QE-17, pp. 15–22, Jan. 1981.

38. S. C. Rashleigh, "Origins and control of polarization effects in single-mode fibers," *J. Lightwave Tech.*, vol. LT-1, pp. 312–331, June 1983.

39. X.-H. Zheng, W. M. Henry, and A. W. Snyder, "Polarization characteristics in the fundamental mode of optical fibers," *J. Lightwave Tech.*, vol. LT-6, pp. 1300–1305, Aug. 1988.

40. D. Gloge and E. Marcatili, "Multimode theory of graded core fibers," *Bell Sys. Tech. J.*, vol. 52, pp. 1563–1578, Nov. 1973.

41. B.E.A. Saleh and M. Teich, *Fundamentals of Photonics*, chap. 8, Wiley, Hoboken, NJ, 2nd ed., 2007.

42. R. H. Doremus, *Glass Science*, Wiley, Hoboken, NJ, 2nd ed., 1994.

43. S. R. Nagel, "Fiber materials and fabrication methods," in S. E. Miller and I. P. Kaminow, eds., *Optical Fiber Telecommunications–II*, Academic, New York, 1988.

44. B. Mysen and P. Richet, *Silicate Glasses and Melts*, Elsevier, 2005.

45. B. J. Ainslie, "A review of the fabrication and properties of erbium-doped fibers for optical amplifiers," *J. Lightwave Tech.*, vol. 9, pp. 220–227, Feb. 1991.

46. W. Miniscalco, "Erbium-doped glasses for fiber amplifiers at 1500 nm," *J. Lightwave Tech.*, vol. 9, pp. 234–250, Feb. 1991.

47. E. Desurvire, *Erbium-Doped Fiber Amplifiers*, Wiley, Hoboken, NJ, 2002.

48. (*a*) J. Zubia and J. Arrue, "Plastic optical fibers: An introduction to their technological processes and applications," *Opt. Fiber Technol.*, vol. 7, no. 2, pp. 101–140, Apr. 2001; (*b*) O. Ziemann, J. Krauser, P. E. Zamzow, and W. Daum, *POF Handbook,* 2nd ed., Springer, Berlin, 2008.

49. I. T. Monroy, H.P.A. vd Boom, A.M.J. Koonen, G. D. Khoe, Y. Watanabe, Y. Koike, and T. Ishigure, "Data transmission over polymer optical fibers," *Optical Fiber Tech.*, vol. 9, no. 3, pp. 159–171, July 2003.

50. Some early papers on the concepts and realizations of photonic crystal fibers: (*a*) T. A. Birks, J. C. Knight, and P. St. J. Russell, "Endlessly single-mode photonic crystal fiber," *Opt. Lett.*, vol. 22, pp. 961–963, July 1997; (*b*) J. C. Knight, J. Broeng, T. A. Birks, and P. St. J.Russell, "Photonic band gap guidance in optical fibers," *Science*, vol. 282, pp. 1476–1478, Nov. 1998; (*c*) J. Broeng, D. Mogilevstev, S. E. Barkou, and A. Bjarklev, "Photonic crystal fibers: A new class of optical waveguides," *Opt. Fiber Technol.*, vol. 5, pp. 305–330, July 1999; (*d*) P. St. J. Russell, "Photonic crystal fibers," *Science*, vol. 299, pp. 358–362, Jan. 2003.

51. Some basic articles and books on photonic crystal fibers: (*a*) F. Poli, A. Cucinotta, and S. Selleri, *Photonic Crystal Fibers*, Springer, New York, 2007; (*b*) M Large, L. Poladian, G. Barton, and M. A. van Eijkelenborg, *Microstructured Polymer Optical Fibres*, Springer, New York, 2008; (*c*) A. Argyros, "Microstructured polymer optical fibers," *J. Lightwave Tech.*, vol. 27, no. 11, pp. 1571–1579, June 2009.

52. K. Kurokawa, K. Nakajima, K. Tsujikawa, T. Yamamoto, and K. Tajima, "Ultra-wideband transmission over low loss PCF," *J. Lightwave Tech.*, vol. 27, no. 11, pp. 1653–1662, June 2009.

53. Papers related to tuning effects in photonic crystal fibers: (*a*) M. Haakestad, T. Alkeskjold, M. Nielsen, L. Scolari, J. Riishede, H. Engan, and A. Bjarklev, "Electrically tunable photonic bandgap guidance in a liquid-crystal-filled photonic crystal fiber," *IEEE Photon. Tech. Lett.*, vol. 17, no. 4, pp. 819–821, Apr. 2005; (*b*) K. Nielsen, D. Noordegraaf, T. Sorensen, A. Bjarklev, and T. Hansen, "Selective filling of photonic crystal fibres," *J. Opt. A Pure Appl. Opt.*, vol. 7, no. 8, pp. L13–L20, Aug. 2005; (*c*) C. P. Yu, J. H. Liou, S. S. Huang, and H. C. Chang, "Tunable dual-core liquid-filled photonic crystal fibers for dispersion compensation," *Opt. Express*, vol. 16, pp. 4443–4451, 2008; (*d*) B. T. Kuhlmey, B. J. Eggleton, and D.K.C. Wu, "Fluid-filled solid-core photonic bandgap fibers," *J. Lightwave Tech.*, vol. 27, no. 11, pp. 1617–1630, June 2009.

54. X. Yu, M. Yan, G. Ren, W. Tong, X. Cheng, J. Zhou, P. Shum, and N. Q. Ngo, "Nanostructure core fiber with enhanced performances: Design, fabrication and devices," *J. Lightwave Tech.*, vol. 27, no. 11, pp. 1548–1555, June 2009.

55. P. C. Schultz, "Progress in optical waveguide processes and materials," *Appl. Opt.*, vol. 18, pp. 3684–3693, Nov. 1979.

56. W. G. French, R. E. Jaeger, J. B. MacChesney, S. R. Nagel, K. Nassau, and A. D. Pearson, "Fiber perform preparation," in S. E. Miller and A. G. Chynoweth, eds., *Optical Fiber Telecommunications*, Academic, New York, 1979.

57. R. E. Jaeger, A. D. Pearson, J. C. Williams, and H. M. Presby, "Fiber drawing and control," in S. E. Miller and A. G. Chynoweth, eds., *Optical Fiber Telecommunications*, Academic, New York, 1979.

58. D. J. DiGiovanni, D. P. Jablonowski, and M. F. Yan, "Advances in fiber design and processing," in I. P. Kaminow and T. L. Koch, eds., *Optical Fiber Telecommunications–IIIA*, Academic, New York, 1997.

59. Q. Jiang, F. Yang, and R. Pitchumani, "Analysis of coating thickness variation during optical fiber processing," *J. Lightwave Tech.*, vol. 23, pp. 1261–1272, Mar. 2005.

60. F. P. Kapron, D. B. Keck, and R. D. Maurer, "Radiation losses in glass optical waveguides," *Appl. Phys. Lett.*, vol. 17, pp. 423–425, Nov. 1970.

61. P. C. Schultz, "Fabrication of optical waveguides by the outside vapor deposition process," *Proc. IEEE*, vol. 68, pp. 1187–1190, Oct. 1980.

62. R. V. VanDewoestine and A. J. Morrow, "Developments in optical waveguide fabrication by the outside vapor deposition process," *J. Lightwave Tech.*, vol. LT-4, pp. 1020–1025, Aug. 1986.

63. T. Izawa and N. Inagaki, "Materials and processes for fiber perform fabrication: Vapor-phase axial deposition," *Proc. IEEE*, vol. 68, pp. 1184–1187, Oct. 1980.

64. H. Murata, "Recent developments in vapor-phase axial deposition," *J. Lightwave Tech.*, vol. LT-4, pp. 1026–1033, Aug. 1986.

65. S. R. Nagel, J. B. MacChesney, and K. L. Walker, "Modified chemical vapor deposition," in T. Li, ed., in *Optical Fiber Communications, Vol. 1, Fiber Fabrication*, Academic, New York, 1985.

66. E. M. Dianov and V. M. Mashinsky, "Germania-based core optical fibers," *J. Lightwave Tech.*, vol. 23, pp. 3500–3508, Nov. 2005.

67. Y. Chigusa, Y. Yamamoto, T. Yokokawa, T. Sasaki, T. Taru, M. Hirano, M. Kakui, M. Onishi, and E. Sasaoka, "Low-loss pure-silica-core fibers and their possible impact on transmission systems," *J. Lightwave Tech.*, vol. 23, pp. 3541–3550, Nov. 2005.

68. P. Geittner and H. Lydtin, "Manufacturing optical fibers by the PCVD process," *Philips Tech. Rev. (Netherlands)*, vol. 44, pp. 241–249, May 1989.

69. T. Hünlich, H. Bauch, R. T. Kersten, V. Paquet, and G. F. Weidmann, "Fiber perform fabrication using plasma technology: A review," *J. Opt. Commun.*, vol. 8, pp. 122–129, Dec. 1987.

70. H. Lydtin, "PCVD: A technique suitable for large-scale fabrication of optical fibers," *J. Lightwave Tech.*, vol. LT-4, pp. 1034–1038, Aug. 1986.

71. S. C. Xue, M.C.J. Large, G. W. Barton, R. I. Tanner, L. Poladian, and R. Lwin, "Role of material properties and drawing conditions in the fabrication of microstructured optical fibers," *J. Lightwave Tech.*, vol. 24, pp. 853–860, Feb. 2006.

72. V. V. Ravi Kumar, A. George, W. Reeves, J. Knight, P. Russell, F. Omenetto, and A. Taylor, "Extruded soft glass photonic crystal fiber for ultrabroad supercontinuum generation," *Optics Express*, vol. 10, pp. 1520–1525, Dec. 2002.

73. J.Y.Y. Leong, P. Petropoulos, J.H.V. Price, H. Ebendorff-Heidepriem, S. Asimakis, R. Moore, K. Frampton, V. Finazzi, X. Feng, T. M. Monro, D. J. Richardson, "High nonlinearity dispersion-shifted lead-silicate holey fibers for efficient 1-μm pumped supercontinuum generation," *J. Lightwave Tech.*, vol. 24, pp. 183–190, Jan. 2006.

74. (*a*) H. Ebendorff-Heidepriem, P. Petropoulos, R. Moore, K. Frampton, D. J. Richardson, and

T. M. Monro, "Fabrication and optical properties of lead silicate glass holey fibers," *J. Non-Crystalline Solids*, vol. 345&346, pp. 293–296, Aug. 2004; (*b*) H. Ebendorff-Heidepriem, T. M. Monro, M. A. van Eijkelenborg, and M.C.J. Large, "Extruded high-NA microstructured polymer optical fiber," *Opt. Commun.*, vol. 273, pp. 133–137, May 2007.

75. G. Barton, M. A. van Eijkelenborg, G. Henry, M.C.J. Large, and J. Zagari, "Fabrication of microstructured polymer optical fibers," *Opt. Fiber Tech.*, vol. 10, pp. 325–335, Oct. 2004.

76. X. Feng, A. K. Mairaj, D. W. Hewak, and T. M. Monro, "Nonsilica glasses for holey fibers," *J. Lightwave Tech.*, vol. 23, pp. 2046–2054, June 2005.

77. D. Kalish, D. L. Key, C. R. Kurkjian, B. K. Turiyal, and T. T. Wang, "Fiber characterization—Mechanical," in S. E. Miller and A. G. Chynoweth, eds., *Optical Fiber Telecommunications*, Academic, New York, 1979.

78. C. R. Kurjian, J. T. Krause, and M. J. Matthewson, "Strength and fatigue of silica optical fibers," *J. Lightwave Tech.*, vol. 7, pp. 1360–1370, Sept. 1989.

79. G. S. Glaesemann and D. J. Walter, "Methods for obtaining long-length strength distributions for reliability prediction," *Optical Eng.*, vol. 30, pp. 746–748, June 1991.

80. K. Yoshida, T. Satoh, N. Enomoto, T. Yagi, H. Hihara, and M. Oku, "Studies on the strength of optical fiber fabricated by a hybridized process," *J. Lightwave Tech.*, vol. 14, pp. 2513–2518, Nov. 1996.

81. K. B. Broberg, *Cracks and Fracture*, Academic, New York, 1999.

82. P.D.T. O'Conner, *Practical Reliability Engineering*, Wiley, Hoboken, NJ, 4th ed., 2002.

83. T. J. Miller, A. C. Hart, W. I. Vroom, and M. J. Bowden, "Silicone and ethylene-vinyl-acetate-coated laser-drawn silica fibers with tensile strengths > 3.5 GN/m^2 (500 kpsi) in > 3 km lengths," *Electron. Lett.*, vol. 14, pp. 603–605, Aug. 1978.

84. J. L. Mrotek, M. J. Matthewson, and C.R. Kurkjian, "The fatigue of high-strength fused silica optical fibers in low humidity,"

J. Non-Crystalline Solids, vol. 297, pp. 91–95, Jan. 2002.

85. Y. Katsuyama, Y. Mitsunaga, H. Kobayashi, and Y. Ishida, "Dynamic fatigue of optical fiber under repeated stress," *J. Appl. Phys.*, vol. 53, pp. 318–321, Jan. 1982.

86. V. Annovazzi-Lodi, S. Donati, S. Merlo, and G. Zapelloni, "Statistical analysis of fiber failures under bending-stress fatigue," *J. Lightwave Tech.*, vol. 15, pp. 288–293, Feb. 1997.

87. TIA/EIA-455-31-C, FOTP-31C, *Proof Testing Optical Fibers by Tension*, 2004 (R 2005).

88. TIA/EIA-455-28-C, FOTP-28C, *Measuring Dynamic Strength and Fatigue Parameters of Optical Fibers by Tension,* 2004 (R 2005).

89. B. Wiltshire and M. H. Reeve, "A review of the environmental factors affecting optical cable design," *J. Lightwave Tech.*, vol. 6, pp. 179–185, Feb. 1988.

90. C. H. Gartside III, P. D. Patel, and M. R. Santana, "Optical fiber cables," in S. E. Miller and I. P Kaminow, eds., *Optical Fiber Telecommunications–II*, Academic, New York, 1988.

91. K. Hogari, S. Furukawa, Y. Nakatsuji, S. Koshio, and K. Nishizawa, "Optical fiber cables for residential and business premises," *J. Lightwave Tech.*, vol. 16, pp. 207–213, Feb. 1998.

92. A. L. Crandall, C. R. Herron, and R. B. Washburn, "Controlling axial load forces on optical fiber cables during installation," Paper NWC3, *OFC/NFOEC 2005 Conf. Program*, Anaheim, CA, March 6–11, 2005.

93. G. Mahlke and P. Gössing, *Fiber Optic Cables: Fundamentals, Cable Design, System Planning*, Wiley, Hoboken, NJ, 4th ed., 2001.

94. O. L. Storaasli, "Compatibility of fiber optic microduct cables with various blowing installation equipment," Paper NWC2, *OFC/NFOEC 2005 Conf. Program*, Anaheim, CA, March 6–11, 2005.

95. TIA/EIA-590-A, *Standard for Physical Location and Protection of Below-Ground Fiber Optic Cable Plant*, July 2001.

96. Special issue on "Undersea communications technology," *AT&T Technical Journal*, vol. 74, no. 1, Jan/Feb. 1995.

97. J. Chesnoy, *Undersea Fiber Communication Systems*, Academic, San Diego, 2002.

98. S. Bigo, "Technologies for global telecommunications using undersea cables," chap. 14, in I. P. Kaminow, T. Li, and A. E. Willner, eds., *Optical Fiber Telecommunications–V*, vol. B, Academic, New York, 2008.

99. L. M. Johnson, *A Practical Installer's Guide to Fiber Optics: Theory, Installation, Maintenance, and Troubleshooting*, Wiley-IEEE Press, Hoboken, NJ, 2010.

100. OFS Fitel, LLC has a series of application documents on their website: www.ofsoptics.com.

101. Draka, Inc., "General optical fiber cable installation considerations," Oct. 2009, www.drakaamericas.com.

102. As some examples, the websites of the following organizations have a series of descriptions and colored photos about recent undersea fiber cable installations: (*a*) Pipe International (www.pipeinternational.com); (*b*) TE SubCom (www.SubCom.com).

103. D. Marcuse, "Loss analysis of single-mode fiber splices," *Bell Sys. Tech. J.*, vol. 56, pp. 703–718, May-June 1977.

CHAPTER 3

Attenuation and Dispersion

In Chapter 2 we showed the structure of optical fibers and examined the concepts of how light propagates along a cylindrical dielectric optical waveguide. Here, we shall continue the discussion of optical fibers by answering two very important questions:

1. What are the loss or signal attenuation mechanisms in a fiber?
2. Why and to what degree do optical signals get distorted as they propagate along a fiber?

Signal attenuation (also known as *fiber loss* or *signal loss*) as is one of the most important properties of an optical fiber because it largely determines the maximum unamplified or repeaterless separation between a transmitter and a receiver. Since amplifiers and repeaters are expensive to fabricate, install, and maintain, the degree of attenuation in a fiber has a large influence on system cost. Of equal importance is signal distortion. The distortion mechanisms in a fiber cause optical signal pulses to broaden as they travel along a fiber. If these pulses travel sufficiently far, they will eventually overlap with neighboring pulses, thereby creating errors in the receiver output. The signal distortion mechanisms thus limit the information-carrying capacity of a fiber.

3.1 Attenuation

Attenuation of a light signal as it propagates along a fiber is an important consideration in the design of an optical communication system; the degree of attenuation plays a major role in determining the maximum transmission distance between a transmitter and a receiver or an in-line amplifier. The basic attenuation mechanisms in a fiber are absorption, scattering, and radiative losses of the optical energy.[1–5] Absorption is related to the fiber material, whereas scattering is associated both with the fiber material and with structural imperfections in the optical waveguide. Attenuation owing to radiative effects originates from perturbations (both microscopic and macroscopic) of the fiber geometry.

This section first discusses the units in which fiber losses are measured and then presents the physical phenomena giving rise to attenuation.

3.1.1 Attenuation Units

As light travels along a fiber, its power decreases exponentially with distance. If $P(0)$ is the optical power in a fiber at the origin (at $z = 0$), then the power $P(z)$ at a distance z farther down the fiber is

$$P(z) = P(0)e^{-\alpha_p z} \tag{3.1a}$$

where

$$\alpha_p = \frac{1}{z} \ln\left[\frac{P(0)}{P(z)}\right] \tag{3.1b}$$

is the fiber *attenuation coefficient* given in units of, for example, km^{-1}. Note that the units for $2z\alpha_p$ can also be designated by *nepers* (see App. D).

For simplicity in calculating optical signal attenuation in a fiber, the common procedure is to express the attenuation coefficient in units of *decibels per kilometer,* denoted by dB/km. Designating this parameter by α, we have

$$\alpha(\text{dB/km}) = \frac{10}{z} \log\left[\frac{P(0)}{P(z)}\right] = 4.343 \, \alpha_p (km^{-1}) \tag{3.1c}$$

This parameter is generally referred to as the *fiber loss* or the *fiber attenuation.* It depends on several variables, as is shown in the following sections, and it is a function of the wavelength.

Example 3.1   An ideal fiber would have no loss so that $P_{out} = P_{in}$. This corresponds to a 0-dB/km attenuation, which, in practice, is impossible. An actual low-loss fiber may have a 3-dB/km average loss at 900 nm, for example.

This means that the optical signal power would decrease by 50 percent over a 1-km length and would decrease by 75 percent (a 6-dB loss) over a 2-km length, since loss contributions expressed in decibels are additive.

Example 3.2   As Sec. 1.3 describes, optical powers are commonly expressed in units of *dBm,* which is the decibel power level referred to 1 mW. Consider a 30-km long optical fiber that has an attenuation of 0.4 dB/km at 1310 nm. Suppose we want to find the optical output power P_{out} if 200 μW of optical power is launched into the fiber. We first express the input power in dBm units:

$$P_{in}(\text{dBm}) = 10 \log\left[\frac{P_{in}(\text{W})}{1\,\text{mW}}\right]$$

$$= 10 \log\left[\frac{200 \times 10^{-6}\,\text{W}}{1 \times 10^{-3}\,\text{W}}\right] = -7.0 \text{ dBm}$$

From Eq. (3.1c) with $P(0) = P_{in}$ and $P(z) = P_{out}$ the output power level (in dBm) at $z = 30$ km is

$$P_{out}(\text{dBm}) = 10 \log\left[\frac{P_{out}(\text{W})}{1\,\text{mW}}\right]$$

$$= 10 \log\left[\frac{P_{in}(\text{W})}{1\,\text{mW}}\right] - \alpha z$$

$$= -7.0 \text{ dBm} - (0.4 \text{ dB/km}) (30 \text{ km})$$

$$= -19.0 \text{ dBm}$$

In unit of watts, the output power is

$$P(30 \text{ km}) = 10^{-19.0/10}(1 \text{ mW}) = 12.6 \times 10^{-3} \text{ mW}$$

$$= 12.6 \, \mu\text{W}$$

3.1.2 Absorption

Absorption is caused by three different mechanisms:

1. Absorption by atomic defects in the glass composition.
2. Extrinsic absorption by impurity atoms in the glass material.
3. Intrinsic absorption by the basic constituent atoms of the fiber material.

Atomic defects are imperfections in the atomic structure of the fiber material. Examples of these defects include missing molecules, high-density clusters of atom groups, or oxygen defects in the glass structure. Usually, absorption losses arising from these defects are negligible compared with intrinsic and impurity absorption effects. However, they can be significant if the fiber is exposed to ionizing radiation, as might occur in a nuclear reactor environment, in medical radiation therapies, in space missions that pass through the earth's Van Allen belts, or in accelerator instrumentation.[6–9] In such applications, high radiation doses may be accumulated over several years.

Radiation damages a material by changing its internal structure. The damage effects depend on the energy of the ionizing particles or rays (e.g., electrons, neutrons, or gamma rays), the radiation flux (dose rate), and the fluence (particles per square centimeter). The total dose a material receives is expressed in units of rad(Si), which is a measure of radiation absorbed in bulk silicon. This unit is defined as

$$1 \text{ rad(Si)} = 100 \text{ erg/g} = 0.01 \text{ J/kg}$$

The basic response of a fiber to ionizing radiation is an increase in attenuation owing to the creation of atomic defects, or attenuation centers, that absorb optical energy. The higher the radiation level, the larger the attenuation, as Fig. 3.1*a* illustrates. However, the attenuation centers will relax or anneal out with time, as shown in Fig. 3.1*b*. The degree of the radiation effects depends on the dopant materials used in the fiber. Pure silica fibers or fibers with a low Ge doping and no other dopants have the lowest radiation-induced losses.
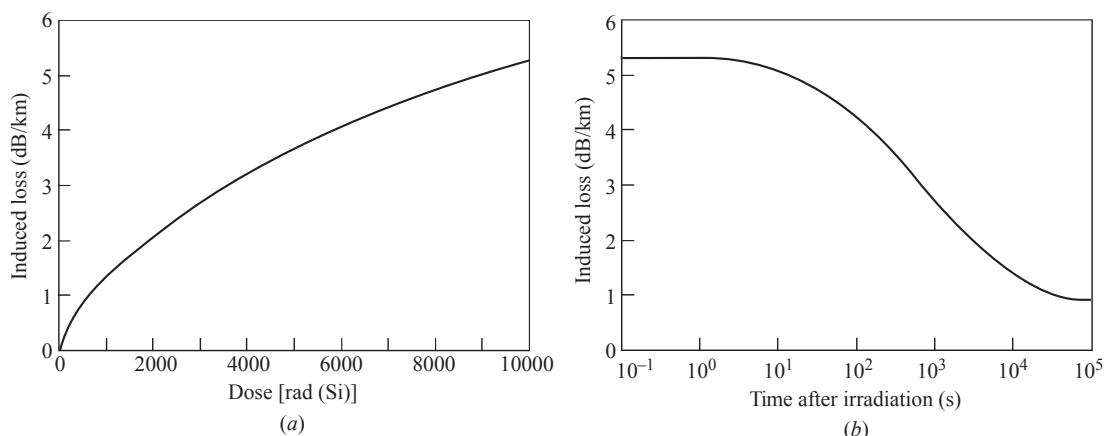
Fig. 3.1 *General trend of the effects of ionizing radiation on optical fiber attenuation. (a) Loss increase during steady irradiation to a total dose of 10^4 rad (SiO_2). (b) Subsequent recovery as a function of time after radiation has stopped. (Modified with permission from West et al.,[7] © 1994, IEEE.)*

Table 3.1 *Examples of absorption loss in silica glass at different wavelengths due to 1 ppm of water-ions and various transition-metal impurities*

Impurity	Loss due to 1 ppm of impurity (dB/km)	Absorption peak (nm)
Iron: Fe^{2+}	0.68	1100
Iron: Fe^{3+}	0.15	400
Copper: Cu^{2+}	1.1	850
Chromium: Cr^{2+}	1.6	625
Vanadium: V^{4+}	2.7	725
Water: OH^-	1.0	950
Water: OH^-	2.0	1240
Water: OH^-	4.0	1380

The dominant absorption factor in silica fibers is the presence of minute quantities of impurities in the fiber material. These impurities include OH^- (water) ions that are dissolved in the glass and transition metal ions such as iron, copper, chromium, and vanadium. Transition metal impurity levels were around 1 part per million (ppm) in glass fibers made in the 1970s, which resulted in losses ranging from 1 to 4 dB/km, as Table 3.1 shows. Impurity absorption losses occur either because of electron transitions between the energy levels within these ions or because of charge transitions between ions. The absorption peaks of the various transition metal impurities tend to be broad, and several peaks may overlap, which further broadens the absorption in a specific region. Modern vapor-phase fiber techniques for producing a fiber preform (see Sec. 2.9) have reduced the transition-metal impurity levels by several orders of magnitude. Such low impurity levels allow the fabrication of low-loss fibers.

The presence of OH ion impurities in a fiber preform results mainly from the oxyhydrogen flame used in the hydrolysis reaction of the $SiCl_4$, $GeCl_4$, and $POCl_3$ starting materials. Water impurity concentrations of less than a few parts per billion (ppb) are required if the attenuation is to be less than 20 dB/km. The high levels of OH ions in early fibers resulted in large absorption peaks at 725, 950, 1240, and 1380 nm. Regions of low attention lie between these absorption peaks.

The peaks and valleys in the attenuation curves resulted in the designation of the various *transmission windows* shown in Fig. 3.2. By reducing the residual OH content of fibers to below 1 ppb, standard commercially available single-mode fibers have nominal attenuations of 0.4 dB/km at 1310 nm (in the O-band) and less than 0.25 dB/km at 1550 nm (in the C-band). Further elimination of water ions diminishes the absorption peak around 1440 nm and thus opens up the E-band for data transmission, as indicated by the dashed line in Fig. 3.2. Optical fibers that can be used in the E-band are known by names such as *low-water-peak* or *full-spectrum fibers*.

Intrinsic absorption is associated with the basic fiber material (e.g., pure SiO_2) and is the principal physical factor that defines the transparency window of a material over a specified spectral region. Intrinsic absorption sets the fundamental lower limit on absorption for any particular material; it is defined as the absorption that occurs when the material is in a perfect state with no density variations, impurities, or material inhomogeneities.

Intrinsic absorption results from electronic absorption bands in the ultraviolet region and from atomic vibration bands in the near-infrared region. The electronic absorption bands are associated with the band gaps of the amorphous glass materials. Absorption occurs when a photon interacts with an electron
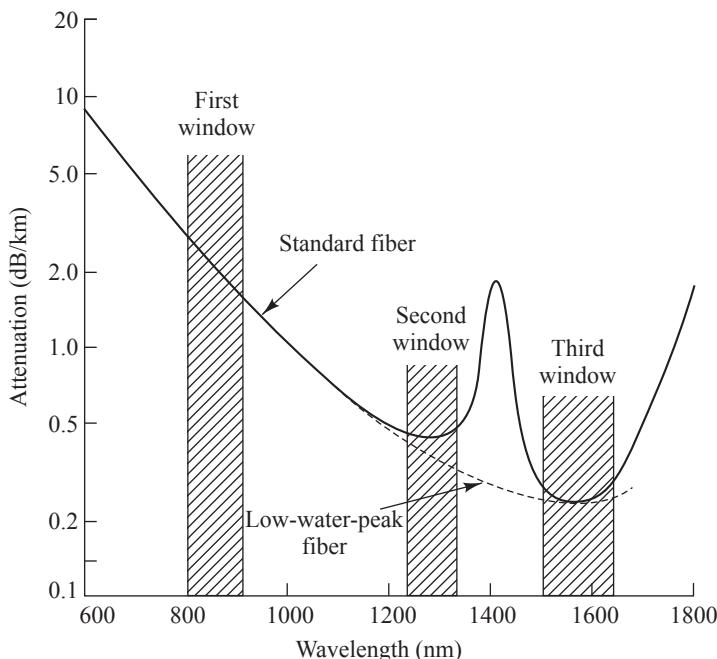
Fig. 3.2 *Optical fiber attenuation as a function of wavelength yields nominal values of 0.40 dB/km at 1310 nm and 0.25 dB/km at 1550 nm for standard single-mode fiber. Absorption by water molecules causes the attenuation peak around 1400 nm for standard fiber. The dashed curve is the attenuation for low-water-peak fiber.*

in the valence band and excites it to a higher energy level, as is described in Sec. 2.1. The ultraviolet edge of the electron absorption bands of both amorphous and crystalline materials follow the empirical relationship[1,3]

$$\alpha_{uv} = Ce^{E/E_0} \tag{3.2a}$$

which is known as Urbach's rule. Here, C and E_0 are empirical constants and E is the photon energy. The magnitude and characteristic exponential decay of the ultraviolet absorption are shown in Fig. 3.3. Since E is inversely proportional to the wavelength λ, ultraviolet absorption decays exponentially with increasing wavelength. In particular, the ultraviolet loss contribution in dB/km at any wavelength (given in μm) can be expressed empirically (derived from observation or experiment) as a function of the mole fraction x of GeO_2 as[10]

$$\alpha_{uv} = \frac{154.2x}{46.6x + 60} \times 10^{-2} \exp\left(\frac{4.63}{\lambda}\right) \tag{3.2b}$$

As shown in Fig. 3.3, the ultraviolet loss is small compared with scattering loss in the near-infrared region.

Example 3.3 Consider two silica fibers that are doped with 6 percent and 18 percent mole fractions of GeO$_2$, respectively. Compare the ultraviolet absorptions at wavelengths of 0.7 μm and 1.3 μm.

Solution: Using Eq. (3.2b) for the ultraviolet absorption, we have the following:

(*a*) For the fiber with $x = 0.06$ and $\lambda = 0.7$ μm

$$\alpha_{uv} = \frac{1.542(0.06)}{46.6(0.06) + 60} \exp\left(\frac{4.63}{0.7}\right) = 1.10 \text{ dB/km}$$

(*b*) For the fiber with $x = 0.06$ and $\lambda = 1.3$ μm

$$\alpha_{uv} = \frac{1.542(0.06)}{46.6(0.06) + 60} \exp\left(\frac{4.63}{1.3}\right) = 0.07 \text{dB/km}$$

(*c*) For the fiber with $x = 0.18$ and $\lambda = 0.7$ μm

$$\alpha_{uv} = \frac{1.542(0.18)}{46.6(0.18) + 60} \exp\left(\frac{4.63}{0.7}\right) = 3.03 \text{ dB/km}$$

(*d*) For the fiber with $x = 0.18$ and $\lambda = 1.3$ μm

$$\alpha_{uv} = \frac{1.542(0.18)}{46.6(0.18) + 60} \exp\left(\frac{4.63}{1.3}\right) = 0.19 \text{ dB/km}$$
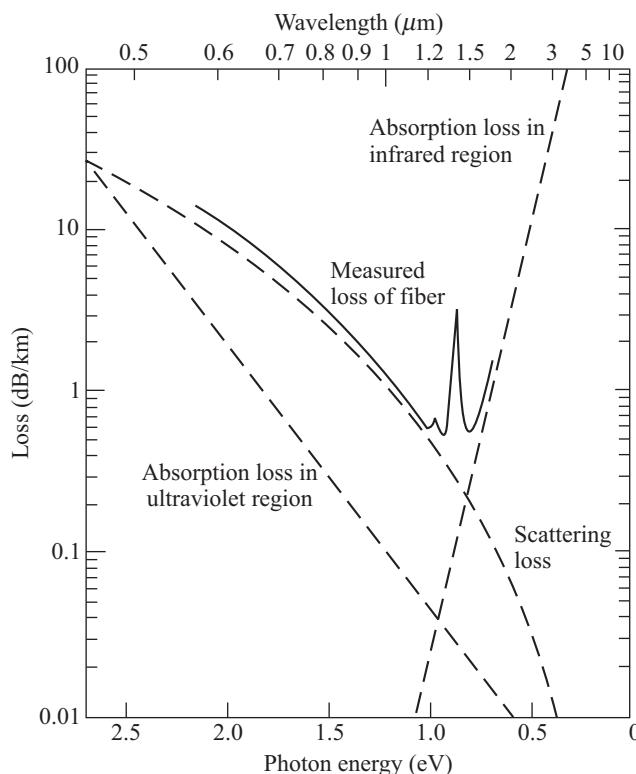
Fig. 3.3 *Optical fiber attenuation characteristics and their limiting mechanisms for a GeO$_2$-doped low-loss low-water-content silica fiber. (Reproduced with permission from Osanai et al.[13])*
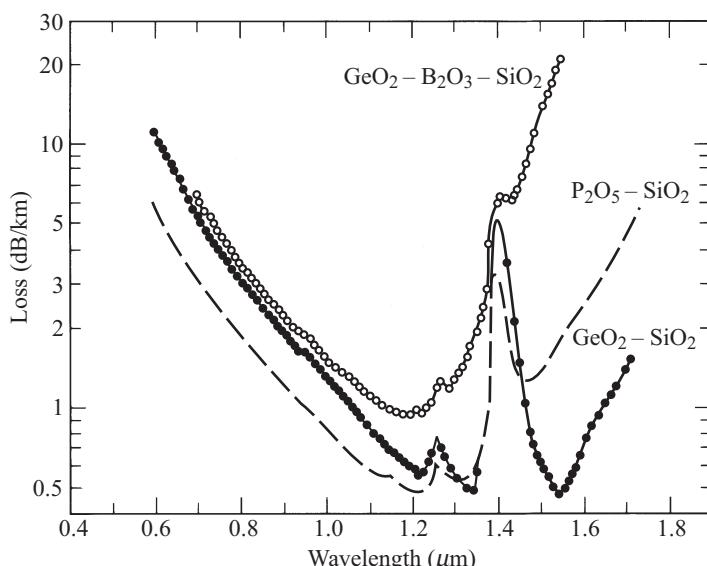
Fig. 3.4 *A comparison of the infrared absorption induced by various doping materials in low-loss silica fibers. (Reproduced with permission from Osanai et al.[13])*

In the near-infrared region above 1.2 μm, the optical waveguide loss is predominantly determined by the presence of OH ions and the inherent infrared absorption of the constituent material. The inherent infrared absorption is associated with the characteristic vibration frequency of the particular chemical bond between the atoms of which the fiber is composed. An interaction between the vibrating bond and the electromagnetic field of the optical signal results in a transfer of energy from the field to the bond, thereby giving rise to absorption. This absorption is quite strong because of the many bonds present in the fiber. An empirical expression for the infrared absorption in dB/km for GeO$_2$-SiO$_2$ glass with λ given in μm is[10]

$$\alpha_{\text{IR}} = 7.81 \times 10^{11} \times \exp\left(\frac{-48.48}{\lambda}\right) \qquad (3.3)$$

These mechanisms result in a wedge-shaped spectral-loss characteristic. Within this wedge, losses as low as 0.148 dB/km at 1.57 μm in a single-mode fiber have been measured.[11,12] A comparison[13] of the infrared absorption induced by various doping materials in low-water content fibers is shown in Fig. 3.4. This indicates that for operation at longer wavelengths GeO$_2$-doped fiber material is the most desirable. Note that the absorption curve shown in Fig. 3.3 is for a GeO$_2$ doped fiber.

3.1.3 Scattering Losses

Scattering losses in glass arise from microscopic variations in the material density, from compositional fluctuations, and from structural inhomogeneities or defects occurring during fiber manufacture. As Sec. 2.7 describes, glass is composed of a randomly connected network of molecules. Such a structure naturally contains regions in which the molecular density is either higher or lower than the average density in the glass. In addition, since glass is made up of several oxides, such as SiO$_2$, GeO$_2$, and P$_2$O$_5$,

compositional fluctuations can occur. These two effects give rise to refractive-index variations that occur within the glass over distances that are small compared with the wavelength. These index variations cause a Rayleigh-type scattering of the light. Rayleigh scattering in glass is the same phenomenon that scatters light from the sun in the atmosphere, thereby giving rise to a blue sky.

The expressions for scattering-induced attenuation are fairly complex owing to the random molecular nature and the various oxide constituents of glass. For single-component glass the scattering loss at a wavelength λ (given in μm) resulting from density fluctuations can be approximated by[3,14] (in base e units)

$$\alpha_{\text{scat}} = \frac{8\pi^3}{3\lambda^4}(n^2-1)^2 \, k_B T_f \beta_T \tag{3.4a}$$

Here, n is the refractive index, k_B is Boltzmann's constant, β_T is the isothermal compressibility of the material, and the fictive temperature T_f is the temperature at which the density fluctuations are frozen into the glass as it solidifies (after having been drawn into a fiber). Alternatively, the relation[3,15] (in base e units)

$$\alpha_{\text{scat}} = \frac{8\pi^3}{3\lambda^4} \, n^8 p^2 k_B T_f \beta_T \tag{3.4b}$$

has been derived, where p is the photoelastic coefficient. A comparison of Eqs. (3.4a) and (3.4b) is given in Prob. 3.6. Note that Eqs. (3.4a) and (3.4b) are given in units of *nepers* (that is, base e units). As shown in Eq. (3.1), to change this to decibels for optical power attenuation calculations, multiply these equations by 10 log e = 4.343.

Example 3.4 For silica the fictive temperature T_f is 1400 K, the isothermal compressibility β_T is 6.8×10^{-12} cm^2/dyn = 6.8×10^{-11} m^2/N, and the photoelastic coefficient is 0.286. Estimate the scattering loss at a 1.30-μm wavelength where $n = 1.450$.

Solution: Using Eq. (3.4b),

$$\alpha_{\text{scat}} = \frac{8\pi^3}{3\lambda^4}n^8 p^2 k_B T_f \beta_T$$

$$= \frac{8\pi^3}{3(1.3)^4}(1.45)^8(0.286)^2$$

$$\times (1.38 \times 10^{-23})(1400)(6.8 \times 10^{-12})$$

$$= 6.08 \times 10^{-2}\,\text{nepers/km} = 0.26\,\text{dB/km}$$

Example 3.5 For pure silica glass an approximate equation for the Rayleigh scattering loss is given by

$$\alpha(\lambda) \simeq \alpha_0 \left(\frac{\lambda_0}{\lambda}\right)^4$$

where $\alpha_0 = 1.64$ dB/km at $\lambda_0 = 850$ nm. This formula predicts scattering losses of 0.291 dB/km at 1310 nm and 0.148 dB/km at 1550 nm.

For multicomponent glasses the scattering at a wavelength λ (measured in μm) is given by[3]

$$\alpha = \frac{8\pi^3}{3\lambda^4}(\delta n^2)^2 \, \delta V \tag{3.5}$$

where the square of the mean-square refractive-index fluctuation $(\delta n^2)^2$ over a volume of δV is

$$(\delta n^2)^2 = \left(\frac{\partial n^2}{\partial \rho}\right)^2 (\delta \rho)^2 + \sum_{i=1}^{m} \left(\frac{\partial n^2}{\partial C_i}\right)(\delta C_i)^2 \tag{3.6}$$
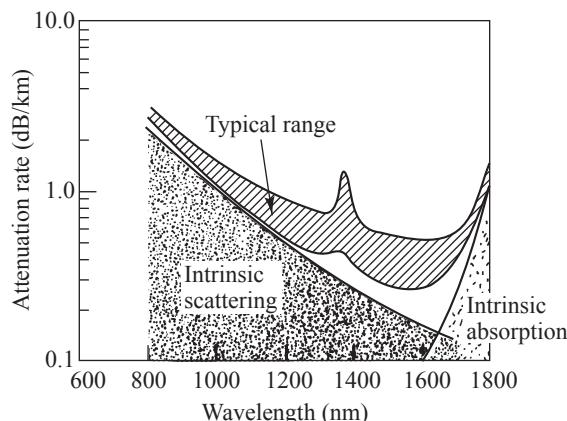
Fig. 3.5 *Typical spectral attenuation range for production-run graded-index multimode fibers. (Reproduced with permission from Keck,[16] © 1985, IEEE.)*
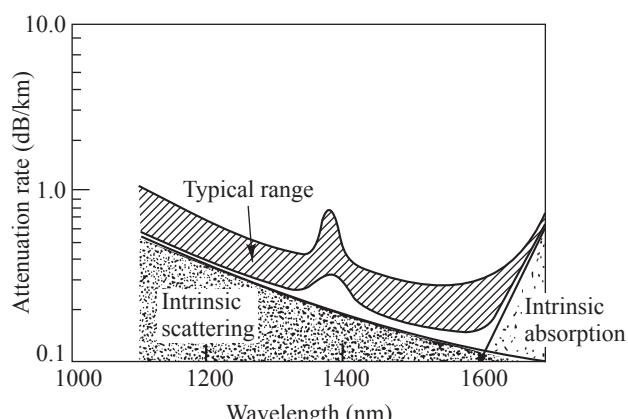
Fig. 3.6 *Typical spectral attenuation range for production-run single-mode fibers. (Reproduced with permission from Keck,[16] © 1985, IEEE.)*

Here, $\delta \rho$ is the density fluctuation and δC_i, is the concentration fluctuation of the ith glass component. Their magnitudes must be determined from experimental scattering data. The factors $\partial n^2/\partial \rho$ and $\partial n^2/\partial C_i$ are the variations of the square of the index with respect to the density and the ith glass component, respectively.

Structural inhomogeneities and defects created during fiber fabrication can also cause scattering of light out of the fiber. These defects may be in the form of trapped gas bubbles, unreacted starting materials, and crystallized regions in the glass. In general, the preform manufacturing methods that have evolved have minimized these extrinsic effects to the point where scattering that results from them is negligible compared with the intrinsic Rayleigh scattering.

Since Rayleigh scattering follows a characteristic λ^{-4} dependence, it decreases dramatically with increasing wavelength, as is shown in Fig. 3.3. For wavelengths below about 1 μm it is the dominant loss mechanism in a fiber and gives the attenuation-versus-wavelength plots their characteristic downward trend with increasing wavelength. At wavelengths longer than 1 μm, infrared absorption effects tend to dominate optical signal attenuation.

Combining the infrared, ultraviolet, and scattering losses, we get the results shown in Fig. 3.5 for multimode fibers and Fig. 3.6 for single-mode fibers.[16] Both of these figures are for typical commercial-grade silica fibers. The losses of multimode fibers are generally higher than those of single-mode fibers. This is a result of higher dopant concentrations and the accompanying larger scattering loss due to greater compositional fluctuation in multimode

fibers. In addition, multimode fibers are subject to higher-order-mode losses owing to perturbations at the core-cladding interface.

3.1.4 Bending Losses

Radiative losses occur whenever an optical fiber undergoes a bend of finite radius of curvature.[17–26] Fibers can be subject to two types of curvatures: (*a*) macroscopic bends having radii that are large compared with the fiber diameter, such as those that occur when a fiber cable turns a corner, and (*b*) random microscopic bends of the fiber axis that can arise when the fibers are incorporated into cables.

Fig. 3.7 *Sketch of the fundamental mode field in a curved optical waveguide. (Reproduced with permission from E.A.J. Marcatili and S. E. Miller,* Bell Sys. Tech. J., *vol. 48, p. 2161, Sept. 1969, © 1969, AT&T.)*

Let us first examine large-curvature radiation losses, which are known as *macrobending losses* or simply *bending losses.* For slight bends the excess loss is extremely small and is essentially unobservable. As the radius of curvature decreases, the loss increases exponentially until at a certain critical radius the curvature loss becomes observable. If the bend radius is made a bit smaller once this threshold point has been reached, the losses suddenly become extremely large.

Qualitatively, these curvature loss effects can be explained by examining the modal electric field distributions shown in Fig. 2.19. Recall that this figure shows that any bound core mode has an evanescent field tail in the cladding that decays exponentially as a function of distance from the core. Since this field tail moves along with the field in the core, part of the energy of a propagating mode travels in the fiber cladding. When a fiber is bent, the field tail on the far side of the center of curvature must move faster to keep up with the field in the core, as is shown in Fig. 3.7 for the lowest-order fiber mode. At a certain critical distance x_c from the center of the fiber, the field tail would have to move faster than the speed of light to keep up with the core field. Since this is not possible, the optical energy in the field tail beyond x_c radiates away.

The amount of optical radiation from a bent fiber depends on the field strength at x_c and on the radius of curvature R. Since higher-order modes are bound less tightly to the fiber core than lower-order modes, the higher-order modes will radiate out of the fiber first. Thus the total number of modes that can be supported by a curved fiber is less than in a straight fiber. The following expression[18] has been derived for the effective number of modes M_{eff} that are guided by a curved multimode fiber of radius *a:*

$$M_{\text{eff}} = M_\infty \left\{ 1 - \frac{\alpha + 2}{2\alpha\Delta} \left[\frac{2a}{R} + \left(\frac{3}{2n_2 kR} \right)^{2/3} \right] \right\} \tag{3.7}$$

where α defines the graded-index profile, Δ is the core-cladding index difference, n_2 is the cladding refractive index, $k = 2\pi/\lambda$ is the wave propagation constant, and
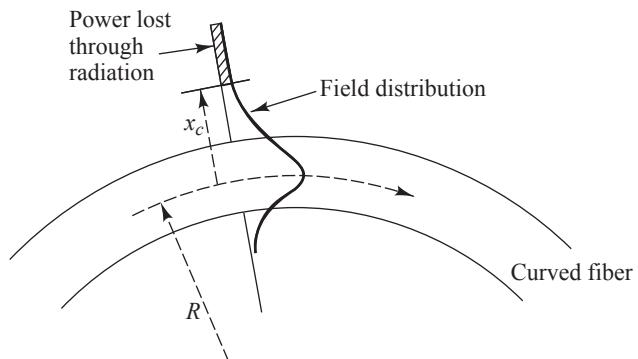
$$M_\infty = \frac{\alpha}{\alpha + 2}(n_1 ka)^2 \Delta \tag{3.8}$$

is the total number of modes in a straight fiber [see Eq. (2.81)].

Example 3.6 Consider a graded-index multimode fiber for which the index profile $\alpha = 2.0$, the core index $n_1 = 1.480$, the core-cladding index difference $\Delta = 0.01$, and the core radius $a = 25$ μm. If the radius of curvature of the fiber is $R = 1.0$ cm, what percentage of the modes remain in the fiber at a 1300-nm wavelength?

Solution: From Eq. (3.7) the percentage of modes at a given curvature R is

$$\frac{M_{eff}}{M_\infty} = 1 - \frac{\alpha + 2}{2\alpha\Delta}\left[\frac{2a}{R} + \left(\frac{3}{2n_2 kR}\right)^{2/3}\right]$$

$$= 1 - \frac{1}{.01}\left[\frac{2(25)}{10000} + \left(\frac{3(1.3)}{2(1.465)2\pi(10000)}\right)^{2/3}\right]$$

$$= 0.42$$

Thus 42 percent of the modes remain in this fiber at a 1.0-cm bend radius.

Another form of radiation loss in optical waveguide results from mode coupling caused by random microbends of the optical fiber.[27–30] *Microbends* are repetitive small-scale fluctuations in the radius of curvature of the fiber axis, as is illustrated in Fig. 3.8. They are caused either by nonuniformities in the manufacturing of the fiber or by nonuniform lateral pressures created during the cabling of the fiber. The latter effect is often referred to as *cabling* or *packaging losses*. An increase in attenuation results from microbending because the fiber curvature causes repetitive coupling of energy between the guided modes and the leaky or nonguided modes in the fiber.

One method of minimizing microbending losses is by extruding a compressible jacket over the fiber. When external forces are applied to this configuration, the jacket will be deformed but the fiber will
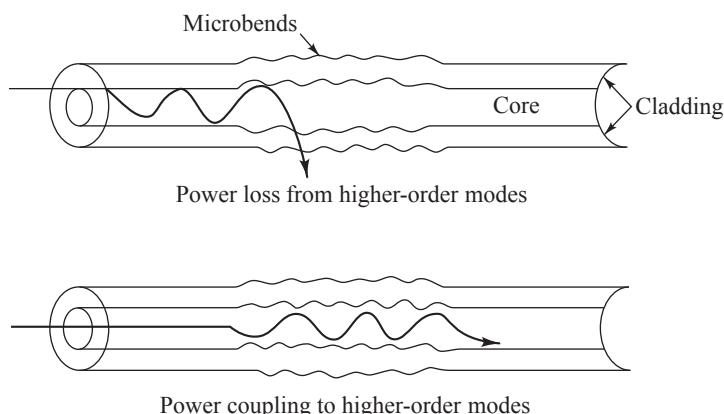
Fig. 3.8 *Small-scale fluctuations in the radius of curvature of the fiber axis lead to microbending losses. Microbends can shed higher-order modes and can cause power from low-order modes to couple to higher-order modes.*

tend to stay relatively straight. For a multimode graded-index fiber having a core radius a, outer radius b (excluding the jacket), and index difference Δ, the microbending loss α_M of a jacketed fiber is reduced from that of an unjacketed fiber by a factor[31]

$$F(\alpha_M) = \left[1 + \pi\Delta^2 \left(\frac{b}{a} \right)^4 \frac{E_f}{E_j} \right]^{-2} \tag{3.9}$$

Here, E_j and E_f are the Young's moduli of the jacket and fiber, respectively. The Young's modulus of common jacket materials ranges from 20 to 500 MPa. The Young's modulus of fused silica glass is about 65 GPa.

3.1.5 Core and Cladding Losses

Upon measuring the propagation losses in an actual fiber, all the dissipative and scattering losses will be manifested simultaneously. Since the core and cladding have different indices of refraction and therefore differ in composition, the core and cladding generally have different attenuation coefficients, denoted α_1 and α_2, respectively. If the influence of modal coupling is ignored,[32] the loss for a mode of order (v, m) for a step-index waveguide is

$$\alpha_{vm} = \alpha_1 \frac{P_{\text{core}}}{P} + \alpha_2 \frac{P_{\text{clad}}}{P} \tag{3.10a}$$

where the fractional powers P_{core}/P and P_{clad}/P are shown in Fig. 2.27 for several low-order modes. Using Eq. (2.71), this can be written as

$$\alpha_{vm} = \alpha_1 + (\alpha_2 - \alpha_1) \frac{P_{\text{clad}}}{P} \tag{3.10b}$$

The total loss of the waveguide can be found by summing over all modes weighted by the fractional power in that mode.

For the case of a graded-index fiber the situation is much more complicated. In this case, both the attenuation coefficients and the modal power tend to be functions of the radial coordinate. At a distance r from the core axis the loss is[32]

$$\alpha(r) = \alpha_1 + (\alpha_2 - \alpha_1) \frac{n^2(0) - n^2(r)}{n^2(0) - n_2^2} \tag{3.11}$$

where α_1 and α_2 are the axial and cladding attenuation coefficients, respectively, and the n terms are defined by Eq. (2.78). The loss encountered by a given mode is then

$$\alpha_{gi} = \frac{\int_0^\infty \alpha(r)\, p(r)\, r\; dr}{\int_0^\infty p(r)\, r\; dr} \tag{3.12}$$

where $p(r)$ is the power density of that mode at r. The complexity of the multimode waveguide has prevented an experimental correlation with a model. However, it has generally been observed that the loss increases with increasing mode number.[26,33]

3.2 Signal Dispersion in Fibers

As shown in Fig. 3.9, an optical signal weakens from attenuation mechanisms and broadens due to dispersion effects as it travels along a fiber. Eventually these two factors will cause neighboring pulses to overlap. After a certain amount of overlap occurs, the receiver can no longer distinguish the individual adjacent pulses and errors arise when interpreting the received signal.

This section first discusses the general factors that cause signal distortion and then examines the various dispersion mechanisms in more detail. Section 3.2.2 addresses modal delay and shows how this delay is related to the information-carrying capacity of a multimode fiber in terms of a transmitted bit rate B. Section 3.2.3 examines the various factors contributing to dispersion in terms of the frequency dependence of the propagation constant β. The next topics include a discussion of group velocity in Sec. 3.2.4 and details of the various dispersion mechanisms in Sec. 3.2.5 through 3.2.8.

3.2.1 Overview of Dispersion Origins

Signal dispersion is a consequence of factors such as intermodal delay (also called intermodal dispersion), intramodal dispersion, polarization-mode dispersion, and higher-order dispersion effects. These distortions can be explained by examining the behavior of the group velocities of the guided modes, where the *group velocity* is the speed at which energy in a particular mode travels along the fiber (see Sec. 3.2.4).
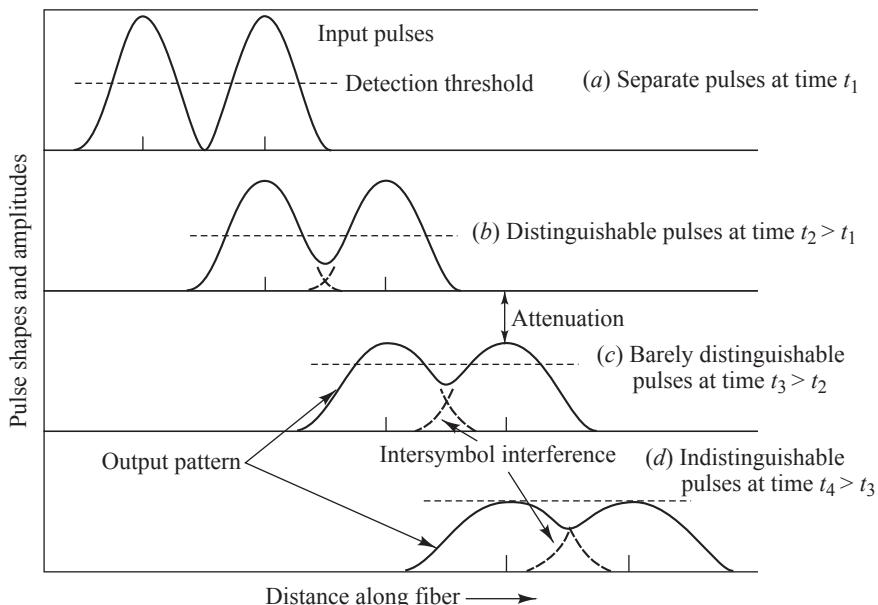
Fig. 3.9 *Broadening and attenuation of two adjacent pulses as they travel along a fiber. (a) Originally the pulses are separate; (b) the pulses overlap slightly and are clearly distinguishable; (c) the pulses overlap significantly and are barely distinguishable; (d) eventually the pulses strongly overlap and are indistinguishable*

Intermodal delay (or simply *modal delay*) appears only in multimode fibers. Modal delay is a result of each mode having a different value of the group velocity at a single frequency. From this effect one can derive an intuitive picture of the information-carrying capacity of a multimode fiber.

Intramodal dispersion or chromatic dispersion is pulse spreading that takes place within a single mode. This spreading arises from the finite spectral emission width of an optical source. The phenomenon also is known as *group velocity dispersion*, since the dispersion is a result of the group velocity being a function of the wavelength. Because intramodal dispersion depends on the wavelength, its effect on signal distortion increases with the spectral width of the light source. The spectral width is the band of wavelengths over which the source emits light. This wavelength band normally is characterized by the root-mean-

Fig. 3.10 *Spectral emission pattern of a representative $Ga_{1-x}Al_xAs$ LED with a peak emission at 850 nm. The width of the spectral pattern at its half-power point is 36 nm.*

square (rms) spectral width σ_λ. Depending on the device structure of a light-emitting diode (LED), the spectral width is approximately 4 to 9 percent of a central wavelength. For example, as Fig. 3.10 illustrates, if the peak wavelength of an LED is 850 nm, a typical source spectral width would be 36 nm; that is, such an LED emits most of its light in the 832-to-868-nm wavelength band. Laser diode optical sources exhibit much narrower spectral widths, with typical values being 1–2 nm for multimode lasers and 10^{-4} nm for single-mode lasers (see Chapter 4).

The two main causes of intramodal dispersion are as follows:

1. *Material dispersion* arises due to the variations of the refractive index of the core material as a function of wavelength. Material dispersion also is referred to as *chromatic dispersion*, since this is the same effect by which a prism spreads out a spectrum. This refractive index property causes a wavelength dependence of the group velocity of a given mode; that is, pulse spreading occurs even when different wavelengths follow the same path.

2. *Waveguide dispersion* causes pulse spreading because only part of the optical power propagation along a fiber is confined to the core. Within a single propagating mode, the cross-sectional distribution of light in the optical fiber varies for different wavelengths. Shorter wavelengths are more completely confined to the fiber core, whereas a larger portion of the optical power at longer wavelengths propagates in the cladding, as shown in Fig. 3.11. The refractive index is lower in the cladding than in the core, so the fraction of light power propagating in the cladding travels faster than the light confined to the core. In addition, we note that the index of refraction depends on the wavelength (see Sec. 3.2.5) so that different spectral components within a single mode have different propagation speeds. Dispersion thus arises because the difference in core-cladding spatial power distributions, together with the speed variations of the various wavelengths, causes a
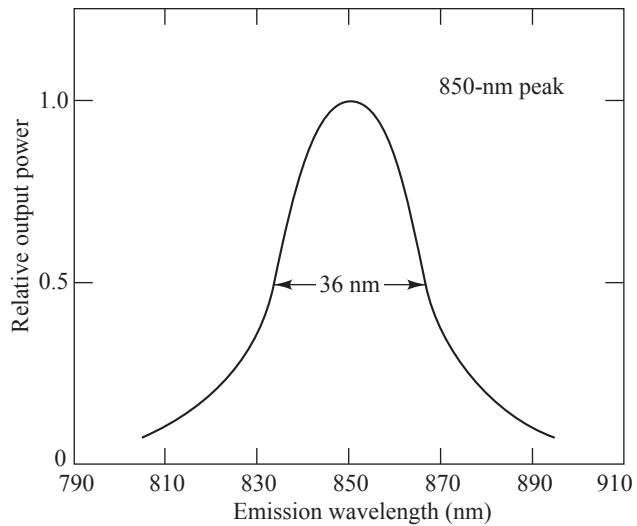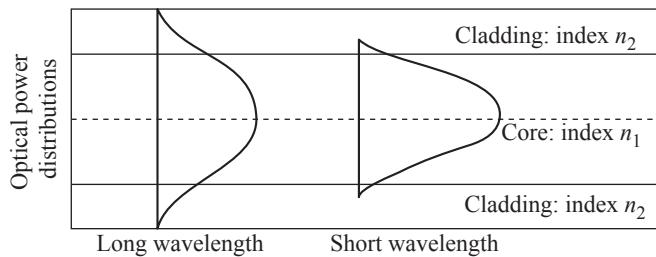
Fig. 3.11 *Shorter wavelengths are confined closer to the center of a fiber core than longer wavelengths.*

change in propagation velocity for each spectral component. The degree of waveguide dispersion depends on the fiber design (see Sec. 3.3.1). Waveguide dispersion usually can be ignored in multimode fibers, but its effect is significant in single-mode fibers.

Polarization-mode dispersion results from the fact that light-signal energy at a given wavelength in a single-mode fiber actually occupies two orthogonal polarization states or modes (see Sec. 2.5). At the start of the fiber the two polarization states are aligned. However, since fiber material is not perfectly uniform throughout its length, each polarization mode will encounter a slightly different refractive index. Consequently each mode will travel at a slightly different velocity. The resulting difference in propagation times between the two orthogonal polarization modes will cause pulse spreading. Section 3.2.8 gives more details on this effect.

3.2.2 Modal Delay

Intermodal dispersion or *modal delay* appears only in multimode fibers. This signal-distorting mechanism is a result of each mode having a different value of the group velocity at a single frequency. To see why the delay arises, consider the meridional ray picture given in Fig. 2.17 for a multimode step-index fiber. The steeper the angle of propagation of the ray congruence, the higher is the mode number and, consequently, the slower the axial group velocity. This variation in the group velocities of the different modes results in a group delay spread, which is the intermodal dispersion. This dispersion mechanism is eliminated by single-mode operation but is important in multimode fibers. The maximum pulse broadening arising from the modal delay is the difference between the travel time $T_{\max}$ of the longest ray congruence paths (the highest-order mode) and the travel time $T_{\min}$ of the shortest ray congruence paths (the fundamental mode). This broadening is simply obtained from ray tracing and for a fiber of length L is given by

$$\Delta T = T_{\max} - T_{\min} = \frac{n_1}{c}\left(\frac{L}{\sin\varphi_c} - L\right) = \frac{L\,n_1^2}{c\,n_2}\Delta \tag{3.13}$$

where from Eq. (2.21) $\sin\varphi_c = n_2/n_1$ and Δ is the index difference.

The question now arises as to what maximum bit rate B can be sent over a multimode step-index fiber. Typically the fiber capacity is specified in terms of the *bit rate-distance product BL*, that is, the bit rate times the possible transmission distance L. In order for neighboring signal pulses to remain distinguishable at the receiver, the pulse spread should be less than $1/B$, which is the width of a bit period. For example,

Example 3.7   Consider a 1-km long multimode step-index fiber in which $n_1 = 1.480$ and $\Delta = 0.01$, so that $n_2 = 1.465$. What is the modal delay per length in this fiber?

Solution: Eq. (3.13) yields

$$\frac{\Delta T}{L} = \frac{n_1^2 \Delta}{c n_2} = 50 \text{ ns/km}$$

This means that a pulse broadens by 50 ns after traveling a distance of 1 km in this type of fiber.

a stringent requirement for a high-performance link might be $\Delta T \leq 0.1/B$. In general, we need to have $\Delta T < 1/B$. Using Eq. (3.13) this inequality gives the bit rate-distance product

$$BL < \frac{n_2}{n_1^2} \frac{c}{\Delta}$$

Taking values of $n_1 = 1.480$, $n_2 = 1.465$, and $\Delta = 0.01$, the capacity of this multimode step-index fiber is $BL = 20$ Mb/s-km.

Example 3.8   Viewed alternatively, as illustrated in Example 3.7, for a multimode step-index fiber with a bandwidth-distance value of $BL = 20$ Mb/s-km the pulse spreading is 50 ns/km. As an example, suppose the pulse width in a transmission system is allowed to widen by at most 25 percent. Then for a 10-Mb/s data rate, in which one pulse is transmitted every 100 ns, this limitation allows a spread of at most 25 ns, which occurs in a transmission distance of 500 m. Now, suppose the data rate is increased to 100 Mb/s, which means that one pulse is transmitted every 10 ns. In this case the 50-ns/km allowable spreading factor will limit the transmission distance to only 50 m in such a multimode step-index fiber.

The root-mean-square (rms) value of the time delay is a useful parameter for assessing the effect of modal delay in a multimode fiber. If it is assumed that the light rays are uniformly distributed over the acceptance angles of the fiber, then the rms impulse response σ_s due to intermodal dispersion in a step-index multimode fiber can be estimated from the expression

$$\sigma_s \approx \frac{L n_1 \Delta}{2\sqrt{3}c} \approx \frac{L(NA)^2}{4\sqrt{3} n_1 c} \tag{3.14a}$$

Here L is the fiber length and NA is the numerical aperture. Equation (3.14a) shows that the pulse broadening is directly proportional to the core-cladding index difference and the length of the fiber.

A successful technique for reducing modal delay in multimode fibers is through the use of a graded refractive index in the fiber core, as shown in Fig. 2.15. In any multimode fiber the ray paths associated with higher-order modes are concentrated near the edge of the core and thus follow a longer path through the fiber than lower-order modes (which are concentrated near the fiber axis). However, if the core has a graded index, then the higher-order modes encounter a lower refractive index near the core edge. Since the speed of light in a material depends on the refractive index value, the higher-order modes travel faster in the outer core region than those modes that propagate through a higher refractive index along the fiber center. Consequently this reduces the delay difference between the fastest and slowest modes. A detailed

analysis using electromagnetic mode theory gives the following absolute modal delay at the output of a graded-index fiber that has a parabolic ($\alpha = 2$) core index profile (see Sec. 2.6):

$$\sigma_s \approx \frac{Ln_1\Delta^2}{20\sqrt{3}c} \tag{3.14b}$$

Thus for an index difference of $\Delta = 0.01$, the theoretical improvement factor for intermodal rms pulse broadening in a graded-index fiber is 1000.

Example 3.9   Consider the following two multimode fibers: (*a*) a step-index fiber with a core index $n_1 = 1.458$ and a core-cladding index difference $\Delta = 0.01$; (*b*) a parabolic-profile graded-index fiber with the same values of n_1 and Δ. Compare the rms pulse broadening per kilometer for these two fibers.

Solution: (*a*) From Eq. (3.14a) we have

$$\frac{\sigma_s}{L} = \frac{n_1\Delta}{2\sqrt{3}c} = \frac{1.458(0.01)}{2\sqrt{3} \times 3 \times 10^8\,\text{m/s}} = 14.0 \text{ ns/km}$$

(*b*) From Eq. (3.14b) we have

$$\frac{\sigma_s}{L} \approx \frac{n_1\Delta^2}{20\sqrt{3}c} = \frac{1.458(.01)^2}{20\sqrt{3} \times 3 \times 10^8\,\text{m/s}} = 14.0 \text{ ps/km}$$

In graded-index fibers, careful selection of the radial refractive-index profile can lead to bit rate-distance products of up to 1 Gb/s-km.

3.2.3   Factors Contributing to Dispersion

This section briefly examines the various factors contributing to dispersion. Sections 3.2.4 through 3.2.8 and Sec. 3.3 describe these factors in more detail.

As Sec. 2.4 notes, the wave propagation constant β is a function of the wavelength, or, equivalently, of the angular frequency ω. Since β is a slowly varying function of this angular frequency, one can see where various dispersion effects arise by expanding β in a Taylor series about a central frequency ω_0. Inserting such an expansion into the waveform equation, for example Eq. (2.1), then shows the effects of variations in β due to modal dispersion and delay effects on the frequency components of a pulse during propagation along a fiber.

Expanding β to third order in a Taylor series yields

$$\beta(\omega) \approx \beta_0(\omega_0) + \beta_1(\omega_0)(\omega - \omega_0) + \frac{1}{2}\beta_2(\omega_0)(\omega - \omega_0)^2 + \frac{1}{6}\beta_3(\omega_0)(\omega - \omega_0)^3 \tag{3.15}$$

where $\beta_m(\omega_0)$ denotes the mth derivative of β with respect to ω evaluated at $\omega = \omega_0$; that is,

$$\beta_m = \left(\frac{\partial^m\beta}{\partial\omega^m}\right)_{\omega=\omega_0} \tag{3.16}$$

Now let us examine the different components of the product βz, where z is the distance traveled along the fiber. The resulting first term $\beta_0 z$ describes a phase shift of the propagating optical wave.

From the second term of Eq. (3.15), the factor $\beta_1(\omega_0)z$ produces a group delay $\tau_g = z/V_g$, where z is the distance traveled by the pulse and $V_g = 1/\beta_1$ is the group velocity [see Eqs. (3.20) and (3.21)]. Assume β_{1x} and β_{1y} are the propagation constants of the polarization components along the x-axis and y-axis, respectively, of a particular mode. If the corresponding group delays of these two polarization components are $\tau_{gx} = z\beta_{1x}$ and $\tau_{gy} = z\beta_{1y}$ in a distance z, then the difference in the propagation times of these two modes

$$\Delta\tau_{\text{PMD}} = z\,|\beta_{1x} - \beta_{1y}| \tag{3.17}$$

is called the *polarization-mode dispersion* (PMD) of the ideal uniform fiber. Note that in a real fiber the PMD varies statistically (see Sec. 3.2.8).

In the third term of Eq. (3.15), the factor β_2 shows that the group velocity of a monochromatic wave depends on the wave frequency. This means that the different group velocities of the frequency components of a pulse cause it to broaden as it travels along a fiber. This spreading of the group velocities is known as *chromatic dispersion* or *group velocity dispersion* (GVD). The factor β_2 is called the *GVD parameter* (see Sec. 3.2.4), and the *dispersion D* is related to β_2 through the expression

$$D = -\frac{2\pi c}{\lambda^2}\beta_2 \tag{3.18}$$

In the fourth term of Eq. (3.15), the factor β_3 is known as the *third-order dispersion*. This term is important around the wavelength at which β_2 equals zero. The third-order dispersion can be related to the dispersion D and the *dispersion slope* $S_0 = \partial D/\partial\lambda$ (the variation in the dispersion D with wavelength) by transforming the derivative with respect to ω into a derivative with respect to λ. Thus we have

$$
\begin{aligned}
\beta_3 &= \frac{\partial\beta_2}{\partial\omega} = -\frac{\lambda^2}{2\pi c}\frac{\partial\beta_2}{\partial\lambda} = -\frac{\lambda^2}{2\pi c}\frac{\partial}{\partial\lambda}\left[-\frac{\lambda^2}{2\pi c}D\right] \\
&= \frac{\lambda^2}{(2\pi c)^2}(\lambda^2 S_0 + 2\lambda D)
\end{aligned}
\tag{3.19}
$$

Section 3.5.3 illustrates how the factors in Eq. (3.19) are specified for commercial fibers.

3.2.4 Group Delay

As Example 3.8 mentions, the information-carrying capacity of a fiber link can be determined by examining the deformation of short light pulses propagating along the fiber. The following discussion on signal distortion thus is carried out primarily from the viewpoint of pulse broadening, which is representative of digital transmission.

First consider an electrical signal that modulates an optical source. For this case, assume that the modulated optical signal excites all modes equally at the input of the fiber. Each waveguide mode thus carries an equal amount of energy through the fiber. Furthermore, each mode contains all the spectral components in the wavelength band over which the source emits. In addition, assume that each of these spectral components is modulated in the same way. As the signal propagates along the fiber, each spectral component can be assumed to travel independently and to undergo a time delay or *group delay* per unit length τ_g/L in the direction of the propagation given by[34]

$$\frac{\tau_g}{L} = \frac{1}{V_g} = \frac{1}{c}\frac{d\beta}{dk} = -\frac{\lambda^2}{2\pi c}\frac{d\beta}{d\lambda} \tag{3.20}$$

Here, L is the distance traveled by the pulse, β is the propagation constant along the fiber axis, $k = 2\pi/\lambda$, and the *group velocity*

$$V_g = c\left(\frac{d\beta}{dk}\right)^{-1} = \left(\frac{\partial\beta}{\partial\omega}\right)^{-1} \tag{3.21}$$

is the velocity at which the energy in a pulse travels along a fiber.

Since the group delay depends on the wavelength, each spectral component of any particular mode takes a different amount of time to travel a certain distance. As a result of this difference in time delays, the optical signal pulse spreads out with time as it is transmitted over the fiber. Thus the quantity we are interested in is the amount of pulse spreading that arises from the group delay variation.

If the spectral width of the optical source is not too wide, the delay difference per unit wavelength along the propagation path is approximately $d\tau_g/d\lambda$. For spectral components that are $\delta\lambda$ apart and which lie $\delta\lambda/2$ above and below a central wavelength λ_0, the total delay difference $\delta\tau$ over a distance L is

$$\delta\tau = \frac{d\tau_g}{d\lambda}\delta\lambda = -\frac{L}{2\pi c}\left(2\lambda\frac{d\beta}{d\lambda} + \lambda^2\frac{d^2\beta}{d\lambda^2}\right)\delta\lambda \tag{3.22}$$

In terms of the angular frequency ω, this is written as

$$\delta\tau = \frac{d\tau_g}{d\omega}\delta\omega = \frac{d}{d\omega}\left(\frac{L}{V_g}\right)\delta\omega = L\left(\frac{d^2\beta}{d\omega^2}\right)\delta\omega \tag{3.23}$$

The factor $\beta_2 \equiv d^2\beta/d\omega^2$ is the GVD parameter, which determines how much a light pulse broadens as it travels along an optical fiber.

If the spectral width $\delta\lambda$ of an optical source is characterized by its rms value σ_λ (see Fig. 3.10 for a typical LED), then the pulse spreading can be approximated by the rms pulse width,

$$\sigma_g \approx \left|\frac{d\tau_g}{d\lambda}\right|\sigma_\lambda = \frac{L\sigma_\lambda}{2\pi c}\left|2\lambda\frac{d\beta}{d\lambda} + \lambda^2\frac{d^2\beta}{d\lambda^2}\right| \tag{3.24}$$

The factor

$$D = \frac{1}{L}\frac{d\tau_g}{d\lambda} = \frac{d}{d\lambda}\left(\frac{1}{V_g}\right) = -\frac{2\pi c}{\lambda^2}\beta_2 \tag{3.25}$$

is designated as the *dispersion.* It defines the pulse spread as a function of wavelength and is measured in picoseconds per kilometer per nanometer [ps/(nm·km)]. It is a result of material and waveguide dispersion. In many theoretical treatments of intramodal dispersion it is assumed, for simplicity, that material dispersion and waveguide dispersion can be calculated separately and then added to give the total

dispersion of the mode. In reality, these two mechanisms are intricately related, since the dispersive properties of the refractive index (which gives rise to material dispersion) also affects the waveguide dispersion. However, an examination[35] of the interdependence of material and waveguide dispersion has shown that, unless a very precise value to a fraction of a percent is desired, a good estimate of the total intramodal dispersion can be obtained by calculating the effect of signal distortion arising from one type of dispersion in the absence of the other. Thus, to a very good approximation, *D* can be written as the sum of the material dispersion D_{mat} and the waveguide dispersion D_{wg}. Material dispersion and waveguide dispersion are therefore considered separately in the next two sections.

3.2.5 Material Dispersion

Material dispersion occurs because the index of refraction varies as a function of the optical wavelength.[36] This is exemplified in Fig. 3.12 for silica. As a consequence, since the group velocity V_g of a mode is a function of the index of refraction, the various spectral components of a given mode will travel at different speeds,

Fig. 3.12 *Variations in the index of refraction as a function of the optical wavelength for silica. (Reproduced with permission from I. H. Malitson, J. Opt. Soc. Amer., vol. 55, pp. 1205–1209, Oct. 1965.)*

depending on the wavelength. Material dispersion is, therefore, an intramodal dispersion effect and is of particular importance for single-mode waveguides and for LED systems (since an LED has a broader output spectrum than a laser diode).

To calculate material-induced dispersion, we consider a plane wave propagating in an infinitely extended dielectric medium that has a refractive index $n(\lambda)$ equal to that of the fiber core. The propagation constant β is thus given as

$$\beta = \frac{2\pi n(\lambda)}{\lambda} \tag{3.26}$$

Substituting this expression for β into Eq. (3.20) with $k = 2\pi/\lambda$ yields the group delay τ_{mat} resulting from material dispersion.

$$\tau_{mat} = \frac{L}{c}\left(n - \lambda\frac{dn}{d\lambda}\right) \tag{3.27}$$

Using Eq. (3.24), the pulse spread σ_{mat} for a source of spectral width σ_λ is found by differentiating this group delay with respect to wavelength and multiplying by σ_λ to yield

Fig. 3.13 *Material dispersion as a function of optical wavelength for pure silica and 13.5-percent GeO$_2$/86.5-percent SiO$_2$. (Reproduced with permission from J. W. Fleming, Electron. Lett., vol. 14, pp. 326–328, May 1978.)*

$$\sigma_{mat} \approx \left| \frac{d\tau_{mat}}{d\lambda} \right| \sigma_\lambda = \frac{\sigma_\lambda L}{c} \left| \lambda \frac{d^2 n}{d\lambda^2} \right| = \sigma_\lambda L \mid D_{mat}(\lambda) \mid \tag{3.28}$$

where $D_{mat}(\lambda)$ is the *material dispersion.*

A plot of the material dispersion for unit length L and unit optical source spectral width σ_λ is given in Fig. 3.13 for the silica material shown in Fig. 3.12. From Eq. (3.28) and Fig. 3.13 it can be seen that material dispersion can be reduced either by choosing sources with narrower spectral output widths (reducing σ_λ) or by operating at longer wavelengths.

Example 3.10 A manufacturer's data sheet lists the material dispersion D_{mat} of a GeO$_2$-doped fiber to be 110 ps/(nm · km) at a wavelength of 860 nm. Find the rms pulse broadening per kilometer due to material dispersion if the optical source is a GaAlAs LED that has a spectral width σ_λ of 40 nm at an output wavelength of 860 nm.

Solution: From Eq. (3.28) we find that the rms material dispersion is

$$\sigma_{mat}/L = \sigma_\lambda D_{mat} = (40 \text{ nm}) \times [110 \text{ ps/(nm · km)}]$$
$$= 4.4 \text{ ns/km}$$

Example 3.11   The manufacturer's data shows that the same fiber as in Example 3.10 has a material dispersion D_{mat} of 15 ps/(nm · km) at a wavelength of 1550 nm. However, now suppose we use a laser source with a spectral width σ_λ of 0.2 nm at an operating wavelength of 1550 nm. What is the rms pulse broadening per kilometer due to material dispersion in this case?

Solution: From Eq. (3.28) we find that the rms material dispersion is

$$\sigma_{\text{mat}}/L = \sigma_\lambda D_{\text{mat}} = (0.2 \text{ nm}) \times [15 \text{ ps/(nm} \cdot \text{km)}]$$
$$= 7.5 \text{ ps/km}$$

This example shows that a dramatic reduction in dispersion can be achieved when operating at longer wavelengths with laser sources.

3.2.6 Waveguide Dispersion

The effect of waveguide dispersion on pulse spreading can be approximated by assuming that the refractive index of the material is independent of wavelength. Let us first consider the group delay—that is, the time required for a mode to travel along a fiber of length *L*. To make the results independent of fiber configuration,[37] we shall express the group delay in terms of the normalized propagation constant *b* defined as

$$b = 1 - \left(\frac{ua}{V}\right)^2 = \frac{\beta^2/k^2 - n_2^2}{n_1^2 - n_2^2} \tag{3.29}$$

For small values of the index difference $\Delta = (n_1 - n_2)/n_1$, Eq. (3.29) can be approximated by

$$b \approx \frac{\beta/k - n_2}{n_1 - n_2} \tag{3.30}$$

Solving Eq. (3.30) for β, we have

$$\beta \approx n_2 k (b\Delta + 1) \tag{3.31}$$

With this expression for β and using the assumption that n_2 is not a function of wavelength, we find that the group delay τ_{wg} arising from waveguide dispersion is

$$\tau_{\text{wg}} = \frac{L}{c} \frac{d\beta}{dk} = \frac{L}{c}\left[n_2 + n_2\Delta \frac{d(kb)}{dk}\right] \tag{3.32}$$

The modal propagation constant β is obtained from the eigenvalue equation expressed by Eq. (2.54) and is generally given in terms of the normalized frequency *V* defined by Eq. (2.57). We shall therefore use the approximation

$$V = ka\left(n_1^2 - n_2^2\right)^{1/2} \approx kan_1\sqrt{2\Delta} \tag{3.33}$$

which is valid for small values of Δ, to write the group delay in Eq. (3.32) in terms of *V* instead of *k*, yielding

$$\tau_{\text{wg}} = \frac{L}{c}\left[n_2 + n_2\Delta \frac{d(Vb)}{dV}\right] \tag{3.34}$$
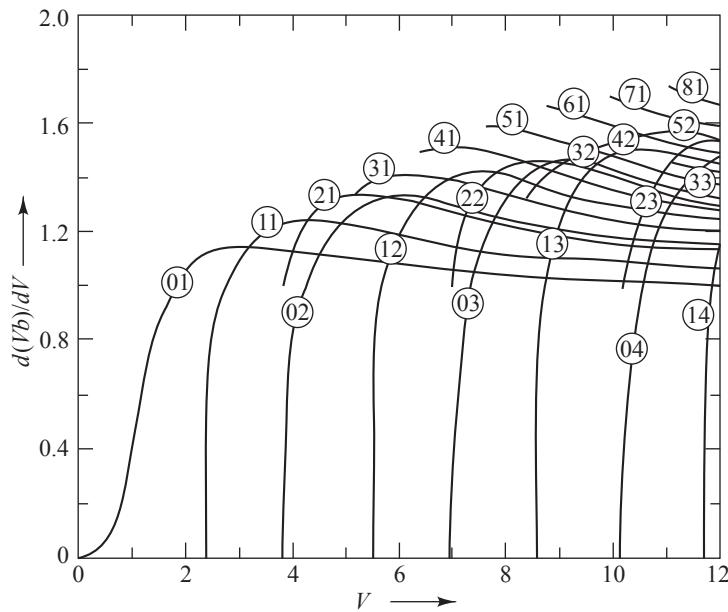
Fig. 3.14 *The group delay arising from waveguide dispersion as a function of the V number for a step-index optical fiber. The curve numbers jm designate the LP$_{jm}$ modes. (Reproduced with permission from Gloge.[37])*

The first term in Eq. (3.34) is a constant and the second term represents the group delay arising from waveguide dispersion. The factor $d(Vb)/dV$ can be expressed as[37]

$$\frac{d(Vb)}{dV} = b\left[1 - \frac{2J_\nu^2(ua)}{J_{\nu+1}(ua)\,J_{\nu-1}(ua)}\right] \tag{3.35}$$

where u is defined by Eq. (2.48) and a is the fiber radius. This factor is plotted in Fig. 3.14 as a function of V for various LP modes. The plots show that, for a fixed value of V, the group delay is different for every guided mode. When a light pulse is launched into a fiber, it is distributed among many guided modes. These various modes arrive at the fiber end at different times depending on their group delay, so that a pulse spreading results. For multimode fibers the waveguide dispersion is generally very small compared with material dispersion and can therefore be neglected.

3.2.7   Dispersion in Single-Mode Fibers

For single-mode fibers, waveguide dispersion is of importance and can be of the same order of magnitude as material dispersion. To see this, let us compare the two dispersion factors. The pulse spread σ_{wg} occurring over a distribution of wavelengths σ_λ is obtained from the derivative of the group delay with respect to wavelength:[37]

$$\sigma_{\mathrm{wg}} \approx \left| \frac{d\tau_{\mathrm{wg}}}{d\lambda} \right| \sigma_\lambda = L \mid D_{\mathrm{wg}}(\lambda) \mid \sigma_\lambda$$

$$= \frac{V}{\lambda} \left| \frac{d\tau_{\mathrm{wg}}}{dV} \right| \sigma_\lambda = \frac{n_2 L \Delta \sigma_\lambda}{c\lambda} V \frac{d^2(Vb)}{dV^2} \tag{3.36}$$

where $D_{\mathrm{wg}}(\lambda)$ is the *waveguide dispersion*.

To see the behavior of the waveguide dispersion, consider the expression of the factor *ua* for the lowest-order mode (i.e., the HE_{11} mode or, equivalently, the LP_{01} mode) in the normalized propagation constant. This can be approximated by[37]

$$ua = \frac{(1 + \sqrt{2})V}{1 + (4 + V^4)^{1/4}} \tag{3.37}$$

Substituting this into Eq. (3.29) yields, for the HE_{11} mode,

$$b(V) = 1 - \frac{(1 + \sqrt{2})^2}{[1 + (4 + V^4)^{1/4}]^2} \tag{3.38}$$

Figure 3.15 shows plots of this expression for *b* and its derivatives $d(Vb)/dV$ and $Vd^2(Vb)/dV^2$ as functions of *V*.

Example 3.12   From Eq. (3.36) we have that the waveguide dispersion is

$$D_{\mathrm{wg}}(\lambda) = -\frac{n_2 \Delta}{c} \frac{1}{\lambda} \left[V \frac{d^2(Vb)}{dV^2} \right]$$

Let $n_2 = 1.48$ and $\Delta = 0.2$ percent. At $V = 2.4$, from Fig. 3.15 the expression in square brackets is 0.26. Choosing $\lambda = 1320$ nm, we then have $D_{\mathrm{wg}}(\lambda) = -1.9$ ps/(nm · km).
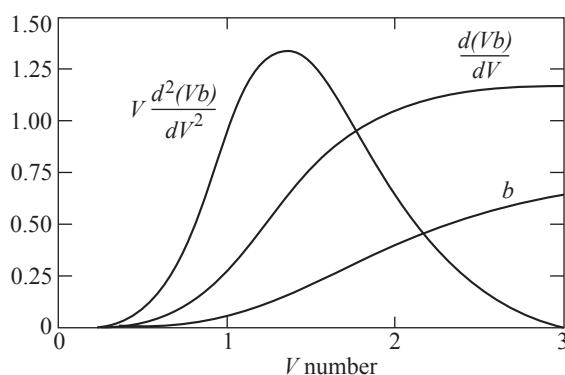
Fig. 3.15   *The waveguide parameter b and its derivatives $d(Vb)/dV$ and $Vd^2(Vb)/dV^2$ plotted as a function of the V number for the HE_{11} mode*
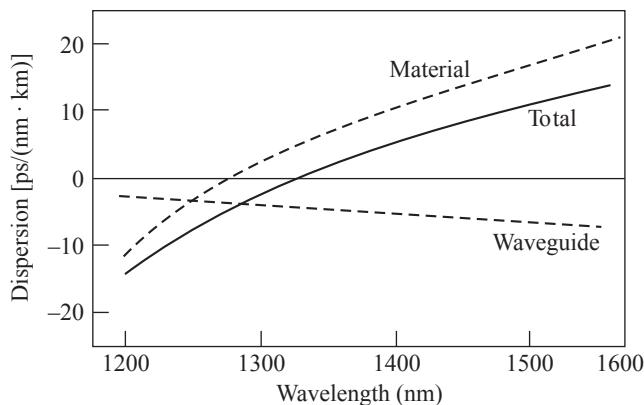
Fig. 3.16   *Examples of the magnitudes of material and waveguide dispersion as a function of optical wavelength for a single-mode fused-silica-core fiber. (Reproduced with permission from Keck,[16] © 1985, IEEE.)*

Figure 3.16 gives examples of the magnitudes of material and waveguide dispersion for a fused-silica-core single-mode fiber having $V = 2.4$. Comparing the waveguide dispersion with the material dispersion, we see that for a standard non-dispersion-shifted fiber, waveguide dispersion is important around 1320 nm. At this point, the two dispersion factors cancel to give a zero total dispersion. However, material dispersion dominates waveguide dispersion at shorter and longer wavelengths; for example, at 900 nm and 1550 nm. This figure used the approximation that material and waveguide dispersions are additive.

3.2.8   Polarization-Mode Dispersion

The effects of fiber birefringence on the polarization states of an optical signal are another source of pulse broadening. This is particularly critical for high-rate long-haul transmission links (e.g., 10 and 40 Gb/s over tens of kilometers). Birefringence can result from intrinsic factors such as geometric irregularities of the fiber core or internal stresses on it. Deviations of less than 1 percent in the circularity of the core can already have a noticeable effect in a high-speed lightwave system. In addition, external factors, such as bending, twisting, or pinching of the fiber, can also lead to birefringence. Since all these mechanisms exist to some extent in any field-installed fiber, there will be a varying birefringence along its length.

A fundamental property of an optical signal is its polarization state. *Polarization* refers to the electric-field orientation of a light signal, which can vary significantly along the length of a fiber. As shown in Fig. 3.17, signal energy at a given wavelength occupies two orthogonal polarization modes. A varying birefringence along its length will cause each polarization mode to travel at a slightly different velocity. The resulting difference in propagation times $\Delta\tau_{\text{PMD}}$ between the two orthogonal polarization modes will result in pulse spreading. This is the *polarization-mode dispersion* (PMD).[38,39] If the group velocities of the two orthogonal polarization modes are V_{gx} and V_{gy}, then the differential time delay $\Delta\tau_{\text{PMD}}$ between the two polarization components during propagation of the pulse over a distance L is

$$\Delta\tau_{\text{PMD}} = \left| \frac{L}{V_{gx}} - \frac{L}{V_{gy}} \right| \tag{3.39}$$
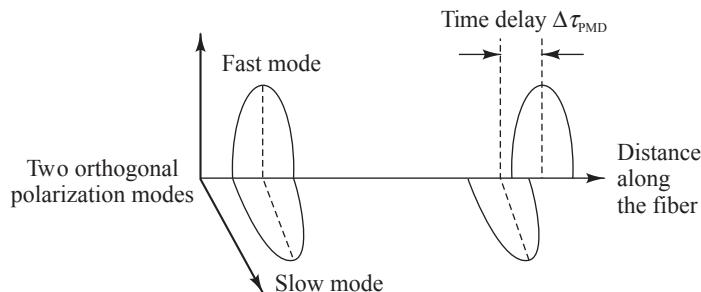
Fig. 3.17 *Differences in the polarization-mode propagation times as an optical pulse passes through a fiber with varying birefringence along its length*

An important point to note is that, in contrast to chromatic dispersion, which is a relatively stable phenomenon along a fiber, PMD varies randomly along a fiber. A principal reason for this is that the perturbations causing the birefringence effects vary with temperature and stress dynamics. In practice, the effect of these perturbations shows up as a random, time-varying fluctuation in the value of the PMD at the fiber output. Thus $\Delta\tau_{PMD}$ given in Eq. (3.39) cannot be used directly to estimate PMD. Instead, statistical estimations are needed to account for its effects.

A useful means of characterizing PMD for long fiber lengths is in terms of the mean value of the differential group delay (see Chapter 14 for PMD measurement techniques). This can be calculated according to the relationship

$$\Delta\tau_{PMD} \approx D_{PMD}\sqrt{L} \tag{3.40}$$

where D_{PMD}, which is measured in ps/$\sqrt{\text{km}}$, is the average PMD parameter. Typical values of D_{PMD} range from 0.05 to 1.0 ps/$\sqrt{\text{km}}$. As an example, one experiment measured values of PMD for three types of cable installations that were subjected to different environments.[40] The setups were a 36-km spooled fiber in a temperature-controlled chamber, a 48.8-km buried cable, and a 48-km aerial cable. Over a 12- to 15-h period, the average PMD parameters were measured to be 0.028, 0.29, and 1.28 ps/$\sqrt{\text{km}}$, respectively. The larger value of PMD for the aerial cable is caused by both gradual and rapid stress variations in the fiber due to temperature fluctuations or from sudden movements of the fiber due to wind.[41]

To keep the probability of errors due to PMD low, a standard limit on the maximum tolerable value of $\Delta\tau_{PMD}$ ranges between 10 and 20 percent of a bit duration. Thus $\Delta\tau_{PMD}$ should be no more than 10 to 20 ps for 10-Gb/s data rates and 3 ps at 40 Gb/s. For example, taking the lower tolerance limit, this means that for a 10-Gb/s link that has 20 spans of 80 km each, the PMD of the transmission fiber must be less than 0.2 ps/$\sqrt{\text{km}}$. Various optical and electronic means to monitor and mitigate PMD in a fiber have been investigated.[42–48] In addition, fibers with low polarization-mode dispersion are being developed and characterized.[49–50]

3.3 | **Characteristics of Single-Mode Fibers**

This section addresses the basic design and operational characteristics of single-mode fibers. These characteristics include index-profile configurations used to produce different fiber types, the concept of cutoff wavelength, signal dispersion designations and calculations, the definition of mode-field diameter, and signal loss due to fiber bending.

3.3.1 Refractive-Index Profiles

When creating single-mode fibers, manufacturers pay special attention to how the fiber design affects both chromatic and polarization-mode dispersions. Such considerations are important because these dispersions set the limits on long-distance and high-speed data transmission. As Fig. 3.16 illustrates, the chromatic dispersion of a step-index silica fiber is lowest at 1310 nm. However, if the goal is to transmit a signal as far as possible, it is better to operate the link at 1550 nm (in the C-band) where the fiber attenuation is lower. For high-speed links the C-band presents a problem because chromatic dispersion is much larger at 1550 nm than at 1310 nm. Consequently, fiber designers devised methods for adjusting the fiber parameters to shift the zero-dispersion point to longer wavelengths.

The basic material dispersion is hard to alter significantly. However, it is possible to modify the waveguide dispersion by changing from a simple step-index design to more complex index profiles for the cladding, thereby creating different chromatic-dispersion characteristics in single-mode fibers.[13,51–57] Figure 3.18 shows representative refractive-index profiles of four fiber-design categories. These are 1310-nm-optimized fibers, dispersion-shifted fibers, dispersion-flattened fibers, and large-effective-core-area fibers. To get a better feeling of what this geometry looks like, Fig. 3.19 gives examples of the three-dimensional index profiles for several different types of single-mode fibers.

Popular single-mode fibers that are used widely in telecommunication networks are near-step-index fibers, which are optimized for use in the O-band around 1310 nm. These *1310-nm-optimized single-mode fibers* are of either the *matched-cladding*[13,51,52] or the *depressed-cladding*[53,54] design, as shown in Figs 3.18*a*, 3.19*a*, and 3.19*b*. Matched-cladding fibers have a uniform refractive index throughout the cladding. Typical mode-field diameters are 9.5 μm and the core-to-cladding index differences are around 0.35 percent. In depressed-cladding fibers the cladding material next to the core has a lower index than the outer cladding region. Mode-field diameters are around 9.0 μm, and typical positive and negative index differences are 0.25 and 0.12 percent, respectively.

As Eqs. (3.28) and (3.36) show, whereas material dispersion depends only on the composition of the material, waveguide dispersion is a function of the core radius, the refractive index difference, and the shape of the refractive index profile. Thus the waveguide dispersion can vary dramatically with the fiber design parameters. By creating a fiber with a larger negative waveguide dispersion and assuming the same values for material dispersion as in a standard single-mode fiber, the addition of waveguide and material dispersion can then shift the zero dispersion point to longer wavelengths. The resulting optical fiber is known as a *dispersion-shifted fiber* (DSF).[52,55–57] Examples of refractive-index profiles for dispersion-shifted fibers are shown in Figs 3.18*b* and 3.19*c*. A typical waveguide dispersion curve for this type of fiber is depicted in Fig. 3.20*a*. The resultant total dispersion curve is shown in Fig. 3.20*b* for fibers with a zero-dispersion wavelength at 1550 nm.

Since the zero-dispersion value of a DSF falls at 1550 nm, the chromatic dispersion is negative for wavelengths less than 1550 nm and positive for longer wavelengths. These positive and negative dispersions will seriously affect closely spaced WDM signals within the C-band because of nonlinear effects in the fiber, as Chapter 12 describes. To reduce the effects of fiber nonlinearities, fiber designers
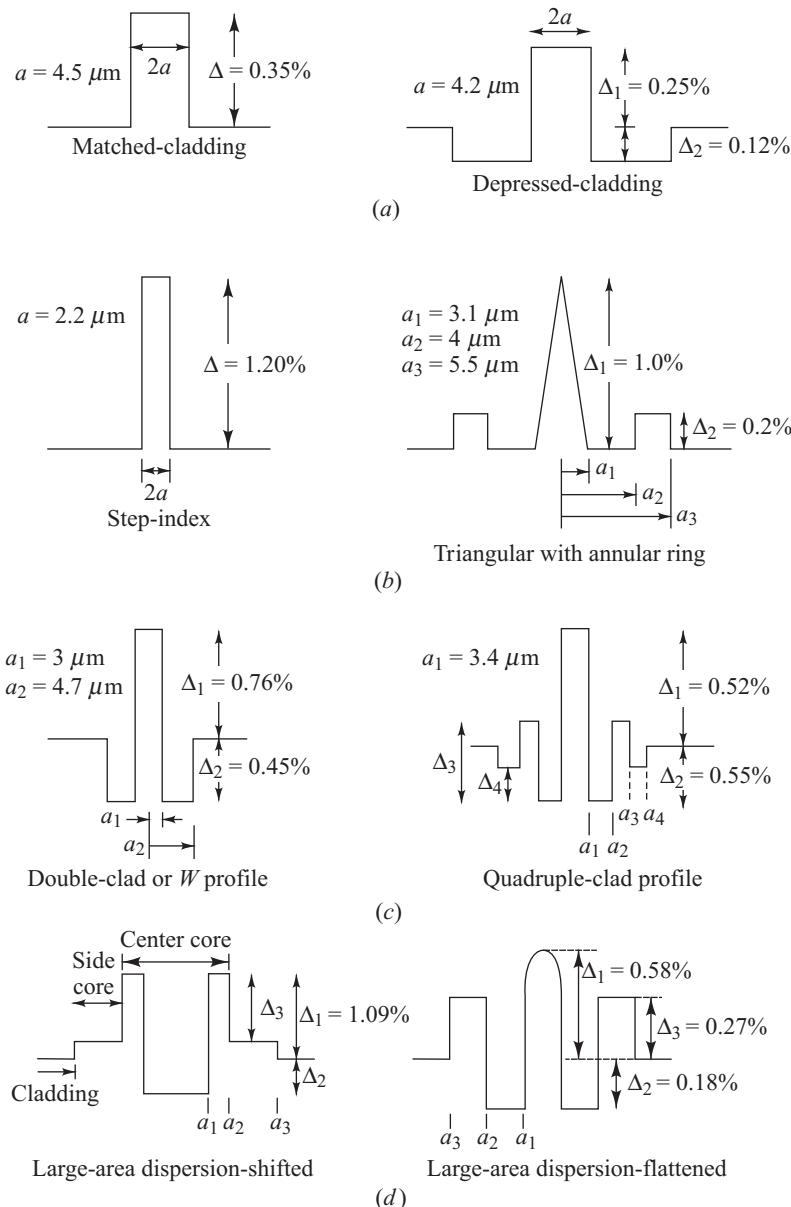
Fig. 3.18 *Representative cross sections of index profiles for (a) 1310-nm-optimized, (b) dispersion-shifted, (c) dispersion-flattened, and (d) large-effective-core-area fibers*
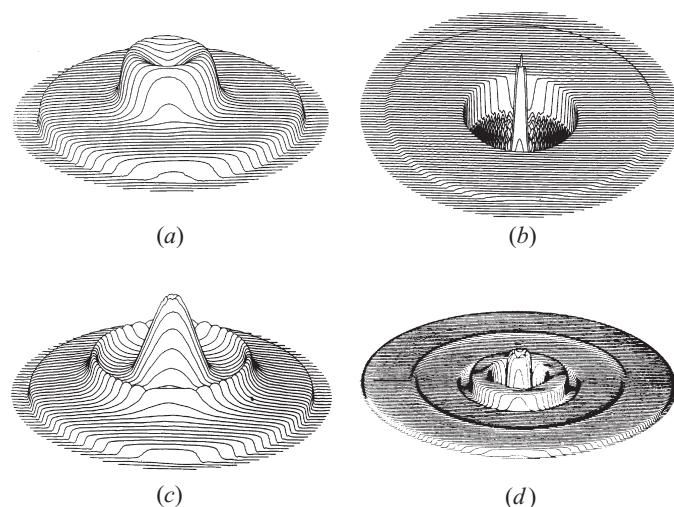
Fig. 3.19 *Three-dimensional refractive index profiles for (a) matched-cladding 1310-nm-optimized, (b) depressed-cladding 1310-nm-optimized, (c) triangular dispersion-shifted, and (d) quadruple-clad dispersion-flattened single-mode fibers. [(a) and (c) Courtesy of Corning, Inc.; (b) Courtesy of York Technology; (d) Reproduced with permission from H. Lydtin, J. Lightwave Tech., vol. LT-4, pp. 1034–1038, Aug. 1986, © 1986, IEEE.]*

developed the *nonzero dispersion-shifted fiber* (NZDSF). These fibers have a small amount of either all positive or all negative dispersion throughout the C-band. A typical positive chromatic dispersion value for a NZDSF is 4.5 ps/(nm-km) at 1550 nm. Section 3.4.2 gives an example of a NZDSF that has positive dispersion throughout the C-band.

Among the NZDSF types is a single-mode optical fiber with a larger effective core area.[58–61] The larger core areas reduce the effects of fiber nonlinearities, which limit system capacities of transmission systems that have densely spaced WDM channels. Figure 3.18*d* gives two examples of the index profile for these large-effective-area (LEA) fibers. Whereas standard single-mode fibers have effective core areas of about 55 μm^2, these profiles yield values greater than 100 μm^2.

An alternative fiber design concept is to distribute the dispersion minimum over a wider spectral range. This approach is known as *dispersion flattening*.[62,63] Dispersion-flattened fibers are more complex to design than dispersion-shifted fibers, because dispersion must be considered over a much broader range of wavelengths. However, they offer desirable characteristics over a wide span of wavelengths. Figures 3.18*c* and 3.19*d* show typical cross-sectional and three-dimensional refractive-index profiles, respectively. A typical waveguide dispersion curve for this type of fiber is depicted in Fig. 3.20*a*. Figure 3.20*b* gives the resultant total flattened dispersion characteristic.

3.3.2   Cutoff Wavelength

The cutoff wavelength of the first higher-order mode (LP_{11}) is an important transmission parameter for single-mode fibers because it separates the single-mode from the multimode regions.[64–65] Recall from Eq. (2.58) that single-mode operation occurs above the theoretical cutoff wavelength given by

(a)

(b)
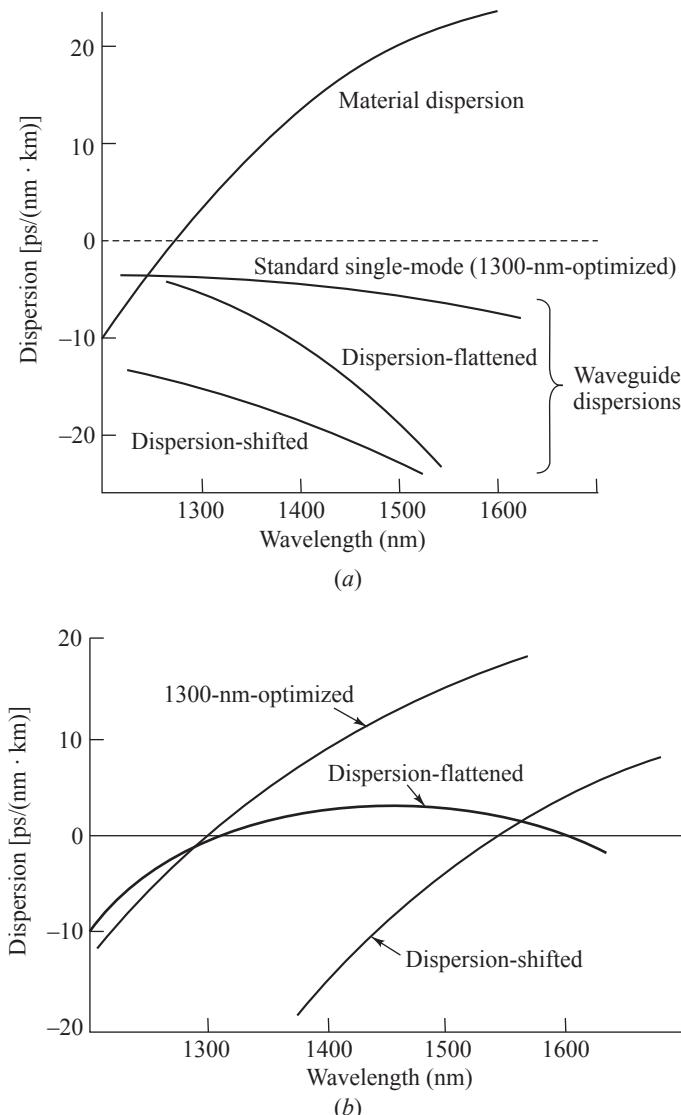
Fig. 3.20 *(a) Typical waveguide dispersions and the common material dispersion for three different single-mode fiber designs; (b) resultant total dispersions*

$$\lambda_c = \frac{2\pi a}{V}\left(n_1^2 - n_2^2\right)^{1/2} \simeq \frac{2\pi a}{V} n_1 \sqrt{2\Delta} \tag{3.41}$$

with $V = 2.405$ for step-index fibers. At this wavelength, only the LP_{01} mode (i.e., the HE_{11} mode) should propagate in the fiber.

Since in the cutoff region the field of the LP_{11} mode is widely spread across the fiber cross section (i.e., it is not tightly bound to the core), its attenuation is strongly affected by fiber bends, length, and

Example 3.13 A given step-index fiber has a core refractive index of 1.480, a core radius equal to 4.5 μm, and a core-cladding index difference of 0.25 percent. What is the cutoff wavelength for this fiber?

Solution: From Eq. (3.41) we have for $V = 2.405$

$$\lambda_c = \frac{2\pi a}{V} n_1 \sqrt{2\Delta}$$

$$= \frac{2\pi(4.5)}{2.405}(1.480)\sqrt{2(0.0025)}$$

$$= 1.23\ \mu\text{m}$$

$$= 1230\ \text{nm}$$

cabling. Recommendation G.650.1 of the ITU-T[65] and the EIA-455–80C Standard[66] specify methods for determining an effective cutoff wavelength λ_c. The setup consists of a 2-m length of fiber that contains a single 14-cm-radius loop or several 14-cm-radius curvatures that add up to one complete loop. Using a tunable light source that has a full-width half-maximum linewidth not exceeding 10 nm, light is launched into the fiber so that both the LP_{01} and the LP_{11} modes are uniformly excited.

First, the output power $P_1(\lambda)$ is measured as a function of wavelength in a sufficiently wide range around the expected cutoff wavelength. Next, the output power $P_2(\lambda)$ is measured over the same wavelength range when a loop of sufficiently small radius is included in the test fiber to filter the LP_{11} mode. A typical radius for this loop is 30 mm. With this method, the logarithmic ratio $R(\lambda)$ between the two transmitted powers $P_1(\lambda)$ and $P_2(\lambda)$ is calculated as

$$R(\lambda) = 10 \log \left[\frac{P_1(\lambda)}{P_2(\lambda)} \right] \tag{3.42}$$
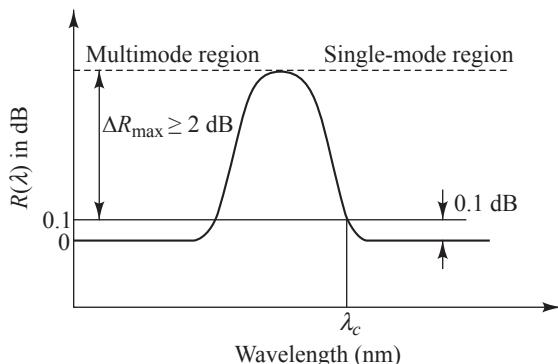
Fig. 3.21 *Typical attenuation-ratio versus wavelength plot for determining the cutoff wavelength using the bend-reference (or single-mode-reference) transmission method. The peak ratio should be at least 2 dB above the cutoff level.*

Figure 3.21 gives a typical curve of the result. The effective cutoff wavelength λ_c is defined as the largest wavelength at which the higher-order LP_{11} mode power relative to the fundamental LP_{01} mode power is reduced to 0.1 dB; that is, when, $R(\lambda) = 0.1$ dB, as is shown in Fig. 3.21. Recommended values of λ_c range from 1100 to 1280 nm, to avoid modal noise and dispersion problems.

3.3.3 Dispersion Calculations

As noted in Sec. 3.3.1, the total chromatic dispersion in single-mode fibers consists mainly of material and waveguide dispersions. The resultant intramodal or chromatic dispersion is represented by[65,67–70]

$$D(\lambda) = \frac{1}{L} \frac{d\tau}{d\lambda} \tag{3.43}$$

where τ is the group delay. The dispersion is commonly expressed in ps/(nm . km). The broadening σ of an optical pulse over a fiber of length L is given by

$$\sigma = D(\lambda)\, L \sigma_\lambda \tag{3.44}$$

where σ_λ is the half-power spectral width of the optical source. To measure the dispersion, one examines the pulse delay over a desired wavelength range.

As illustrated in Fig. 3.20, the dispersion behavior varies with wavelength and also with fiber type. Thus, the EIA and the ITU-T have recommended different formulas to calculate the chromatic dispersion for specific fiber types operating in a given wavelength region. To calculate the dispersion for a non-dispersion-shifted fiber (called a Class IVa fiber by the EIA) in the 1270-to-1340-nm region, the standards recommend fitting the measured group delay per unit length to a three-term Sellmeier equation of the form[65]

$$\tau = A + B\lambda^2 + C\lambda^{-2} \tag{3.45}$$

to the measured pulse date. Here, A, B, and C are the curve-fitting parameters. An equivalent expression is

$$\tau = \tau_0 + \frac{S_0}{8}\left(\lambda - \frac{\lambda_0^2}{\lambda}\right)^2 \tag{3.46}$$

where τ_0 is the relative delay minimum at the zero-dispersion wavelength λ_0, and S_0 is the value of the *dispersion slope* $S(\lambda) = dD/d\lambda$ at λ_0, which is given in ps/(nm^2·km). Using Eq. (3.43), the dispersion for a non-dispersion-shifted fiber is

$$D(\lambda) = \frac{\lambda S_0}{4}\left[1 - \left(\frac{\lambda_0}{\lambda}\right)^4\right] \tag{3.47}$$

Example 3.14 A manufacturer's data sheet states that a non-dispersion-shifted fiber has a zero-dispersion wavelength of 1310 nm and a dispersion slope of 0.092 ps/(nm^2 · km). Compare the dispersions for this fiber at wavelengths of 1280 nm and 1550 nm.

Solution: Using Eq. (3.47) we find that

$$D(1280) = \frac{\lambda S_0}{4}\left[1 - \left(\frac{\lambda_0}{\lambda}\right)^4\right]$$

$$= \frac{(1280)(0.092)}{4}\left[1 - \left(\frac{1310}{1280}\right)^4\right]$$

$$= -2.86 \text{ ps/(nm · km)}$$

$$D(1550) = \frac{\lambda S_0}{4}\left[1 - \left(\frac{\lambda_0}{\lambda}\right)^4\right]$$

$$= \frac{(1550)(0.092)}{4}\left[1 - \left(\frac{1310}{1550}\right)^4\right]$$

$$= 17.5 \text{ ps/(nm · km)}$$

To calculate the dispersion for a dispersion-shifted fiber (called a Class IVb fiber by the EIA) in the 1500-to-1600-nm region, the standards recommend using the quadratic expression[65]

$$\tau = \tau_0 + \frac{S_0}{2}(\lambda - \lambda_0)^2 \tag{3.48}$$

which results in the dispersion expression

$$D(\lambda) = (\lambda - \lambda_0)\, S_0 \tag{3.49}$$

Finally, recall from Eq. (3.19) that the *third-order dispersion β_3* can be given as

$$\beta_3 = \frac{\lambda^2}{(2\pi c)^2}[\lambda^2 S_0 + 2\lambda D] \tag{3.50}$$

When measuring a set of fibers, one will get values of λ_0 ranging from $\lambda_{0,\,min}$ to $\lambda_{0,max}$. Figure 3.22 shows the range of expected dispersion values for a set of non-dispersion-shifted fibers in the 1270-to-1340-nm
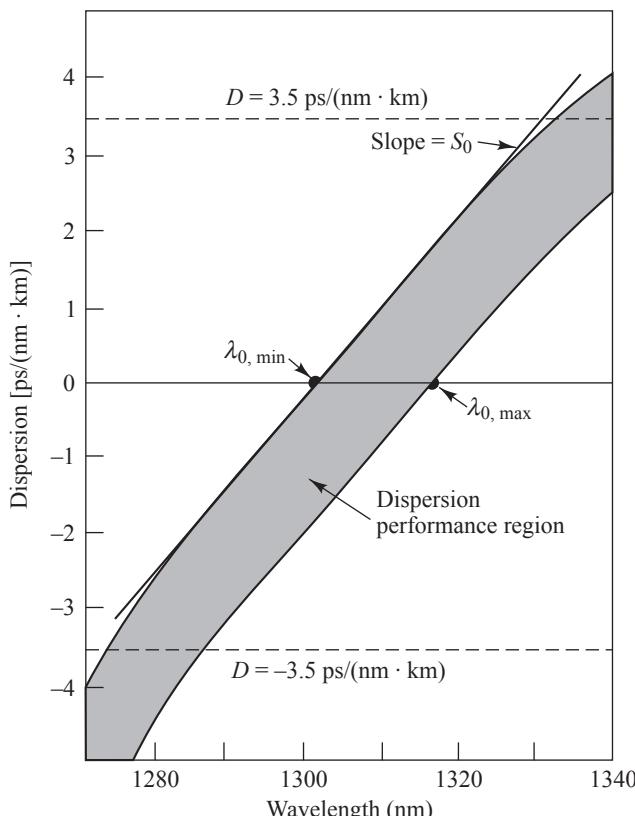
Fig. 3.22 *Example of a dispersion performance curve for a set of single-mode fibers. The two slightly curved lines are found by solving Eq. (3.47). S_0 is the slope of $D(\lambda)$ at the zero-dispersion wavelength λ_0.*
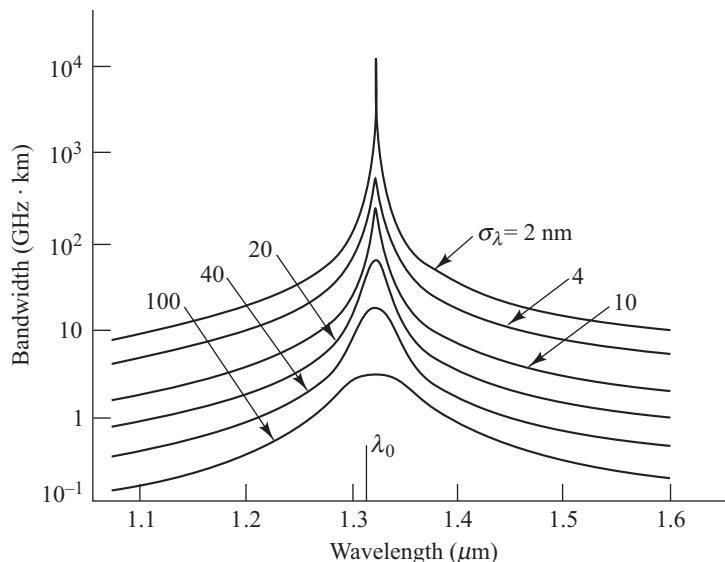
Fig. 3.23 *Examples of bandwidth versus wavelength for different source spectral widths σ_λ in a single-mode fiber having a dispersion minimum at 1300-nm. (Reproduced with permission from Reed, Cohen, and Shang,[56] © 1987, AT&T.)*

region. Typical values of S_0 are 0.092 ps/(nm² · km) for standard non-dispersion-shifted fibers, and are between 0.06 and 0.08 ps/(nm² · km) for dispersion-shifted fibers. Alternatively, the ITU-T Rec. G.652 has specified this as a maximum dispersion of 3.5 ps/(nm · km) in the 1285-to-1330-nm region, as denoted by the dashed lines in Fig. 3.22.

Figure 3.23 illustrates the importance of controlling dispersion in single-mode fibers. As optical pulses travel down a fiber, temporal broadening occurs because material and waveguide dispersion cause different wavelengths in the optical pulse to propagate with different velocities. Thus, as Eq. (3.44) implies, the broader the spectral width σ_λ of the source, the greater the pulse dispersion will be. This effect is clearly seen in Fig. 3.23.

3.3.4 Mode-Field Diameter

Section 2.5.1 gives the definition of the mode-field diameter in single-mode fibers. One uses the mode-field diameter in describing the functional properties of a single-mode fiber, since it takes into account the wavelength-dependent field penetration into the cladding. This is shown in Fig. 3.24 for 1300-nm-optimized, dispersion-shifted, and dispersion-flattened single-mode fibers.

3.3.5 Bending Loss

Macrobending and microbending losses are important in the design of single-mode fibers.[19–25] These losses are principally evident in the 1550-nm region and show up as a rapid increase in attenuation when the fiber is bent smaller than a certain bend radius. The lower the cutoff wavelength relative to the operating wavelength, the more susceptible single-mode fibers are to bending. For example, in a fiber optimized for operation at 1300 nm, both the microbending and macrobending losses are greater at 1550 nm than at
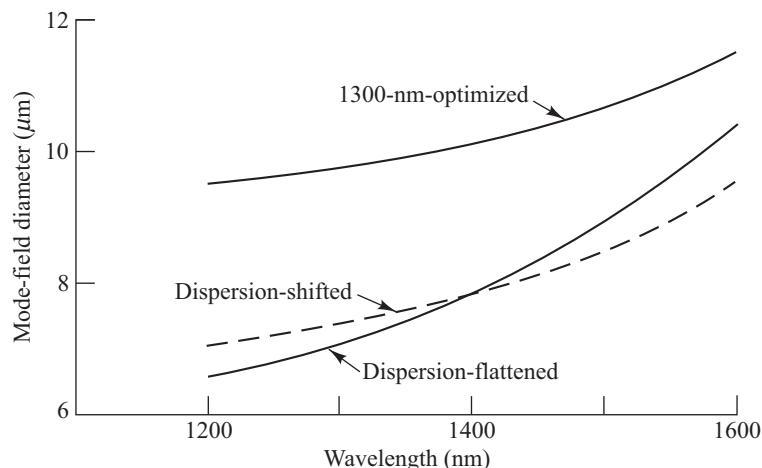
Fig. 3.24 *Typical mode-field diameter variations with wavelength for (a) 1300 nm-optimized, (b) dispersion-shifted, and (c) dispersion-flattened single-mode fibers*

1300 nm by a factor of 3 to 5, as Fig. 3.25 illustrates. A fiber thus might be transmitting well at 1300 nm but have a significant loss at 1550 nm.

The bending losses are primarily a function of the mode-field diameter. Generally, the smaller the mode-field diameter (i.e., the tighter the confinement of the mode to the core), the smaller the bending loss. This is true for both matched-clad and depressed-clad fibers, as Fig. 3.26 shows.

In examining the bending loss, early theories assume a simple model of a fiber with an infinitely extending cladding. This results in the prediction of a smooth exponential increase of bending loss with
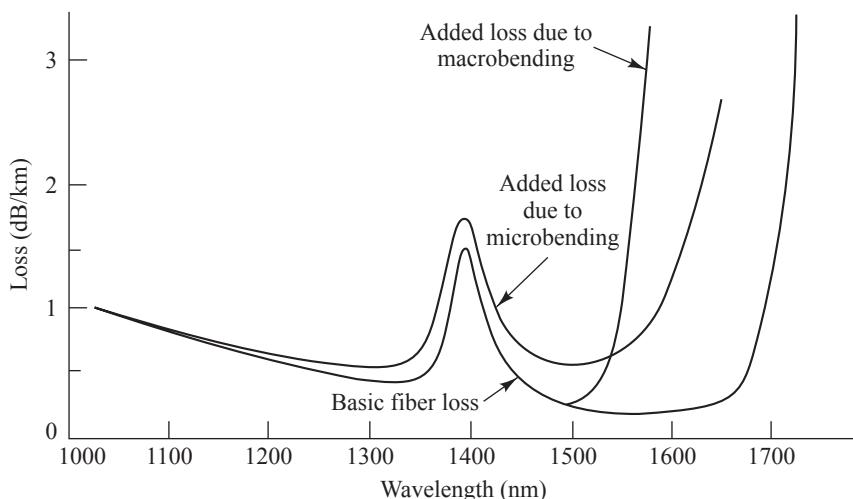
Fig. 3.25 *Representative increases in single-mode fiber attenuation owing to microbending and macrobending effects. (Reproduced with permission from Kalish and Cohen,[57] © 1987, AT&T.)*
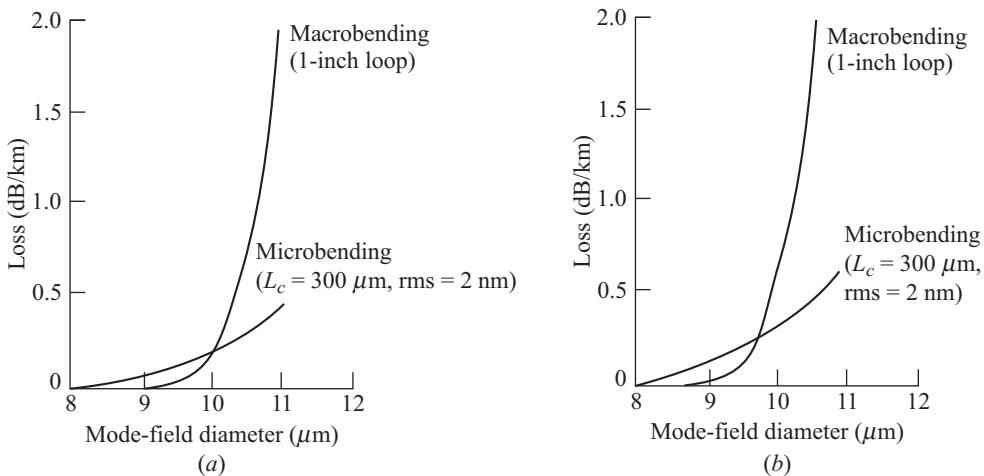
Fig. 3.26 *Calculated increase in attenuation at 1310 nm from microbending and macrobending effects as a function of mode-field diameter for (a) depressed-cladding single-mode fiber (V = 2.514) and (b) matched-cladding single-mode fiber (V = 2.373). The microbending calculations assume a correlation length L_c (microbending repetition rate) of 300 nm and a 2-nm deformation amplitude. (Reproduced with permission from Kalish and Cohen,[57] © 1987, AT&T.)*

increasing wavelength or radius of curvature. In an actual fiber, oscillations in the bend loss versus both the wavelength and bending radius are observed. These oscillations can be attributed to coherent coupling between the field propagating in the core and the fraction of the radiated field that is reflected at the boundary between the cladding and the fiber-coating material. Figure 3.27 gives an example of calculated bend loss as a function of bend radius at a 1300-nm wavelength. The fiber parameters were core radius $a = 3.6$ μm, cladding radius $b = 60$ μm, $(n_1 - n_2)/n_2 = 3.56 \times 10^{-3}$ and $(n_3 - n_2)/n_2 = 0.07$, where n_1, n_2, and n_3 are the core, cladding, and coating indices of refraction, respectively.

By specifying bend-radius limitations when installing standard single-mode fibers, one can largely avoid high microbending losses. Manufacturers usually recommend a fiber or cable bend diameter no smaller than 40–50 mm (1.6–2.0 in.). This is consistent with bend diameter limitations of 50–75 mm specified by installation guides for cable placement in ducts, fiber-splice enclosures, and equipment racks. Furthermore, as Sec. 3.5 describes, the development of
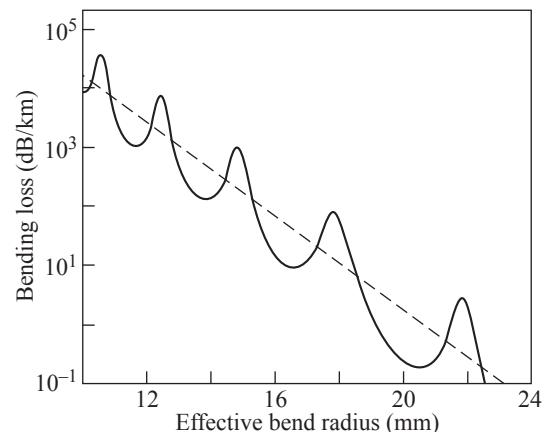
Fig. 3.27 *Calculated bend loss as a function of bend radius at 1300 nm. The dashed line represents the infinite-cladding case, i.e., $n_2 = n_3$. (Modified with permission from Renner,[19] © 1992, IEEE.)*

bend-insensitive fibers allows much tighter coiling of these fibers in optoelectronic packages. In addition, use of these bend-insensitive fibers in jumper cables greatly reduces bending loss effects when they are installed in highly confined equipment racks.

3.4 International Standards

As noted in Sec. 1.7, the ITU-T, the TIA, and the EIA are leading organizations that develop and publish a wide range of internationally recognized recommendations and standards. In particular, the ITU-T has created a series of recommendations for manufacturing and testing various classes of multimode and single-mode optical fibers used in telecommunications. These documents give guidelines for bounds on fiber parameters, such as core and cladding sizes and circularity, attenuation, cutoff wavelength, and chromatic dispersion. The recommendations allow a reasonable degree of design flexibility, so that fiber manufacturers can improve products and develop new ones within the guidelines given in the performance specifications.

Table 3.2 summarizes the ITU-T recommendations for multimode and single-mode optical fibers used in telecommunication, access, and enterprise networks. The following subsections describe the basic characteristics of these fibers.[71]

3.4.1 Recommendation G.651.1

The economic demand for low-cost installations of high-speed short-distance optical fiber links created an extensive market for multimode fibers. These fiber links use moderately priced light sources that operate in either the short-wavelength region (770 to 860 nm) or in the O-band (around 1310 nm). Applications include links in locations such as an office or government building, a medical facility, a university campus, or a manufacturing plant, where the desired transmission distance is typically 2 km or less.

Recommendation G.651.1 replaces the original G.651 document. This new version gives the requirements of a 50/125 μm multimode graded-index silica-optical-fiber cable for use in the 850-nm or 1300-nm regions. System operation for this fiber is allowed either in each wavelength band individually or simultaneously in both spectral bands. The applications are intended for access and enterprise networks in multiple-tenant building environments in which broadband services have to be delivered to individual apartments or offices of individual businesses. The recommended multimode fiber supports the cost-effective use of 1-Gb/s Ethernet systems over link lengths up to 550 m. These Ethernet links usually employ 850-nm transceivers. The optical fiber attenuation values range from 2.5 dB/km at 850 nm to less than 0.6 dB/km at 1310 nm. Amendment 1 to Rec. G.651.1 gives background information on the evolution of the specification of multimode optical fiber cable in the ITU-T recommendations.

Table 3.3 lists the possible transmission distances when using fibers with different core sizes and bandwidths for Ethernet, Fibre Channel, and SONET/SDH applications. The light source used for these examples is a vertical-cavity surface-emitting laser (VCSEL) operating at 850 nm (see Chapter 4). In particular, Ethernet links running at data rates up to 10 Gb/s over distances up to 550 m can use multimode fibers.

3.4.2 Recommendation G.652

Recommendation G.652 deals with the geometrical, mechanical, and transmission characteristics of a single-mode fiber that has a zero-dispersion value at 1310 nm. Figure 3.28 compares the dispersion of the G.652 fiber with other single-mode fiber types. This fiber consists of a germanium-doped silica core that has a diameter between 5 and 8 μm, and a 125-μm silica cladding diameter. The nominal attenuation

Table 3.2 *Recommendations for fibers used in telecom, access, and enterprise networks*
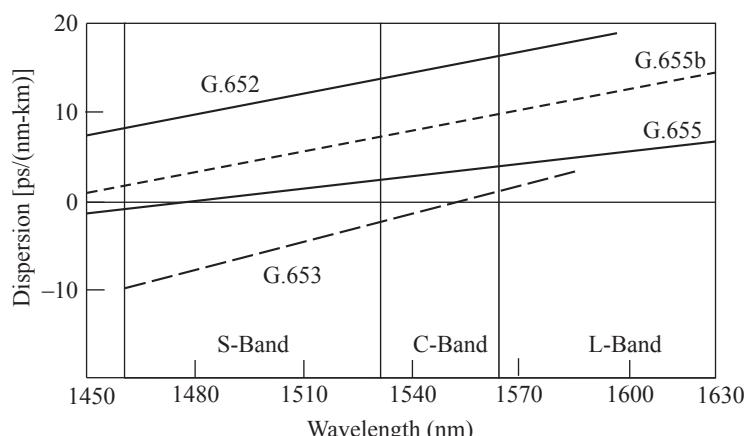
ITU-T rec. no.	Title and description
G.651.1 (Edition 1, July 2007); Addendum (Dec. 2008)	Title: *Characteristics of a 50/125 μm multimode graded index optical fiber cable for the optical access network*
	Description: Gives the requirements of a silica 50/125 μm multimode graded index optical fiber cable for use in the 850-nm or 1300-nm regions, either individually or simultaneously
G.652 (Edition 8, Nov. 2009)	Title: *Characteristics of a Single-Mode Optical Fiber and Cable*
	Description: Discusses single-mode fiber optimized for O-band (1310-nm) use, but which also can be used in the 1550-nm region
G.653 (Edition 6, Dec. 2006)	Title: *Characteristics of a Dispersion-Shifted Single-Mode Optical Fiber and Cable*
	Description: Discusses single-mode optical fiber with the zero-dispersion wavelength shifted into the 1550 nm region. Describes chromatic dispersion for the 1460-to-1625-nm range for CWDM applications
G.654 (Edition 7, Dec. 2006)	Title: *Characteristics of a Cut-Off Shifted Single-Mode Optical Fiber and Cable*
	Description: Undersea applications; discusses single-mode optical fiber with a zero-dispersion wavelength around 1300 nm and with cutoff wavelength shifted to around 1550 nm
G.655 (Edition 5, Nov. 2009)	Title: *Characteristics of a Non-Zero Dispersion-Shifted Single-Mode Optical Fiber and Cable*
	Description: For applications in long-haul links; describes single-mode optical fiber with chromatic dispersion greater than zero throughout the 1530-to-1565-nm wavelength range
G.656 (Edition 2, Dec. 2006)	Title: *Characteristics of a Fiber and Cable with Non-Zero Dispersion for Wideband Optical Transport*
	Description: Low chromatic dispersion fiber for expanded WDM applications; can be used for both CWDM and DWDM systems throughout the wavelength region between 1460 and 1625 nm
G.657 (Edition 2, Nov. 2009)	Title: *Characteristics of a bending loss insensitive single-mode optical fiber and cable for the access network*
	Description: Addresses use of single-mode fiber for broadband access networks; includes issues such as sensitivity to tight bending conditions for in-building use

Table 3.3   *Transmission distances in meters in multimode fibers using an 850-nm VCSEL*

| | | 50-μm core | | 62.5-μm core | |
Application	Data rate (Gb/s)	500 MHz.km	2000 MHz.km	160 MHz.km	200 MHz.km
Ethernet	1	550	860	220	275
	10	82	300	26	33
Fibre Channel	1	500	860	250	300
	2	300	500	120	150
	10	82	300	26	33
SONET/SDH	10	85	300	25	33

is 0.4 dB/km at 1310 nm and 0.35 dB/km at 1550 nm. The maximum polarization mode dispersion is 0.2 ps/$\sqrt{\text{km}}$. Four subsets ranging from G.652a to G.652d describe different variations of this type of fiber. Since G.652a/b fibers were installed widely in telecommunication networks in the 1990s, they are commonly known as *standard single-mode fibers* or *1310-nm optimized fibers*. The G.652c/d fibers allow operation in the *E*-band and are used widely for *fiber-to-the-premises* (FTTP) installations (See Sec. 13.8 for a description of FTTP networks).

Although many long-distance cable plant installations now are using nonzero-dispersion-shifted fiber, the huge base of G.652 fiber that is installed worldwide will be in service for many years. If the G.652 fiber is used at 1550 nm, the chromatic dispersion value of about 17 ps/(nm-km) must be taken into account. This requires implementation of chromatic dispersion-compensation techniques or special data formats at high data rates. As an example, a number of field experiments have demonstrated the ability to transmit 160-Gb/s data rates over long distances of installed G.652a/b fiber.

Fig. 3.28   *Chromatic dispersion as a function of wavelength in various spectral bands for several different optical fiber types*

The G.652c/d fibers are created by reducing the water ion concentration in order to eliminate the attenuation spike in the 1360-to-1460-nm *E*-band. They are called *low-water-peak fiber* and allow operation over the entire wavelength range from 1260 to 1625 nm. One use of this fiber is for low-cost short-reach CWDM (coarse wavelength division multiplexing) applications in the *E*-band. In CWDM the wavelength channels are spaced by 20 nm, so that minimum wavelength stability control is needed for the optical sources, as described in Chapter 10. Another important application is in a *passive optical network* (PON) for FTTP access networks. Typically a FTTP link transmits three independent bidirectional channels at 1310, 1490, and 1550 nm over the same fiber. Chapter 13 has more details on FTTP networks.

3.4.3 Recommendation G.653

Dispersion-shifted fiber (DSF) was developed for use with 1550-nm lasers. As Fig. 3.28 illustrates, in this fiber type the zero-dispersion point is shifted to 1550 nm where the fiber attenuation is about half that at 1310 nm. Therefore, this fiber allows a high-speed data stream of a single-wavelength channel at or near 1550 nm to maintain its fidelity over long distances. However, it presents problems associated with nonlinear effects in dense wavelength division multiplexing (DWDM) applications in the center of the *C*-band where many wavelengths are packed tightly into one or more of the operational bands. As noted in Chapter 10, to prevent undesirable nonlinear effects in DWDM systems, the chromatic dispersion values should be positive (or negative) over the entire operational band. Figure 3.28 shows that for G.653 fibers the chromatic dispersion has a different sign above and below 1550 nm. Therefore, the use of G.653 fibers for DWDM should be restricted to either the *S*-band (wavelengths lower than 1550 nm) or the *L*-band (wavelengths higher than 1550 nm). These fibers are seldom deployed anymore because G.655 fibers offer a better solution.

3.4.4 Recommendation G.654

This recommendation deals with *cutoff-wavelength-shifted fiber* that is designed for long-distance high-power signal transmission. It describes the geometrical, mechanical, and transmission characteristics of a single-mode optical fiber, which has the zero-dispersion wavelength around 1300 nm. The fiber has a very low loss in the 1550-nm band, which is achieved by using a pure silica core. Since it has a high cutoff wavelength of 1500 nm, this fiber is restricted to operation in the 1500-to-1600-nm region. It typically is used only in long-distance undersea applications.

3.4.5 Recommendation G.655

Nonzero-dispersion-shifted fiber (NZDSF) was introduced in the mid-1990s for WDM applications. Its principal characteristic is that it has a positive nonzero dispersion value over the entire *C*-band, which is the spectral operating region for erbium-doped optical fiber amplifiers (see Chapter 11). This is in contrast to G.653 fibers in which the dispersion varies from negative through zero to positive values in the *C*-band. Version G.655b was introduced to extend WDM applications into the *S*-band. As shown in Fig. 3.28, the principal characteristic of a G.655b fiber is that it has a nonzero dispersion value over the entire *S*-band and the *C*-band. This is in contrast to standard G.655 fibers in which the dispersion varies from negative through zero to positive values in the *S*-band. Version G.655c specifies a lower PMD value of 0.2 ps/$\sqrt{\text{km}}$ than the 0.5 ps/$\sqrt{\text{km}}$ value of G.655a/b.

3.4.6 Recommendation G.656

This recommendation describes the characteristics of a single-mode optical fiber that has a positive chromatic dispersion value ranging from 2 to 14 ps/(nm-km) in the 1460-to-1625-nm wavelength band.

This means that the dispersion slope is significantly lower than in G.655 fibers for which the chromatic dispersion ranges from 1 to 10 ps/(nm-km) in the 1530-to-1565 band. The consequence of a lower dispersion slope means that the chromatic dispersion changes slower with wavelength so that dispersion compensation is simpler or not needed. This allows the use of CWDM without chromatic dispersion compensation and also means that 40 additional DWDM channels can be implemented in this wavelength band. Other G.656 attributes are similar to those of G.655 fibers. For example, the mode-field diameter ranges from 7 to 11 μm (compared to 8 to 11 μm for G.655 fibers), the maximum PMD value of cabled fiber is 0.2 ps/$\sqrt{\text{km}}$, and the cutoff wavelength is 1310 nm (the same as for G.655).

3.4.7 Recommendation G.657

The rapidly growing worldwide demand for broadband services in high-capacity access networks and enterprise networks put new demands on the performance characteristics of single-mode fibers that are different from telecom applications (see Sec. 13.8 for descriptions of these network types). These performance differences are due mainly to the high density of localized distribution and drop cables in the access and enterprise networks compared to metro and long-haul telecom networks. In particular, in-building applications typically are done in small and crowded spaces. These conditions and the many installation manipulations imposed on the cabling system call for optical fibers with low bending sensitivity.

Thus the aim of ITU-T Recommendation G.657 is to describe requirements for a fiber type that exhibits improved bending performance compared with existing G.652 single-mode fiber and cables. Two categories of single-mode fibers are specified. Fibers in Category A are fully compliant with the G.652 single-mode fibers and also can be used in other parts of the network. Single-mode fibers in Category B are not necessarily compliant with G.652 but exhibit low values of bending losses at very-low-bend radii. Fibers in Category B are predominantly intended for in-building use.

3.5 | Specialty Fibers

Telecommunication fibers, such as those described in Sec. 3.4, are designed to transmit light with minimal change in the signal fidelity. In contrast, *specialty fibers* are designed to interact with light and thereby manipulate or control some characteristic of an optical signal.[72–74] The light manipulation applications include optical signal amplification, optical power coupling, dispersion compensation, wavelength conversion, and sensing of physical parameters such as temperature, stress, pressure, vibration, and fluid levels. For light-control applications a specialty fiber can be insensitive to bends, maintain polarization states, redirect specific wavelengths, or provide a very high attenuation for fiber terminations.

Specialty fibers can be of either a multimode or a single-mode design. Among the optical devices that may use a specialty fiber are light transmitters, light signal modulators, optical receivers, wavelength multiplexers, light couplers and splitters, optical amplifiers, optical switches, wavelength add/drop modules, and optical power attenuators. Table 3.4 gives a summary of some specialty fibers and their applications.

Erbium-Doped Fiber   These fibers have small amounts of erbium ions (for example, 1000 parts per million weight) added to the silica material to form a basic building block for optical fiber amplifiers. As described in Chapter 11, a length of Er-doped fiber ranging from 10 to 30 m serves as a gain medium for amplifying optical signals in either the *C*-band (1530 to 1560 nm) or the *L*-band (1560 to 1625 nm). There are many variations on the doping level, cutoff wavelength, mode-field diameter, numerical

Table 3.4 *Examples of specialty fibers and their applications*

Specialty fiber type	Application
Erbium-doped fiber	Gain medium for optical fiber amplifiers
Photosensitive fibers	Fabrication of fiber Bragg gratings
Bend-insensitive fibers	Tightly looped connections in device packages
Termination fiber	Termination of open optical fiber ends
Polarization-preserving fibers	Pump lasers, polarization-sensitive devices, sensors
High-index fibers	Fused couplers, short-λ sources, DWDM devices
Photonic crystal fibers	Switches; dispersion compensation

aperture, and cladding diameter for these fibers. Specific erbium-doped fiber configurations will yield a variety of optical amplifier designs that can be selected according to pump laser power requirement, noise figure, signal gain, and flatness of the output spectrum. Higher erbium concentrations allow the use of shorter fiber lengths, smaller claddings are useful for compact packages, and a higher numerical aperture allows the fiber to be coiled tightly in small packages. Table 3.5 lists some generic parameter values of an erbium-doped fiber for use in the *C*-band.

Photosensitive Sensitive Fiber   The refractive index of a photosensitive fiber changes when it is exposed to ultraviolet light. This sensitivity may be provided by doping the fiber material with germanium and boron ions. The main application is to create a fiber Bragg grating, which is a periodic variation of the refractive index along the fiber axis (see Chapter 10). Applications of fiber Bragg gratings include light-coupling mechanisms for pump lasers used in optical amplifiers, wavelength add/drop modules, optical filters, and chromatic dispersion compensation modules.

Bend-Insensitive Fiber   Increasing the numerical aperture (NA) reduces the sensitivity of a single-mode fiber to bending loss.[75] The higher NA decreases the *mode-field diameter* (MFD), thereby

Table 3.5 *Generic parameter values of an erbium-doped fiber for use in the C-band*

Parameter	Specification
Peak absorption at 1530 nm	5 to 10 dB/m
Effective numerical aperture	0.14 to 0.31
Cutoff wavelength	900 ± 50 nm; or 1300 nm
Mode-field diameter at 1550 nm	5.0 to 7.3 μm
Cladding diameter	125 μm standard; 80 μm for tight coils
Coating material	UV-cured acrylic

confining optical power more tightly within the core than in conventional single-mode fibers. For example, at 1310 nm a fiber designed with an NA of 0.16 can have a MFD of 6.7 μm. An increase of the NA to 0.21 decreases the MFD to 5.1 μm. Bend insensitive fibers are available commercially in a range of core diameters to provide optimum performance at specific operating wavelengths, such as 820, 1310 nm, or 1550 nm. These fibers are offered with either an 80-μm or a 125-μm cladding diameter as standard products. The 80-μm *reduced-cladding fiber* results in a much smaller volume compared with a 125-μm cladding diameter when a fiber length is coiled up within a device package. Whereas there is a high bending loss for tightly wound conventional single-mode fibers, the induced attenuation when bend-insensitive fiber is wound into five coils with a 10-mm radius is less than 0.01 dB at 1310 nm and less than 0.5 dB at 1550 nm.

Termination Fiber   Often an optical device with multiple ports will have one or more unused or open branches. Back reflections from these ports can cause instabilities and need to be suppressed. This can be achieved by using a termination fiber. An example of such a fiber is a coreless silica construction. A termination that has a return loss of better than 65 dB can be achieved by splicing about 25 cm of a termination fiber onto the end of unused fiber branches.

Polarization-Preserving Fiber   In contrast to standard single-mode optical fibers in which the state of polarization fluctuates as a light signal propagates through the fiber, *polarization-preserving fibers* have a special core design that maintains the state of polarization. Applications of these fibers include light signal modulators fabricated from lithium niobate, optical amplifiers for polarization multiplexing, light-coupling fibers for pump lasers, and polarization-mode dispersion compensators. Figure 3.29 illustrates the cross-sectional geometry of four different polarization-maintaining fibers. The light circles represent the cladding and the dark areas are the core configurations. The goal in each design is to use stress-applying parts to create slow and fast axes in the core. Each of these axes will guide light at a different velocity. Crosstalk between the two axes is suppressed so that polarized light launched into either of the axes will maintain its state of polarization as it travels along the fiber.
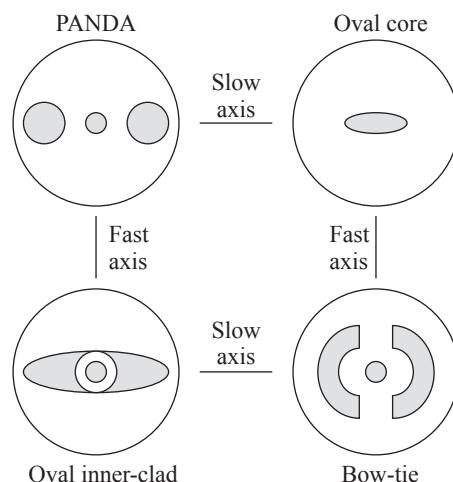
Fig. 3.29   *Cross-sectional geometry of four different polarization-maintaining fibers*

PROBLEMS

3.1 Verify the expression given in Eq. (3.1c) that relates α, which is in units of dB/km, to α_p, which is in units of km^{-1}.

3.2 A certain optical fiber has an attenuation of 0.6 dB/km at 1310 nm and 0.3 dB/km at 1550 nm. Suppose the following two optical signals are launched simultaneously into the fiber: an optical power of 150 μW at 1310 nm and an optical power of 100 μW at 1550 nm. What are the power levels in μW of these two signals at (*a*) 8 km and (*b*) 20 km?

3.3 An optical signal at a specific wavelength has lost 55 percent of its power after traversing 7.0 km of fiber. What is the attenuation in dB/km of this fiber?

3.4 A continuous 40-km-long optical fiber link has a loss of 0.4 dB/km.
   (*a*) What is the minimum optical power level that must be launched into the fiber to maintain an optical power level of 2.0 μW at the receiving end?
   (*b*) What is the required input power if the fiber has a loss of 0.6 dB/km?

3.5 Consider a step-index fiber with a SiO_2–GeO_2 core having a mole fraction 0.08 of GeO_2. Plot Eqs (3.2b) and (3.3) from 500 nm to 5 μm, and compare the results with the curves in Fig. 3.5.

3.6 The optical power loss resulting from Rayleigh scattering in a fiber can be calculated from either Eq. (3.4a) or Eq. (3.4b). Compare these two equations for silica ($n = 1.460$ at 630 nm), given that the fictive temperature T_f is 1400 K, the isothermal compressibility β_T is 6.8×10^{-12} cm^2/dyn, and the photoelastic coefficient is 0.286. How does this agree with measured values ranging from 3.9 to 4.8 dB/km at 633 nm?

3.7 Solve Eq. (3.7) to make plots of the radius of bend curvature versus the fiber core radius a for values of $M_{eff}/M_\infty = 10$, 50, and 75 percent at wavelengths of 1300 nm and 1550 nm. Let a be in the range 5 μm $\le a \le$ 30 μm.

3.8 Consider graded-index fibers having index profiles $\alpha = 2.0$, cladding refractive indices $n_2 = 1.50$, and index differences $\Delta = 0.01$. Using Eq. (3.7), plot the ratio M_{eff}/M_∞ for bend radii less than 10 cm at $\lambda = 1$ μm for fibers having core radii of 4, 25, and 100 μm.

3.9 Two common fiber jacket materials are Elvax® 265 ($E_j = 21$ MPa) and Hytrel® 4056 ($E_j = 58$ MPa), both made by DuPont. If the Young's modulus of a glass fiber is 64 GPa, plot the reduction in microbending loss as a function of the index difference Δ when fibers are coated with these materials. Make these plots for Δ values ranging from 0.1 to 1.0 percent and for a fiber cladding-to-core ratio of $b/a = 2$.

3.10 Assume that a step-index fiber has a V number of 6.0.
   (*a*) Using Fig. 2.27, estimate the fractional power P_{clad}/P traveling in the cladding for the six lowest-order LP modes.
   (*b*) If the fiber in (*a*) is a glass-core glass-clad fiber having core and cladding attenuations of 3.0 and 4.0 dB/km, respectively, find the attenuations for each of the six lowest-order modes.

3.11 Assume a given mode in a graded-index fiber has a power density $p(r) = P_0 \exp(-Kr^2)$, where the factor K depends on the modal power distribution.
   (*a*) Letting $n(r)$ in Eq. (3.11) be given by Eq. (2.78) with $\alpha = 2$, show that the loss in this mode is

$$\alpha_{gi} = \alpha_1 + \frac{\alpha_2 - \alpha_1}{Ka^2}$$

   Since $p(r)$ is a rapidly decaying function of r and since $\Delta \ll 1$, for ease of calculation assume that the top relation in Eq. (2.78) holds for all values of r.
   (*b*) Choose K such that $p(a) = 0.1 P_0$; that is, 10 percent of the power flows in the cladding. Find α_{gi} in terms of α_1 and α_2.

3.12 For wavelengths less than 1.0 μm the refractive index n satisfies a Sellmeier relation of the form[69]

$$n^2 = 1 + \frac{E_0 E_d}{E_0^2 - E^2}$$

where $E = hc/\lambda$ is the photon energy and E_0 and E_d are, respectively, material oscillator energy and dispersion energy parameters. In SiO$_2$ glass, $E_0 = 13.4$ eV and $E_d = 14.7$ eV. Show that, for wavelengths between 0.20 and 1.0 μm, the values of n found from the Sellmeier relation are in good agreement with those shown in Fig. 3.12. To make the comparison, select three representative points, for example, at 0.2, 0.6, and 1.0 μm.

3.13 (*a*) An LED operating at 850 nm has a spectral width of 45 nm. What is the pulse spreading in ns/km due to material dispersion? What is the pulse spreading when a laser diode having a 2-nm spectral width is used?

(*b*) Find the material-dispersion-induced pulse spreading at 1550 nm for an LED with a 75-nm spectral width. Use Fig. 3.13 to estimate $d\tau/d\lambda$.

3.14 Verify the plots for b, $d(Vb)/dV$, and $V\, d^2(Vb)/dV^2$ shown in Fig. 3.15. Use the expression for b given by Eq. (3.38).

3.15 Derive Eq. (3.13) by using a ray-tracing method.

3.16 Consider a step-index fiber with core and cladding diameters of 62.5 and 125 μm, respectively. Let the core index $n_1 = 1.48$ and let the index difference $\Delta = 1.5$ percent. Compare the modal dispersion in units of ns/km at 1310 nm of this fiber as given by Eq. (3.13) with the more exact expression

$$\frac{\sigma_{\text{mod}}}{L} = \frac{n_1 - n_2}{c}\left(1 - \frac{\pi}{V}\right)$$

where L is the length of the fiber and n_2 is the cladding index.

3.17 Consider a standard G.652 non-dispersion-shifted single-mode optical fiber that has a zero-dispersion wavelength at 1310 nm with a dispersion slope of $S_0 = 0.0970$ ps/(nm$^2 \cdot$ km). Plot the dispersion in the wavelength range 1270 nm $\leq \lambda \leq$ 1340 nm. Use Eq. (3.47).

3.18 A typical G.653 dispersion-shifted single-mode optical fiber has a zero-dispersion

wavelength at 1550 nm with a dispersion slope of $S_0 = 0.070$ ps/(nm$^2 \cdot$ km).

(*a*) Plot the dispersion in the wavelength range 1500 nm $\leq \lambda \leq$ 1600 nm using Eq. 3.49.

(*b*) Compare the dispersion at 1500 nm with the dispersion value for the non-dispersion-shifted fiber described in Prob. 3.17.

3.19 Starting with Eq. (3.45), derive the dispersion expression given in Eq. (3.47).

3.20 Renner[19] derived a simplified approximation to describe the bend losses of single-mode optical fibers. This expression for the bending loss is

$$\alpha_{\text{simp}} = \alpha_{\text{conv}}\frac{2(Z_3 Z_2)^{1/2}}{(Z_3 + Z_2) - (Z_3 - Z_2)\cos(2\Theta)}$$

where the conventional bending loss is

$$\alpha_{\text{conv}} = \frac{1}{2}\left(\frac{\pi}{\gamma^3 R}\right)^{1/2}\frac{\kappa^2}{V^2 K_1^2(\gamma a)}\exp\left(-\frac{2\gamma^3 R}{3\beta_0^3}\right)$$

where V is given by Eq. (2.57), β_0 is the propagation constant in a straight fiber with an infinite cladding given by Eq. (2.46), K_1 is the modified Bessel function (see App. C), and

$$Z_q \approx k^2 n_q^2(1 + 2b/R) - \beta_0^2$$
$$\approx k^2 n_q^2(1 + 2b/R) - k^2 n_2^2$$
$$\text{for } q = 2, 3$$

$$\Theta = \frac{\gamma^3 R}{3k^2 n_2^2}\left(\frac{R_c}{R} - 1\right)^{3/2}$$

$$\gamma = \left(\beta_0^2 - k^2 n_2^2\right)^{1/2} \approx k\left(n_1^2 - n_2^2\right)^{1/2}$$

$$\kappa^2 = k^2 n_1^2 - \beta_0^2 \approx k^2(n_1^2 - n_2^2)$$

$$R_c = 2k^2 n_2^2\, b/\gamma^2 = \text{the critical bend radius}$$

Using a computer, (*a*) verify the plot given in Fig. 3.27 at 1300 nm, and (*b*) calculate and plot the bend loss as a function of wavelength for 800 nm $\leq \lambda \leq$ 1600 nm at several different bend radii (e.g., 15 and 20 mm). Let $n_1 = 1.480$, $n_2 = 1.475$, $n_3 = 1.07 n_2 = 1.578$, and $b = 60\,\mu$m.

3.21 Faustini and Martini[20] developed a more detailed formula for describing the oscillatory behavior of bend loss as a function of the bend radius and wavelength. Using a computer, use their formulation to reproduce the three-dimensional plots of bend loss versus radius of curvature and wavelength given in Fig. 5 of their paper (*J. Lightwave Tech.,* vol. 15, pp. 671–679, April 1997).

3.22 Calculate the waveguide dispersion at 1320 nm in units of [ps/(nm · km)] for a single-mode fiber with core and cladding diameters of 9 μm and 125 μm, respectively. Let the core index n_1 = 1.48 and let the index difference Δ = 0.22 percent.

REFERENCES

1. B. C. Bagley, C. R. Kurkjian, J. W. Mitchell, G. E. Peterson, and A. R. Tynes, "Materials, properties, and choices," in *Optical Fiber Telecommunications*, S. E. Miller and A. G. Chynoweth, eds., Academic, New York, 1979.

2. D. Gloge, "The optical fibre as a transmission medium," *Rpts. Prog. Phys.*, vol. 42, pp. 1777–1824, Nov. 1979.

3. R. Olshansky, "Propagation in glass optical waveguides," *Rev. Mod. Phys.*, vol. 51, pp. 341–367, Apr. 1979.

4. K. Tsujikawa, K. Tajima, and J. Zhou, "Intrinsic loss of optical fibers," *Optical Fiber Tech.*, vol. 11, no. 4, pp. 319–331, Oct. 2005.

5. E. M. Dianov and V. M. Mashinsky, "Germania–based core optical fibers," *J. Lightwave Tech.*, vol. 23, pp. 3500–3508, Nov. 2005.

6. A. Iino and J. Tamura, "Radiation resistivity in silica optical fibers*," J. Lightwave Tech.,* vol. 6, pp. 145–149, Feb. 1988.

7. R. H. West, H. Buker, E. J. Friebele, H. Henschel, and P. B. Lyons, "The use of optical time-domain reflectometers to measure radiation-induced losses in optical fibers," *J. Lightwave Tech.,* vol. 12, pp. 614–620, Apr. 1994.

8. H. Henschel and O. Köhn, "Regeneration of irradiated optical fibers by photobleaching," *IEEE Trans. Nucl. Sci.,* vol. 47, pp. 699–704, June 2000.

9. S. Girard, J. Baggio, and J.-L. Leray, "Radiation-induced effects in a new class of optical waveguides: The air-guiding photonic crystal fibers," *IEEE Trans. Nucl. Sci.,* vol. 52, pp. 2683–2688, Dec. 2005.

10. S. R. Nagel, "Fiber materials and fabrication methods," in S. E. Miller and I. P Kaminow, eds., *Optical Fiber Telecommunications–II*, Academic, New York, 1988.

11. K. Nagayama, M. Matsui, M. Kakui, T. Saitoh, K. Kawasaki, H. Takamizawa, Y. Ooga, I. Tsuchiya, and Y. Chigusa, "Ultra low loss (0.1484 dB/km) pure silica core fiber," *SEI Technical Review*, vol. 57, pp. 3–6, Jan. 2003.

12. Y. Chigusa, Y. Yamamoto, T. Yokokawa, T. Sasaki, T. Taru, M. Hirano, M. Kakui, M. Onishi, and E. Sasaoka, "Low-loss pure-silica-core fibers and their possible impact on transmission systems," *J. Lightwave Tech.*, vol. 23, pp. 3541–3550, Nov. 2005.

13. H. Osanai, T. Shioda, T. Moriyama, S. Araki, M. Horiguchi, T. Izawa, and H. Takata, "Effects of dopants on transmission loss of low OH content optical fibers," *Electron. Lett.*, vol. 12, pp. 549–550, Oct. 1976.

14. R. Maurer, "Glass fibers for optical communications," *Proc. IEEE*, vol. 61, pp. 452–462, Apr. 1973.

15. D. A. Pinnow, T. C. Rich, F. W. Ostermeyer, and M. DiDomenico Jr., "Fundamental optical attenuation limits in the liquid and gassy state with application to fiber optical waveguide material," *Appl. Phys. Lett.*, vol. 22, pp. 527–529, May 1973.

16. D. B. Keck, "Fundamentals of optical waveguide fibers," *IEEE Commun. Magazine*, vol. 23, pp. 17–22, May 1985.

17. W. A. Gambling, H. Matsumura, and C. M. Ragdale, "Curvature and microbending losses

in single-mode optical fibers," *Opt. Quantum Electron.*, vol. 11, no. 1, pp. 43–59, Jan. 1979.

18. D. Gloge, "Bending loss in multimode fibers with graded and ungraded core index," *Appl. Opt.*, vol. 11, pp. 2506–2512, Nov. 1972.

19. H. Renner, "Bending losses of coated single-mode fibers: A simple approach," *J. Lightwave Tech.*, vol. 10, pp. 544–551, May 1992.

20. L. Faustini and G. Martini, "Bend loss in single-mode fibers," *J. Lightwave Tech.,* vol. 15, pp. 671–679, Apr. 1997.

21. S.-L. Tsao and W. M. Cheng, "Realization of an on-line fiber-optic bending loss measurement system," *IEEE Trans. Instrum. Meas.,* vol. 53, pp. 72–79, Feb. 2004.

22. K. Makino, T. Nakamura, T. Ishigure, and Y. Koike, "Analysis of graded-index polymer optical fiber link performance under fiber bending," *J. Lightwave Tech.,* vol. 23, pp. 2062–2072, June 2005.

23. J. Koning, R. N. Rieben, and G. H. Rodrigue, "Vector finite-element modeling of the full-wave Maxwell equations to evaluate power loss in bent optical fibers," *J. Lightwave Tech.,* vol. 23, pp. 4147–4154, Dec. 2005.

24. K. Himeno, S. Matsuo, N. Guan, and A. Wada, "Low-bending-loss single-mode fibers for fiber-to-the-home," *J. Lightwave Tech.*, vol. 23, pp. 3494–3499, Nov. 2005.

25. P. Matthijsse and W. Griffioen, "Matching optical fiber lifetime and bend-loss limits for optimized local loop fiber storage," *Optical Fiber Tech.,* vol. 11, no. 1, pp. 92–99, Jan. 2005.

26. T. P. Hansen, J. Broeng, C. Jakobsen, G. Vienne, H. R. Simonsen, M. D. Nielsen, P.M.W. Skovgaard, J. R. Folkenberg, and A. Bjarklev, "Air-guiding photonic bandgap fibers: Spectral properties, macrobending loss, and practical handling," *J. Lightwave Tech.*, vol. 22, pp. 11–15, Jan. 2005.

27. W. B. Gardner, "Microbending loss in optical fibers," *Bell Sys. Tech. J.*, vol. 54, pp. 457–465, Feb. 1975.

28. J. Sakai and T. Kimura, "Practical microbending loss formula for single mode optical fibers," *IEEE J. Quantum Electron.*, vol. QE-15, pp. 497–500, June 1979.

29. S.-T. Shiue and Y.-K. Tu, "Design of single-coated optical fibers to minimize thermally and mechanically induced microbending losses," *J. Optical Commun.*, vol. 15, pp. 16–19, Jan. 1994.

30. V. Arya, K. A. Murphy, A. Wang, and R. O. Claus, "Microbend losses in single-mode optical fibers: Theoretical and experimental investigation," *J. Lightwave Tech.*, vol. 13, pp. 1998–2002, Oct. 1995.

31. D. Gloge, "Optical fiber packaging and its influence on fiber straightness and loss," *Bell Sys. Tech. J.*, vol. 54, pp. 245–262, Feb. 1975.

32. D. Gloge, "Propagation effects in optical fibers," *IEEE Trans. Microwave Theory Tech.*, vol. MTT–23, pp. 106–120, Jan. 1975.

33. D. Marcuse, *Theory of Dielectric Optical Waveguides*, Academic, New York, 2nd. ed., 1991.

34. D. Gloge, E.A.J. Marcatili, D. Marcuse, and S. D. Personick, "Dispersion properties of fibers," in *Optical Fiber Telecommunications*, S. E. Miller and A. G. Chynoweth, eds., Academic, New York, 1979.

35. D. Marcuse, "Interdependence of waveguide and material dispersion," *Appl. Opt.*, vol. 18, pp. 2930–2932, Sept. 1979.

36. B.E.A. Saleh and M. Teich, *Fundamentals of Photonics*, Wiley, Hoboken, NJ, 2nd ed., 2007.

37. D. Gloge, "Weakly guiding fibers," *Appl. Opt.*, vol. 10, pp. 2252–2258, Oct. 1971; "Dispersion in weakly guiding fibers," *Appl. Opt.*, vol. 10, pp. 2442–2445, Nov. 1971.

38. C. D. Poole and J. Nagel, "Polarization effects in lightwave systems," chap. 6, pp. 114–161, in I. P. Kaminow and T. L. Koch, eds., *Optical Fiber Telecommunications–III*, Vol. A, Academic, New York, 1997.

39. A. E. Willner, S.M.R. Motaghian Nezam, L. Yan, Z. Pan, and M. C. Hauer, "Monitoring and control of polarization-related impairments in optical fiber systems," *J. Lightwave Tech.*, vol. 22, pp. 106–125, Jan. 2004.

40. J. Cameron, L. Chen, X. Bao, and J. Stears, "Time evolution of polarization mode dispersion in optical fibers," *IEEE Photonics Tech. Lett.*, vol. 10, pp. 1265–1267, Sept. 1998.

41. D. S. Waddy, L. Chen, and X. Bao, "Polarization effects in aerial fibers," *Optical Fiber Tech.*, vol. 11, no. 1, pp. 1–19, Jan. 2005.

42. S. Lanne and E. Corbel, "Practical considerations for optical polarization-mode dispersion compensators," *J. Lightwave Tech.*, vol. 22, pp. 1033–1040, Apr. 2004.

43. H. F. Haunstein, W. Sauer-Greff, A. Dittrich, K. Sticht, and R. Urbansky, "Principles of electronic equalization of polarization-mode dispersion," *J. Lightwave Tech.*, vol. 22, pp. 1169–1182, Apr. 2004.

44. R. Noe, D. Sandel, and V. Mirvoda, "PMD in high-bit-rate transmission and means for its mitigation," *IEEE J. Sel. Topics Quantum Electron.*, vol. 10, no. 2, pp. 341–355, Mar./Apr. 2004.

45. F. Buchali and H. Bülow, "Adaptive PMD compensation by electrical and optical techniques," *J. Lightwave Tech.*, vol. 22, pp. 1116–1126, Apr. 2004.

46. M. Jäger, T. Rankl, J. Speidel, H. Bülow, and F. Buchali, "Performance of turbo equalizers for optical PMD channels," *J. Lightwave Tech.*, vol. 24, pp. 1226–1236, Mar. 2006.

47. L. S. Yan, X. Steve Yao, Y. Shi, and A.E. Willner, "Simultaneous monitoring of both optical signal-to-noise-ratio and polarization-mode-dispersion using polarization scrambling and polarization-beam-splitting," *J. Lightwave Tech.*, vol. 23, pp. 3290–3294, Oct. 2005.

48. L. S. Yan, X. Steve Yao, C. Yu, G. Xie, Y. Wang, L. Lin, Z. Chen, and A. E. Willner, "High-speed and highly repeatable polarization-state generator and analyzer for 40-Gb/s system performance monitoring," *IEEE Photonics Tech. Lett.*, vol. 18, pp. 643–645, 2006.

49. D. A. Nolan, X. Chen, and M.-J. Li, "Fibers with low polarization-mode dispersion," *J. Lightwave Tech.*, vol. 22, pp. 1066–1077, Apr. 2004.

50. A. Galtarossa, Y. Jung, M. J. Kim, B. H. Lee, K. Oh, U.-C. Paek, L. Palmieri, A. Pizzanat, and L. Schenato, "Effects of spin inaccuracy on PMD reduction in spun fibers," *J. Lightwave Tech.*, vol. 23, pp. 4184–4191, Dec. 2005.

51. J. C. Lapp, V. A. Bhagavatula, and A. J. Morrow, "Segmented-core single-mode fiber optimized for bending performance," *J. Lightwave Tech.*, vol. 6, pp. 1462–1465, Oct. 1988.

52. B. J. Ainsle and C. R. Day, "A review of single-mode fibers with modified dispersion characteristics," *J. Lightwave Tech.*, vol. LT–4, pp. 967–979, Aug. 1986.

53. D. P. Jablonowski, U. C. Paek, and L. S. Watkins, "Optical fiber manufacturing techniques," *AT&T Tech. J.*, vol. 66, pp. 33–44, Jan./Feb. 1987.

54. H. J. Hagemann, H. Lade, J. Warnier, and D. H. Wiechert, "The performance of depressed-cladding single-mode fibers with different b/a ratios," *J. Lightwave Tech.*, vol. 9, pp. 689–694, June 1991.

55. Y. W. Li, C. D. Hussey, and T. A. Birks, "Triple-clad single-mode fibers for dispersion-shifting," *J. Lightwave Tech.*, vol. 11, pp. 1812–1819, Nov. 1993.

56. W. A. Reed, L. G. Cohen, and H. T. Shang, "Tailoring optical characteristics of dispersion-shifted lightguides for applications near 1.55 μm," *AT&T Tech. J.*, vol. 65, pp. 105–122, Sept./Oct. 1986.

57. D. Kalish and L. G. Cohen, "Single-mode fiber: From research and development to manufacturing," *AT&T Tech. J.*, vol. 66, pp. 19–32, Jan./Feb. 1987.

58. M. Kato, K. Kurokawa, and Y. Miyajima, "A new design for dispersion-shifted fiber with an effective core area larger than 100 μm^2 and good bending characteristics," *1998 OSA Tech. Digest–Opt. Fiber Comm. Conf (OFC 98)*, pp. 301–302, Feb. 1998.

59. C. Weinstein, "Fiber design improves long-haul performance," *Laser Focus World*, vol. 33, pp. 215–220, May 1997.

60. A. V. Belov, "Fiber design with effective area over 100 μm^2 for long-haul communication lines," *J. Opt. Commun.*, vol. 22, no. 2, pp. 64–66, 2001.

61. X. Jiang and R. Wang, "Non-zero dispersion-shifted optical fiber with ultra-large effective area and low dispersion slope for terabit communication system," *Opt. Commun.*, vol. 236, no. 1–3, pp. 69–74, 2004.

62. V. A Bhagavatula, M. S. Spotz, W. F. Love, and D. B. Keck, "Segmented-core single-mode fibers

with low loss and low dispersion," *Electron. Lett.*, vol. 19, pp. 317–318, Apr. 28, 1983.

63. P. K. Bachmann, D. Leers, H. Wehr, D. U. Wiechert, J. A. Van Steenwijk, D.L.A. Tjaden, and E. R. Wehrhahn, "Dispersion-flattened single-mode fibers prepared with PCVD: Performance, limitations, design optimization*," J. Lightwave Tech.*, vol. LT-4, pp. 858–863, July 1986.

64. D. L. Franzen, "Determining the effective cutoff wavelength of single-mode fibers: An interlaboratory comparison," *J. Lightwave Tech.*, vol. 3, pp. 128–134, Feb. 1985.

65. ITU-T Recommendation G.650.1, *Definitions and Test Methods for Linear, Deterministic Attributes of Single-Mode Fibre and Cable*, Apr. 2004.

66. TIA/EIA-455-80C (FOTP-80), *Cutoff Wavelength*, 2003.

67. A. J. Barlow, R. S. Jones, and K. W. Forsyth, "Technique for direct measurement of single-mode fiber chromatic dispersion," *J. Lightwave Tech.*, vol. 5, pp. 1207–1217, Sept. 1987.

68. ITU-T Recommendation G.652, *Characteristics of a Single-Mode Optical Fiber and Cable*, ed. 8, Nov. 2009.

69. M. DiDomenico Jr., "Material dispersion in optical fiber waveguides," *Appl. Opt.,* vol. 11, pp. 652–654, Mar. 1972.

70. R. Hui and M. O'Sullivan, *Fiber Optic Measurement Techniques*, Academic, New York, 2009.

71. L.-A. de Montmorillon, G. Kuyt, P. Nouchi, and A. Bertaina, "Latest advances in optical fibers," *Comptes Rendus Physique*, vol. 9, pp. 1045–1054, Nov. 2008.

72. A. Mendez and T. F. Morse, *Specialty Optical Fibers Handbook*, Academic, New York, 2008.

73. A. A. Stolov, D. A. Simoff, and J. Li, "Thermal stability of specialty optical fibers," *J. Lightwave Tech.*, vol. 26, no. 20, pp. 3443–3451, Oct. 2008.

74. U.-C. Paek and K. Oh, *Specialty Optical Fiber Technology for Optical Devices and Components*, Wiley, Hoboken, NJ, 2010.

75. M.-J. Li, P. Tandon, D. C. Bookbinder, S. R. Bickham, M. A. McDermott, R. B. Desorcie, D. A. Nolan, J. J. Johnson, K. A. Lewis, and J. J. Englebert, "Ultra-low bending loss single-mode fiber for FTTH," *J. Lightwave Tech.*, vol. 27, no. 3, pp. 376–382, Feb. 2009.

CHAPTER 4

Optical Sources

Two classes of light sources that are widely used for fiber optic communications are heterojunction-structured semiconductor *laser diodes* (also referred to as *injection laser diodes* or ILDs) and *light-emitting diodes* (LEDs). A *heterojunction* consists of two adjoining semiconductor materials with different bandgap energies. These devices are suitable for fiber transmission systems because they have adequate output power for a wide range of applications, their optical power output can be directly modulated by varying the input current to the device, they have a high efficiency, and their dimensions are compatible with those of the optical fiber. Comprehensive treatments of the major aspects of LEDs and laser diodes are presented in various books and review articles.[1–12] Chapter 11 addresses pump lasers.

The intent of this chapter is to give an overview of the pertinent characteristic of fiber-compatible luminescent sources. The first section discusses semiconductor material fundamentals that are relevant to light source operation. The next two sections present the output and operating characteristics of LEDs and laser diodes, respectively. These are followed by sections discussing the temperature responses of optical sources, their linearity characteristics, and their reliability under various operating conditions.

We shall see in this chapter that the light-emitting region of both LEDs and laser diodes consists of a *pn* junction constructed of direct-bandgap III-V semiconductor materials. When this junction is forward biased, electrons and holes are injected into the *p* and *n* regions, respectively. These injected minority carriers can recombine either radiatively, in which case a photon of energy $h\nu$ is emitted, or nonradiatively, whereupon the recombination energy is dissipated in the form of heat. This *pn* junction is thus known as the *active* or *recombination region.*

A major difference between LEDs and laser diodes is that the optical output from an LED is incoherent, whereas that from a laser diode is coherent. In a coherent source, the optical energy is produced in an optical resonant cavity. The optical energy released from this cavity has spatial and temporal coherence, which means it is highly monochromatic and the output beam is very directional. In an incoherent LED source, no optical cavity exists for wavelength selectivity. The output radiation has a broad spectral width, since the emitted photon energies range over the energy distribution of the recombining electrons and holes, which usually lie between 1 and $2k_B T$ (k_B is Boltzmann's constant and T is the absolute temperature at the *pn* junction). In addition, the incoherent optical energy is emitted into a broad elliptical region according to a cosine power distribution and thus has a large beam divergence.

In choosing an optical source compatible with the optical waveguide, various characteristics of the fiber, such as its geometry, its attenuation as a function of wavelength, its group delay distortion (bandwidth),

and its modal characteristics, must be taken into account. The interplay of these factors with the optical source power, spectral width, radiation pattern, and modulation capability needs to be considered. The spatially directed coherent optical output from a laser diode can be coupled into either single-mode or multimode fibers. In general, LEDs are used with multimode fibers, since normally it is only into a multimode fiber that the incoherent optical power from an LED can be coupled in sufficient quantities to be useful. However, LEDs have been employed in high-speed local-area applications in which one wants to transmit several wavelengths on the same fiber. Here, a technique called *spectral slicing* is used.[13–15] This entails using a passive device such as a waveguide grating array (see Chapter 10) to split the broad spectral emission of the LED into narrow spectral slices. Because these slices are each centered at a different wavelength, they can be individually modulated externally with independent data streams and simultaneously sent on the same fiber.

4.1 | Topics from Semiconductor Physics

The material in this chapter assumes a rudimentary knowledge of semiconductor physics, but various relevant definitions are given here for semiconductor material properties, including the concepts of energy bands, intrinsic and extrinsic materials, *pn* junctions, and direct and indirect bandgaps. Further details can be found in Refs. 16–18.

4.1.1 Energy Bands

Semiconductor materials have conduction properties that lie somewhere between those of metals and insulators. As an example material, we consider silicon (Si), which is located in the fourth column (group IV) of the periodic table of elements. A Si atom has four electrons in its outer shell, by which it makes covalent bonds with its neighboring atoms in a crystal. Such outer-shell electrons are called *valence electrons.*

The conduction properties of a semiconductor can be interpreted with the aid of the *energy-band diagrams* shown in Fig. 4.1a. In a semiconductor the valence electrons occupy a band of energy levels called the *valence band.* This is the lowest band of allowed states. The next higher band of allowed energy levels for the electrons is called the *conduction band.* In a pure crystal at low temperatures, the *conduction band* is completely empty of electrons and the *valence band* is completely full. These two bands are separated by an *energy gap,* or *bandgap,* in which no energy levels exist. As the temperature is raised, some electrons are thermally excited across the bandgap. For Si this excitation energy must be greater than 1.1 eV, which is the bandgap energy. This electron excitation process gives rise to a concentration n of free electrons in the conduction band, which leaves behind an equal concentration p of vacancies, or *holes,* in the valence band, as is shown schematically in Fig. 4.1b. Both the free electrons and the holes are mobile within the material, so that both can contribute to electrical conductivity; that is, an electron in the valance band can move into a vacant hole. This action makes the hole move in the opposite direction to the electron flow, as is shown in Fig. 4.1a.

When an electron propagates in a semiconductor, it interacts with the periodically arranged constituent atoms of the material and thus experiences *external forces*. As a result, to describe its acceleration a_{crys} in a semiconductor crystal under an external force F_{ext} its mass needs to be described by a quantum mechanical quantity m_e called the *effective mass*. That is, when using the relationship $F_{ext} = m_e a_{crys}$ (force equals mass times acceleration), the effects of all the forces exerted on the electron within the material are incorporated into m_e.
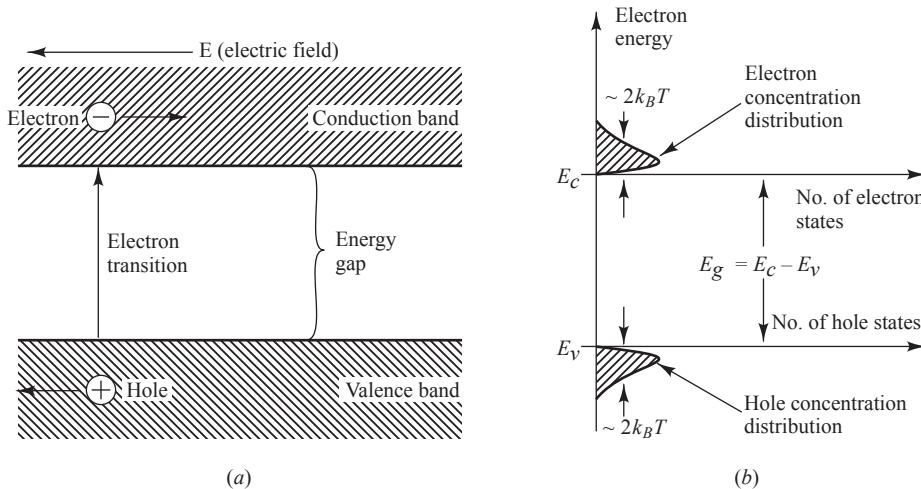
Fig. 4.1 *(a) Energy level diagrams showing the excitation of an electron from the valence band to the conduction band. The resultant free electrons and holes move under the influence of an external electric field E; (b) equal electron and hole concentrations in an intrinsic semiconductor created by the thermal excitation of electrons across the bandgap.*

The concentration of electrons and holes is known as the *intrinsic carrier concentration n_i,* and for a perfect material with no imperfections or impurities it is given by

$$n = p = n_i = K \exp\left(-\frac{E_g}{2k_B T}\right) \tag{4.1}$$

where

$$K = 2(2\pi k_B T/h^2)^{3/2} (m_e m_h)^{3/4}$$

Example 4.1 Consider the following parameter values for GaAs at 300 K:

Electron rest mass $m = 9.11 \times 10^{-31}$ kg

Effective electron mass $m_e = 0.068\, m = 6.19 \times 10^{-32}$ kg

Effective hole mass $m_h = 0.56\, m = 5.10 \times 10^{-31}$ kg

Bandgap energy $E_g = 1.42$ eV

What is the intrinsic carrier concentration?

Solution: First we need to change the bandgap energy to units of joules:

$$E_g = 1.42 \text{ eV} \times 1.60 \times 10^{-19} \text{ J/eV}$$

Then from Eq. (4.1) we find that the intrinsic carrier concentration is

$$n_i = 2\left(\frac{2\pi(1.381 \times 10^{-23})300}{(6.626 \times 10^{-34})^2}\right)^{3/2} [(6.19 \times 10^{-32})$$

$$\times (5.10 \times 10^{-31})]^{3/4} \exp\left(-\frac{1.42 \times 1.60 \times 10^{-19}}{2(1.381 \times 10^{-23})300}\right)$$

$$= 2.62 \times 10^{12} \text{ m}^{-3} = 2.62 \times 10^{6} \text{ cm}^{-3}$$
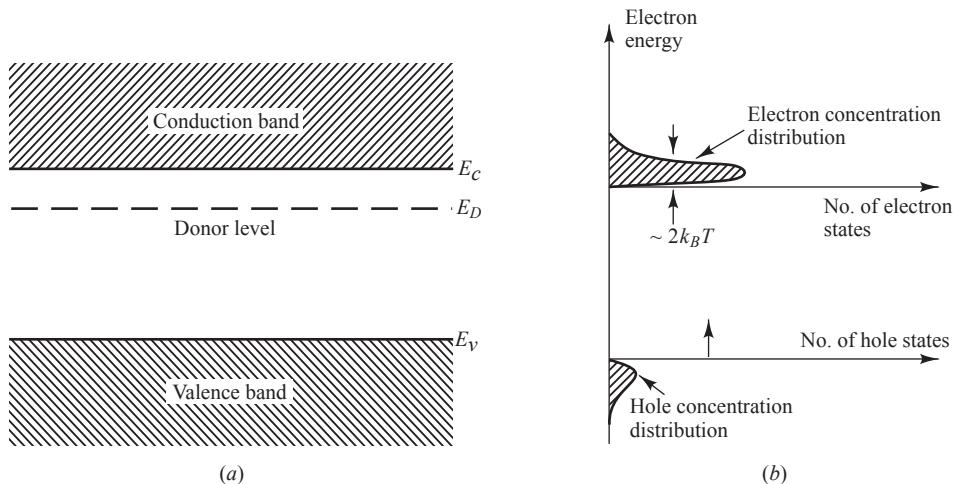
Fig. 4.2 *(a) Donor level in an n-type material; (b) the ionization of donor impurities increases the electron concentration distribution in the conduction band.*

is a constant that is characteristic of the material. Here, T is the temperature in degrees Kelvin, k_B is Boltzmann's constant, h is Planck's constant, and m_e and m_h are the effective masses of the electrons and holes, respectively, which can be smaller by a factor of 10 or more than the free-space electron rest mass of 9.11×10^{-31} kg.

The conduction can be greatly increased by adding traces of impurities from the group V elements (e.g., P, As, Sb). This process is called *doping*, and the doped semiconductor is called an *extrinsic material*. These elements have five electrons in the outer shell. When they replace a Si atom, four electrons are used for covalent bonding, and the fifth, loosely bound electron is available for conduction. As shown in Fig. 4.2*a*, this gives rise to an occupied level, just below the conduction band, called the *donor level*. The impurities are called *donors* because they can give up (donate) an electron to the conduction band. This is reflected by the increase in the free-electron concentration in the conduction band, as shown in Fig. 4.2*b*. Since in this type of material the current is carried by (negative) electrons (because the electron concentration is much higher than that of holes), it is called *n-type* material.

The conduction can also be increased by adding group III elements, which have three electrons in the outer shell. In this case, three electrons make covalent bonds, and a hole with properties identical to that of the donor electron is created. As shown in Fig. 4.3*a*, this gives rise to an unoccupied level just above the valence band. Conduction occurs when electrons are excited from the valence band to this *acceptor level* (so called because the impurity atoms have accepted electrons from the valence band). Correspondingly, the free-hole concentration increases in the valence band, as shown in Fig. 4.3*b*. This is called *p*-type material because the conduction is a result of (positive) hole flow.

4.1.2   Intrinsic and Extrinsic Material

A perfect material containing no impurities is called an *intrinsic material*. Because of thermal vibrations of the crystal atoms, some electrons in the valence band gain enough energy to be excited to the conduction band. This *thermal generation process* produces free electron–hole pairs because every electron that moves to the conduction band leaves behind a hole. Thus for an intrinsic material the number of electrons and holes are both equal to the intrinsic carrier density, as denoted by Eq. (4.1). In the opposite *recombination process*, a free electron releases its energy and drops into a free hole in the valence band. For an extrinsic
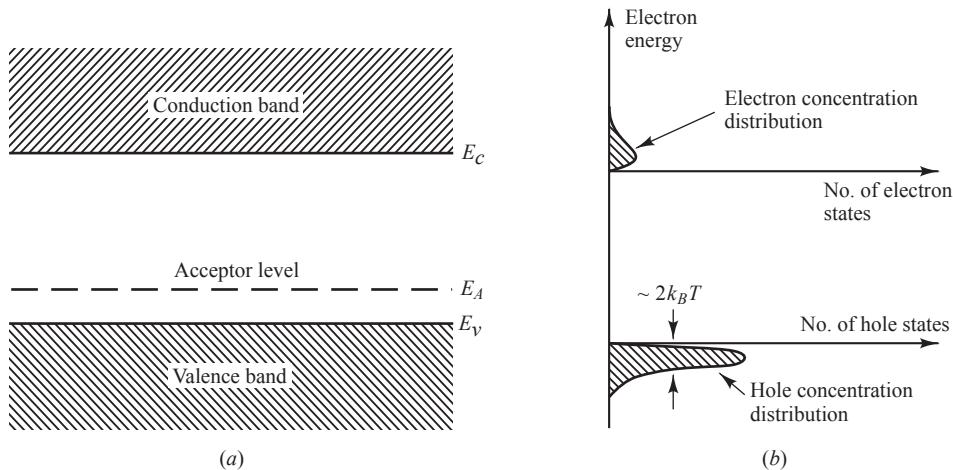
Fig. 4.3 *(a) Acceptor level in a p-type material; (b) the ionization of acceptor impurities increases the hole concentration distribution in the valence band.*

semiconductor, the increase of one type of carrier reduces the number of the other type. In this case, the product of the two types of carriers remains constant at a given temperature. This gives rise to the *mass-action law*

$$pn = n_i^2 \tag{4.2}$$

which is valid for both intrinsic and extrinsic materials under thermal equilibrium.

Since the electrical conductivity is proportional to the carrier concentration, two types of charge carriers are defined for this material:

1. *Majority carriers* refer either to electrons in *n*-type material or to holes in *p*-type material.
2. *Minority carriers* refer either to holes in *n*-type material or to electrons in *p*-type material.

The operation of semiconductor devices is essentially based on the *injection* and *extraction* of minority carriers.

Example 4.2 Consider an *n*-type semiconductor that has been doped with a net concentration of N_D donor impurities. Let n_N and p_N be the electron and hole concentrations, respectively, where the subscript N is used to denote *n*-type semiconductor characteristics. In this case, holes are created exclusively by thermal ionization of intrinsic atoms. This process generates equal concentrations of electrons and holes, so that the hole concentration in an *n*-type semiconductor is

$$p_N = p_i = n_i$$

Since conduction electrons are generated by both impurity and intrinsic atoms, the total conduction-electron concentration n_N is

$$n_N = N_D + n_i = N_D + p_N$$

Substituting Eq. (4.2) for p_N (which states that, in equilibrium, the product of the electron and hole concentrations equals the square of the intrinsic carrier density, so that $p_N = n_i^2 / n_N$), we have

$$n_N = \frac{N_D}{2} \left(\sqrt{1 + \frac{4n_i^2}{N_D^2}} + 1 \right)$$

If $n_i \ll N_D$, which is generally the case, then to a good approximation

$$n_N = N_D \quad \text{and} \quad p_N = n_i^2 / N_D$$
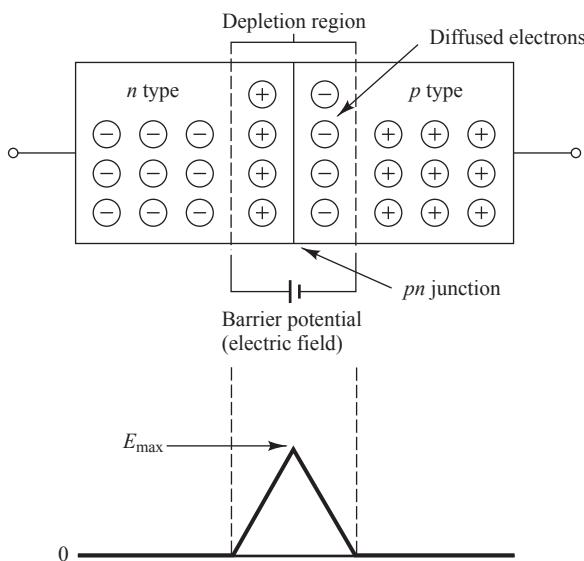
Fig. 4.4 *Electron diffusion across a pn junction creates a barrier potential (electric field) in the depletion region.*

4.1.3 The *pn* Junctions

Doped *n-* or *p*-type semiconductor material by itself serves only as a conductor. To make devices out of these semiconductors, it is necessary to use both types of materials (in a single, continuous crystal structure). The junction between the two material regions, which is known as the *pn junction,* is responsible for the useful electrical characteristics of a semiconductor device.

When a *pn* junction is created, the majority carriers diffuse across it. This causes electrons to fill holes in the *p* side of the junction and causes holes to appear on the *n* side. As a result, an electric field (or *barrier potential*) appears across the junction, as is shown in Fig. 4.4. This field prevents further net movements of charges once equilibrium has been established. The junction area now has no mobile carriers because its electrons and holes are locked into a covalent bond structure. This region is called either the *depletion region* or the *space charge region.*

When an external battery is connected to the *pn* junction with its positive terminal to the *n*-type material and its negative terminal to the *p*-type material, the junction is said to be *reverse-biased.* This is shown in Fig. 4.5. As a result of the reverse bias, the width of the depletion region will increase on both the *n* side and the *p* side. This effectively increases the barrier potential and prevents any majority carriers from flowing across the junction. However, minority carriers can move with the field across the junction. The minority carrier flow is small at normal temperatures and operating voltages, but it can be significant when excess carriers are created as, for example, in an illuminated photodiode.

When the *pn* junction is *forward-biased,* as shown in Fig. 4.6, the magnitude of the barrier potential is reduced. Conduction-band electrons on the *n* side and valence-band holes on the *p* side are,

Fig. 4.5 *A reverse bias widens the depletion region but allows minority carriers to move freely with the applied field.*

Fig. 4.6 *Lowering the barrier potential with a forward bias allows majority carriers to diffuse across the junction.*

thereby, allowed to diffuse across the junction. Once across, they significantly increase the minority carrier concentrations, and the excess carriers then recombine with the oppositely charged majority carriers. The recombination of excess minority carriers is the mechanism by which optical radiation is generated.

4.1.4 Direct and Indirect Bandgaps

In order for electron transitions to take place to or from the conduction band with the absorption or emission of a photon, respectively, both energy and momentum must be conserved. Although a photon can have considerable energy, its momentum hv/c is very small.

(a)

(b)
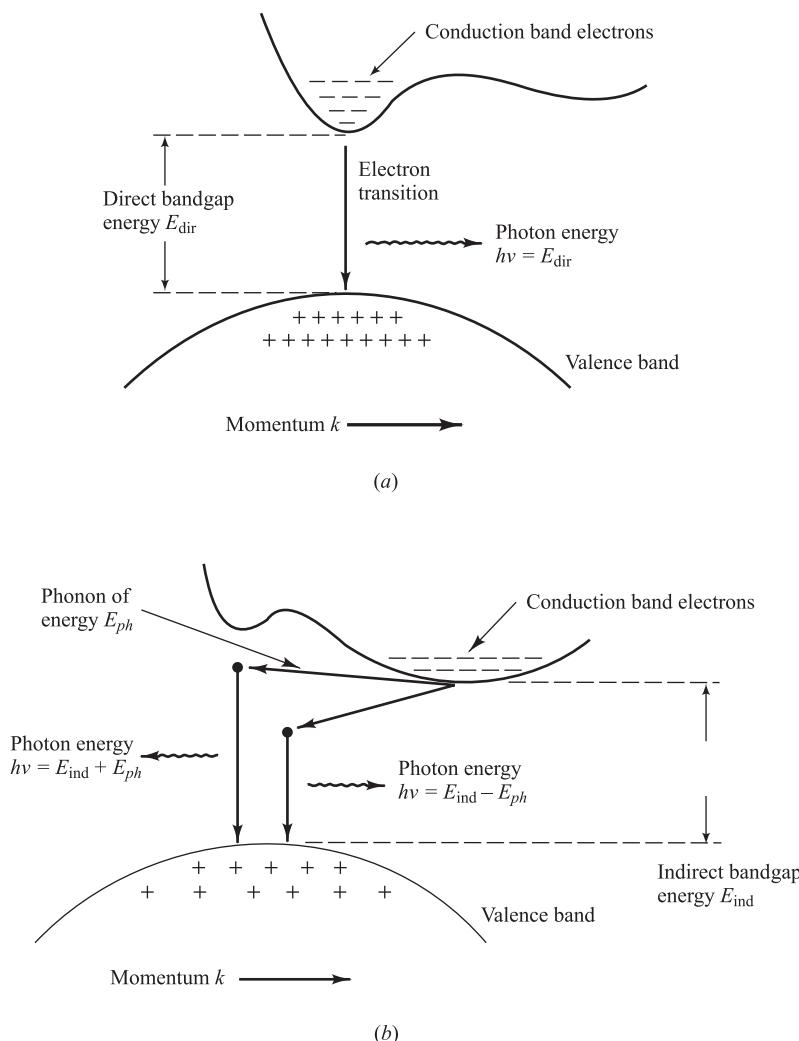
Fig. 4.7 *(a) Electron recombination and the associated photon emission for a direct-bandgap material; (b) electron recombination for indirect-bandgap materials requires a phonon of energy E_{ph} and momentum k_{ph}.*

Semiconductors are classified as either *direct-bandgap* or *indirect-bandgap* materials depending on the shape of the bandgap as a function of the momentum k, as shown in Fig. 4.7. Let us consider recombination of an electron and a hole, accompanied by the emission of a photon. The simplest and most probable recombination process will be that where the electron and hole have the same momentum value (see Fig. 4.7*a*). This is a direct-bandgap material.

For indirect-bandgap materials, the conduction-band minimum and the valence-band maximum energy levels occur at different values of momentum, as shown in Fig. 4.7*b*. Here, band-to-band recombination must involve a third particle to conserve momentum because the photon momentum is very small. *Phonons* (i.e., crystal lattice vibrations) serve this purpose.

4.1.5 Semiconductor Device Fabrication

In fabricating semiconductor devices, the crystal structure of the various material regions must be carefully taken into account. In any crystal structure, single atoms (e.g., Si or Ge) or groups of atoms (e.g., NaCl or GaAs) are arranged in a repeated pattern in space. This periodic arrangement defines a *lattice,* and the spacing between the atoms or groups of atoms is called the *lattice spacing* or the *lattice constant.* Typical lattice spacings are a few angstroms.

Semiconductor devices are generally fabricated by starting with a crystalline substrate that provides mechanical strength for mounting the device and for making electric contacts. A technique of crystal growth by chemical reaction is then used to grow thin layers of semiconductor materials on the substrate. These materials must have lattice structures that are identical to those of the substrate crystal. In particular, the lattice spacings of adjacent materials should be closely matched to avoid temperature-induced stresses and strains at the material interfaces. This type of growth is called *epitaxial,* which is derived from the Greek words *epi* meaning "on" and *taxis* meaning "arrangement"; that is, it is an arrangement of atoms from one material on another material. An important characteristic of epitaxial growth is that it is relatively simple to change the impurity concentration of successive material layers, so that a layered semiconductor device can be fabricated in a continuous process. Epitaxial layers can be formed by growth techniques of either vapor phase, liquid phase, or molecular beam.[16,17]

4.2 Light-Emitting Diodes (LEDs)

For optical communication systems requiring bit rates less than approximately 100–200 Mb/s together with multimode fiber-coupled optical power in the tens of microwatts, semiconductor light-emitting diodes (LEDs) are usually the best light source choice. These LEDs require less complex drive circuitry than laser diodes as no thermal or optical stabilization circuits are needed (see Sec. 4.3.6), and they can be fabricated less expensively with higher yields.

4.2.1 LED Structures

To be useful in fiber transmission applications, an LED must have a high radiance output, a fast emission response time, and a high quantum efficiency. Its *radiance* (or *brightness*) is a measure, in watts, of the optical power radiated into a unit solid angle per unit area of the emitting surface. High radiances are necessary to couple sufficiently high optical power levels into a fiber, as shown in detail in Chapter 5. The emission response time is the time delay between the application of a current pulse and the onset of optical emission. As we discuss in Secs. 4.2.4 and 4.3.7, this time delay is the factor limiting the bandwidth with which the source can be modulated directly by varying the injected current. The quantum efficiency

is related to the fraction of injected electron–hole pairs that recombine radiatively. This is defined and described in detail in Sec. 4.2.3.

To achieve a high radiance and a high quantum efficiency, the LED structure must provide a means of confining the charge carriers and the stimulated optical emission to the active region of the *pn* junction where radiative recombination takes place. Carrier confinement is used to achieve a high level of radiative recombination in the active region of the device, which yields a high quantum efficiency. Optical confinement is of importance for preventing absorption of the emitted radiation by the material surrounding the *pn* junction.

To achieve carrier and optical confinement, LED configurations such as homojunctions and single and double heterojunctions have been widely investigated.[19] The most effective of these structures is the configuration shown in Fig. 4.8. This is referred to as a *double-heterostructure* (or *heterojunction*) device because of the two different alloy layers on each side of the active region. This configuration evolved from studies on laser diodes. By means of this sandwich structure of differently composed
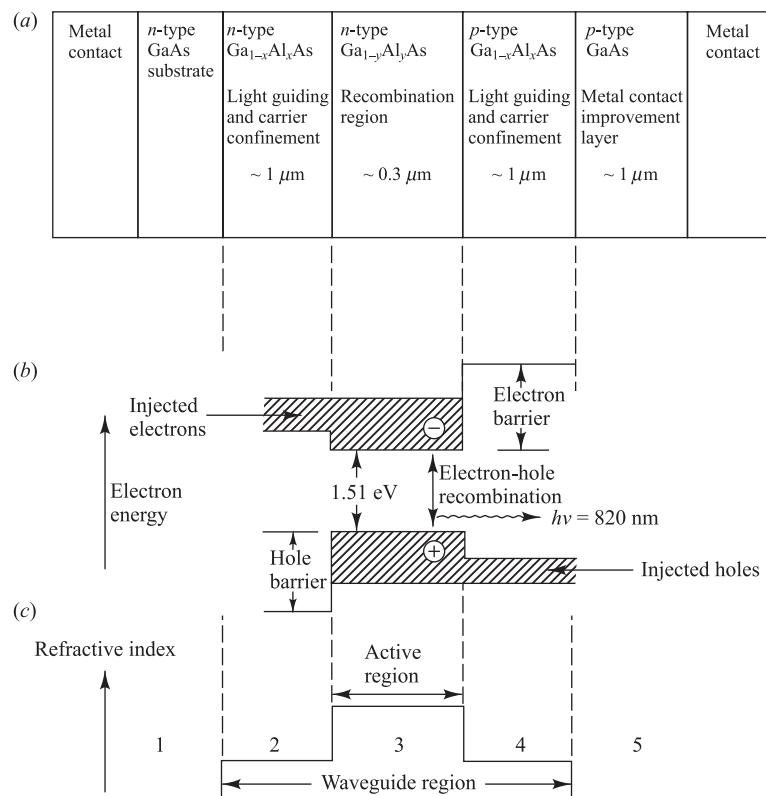
Fig. 4.8 *(a) Cross-sectional drawing (not to scale) of a typical GaAlAs double-heterostructure light emitter. In this structure, x > y to provide for both carrier confinement and optical guiding; (b) energy band diagram showing the active region, and the electron and hole barriers that confine the charge carriers to the active layer; (c) variations in the refractive index; the lower index of refraction of the material in regions 1 and 5 creates an optical barrier around the waveguide region.*
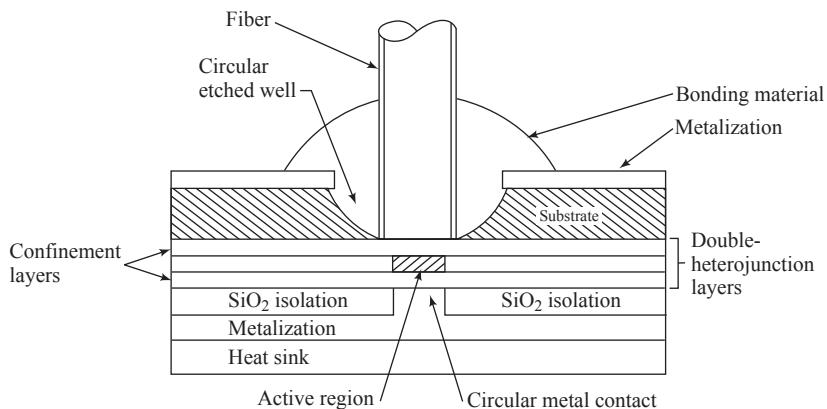
Fig. 4.9 *Schematic (not to scale) of a high-radiance surface-emitting LED. The active region is limited to a circular section having an area compatible with the fiber-core end face.*

alloy layers, both the carriers and the optical field are confined in the central active layer. The bandgap differences of adjacent layers confine the charge carriers (Fig. 4.8*b*), while the differences in the indices of refraction of adjoining layers confine the optical field to the central active layer (Fig. 4.8*c*). This dual confinement leads to both high efficiency and high radiance. Other parameters influencing the device performance include optical absorption in the active region (self-absorption), carrier recombination at the heterostructure interfaces, doping concentration of the active layer, injection carrier density, and active-layer thickness. We shall see the effects of these parameters in the following sections.

The two basic LED configurations being used for fiber optics are *surface emitters* (also called *Burrus* or *front emitters*) and *edge emitters*.[20] In the surface emitter, the plane of the active light-emitting region is oriented perpendicularly to the axis of the fiber, as shown in Fig. 4.9. In this configuration, a well is etched through the substrate of the device, into which a fiber is then cemented in order to accept the emitted light. The circular active area in practical surface emitters is nominally 50 μm in diameter and up to 2.5 μm thick. The emission pattern is essentially isotropic with a 120° half-power beam width.

This isotropic pattern from such a surface emitter is called a *lambertian pattern*. In this pattern, the source is equally bright when viewed from any direction, but the power diminishes as cos θ, where θ is the angle between the viewing direction and the normal to the surface (this is because the projected area one sees decreases as cos θ). Thus, the power is down to 50 percent of its peak when $\theta = 60°$, so that the total half-power beam width is 120°.

The edge emitter depicted in Fig. 4.10 consists of an active junction region, which is the source of the incoherent light, and two guiding layers. The guiding layers both have a refractive index lower than that of the active region but higher than the index of the surrounding material. This structure forms a waveguide channel that directs the optical radiation toward the fiber core. To match the typical fiber-core diameters (50–100 μm), the contact stripes for the edge emitter are 50–70 μm wide. Lengths of the active regions usually range from 100 to 150 μm. The emission pattern of the edge emitter is more directional than that of the surface emitter, as is illustrated in Fig. 4.10. In the plane parallel to the junction, where there is no waveguide effect, the emitted beam is lambertian (varying as cos θ) with a half-power width of $\theta_{||} = 120°$. In the plane perpendicular to the junction, the half-power beam $\theta_{\perp}$ has been made as small as 25–35° by a proper choice of the waveguide thickness.[2, 20]
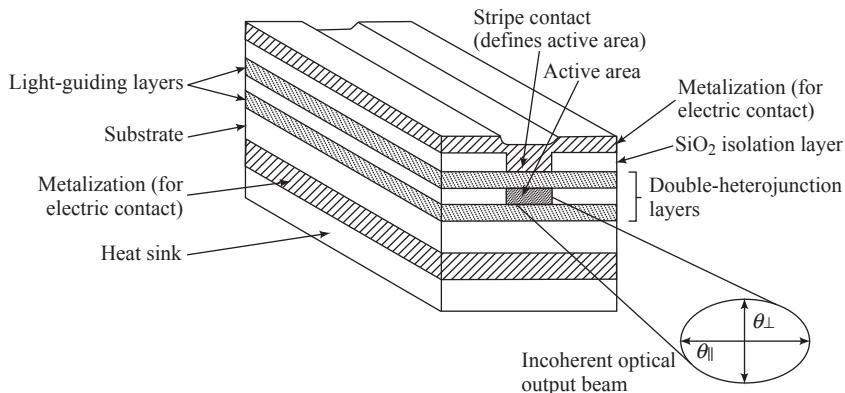
Fig. 4.10 *Schematic (not to scale) of an edge-emitting double-heterojunction LED. The output beam is lambertian in the plane of the pn junction ($\theta_{||}$ = 120°) and highly directional perpendicular to the pn junction ($\theta_\perp \approx 30°$).*

4.2.2 Light Source Materials

The semiconductor material that is used for the active layer of an optical source must have a direct bandgap. In a direct-bandgap semiconductor, electrons and holes can recombine directly across the bandgap without needing a third particle to conserve momentum. Only in direct-bandgap material is the radiative recombination sufficiently high to produce an adequate level of optical emission. Although none of the normal single-element semiconductors are direct-gap materials, many binary compounds are. The most important of these compounds are made from III-V materials. That is, the compounds consist of selections from a group III element (e.g., Al, Ga, or In) and a group V element (e.g., P, As, or Sb). Various ternary and quaternary combinations of binary compounds of these elements are also direct-gap materials and are suitable candidates for optical sources.

For operation in the 800-to-900-nm spectrum, the principal material used is the ternary alloy $Ga_{1-x}Al_xAs$. The ratio x of aluminum arsenide to gallium arsenide determines the bandgap of the alloy and, correspondingly, the wavelength of the peak emitted radiation. This is illustrated in Fig. 4.11. The value of x for the active-area material is usually chosen to give an emission wavelength of 800–850 nm. An example of the emission spectrum of a $Ga_{1-x}Al_xAs$ LED with $x = 0.08$ is shown in Fig. 4.12. The peak output power occurs at 810 nm. The width of the spectral pattern at its half-power point is known as the *full-width half-maximum* (FWHM) spectral width. As shown in Fig. 4.12, this FWHM spectral width σ_λ is 36 nm.

At longer wavelengths the quaternary alloy $In_{1-x}Ga_xAs_yP_{1-y}$ is one of the primary material candidates. By varying the mole fractions x and y in the active area, LEDs with peak output powers at any wavelength between 1.0 and 1.7 μm can be constructed. For simplicity, the notations GaAlAs and InGaAsP are generally used unless there is an explicit need to know the values of x and y. Other notations such as AlGaAs, (Al, Ga)As, (GaAl)As, GaInPAs, and $In_xGa_{1-x}As_yP_{1-y}$ are also found in the literature. From the last notation, it is obvious that, depending on the preference of the particular author, the values of x and $1 - x$ for the same material could be interchanged in different articles in the literature.

The alloys GaAlAs and InGaAsP are chosen to make semiconductor light sources because it is possible to match the lattice parameters of the heterostructure interfaces by using a proper combination of binary, ternary, and quaternary materials. A very close match between the crystal lattice parameters of the two adjoining heterojunctions is required to reduce interfacial defects and to minimize
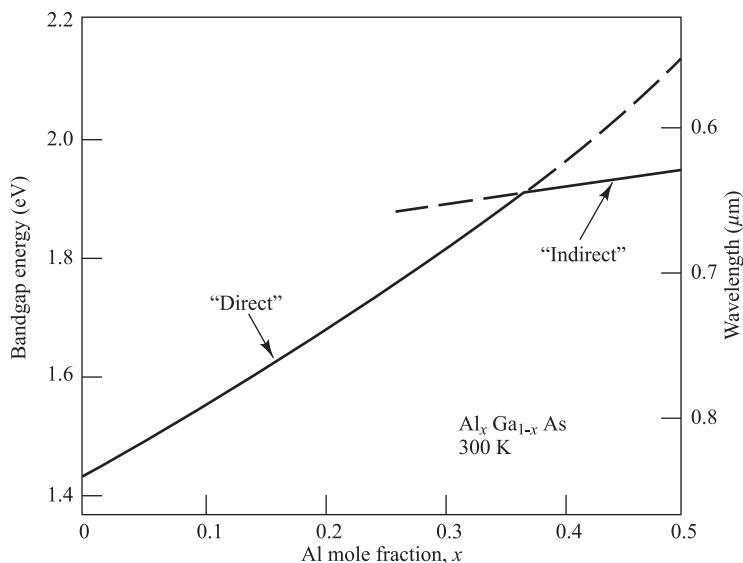
Fig. 4.11 *Bandgap energy and output wavelength as a function of aluminum mole fraction x for Al_xGa_{1-x} As at room temperature. (Reproduced with permission from Miller, Marcatili, and Lee,* Proc. IEEE, *vol. 61, pp. 1703–1751, Dec. 1973, ©1973, IEEE.)*
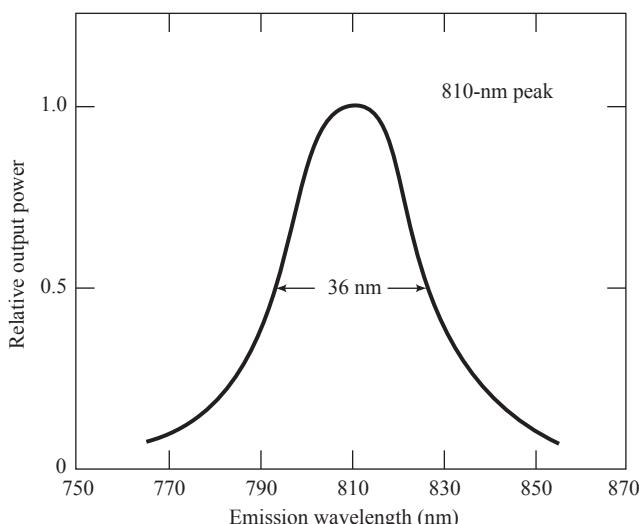
Fig. 4.12 *Spectral emission pattern of a representative $Ga_{1-x}Al_x$ As LED with x = 0.08. The width of the spectral pattern at its half-power point is 36 nm.*

strains in the device as the temperature varies. These factors directly affect the radiative efficiency and lifetime of a light source. Using the fundamental quantum-mechanical relationship between energy E and frequency ν,

$$E = h\nu = \frac{hc}{\lambda}$$

the peak emission wavelength λ in micrometers can be expressed as a function of the bandgap energy E_g in electron volts by the equation

$$\lambda(\mu m) = \frac{1.240}{E_g(eV)} \qquad (4.3)$$

The relationships between the bandgap energy E_g and the crystal lattice spacing (or lattice constant) a_0 for various III–V compounds are plotted in Fig. 4.13.

One can determine the bandgap of a semiconductor by measuring the energy

required to excite electrons from the valence band to the conduction band. Table 4.1 lists the bandgap energies of some common device materials used in various aspects of optical fiber communication technology.

A heterojunction with matching lattice parameters is created by choosing two material compositions that have the same lattice constant but different bandgap energies (the bandgap differences are used to confine the charge carriers). In the ternary alloy GaAlAs the bandgap energy E_g and the crystal lattice spacing a_0 are determined by the dashed line in Fig. 4.13 that connects the materials GaAs ($E_g = 1.43$ eV and $a_0 = 5.64$ Å) and AlAs ($E_g = 2.16$ eV and $a_0 = 5.66$ Å). The energy gap in electron volts for values of x between zero and 0.37 (the direct-bandgap region) can be found from the empirical equation[1,21]

$$E_g = 1.424 + 1.266x + 0.266x^2 \quad (4.4)$$

Given the value of E_g in electron volts, the peak emission wavelength in micrometers is found from Eq. (4.3).
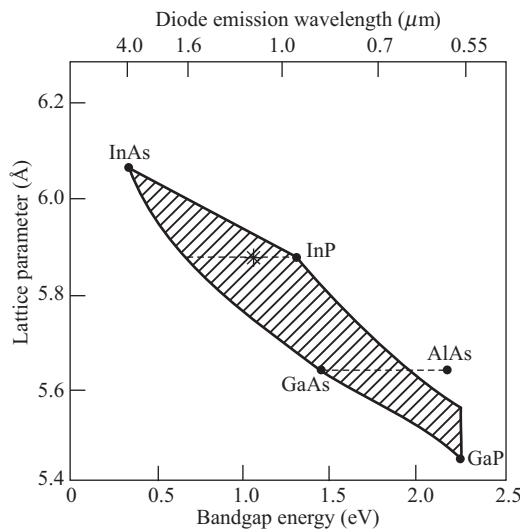
Fig. 4.13 *Relationships between the crystal lattice spacing, energy gap, and diode emission wavelength at room temperature. The shaded area is for the quaternary alloy InGaAsP. The asterisk (*) is for $In_{0.8}Ga_{0.2}As_{0.35}P_{0.65}$ (Eg ≈ 1.1 eV) lattice-matched to InP. (Used with permission from* Optical Fibre Communications by Tech. Staff of CSELT, © 1980, McGraw-Hill Book Co.)*

Table 4.1 *Bandgap energies of some common semiconductor materials*

Semiconductor material	Bandgap energy (eV)
Silicon (Si)	1.12
GaAs	1.43
Germanium (Ge)	0.67
InP	1.35
$Ga_{0.93}Al_{0.03}As$	1.51

Example 4.3 A particular $Ga_{1-x}Al_xAs$ laser is constructed with a material ratio $x = 0.07$. Find (*a*) the bandgap of this material; (*b*) the peak emission wavelength.

Solution: (*a*) From Eq. (4.4), we have $E_g = 1.424 + 1.266(0.07) + 0.66(0.07)2 = 1.51$ eV

(*b*) Using this value of the bandgap energy in Eq. (4.3) yields (in micrometers)

$$\lambda(\mu m) = 1.240/1.51 = 0.82 \ \mu m = 820 \text{ nm}$$

The bandgap energy and lattice-constant range for the quaternary alloy InGaAsP are much larger, as shown by the shaded area in Fig. 4.13. These materials are generally grown on an InP substrate, so that lattice-matched configurations are obtained by selecting a compositional point along the top dashed line in Fig. 4.13, which passes through the InP point. Along this line, the compositional parameters x and y follow the relationship $y \simeq 2.20x$ with $0 \le x \le 0.47$. For $In_{1-x}Ga_xAs_yP_{1-y}$ compositions that are lattice-matched to InP, the bandgap in eV varies as

$$E_g = 1.35 - 0.72y + 0.12y^2 \tag{4.5}$$

Bandgap wavelengths from 0.92 to 1.65 μm are covered by this material system.

Example 4.4 Consider the material alloy $In_{0.74}Ga_{0.26}As_{0.57}P_{0.43}$, that is, $x = 0.26$ and $y = 0.57$ in the general formula $In_{1-x}Ga_xAs_yP_{1-y}$. Find (*a*) the bandgap of this material; (*b*) the peak emission wavelength.

Solution: (*a*) From Eq. (4.5), we have $E_g = 1.35 - 0.72(0.57) + 0.12(0.57)^2 = 0.97$ eV

(*b*) Using this value of the bandgap energy in Eq. (4.3) yields (in micrometers)

$$\lambda(\mu m) = 1.240/0.97 = 1.27 \ \mu m = 1270 \text{ nm}$$

Whereas the FWHM power spectral widths of LEDs in the 800-nm region are around 35 nm, this increases in longer-wavelength materials. For devices operating in the 1300-to-1600-nm region, the spectral widths vary from around 70 to 180 nm. Figure 4.14 shows an example for devices emitting at 1300 nm. In addition, as Fig. 4.14 shows, the output spectral widths of surface-emitting LEDs tend to be broader than those of edge-emitting LEDs because of different internal-absorption effects of the emitted light in the two device structures.

Table 4.2 lists typical characteristics of surface-emitting LEDs (SLED) and edge-emitting LEDS (ELED). The materials used in these examples are GaAlAs for operation at 850 nm and InGaAsP for 1310-nm devices. The fiber-coupled power listing is the amount of light that can be accepted by a 50-μm core diameter multimode fiber.
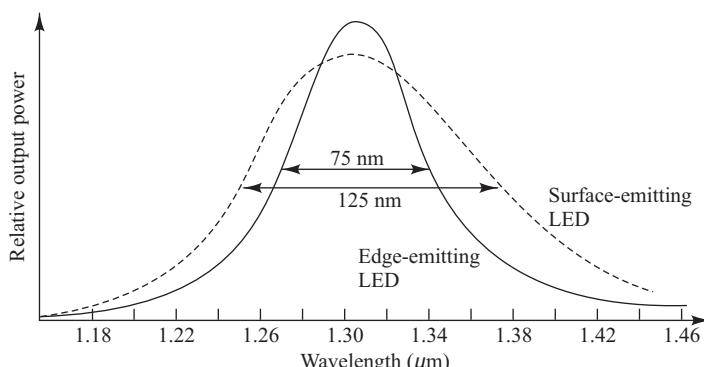
Fig. 4.14 *Typical spectral patterns for edge-emitting and surface-emitting LEDs at 1310 nm. The patterns broaden with increasing wavelength and are wider for surface emitters*

Table 4.2 *Typical characteristics of surface- and edge-emitting LEDs*

LED type	Material	Wavelength (nm)	Operating current (mA)	Fiber-coupled power (μW)	Nominal FWHM (nm)
SLED	GaAlAs	850	110	40	35
ELED	InGaAsP	1310	100	15	80
SLED	InGaAsP	1310	110	30	150

4.2.3   Quantum Efficiency and LED Power

An excess of electrons and holes in *p*- and *n*-type material, respectively (referred to as *minority carriers*) is created in a semiconductor light source by carrier injection at the device contacts. The excess densities of electrons *n* and holes *p* are equal, since the injected carriers are formed and recombine in pairs in accordance with the requirement for charge neutrality in the crystal. When carrier injection stops, the carrier density returns to the equilibrium value. In general, the excess carrier density decays exponentially with time according to the relation

$$n = n_0 e^{-t/\tau} \tag{4.6}$$

where n_0 is the initial injected excess electron density and the time constant τ is the carrier lifetime. This lifetime is one of the most important operating parameters of an electro-optic device. Its value can range from milliseconds to fractions of a nanosecond depending on material composition and device defects.

The excess carriers can recombine either radiatively or nonradiatively. In radiative recombination a photon of energy $h\nu$, which is approximately equal to the bandgap energy, is emitted. Nonradiative recombination effects include optical absorption in the active region (self-absorption), carrier recombination at the heterostructure interfaces, and the Auger process in which the energy released during an electron–hole recombination is transferred to another carrier in the form of kinetic energy.

When there is a constant current flow into an LED, an equilibrium condition is established. That is, the excess density of electrons *n* and holes *p* is equal since the injected carriers are created and recombined in pairs such that charge neutrality is maintained within the device. The total rate at which carriers are generated is the sum of the externally supplied and the thermally generated rates. The externally supplied rate is given by J/qd, where J is the current density in A/cm^2, q is the electron charge, and d is the thickness of the recombination region. The thermal generation rate is given by n/τ. Hence, the rate equation for carrier recombination in an LED can be written as

$$\frac{dn}{dt} = \frac{J}{qd} - \frac{n}{\tau} \tag{4.7}$$

The equilibrium condition is found by setting Eq. (4.7) equal to zero, yielding

$$n = \frac{J\tau}{qd} \tag{4.8}$$

This relationship gives the steady-state electron density in the active region when a constant current is flowing through it.

The *internal quantum efficiency* in the active region is the fraction of the electron–hole pairs that recombine radiatively. If the radiative recombination rate is R_r and the nonradiative recombination rate is R_{nr}, then the internal quantum efficiency η_{int} is the ratio of the radiative recombination rate to the total recombination rate:

$$\eta_{int} = \frac{R_r}{R_r + R_{nr}} \tag{4.9}$$

For exponential decay of excess carriers, the radiative recombination lifetime is $\tau_r = n/R_r$ and the nonradiative recombination lifetime is $\tau_{nr} = n/R_{nr}$. Thus the internal quantum efficiency can be expressed as

$$\eta_{int} = \frac{1}{1 + \tau_r / \tau_{nr}} = \frac{\tau}{\tau_r} \tag{4.10}$$

where the *bulk recombination lifetime τ* is

$$\frac{1}{\tau} = \frac{1}{\tau_r} + \frac{1}{\tau_{nr}} \tag{4.11}$$

In general, τ_r and τ_{nr} are comparable for direct-bandgap semiconductors, such as GaAlAs and InGaAsP. This also means that R_r and R_{nr} are similar in magnitude, so that the internal quantum efficiency is about 50 percent for simple homojunction LEDs. However, LEDs having double-heterojunction structures can have quantum efficiencies of 60–80 percent. This high efficiency is achieved because the thin active regions of these devices mitigate the self-absorption effects, which reduces the nonradiative recombination rate.

If the current injected into the LED is I, then the total number of recombinations per second is

$$R_r + R_{nr} = I/q \tag{4.12}$$

Substituting Eq. (4.12) into Eq. (4.9) then yields $R_r = \eta_{int} I/q$. Noting that R_r is the total number of photons generated per second and that each photon has an energy $h\nu$, then the optical power generated internally to the LED is

$$P_{int} = \eta_{int} \frac{I}{q} h\nu = \eta_{int} \frac{hcI}{q\lambda} \tag{4.13}$$

Example 4.5 A double-heterojunction InGaAsP LED emitting at a peak wavelength of 1310 nm has radiative and nonradiative recombination times of 30 and 100 ns, respectively. The drive current is 40 mA. Find (*a*) the bulk recombination time; (*b*) the internal quantum efficiency; and (*c*) the internal power level.

Solution: (*a*) From Eq. (4.11), the bulk recombination lifetime is

$$\tau = \frac{\tau_r \tau_{nr}}{\tau_r + \tau_{nr}} = \frac{30 \times 100}{30 + 100} \, ns = 23.1 \, ns$$

(*b*) Using Eq. (4.10), the internal quantum efficiency is

$$\eta_{int} = \frac{\tau}{\tau_r} = \frac{23.1}{30} = 0.77$$

(*c*) Substituting this into Eq. (4.13) yields an internal power level of

$$P_{int} = \eta_{int} \frac{hcI}{q\lambda}$$

$$= 0.77 \frac{(6.6256 \times 10^{-34} \, J \cdot s)(3 \times 10^8 m/s)(0.040A)}{(1.602 \times 10^{-19} \, C)(1.31 \times 10^{-6} \, m)}$$

$$= 29.2 \, mW$$
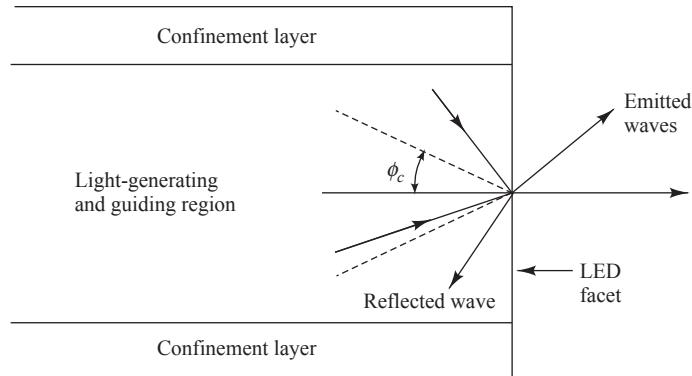
Fig. 4.15 *Only light falling within a cone defined by the critical angle ϕ_c will be emitted from an optical source.*

Not all internally generated photons will exit the device. To find the emitted power, one needs to consider the *external quantum efficiency* η_{ext}. This is defined as the ratio of the photons emitted from the LED to the number of internally generated photons. To find the external quantum efficiency, we need to take into account reflection effects at the surface of the LED. As shown in Fig. 4.15 and described in Sec. 2.2, at the interface of a material boundary only that fraction of light falling within a cone defined by the critical angle ϕ_c will cross the interface. From Eq. (2.18), we have that $\phi_c = \sin^{-1}(n_2/n_1)$. Here, n_1 is the refractive index of the semiconductor material and n_2 is the refractive index of the outside material, which nominally is air with $n_2 = 1.0$. The external quantum efficiency can then be calculated from the expression

$$\eta_{ext} = \frac{1}{4\pi} \int_0^{\phi_c} T(\phi)(2\pi \sin \phi)\, d\phi \qquad (4.14)$$

where $T(\phi)$ is the *Fresnel transmission coefficient* or *Fresnel transmissivity*. This factor depends on the incidence angle ϕ, but, for simplicity, we can use the expression for normal incidence, which is[18,22]

$$T(0) = \frac{4 n_1 n_2}{(n_1 + n_2)^2} \qquad (4.15)$$

Assuming the outside medium is air and letting $n_1 = n$, we have $T(0) = 4n/(n+1)^2$. The external quantum efficiency is then approximately given by

$$\eta_{ext} \approx \frac{1}{n(n+1)^2} \qquad (4.16)$$

From this, it follows that the optical power emitted from the LED is

$$P = \eta_{ext} P_{int} \approx \frac{P_{int}}{n(n+1)^2} \qquad (4.17)$$

Example 4.6   Assume a typical value of $n = 3.5$ for the refractive index of an LED material. What percent of the internally generated optical power is emitted into an air medium?

Solution: Taking the condition for normal incidence, then from Eq. (4.16) the percent of the optical power that is generated internally in the device that is emitted into

an air medium is

$$\eta_{ext} = \frac{1}{n(n+1)^2} = \frac{1}{3.5(3.5+1)^2} = 1.41 \text{ percent}$$

This shows that only a small fraction of the internally generated optical power is emitted from the device.[23]

4.2.4   Modulation of an LED

The *response time* or *frequency response* of an optical source dictates how fast an electrical input drive signal can vary the light output level. The following three factors largely determine the response time: the doping level in the active region, the injected carrier lifetime τ_i in the recombination region, and the parasitic capacitance of the LED. If the drive current is modulated at a frequency ω, the optical output power of the device will vary as[24,25]

$$P(\omega) = P_0[1 + (\omega\tau_i)^2]^{-1/2} \tag{4.18}$$

where P_0 is the power emitted at zero modulation frequency. The parasitic capacitance can cause a delay of the carrier injection into the active junction, and, consequently, could delay the optical output.[26,27] This delay is negligible if a small, constant forward bias is applied to the diode. Under this condition, Eq. (4.18) is valid and the modulation response is limited only by the carrier recombination time.

Example 4.7   A particular LED has a 5-ns injected carrier lifetime. When no modulation current is applied to the device, the optical output power is 0.250 mW for a specified dc bias. Assuming parasitic capacitances are negligible, what are the optical outputs at modulation frequencies of (*a*) 10 MHz and (*b*) 100 MHz?

Solution: (*a*) From Eq. (4.18) the optical output at 10 MHz is

$$P(\omega) = \frac{0.250}{\sqrt{1 + [2\pi(10 \times 10^6)(5 \times 10^{-9})]^2}}$$
$$= 0.239 \text{ mW} = 239 \ \mu\text{W}$$

(*b*) Similarly, the optical output at 10 MHz is

$$P(\omega) = \frac{P_0}{\sqrt{1 + (\omega\tau_i)^2}}$$
$$= \frac{0.250}{\sqrt{1 + [2\pi(100 \times 10^6)(5 \times 10^{-9})]^2}}$$
$$= 0.076 \text{ mW} = 76 \ \mu\text{W}$$

Thus the output of this particular device decreases at higher modulation rates.

The modulation bandwidth of an LED can be defined in either electrical or optical terms. Normally, electrical terms are used because the bandwidth is actually determined via the associated electrical circuitry. Thus the modulation bandwidth is defined as the point where the electrical signal power, designated by $p(\omega)$, has dropped to half its constant value resulting from the modulated portion of the optical signal. This is the electrical 3-dB point; that is, the frequency at which the output electrical power is reduced by 3 dB with respect to the input electrical power, as is illustrated in Fig. 4.16.
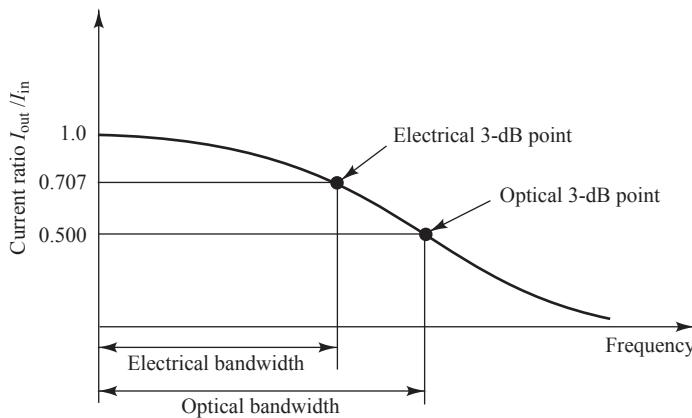
Fig. 4.16 *Frequency response of an optical source showing the electrical and optical 3-dB-bandwidth points*

An optical source exhibits a linear relationship between light power and current, so currents rather than voltages (which are used in electrical systems) are compared in optical systems. Thus, since $p(\omega) = I^2(\omega)/R$, the ratio of the output electrical power at the frequency ω to the power at zero modulation is

$$\text{Ratio}_{\text{elec}} = 10 \log \left[\frac{p(\omega)}{p(0)} \right] = 10 \log \left[\frac{I^2(\omega)}{I^2(0)} \right] \tag{4.19}$$

where $I(\omega)$ is the electrical current in the detection circuitry. The electrical 3-dB point occurs at that frequency point where the detected electrical power $p(\omega) = p(0)/2$. This happens when

$$\frac{I^2(\omega)}{I^2(0)} = \frac{1}{2} \tag{4.20}$$

or $I(\omega)/I(0) = 1/\sqrt{2} = 0.707$

Example 4.8 Consider the particular LED described in Example 4.7, which has a 5-ns injected carrier lifetime. (*a*) What is the 3-dB optical bandwidth of this device? (*b*) What is the 3-dB electrical bandwidth of this device?

Solution: (*a*) The 3-dB optical bandwidth occurs at the modulation frequency for which $P(\omega) = 0.5P_0$. Using Eq. (4.18) we find

$$\frac{1}{[1 + (\omega\tau_i)^2]^{1/2}} = \frac{1}{2}$$

so that $1 + (\omega\tau_i)^2 = 4$, or $\omega\tau_i = \sqrt{3}$. Solving this expression for the frequency $f = 2\pi\omega$, we find

$$f = \frac{\sqrt{3}}{2\pi\tau} = \frac{\sqrt{3}}{2\pi \times 5 \times 10^{-9}} = 55.1 \text{ MHz}$$

(*b*) The 3-dB electrical bandwidth is $f/\sqrt{2} = 0.707$ (55.1 MHz) = 39.0 MHz

Sometimes, the modulation bandwidth of an LED is given in terms of the 3-dB bandwidth of the modulated optical power $P(\omega)$; that is, it is specified at the frequency where $P(\omega) = P_0/2$. In this case, the 3-dB bandwidth is determined from the ratio of the optical power at frequency ω to the unmodulated value of the optical power. Since the detected current is directly proportional to the optical power, this ratio is

$$\text{Ratio}_{\text{optical}} = 10 \log \left[\frac{P(\omega)}{P(0)} \right] = 10 \log \left[\frac{I(\omega)}{I(0)} \right] \qquad (4.21)$$

The optical 3-dB point occurs at that frequency where the ratio of the currents is equal to 1/2. As shown in Fig. 4.16, this gives an inflated value of the modulation bandwidth, which corresponds to an electrical power attenuation of 6 dB.

4.3 | Laser Diodes

Lasers come in many forms with dimensions ranging from the size of a grain of salt to one that will occupy an entire room. The lasing medium can be a gas, a liquid, an insulating crystal (solid state), or a semiconductor. For optical fiber systems the laser sources used almost exclusively are semiconductor laser diodes. They are similar to other lasers, such as the conventional solid-state and gas lasers, in that the emitted radiation has spatial and temporal coherence; that is, the output radiation is highly monochromatic and the light beam is very directional.

Despite their differences, the basic principle of operation is the same for each type of laser. Laser action is the result of three key processes: photon absorption, spontaneous emission, and stimulated emission. These three processes are represented by the simple two-energy-level diagrams in Fig. 4.17, where E_1 is the ground-state energy and E_2 is the excited-state energy. According to Planck's law, a transition between these two states involves the absorption or emission of a photon of energy $h\nu_{12} = E_2 - E_1$. Normally, the system is in the ground state. When a photon of energy $h\nu_{12}$ impinges on the system, an electron in state E_1 can absorb the photon energy and be excited to state E_2, as shown in Fig. 4.17a. Since this is an unstable state, the electron will shortly return to the ground state, thereby emitting a photon of energy $h\nu_{12}$. This occurs without any external stimulation and is called *spontaneous emission*. These emissions are isotropic and of random phase, and thus appear as a narrowband gaussian output.

The electron can also be induced to make a downward transition from the excited level to the ground-state level by an external stimulation. As shown in Fig. 4.17c, if a photon of energy $h\nu_{12}$ impinges on

(a) Absorption   (b) Spontaneous emission   (c) Stimulated emission
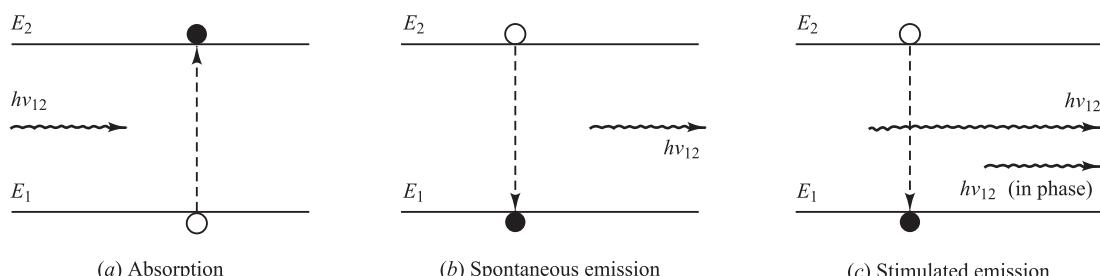
Fig. 4.17 *The three key transition processes involved in laser action. The open circle represents the initial state of the electron and the heavy dot represents the final state; incident photons are shown on the left of each diagram and emitted photons are shown on the right.*

the system while the electron is still in its excited state, the electron is immediately stimulated to drop to the ground state and give off a photon of energy $h\nu_{12}$. This emitted photon is in phase with the incident photon, and the resultant emission is known as *stimulated emission*.

In thermal equilibrium the density of excited electrons is very small. Most photons incident on the system will therefore be absorbed, so that stimulated emission is essentially negligible. Stimulated emission will exceed absorption only if the population of the excited states is greater than that of the ground state. This condition is known as *population inversion*. Since this is not an equilibrium condition, population inversion is achieved by various "pumping" techniques. In a semiconductor laser, population inversion is accomplished by injecting electrons into the material at the device contacts or through an optical absorption method by means of externally injected photons.

4.3.1 Laser Diode Modes and Threshold Conditions

For optical fiber communication systems requiring bandwidths greater than approximately 200 MHz, the semiconductor injection laser diode is preferred over the LED. Laser diodes typically have response times less than 1 ns, can have spectral widths of 2 nm or less, and are capable of coupling from tens to hundreds of milliwatts of useful luminescent power into optical fibers with small cores and small mode-field diameters. Virtually all laser diodes in use are multilayered heterojunction devices. As mentioned in Sec. 4.2, the double-heterojunction LED configuration evolved from the successful demonstration of both carrier and optical confinement in heterojunction injection laser diodes. The more rapid evolvement and utilization of LEDs as compared with laser diodes lies in the inherently simpler construction, the smaller temperature dependence of the emitted optical power, and the absence of catastrophic degradation in LEDs (see Sec. 4.6). The construction of laser diodes is more complicated, mainly because of the additional requirement of current confinement in a small lasing cavity.

Stimulated emission in semiconductor lasers arises from optical transitions between distributions of energy states in the valence and conduction bands. This differs from gas and solid-state lasers, in which radiative transitions occur between discrete isolated atomic or molecular levels. The radiation in one type of laser diode configuration is generated within a Fabry-Perot resonator cavity,[16–18] shown in Fig. 4.18, as in most other types of lasers. Here the cavity is approximately 250–500 μm long, 5–15 μm wide, and 0.1–0.2 μm thick. These dimensions commonly are referred to as the *longitudinal*, *lateral*, and *transverse dimensions* of the cavity, respectively.

As illustrated in Fig. 4.19, two flat, partially reflecting mirrors are directed toward each other to enclose the Fabry-Perot resonator cavity. The mirror facets are constructed by making two parallel clefts along natural cleavage planes of the semiconductor crystal. The purpose of the mirrors is to establish a strong optical feedback in the longitudinal direction. This feedback mechanism converts the device into an oscillator (and hence a light emitter) with a gain mechanism that compensates for optical losses in the cavity at certain resonant optical frequencies. The sides of the cavity are simply formed by roughing the edges of the device to reduce unwanted emissions in the lateral directions.

As the light reflects back and forth within the Fabry-Perot cavity, the electric fields of the light interfere on successive round trips. Those wavelengths that are integer multiples of the cavity length interfere constructively so that their amplitudes add when they exit the device through the right-hand facet. All other wavelengths interfere destructively and thus cancel themselves out. The optical frequencies at which constructive interference occurs are the *resonant frequencies* of the cavity. Consequently, spontaneously emitted photons that have wavelengths at these resonant frequencies reinforce themselves after multiple trips through the cavity so that their optical field becomes very strong. The resonant wavelengths are called the *longitudinal modes* of the cavity because they resonate along the length of the cavity.
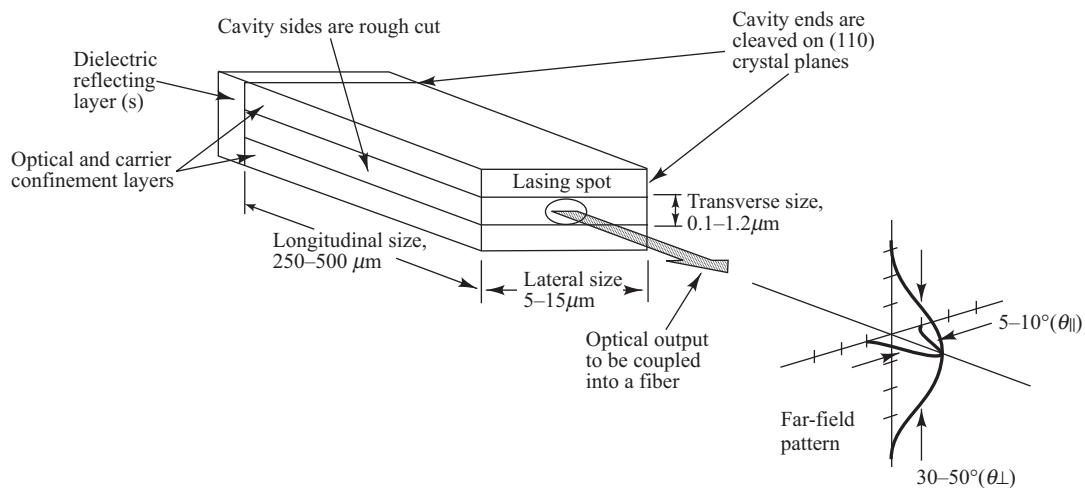
Fig. 4.18   *Fabry-Perot resonator cavity for a laser diode. The cleaved crystal ends function as partially reflecting mirrors. The unused end (the rear facet) can be coated with a dielectric reflector to reduce optical loss in the cavity. Note that the light beam emerging from the laser forms a vertical ellipse, even though the lasing spot at the active-area facet is a horizontal ellipse.*

Figure 4.20 illustrates the behavior of the resonant wavelengths for three values of the mirror reflectivity. The plots give the relative intensity as a function of the wavelength relative to the cavity length. As can be seen from Fig. 4.20, the width of the resonances depends on the value of the reflectivity. That is, the resonances become sharper as the reflectivity increases. Chapter 10 provides further details on the operational theory of Fabry-Perot cavities or etalons.

In another laser diode type, commonly referred to as the *distributed-feedback* (DFB) laser,[1,28,29] the cleaved facets are not required for optical feedback. A typical DFB laser configuration is shown in Fig. 4.21. The fabrication of this device is similar to the Fabry-Perot types, except that the lasing action is obtained from Bragg reflectors (gratings) or periodic variations of the refractive index (called *distributed-feedback corrugations*), which are incorporated into the multilayer structure along the length of the diode. This is discussed in more detail in Sec. 4.3.6.

In general, the full optical output is needed only from the front facet of the laser—that is, the one to be aligned with an optical fiber. In this case, a dielectric reflector can be deposited on the rear laser facet to reduce the optical loss in the cavity, to reduce the threshold current density (the point at which lasing

Fig. 4.19   *Two parallel light-reflecting mirrored surfaces define a Fabry-Perot resonator cavity.*
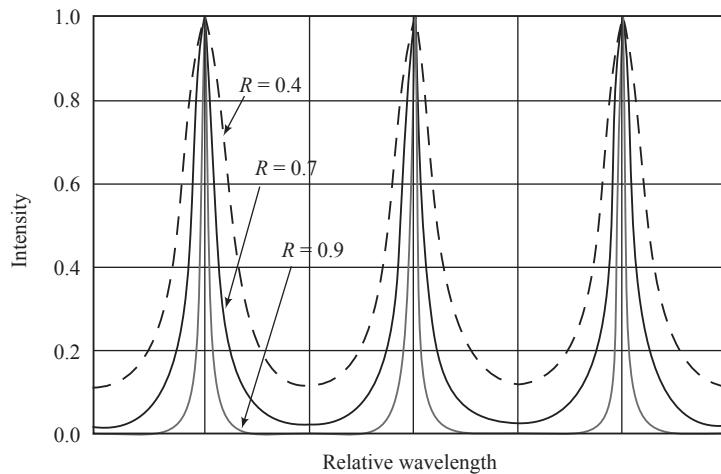
Fig. 4.20 *Behavior of the resonant wavelengths in a Fabry-Perot cavity for three values of the mirror reflectivity*

Example 4.9 As Sec. 10.5 describes, the distance between the adjacent peaks of the resonant wavelengths in a Fabry-Perot cavity, shown in Fig. (4.20), is called the *free spectral range* (FSR). If D is the distance between the reflecting mirrors in a device of refractive index n, then at a peak wavelength λ the FSR is given by the expression

$$FSR = \frac{\lambda^2}{2nD}$$

What is the FSR at an 850-nm wavelength for a 0.8-mm long GaAs Fabry-Perot cavity in which the refractive index is 3.5?

Solution: From the above expression we have

$$FSR = \frac{\lambda^2}{2nD} = \frac{(0.85 \times 10^{-6})^2}{2(3.5)(0.80 \times 10^{-3})} = 0.129 \text{ nm}$$
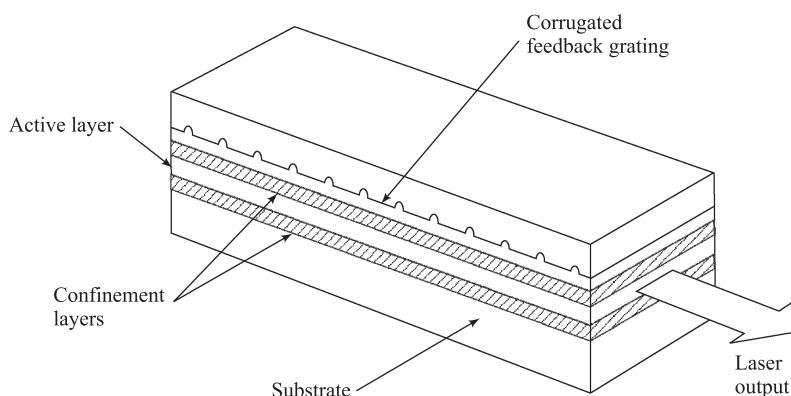
Fig. 4.21 *Structure of a distributed-feedback (DFB) laser diode*

starts), and to increase the external quantum efficiency. Reflectivities greater than 98 percent have been achieved with a six-layer reflector.

The optical radiation within the resonance cavity of a laser diode sets up a pattern of electric and magnetic field lines called the *modes of the cavity* (see Secs. 2.3 and 2.4 for details on modes). These can conveniently be separated into two independent sets of transverse electric (TE) and transverse magnetic (TM) modes. Each set of modes can be described in terms of the longitudinal, lateral, and transverse half-sinusoidal variations of the electromagnetic fields along the major axes of the cavity. The *longitudinal modes* are related to the length L of the cavity and determine the principal structure of the frequency spectrum of the emitted optical radiation. Since L is much larger than the lasing wavelength of approximately 1 μm, many longitudinal modes can exist. *Lateral modes* lie in the plane of the *pn* junction. These modes depend on the side wall preparation and the width of the cavity, and determine the shape of the lateral profile of the laser beam. *Transverse modes* are associated with the electromagnetic field and beam profile in the direction perpendicular to the plane of the *pn* junction. These modes are of great importance as they largely determine such laser characteristics as the radiation pattern (the transverse angular distribution of the optical output power) and the threshold current density.

To determine the lasing conditions and the resonant frequencies, we express the electromagnetic wave propagating in the longitudinal direction (along the axis normal to the mirrors) in terms of the electric field phasor

$$E(z, t) = I(z)\, e^{j(\omega t - \beta z)} \tag{4.22}$$

where $I(z)$ is the optical field intensity, ω is the optical radian frequency, and β is the propagation constant (see Sec. 2.3.2).

Lasing is the condition at which light amplification becomes possible in the laser diode. The requirement for lasing is that a population inversion be achieved. This condition can be understood by considering the fundamental relationship between the optical field intensity I, the absorption coefficient α_λ, and the gain coefficient g in the Fabry-Perot cavity. The stimulated emission rate into a given mode is proportional to the intensity of the radiation in that mode. The radiation intensity at a photon energy $h\nu$ varies exponentially with the distance z that it traverses along the lasing cavity according to the relationship

$$I(z) = I(0) \exp\left\{ [\Gamma g\,(h\nu) - \bar{\alpha}\,(h\nu)]z \right\} \tag{4.23}$$

where $\bar{\alpha}$ is the effective absorption coefficient of the material in the optical path and Γ is the *optical-field confinement factor*—that is, the fraction of optical power in the active layer (see Prob. 4.11 concerning details of transverse and lateral optical-field confinement factors).

Optical amplification of selected modes is provided by the feedback mechanism of the optical cavity. In the repeated passes between the two partially reflecting parallel mirrors, a portion of the radiation associated with those modes that have the highest optical gain coefficient is retained and further amplified during each trip through the cavity.

Lasing occurs when the gain of one or several guided modes is sufficient to exceed the optical loss during one roundtrip through the cavity; that is, $z = 2L$. During this roundtrip, only the fractions R_1 and R_2 of the optical radiation are reflected from the two laser ends 1 and 2, respectively, where R_1 and R_2 are the mirror reflectivities or Fresnel reflection coefficients, which are given by

$$R = \left(\frac{n_1 - n_2}{n_1 + n_2} \right)^2 \tag{4.24}$$

Example 4.10 Assume that the cleaved mirror end faces of a GaAs laser are uncoated and that the outside medium is air. What is the reflectivity for normal incidence of a plane wave on the GaAs-air interface if the GaAs refractive index is 3.6?

Solution: From Eq. (4.24), with $n_1 = 3.6$ for GaAs and $n_2 = 1.0$ for air, we have for both interfaces

$$R_1 = R_2 = \left(\frac{3.6 - 1}{3.6 + 1}\right)^2 = 0.32$$

for the reflection of light at an interface between two materials having refractive indices n_1 and n_2. From this lasing condition, Eq. (4.23) becomes

$$I(2L) = I(0) \, R_1 R_2 \, \exp\{2L\,[\Gamma g(h\nu) - \bar{\alpha}(h\nu)]\} \tag{4.25}$$

For an uncoated cleaved facet the reflectivity is only about 30 percent. To reduce the loss in the cavity and to make the optical feedback stronger, the facets typically are coated with a dielectric material. This can produce a reflectivity of about 99 percent for the rear facet and 90 percent for the front facet through which the lasing light emerges.

At the lasing threshold, a steady-state oscillation takes place, and the magnitude and phase of the returned wave must be equal to those of the original wave. This gives the conditions

$$I(2L) = I(0) \tag{4.26}$$

for the amplitude and

$$e^{-j2\beta L} = 1 \tag{4.27}$$

for the phase. Equation (4.27) gives information concerning the resonant frequencies of the Fabry-Perot cavity. This is discussed further in Sec. 4.3.2. From Eq. (4.26) we can find which modes have sufficient gain for sustained oscillation, and we can find the amplitudes of these modes. The condition to just reach the lasing threshold is the point at which the optical gain is equal to the total loss α_t, in the cavity. From Eq. (4.26), this condition is

$$g_{th} = \alpha_t = \bar{\alpha} + \frac{1}{2L}\ln\left(\frac{1}{R_1 R_2}\right) = \bar{\alpha} + \alpha_{end} \tag{4.28}$$

where α_{end} is the mirror loss in the lasing cavity. Thus, for lasing to occur, we must have the gain $g \geq g_{th}$. This means that the pumping source that maintains the population inversion must be sufficiently strong to support or exceed all the energy-consuming mechanisms within the lasing cavity.

The mode that satisfies Eq. (4.28) reaches threshold first. Theoretically, at the onset of this condition, all additional energy introduced into the laser should augment the growth of this particular mode. In practice, various phenomena lead to the excitation of more than one mode.[1] Studies on the conditions

Example 4.11 Assume for GaAs that $R_1 = R_2 = R = 0.32$ for uncoated facets (i.e., 32 percent of the radiation is reflected at a facet) and $\bar{\alpha} \approx 10$ cm^{-1}. What is the gain threshold for a 500-μm long laser diode?

Solution: From Eq. (4.28) we have

$$g_{th} = \bar{\alpha} + \frac{1}{2L}\ln\left(\frac{1}{R^2}\right)$$

$$= 10 + \frac{1}{2(500 \times 10^{-4})}\ln\left[\frac{1}{(0.32)^2}\right] = 33 \text{ cm}^{-1}$$
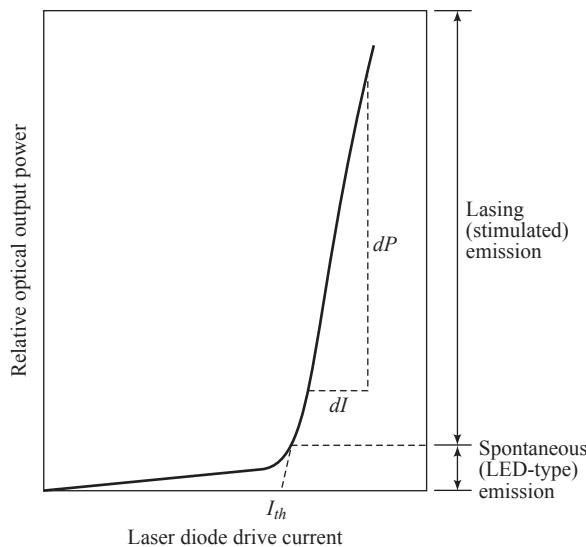
Fig. 4.22 *Relationship between optical output power and laser diode drive current. Below the lasing threshold the optical output is a spontaneous LED-type emission.*

of longitudinal single-mode operation have showed that important factors are thin active regions and a high degree of temperature stability.

The relationship between optical output power and diode drive current is presented in Fig. 4.22. At low diode currents, only spontaneous radiation is emitted. Both the spectral range and the lateral beam width of this emission are broad like that of an LED. A dramatic and sharply defined increase in the power output occurs at the lasing threshold. As this transition point is approached, the spectral range and the beam width both narrow with increasing drive current. The final spectral width of approximately 1 nm and the fully narrowed lateral beam width of nominally 5–10° are reached just past the threshold point. The *threshold current* I_{th} is conventionally defined by extrapolation of the lasing region of the power-versus-current curve, as shown in Fig. 4.22. At high power outputs, the slope of the curve decreases because of junction heating.

Example 4.12 A given GaAlAs laser has an optical cavity length of 300 μm and a 100-μm width. At a normal operating temperature, the gain factor $\bar{\beta} = 21\times10^{-3}$ A $\cdot$ cm^3 and the loss coefficient $\bar{\alpha} \approx 10$ cm^{-1}. Assume the reflectivity is $R_1 = R_2 = R = 0.32$ for each end face. Find (*a*) the threshold current density and (*b*) the threshold current for this device.

Solution: (*a*) From Eqs. (4.28) and (4.29) we have

$$J_{th} = \frac{1}{\bar{\beta}}\left[\bar{\alpha} + \frac{1}{L}\ln\left(\frac{1}{R}\right)\right]$$

$$= \frac{1}{21\times10^{-3}}\left[10 + \frac{1}{(300\times10^{-4})}\ln\left(\frac{1}{0.32}\right)\right]$$

$$= 2.28\times10^3 \, \text{A/cm}^2$$

(*b*) The threshold current I_{th} is given by

$I_{th} = J_{th} \times$ cross-sectional area of the optical cavity

$$= (2.28\times10^3 \, \text{A/cm}^2) \times (300\times10^{-4} \, \text{cm})$$

$$\times (100\times10^{-4} \, \text{cm}) = 684 \, \text{mA}$$

For laser structures that have strong carrier confinement, the *threshold current density* for stimulated emission J_{th} can to a good approximation be related to the lasing-threshold optical gain by

$$g_{th} = \overline{\beta} J_{th} \tag{4.29}$$

where the gain factor $\overline{\beta}$ is a constant that depends on the specific device construction.

4.3.2 Laser Diode Rate Equations

The relationship between optical output power and the diode drive current can be determined by examining the rate equations that govern the interaction of photons and electrons in the active region. As noted earlier, the total carrier population is determined by carrier injection, spontaneous recombination, and stimulated emission. For a *pn* junction with a carrier-confinement region of depth *d*, the *rate equations* are given by

$$\frac{d\Phi}{dt} = Cn\Phi + R_{sp} - \frac{\Phi}{\tau_{ph}} \tag{4.30}$$

$$= \text{stimulated emission} + \text{spontaneous emission} + \text{photon loss}$$

which governs the number of photons Φ, and

$$\frac{dn}{dt} = \frac{J}{qd} - \frac{n}{\tau_{sp}} - Cn\Phi \tag{4.31}$$

$$= \text{injection} + \text{spontaneous recombination} + \text{stimulated emission}$$

which governs the number of electrons *n*. Here, *C* is a coefficient describing the strength of the optical absorption and emission interactions, R_{sp} is the rate of spontaneous emission into the lasing mode (which is much smaller than the total spontaneous-emission rate), τ_{ph} is the photon lifetime, τ_{sp} is the spontaneous-recombination lifetime, and *J* is the injection-current density.

Equations (4.30) and (4.31) may be balanced by considering all the factors that affect the number of carriers in the laser cavity. The first term in Eq. (4.30) is a source of photons resulting from stimulated emission. The second term, describing the number of photons produced by spontaneous emission, is relatively small compared with the first term. The third term in Eq. (4.30) indicates the decay in the number of photons caused by loss mechanisms in the lasing cavity. In Eq. (4.31), the first term represents the increase in the electron concentration in the conduction band as current flows into the device. The second and third terms give the number of electrons lost from the conduction band owing to spontaneous and stimulated transitions, respectively.

Solving these two equations for a steady-state condition will yield an expression for the output power. The steady state is characterized by the left-hand sides of Eqs (4.30) and (4.31) being equal to zero. First, from Eq. (4.30), assuming R_{sp} is negligible and noting that $d\Phi/dt$ must be positive when Φ is small, we have

$$Cn - \frac{1}{\tau_{ph}} \geq 0 \tag{4.32}$$

This shows that *n* must exceed a threshold value n_{th} in order for Φ to increase. Using Eq. (4.31), this threshold value can be expressed in terms of the threshold current J_{th} needed to maintain an inversion

level $n = n_{th}$ in the steady state when the number of photons $\Phi = 0$:

$$\frac{n_{th}}{\tau_{sp}} = \frac{J_{th}}{qd} \tag{4.33}$$

This expression defines the current required to sustain an excess electron density in the laser when spontaneous emission is the only decay mechanism.

Next, consider the photon and electron rate equations in the steady-state condition at the lasing threshold. Respectively, Eqs (4.30) and (4.31) become

$$0 = Cn_{th}\Phi_s + R_{sp} - \frac{\Phi_s}{\tau_{ph}} \tag{4.34}$$

and

$$0 = \frac{J}{qd} - \frac{n_{th}}{\tau_{sp}} - Cn_{th}\Phi_s \tag{4.35}$$

where Φ_s is the steady-state photon density. Adding Eqs (4.34) and (4.35), using Eq. (4.33) for the term n_{th}/τ_{sp}, and solving for Φ_s yields the number of photons per unit volume:

$$\Phi_s = \frac{\tau_{ph}}{qd}(J - J_{th}) + \tau_{ph}\, R_{sp} \tag{4.36}$$

The first term in Eq. (4.36) is the number of photons resulting from stimulated emission. The power from these photons is generally concentrated in one or a few modes. The second term gives the spontaneously generated photons. The power resulting from these photons is not mode-selective, but is spread over all the possible modes of the volume, which are on the order of 10^8 modes.

4.3.3 External Quantum Efficiency

The *external differential quantum efficiency* η_{ext} is defined as the number of photons emitted per radiative electron–hole pair recombination above threshold. Under the assumption that above threshold the gain coefficient remains fixed at g_{th}, η_{ext} is given by[1]

$$\eta_{ext} = \frac{\eta_i(g_{th} - \bar{\alpha})}{g_{th}} \tag{4.37}$$

Here, η_i is the internal quantum efficiency. This is not a well-defined quantity in laser diodes, but most measurements show that $\eta_i \simeq 0.6$–0.7 at room temperature. Experimentally, η_{ext} is calculated from the straight-line portion of the curve for the emitted optical power P versus drive current I, which gives

$$\eta_{ext} = \frac{q}{E_g}\frac{dP}{dI} = 0.8065\lambda\ (\mu m)\frac{dP(mW)}{dI(mA)} \tag{4.38}$$

where E_g is the bandgap energy in electron volts, dP is the incremental change in the emitted optical power in milliwatts for an incremental change dI in the drive current (in milliamperes), and λ is the emission

wavelength in micrometers. For standard semiconductor lasers, external differential quantum efficiencies of 15–20 percent per facet are typical. High-quality devices have differential quantum efficiencies of 30–40 percent.

4.3.4 Resonant Frequencies

Now let us return to Eq. (4.27) to examine the resonant frequencies of the laser. The condition in Eq. (4.27) holds when

$$2\beta L = 2\pi m \tag{4.39}$$

where m is an integer. Using $\beta = 2\pi n/\lambda$ for the propagation constant from Eq. (2.46), we have

$$m = \frac{L}{\lambda/2n} = \frac{2Ln}{c}\nu \tag{4.40}$$

where $c = \nu\lambda$. This states that the cavity resonates (i.e., a standing-wave pattern exists within it) when an integer number m of half-wavelengths spans the region between the mirrors.

Since in all lasers the gain is a function of frequency (or wavelength, since $c = \nu\lambda$), there will be a range of frequencies (or wavelengths) for which Eq. (4.40) holds. Each of these frequencies corresponds to a mode of oscillation of the laser. Depending on the laser structure, any number of frequencies can satisfy Eqs. (4.26) and (4.27). Thus some lasers are single-mode and some are multimode. The relationship between gain and frequency can be assumed to have the gaussian form

$$g(\lambda) = g(0)\exp\left[-\frac{(\lambda - \lambda_0)^2}{2\sigma^2}\right] \tag{4.41}$$

where λ_0 is the wavelength at the center of the spectrum, σ is the spectral width of the gain, and the maximum gain $g(0)$ is proportional to the population inversion.

Let us now look at the frequency, or wavelength, spacing between the modes of a multimode laser. Here, we consider only the longitudinal modes. Note, however, that for each longitudinal mode there may be several transverse modes that arise from one or more reflections of the propagating wave at the sides of the resonator cavity.[1,3] To find the frequency spacing, consider two successive modes of frequencies ν_{m-1} and ν_m represented by the integers $m - 1$ and m. From Eq. (4.40), we have

$$m - 1 = \frac{2Ln}{c}\nu_{m-1} \tag{4.42}$$

and

$$m = \frac{2Ln}{c}\nu_m \tag{4.43}$$

Subtracting these two equations yields

$$1 = \frac{2Ln}{c}\left(\nu_m - \nu_{m-1}\right) = \frac{2Ln}{c}\Delta\nu \tag{4.44}$$

from which we have the frequency spacing

Example 4.13 A GaAs laser operating at 850 nm has a 500-μm length and a refractive index $n = 3.7$. (*a*) What are the frequency spacing and the wavelength spacing? (*b*) If, at the half-power point, $\lambda - \lambda_0 = 2$ nm, what is the spectral width σ of the gain?

Solution: (*a*) From Eq. (4.45) the frequency spacing is

$$\Delta v = \frac{3 \times 10^3 \, m/s}{2(500 \times 10^{-6} \, m)(3.7)} = 81 \, \text{GHz}$$

From Eq. (4.46) the wavelength spacing is

$$\Delta \lambda = \frac{(850 \times 10^{-9} \, m)^2}{2(500 \times 10^{-6} \, m)(3.7)} = 0.195 \, \text{nm}$$

(*b*) Using Eq. (4.41) with $g(\lambda) = 0.5 \, g(0)$ and then solving for σ with $\lambda - \lambda_0 = \Delta \lambda = 0.195$ nm yields

$$\sigma = \frac{\lambda - \lambda_0}{\sqrt{2 \ln 2}} = \frac{0.195 \, \text{nm}}{\sqrt{2 \ln 2}} = 0.170 \, \text{nm}$$

Example 4.14 Consider a double-heterostructure edge-emitting Fabry-Perot AlGaAs laser, which emits at 900 nm. Suppose that the laser chip is 300 μm long and the refractive index of the laser material is 4.3. (*a*) How many half-wavelengths span the region between the Fabry-Perot mirror surfaces? (*b*) What is the spacing between the lasing modes?

Solution: (*a*) From Eq. (4.40) the number of half-wavelengths that span the region between the Fabry-Perot mirror surfaces is

$$m = \frac{2nL}{\lambda} = \frac{2(4.3) \times 300 \, \mu m}{0.90 \, \mu m} = 2866$$

(*b*) From Eq. (4.46) the spacing between the lasing modes is

$$\Delta \lambda = \frac{(900 \times 10^{-9} \, m)^2}{2(300 \times 10^{-6} \, m)(4.3)} = 0.314 \, \text{nm}$$

$$\Delta v = \frac{c}{2Ln} \tag{4.45}$$

This can be related to the wavelength spacing $\Delta \lambda$ through the relationship $\Delta v / v = \Delta \lambda / \lambda$, yielding

$$\Delta \lambda = \frac{\lambda^2}{2Ln} \tag{4.46}$$

Thus, given Eqs. (4.41) and (4.46), the output spectrum of a multimode laser follows the typical gain-versus-frequency plot given in Fig. 4.23, where the exact number of modes, their heights, and their spacings depend on the laser construction.

4.3.5 Laser Diode Structures and Radiation Patterns

A basic requirement for efficient operation of laser diodes is that, in addition to transverse optical confinement and carrier confinement between heterojunction layers, the current flow must be restricted laterally to a narrow stripe along the length of the laser. Numerous novel methods of achieving this, with varying degrees of success, have been proposed, but all strive for the same goals of limiting the number of lateral modes so that lasing is confined to a single filament, stabilizing the lateral gain, and ensuring a relatively low threshold current.

Figure 4.24 shows the three basic *optical-confinement methods* used for bounding laser light in the lateral direction.[30] In the first structure, a narrow electrode stripe (less than 8 μm wide) runs along the length of the diode. The injection of electrons and holes into the device alters the refractive index of the active layer directly below the stripe. The profile of these injected carriers creates a weak, complex waveguide that confines the light laterally. This type of device is commonly referred to as a *gain-guided*
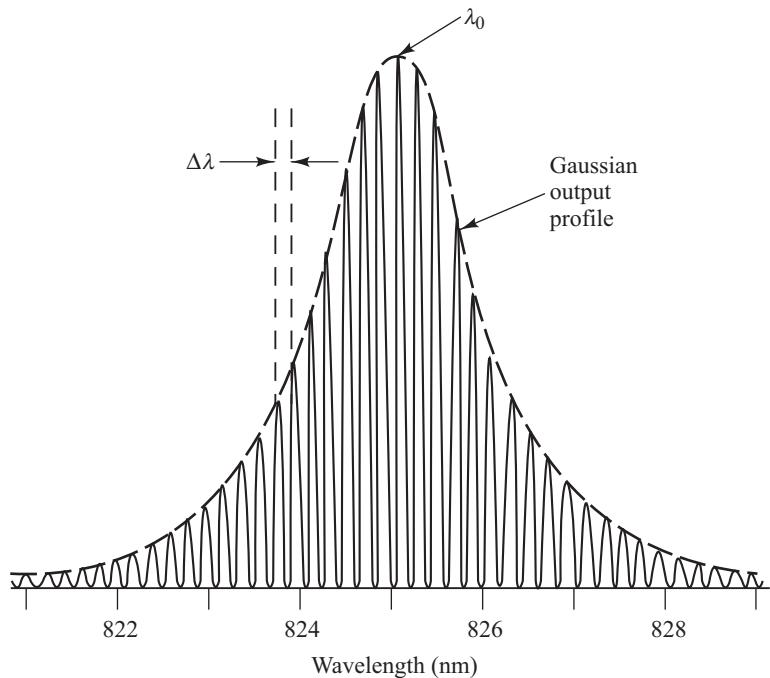
Fig. 4.23 *Typical spectrum from a Fabry-Perot GaAlAs/GaAs laser diode. (Reproduced with permission from K. Peterman and G. Arnold,* IEEE J. Quantum Electron.*, vol. 18, pp. 543–555, Apr. 1982, © 1982, IEEE.)*
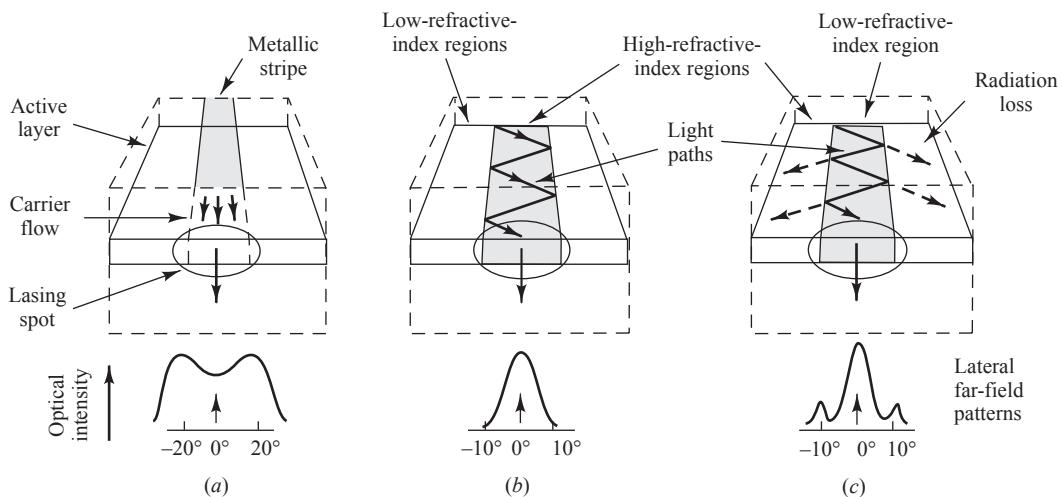
Fig. 4.24 *Three fundamental structures for confining optical waves in the lateral direction: (a) in the gain-induced guide, electrons injected via a metallic stripe contact alter the index of refraction of the active layer; (b) the positive-index waveguide has a higher refractive index in the central portion of the active region; (c) the negative-index waveguide has a lower refractive index in the central portion of the active region. (Reproduced with permission from Botez,[30] © 1985, IEEE.)*

laser. Although these lasers can emit optical powers exceeding 100 mW, they have strong instabilities and can have highly astigmatic, two-peaked beams as shown in Fig. 4.24*a*.

More stable structures use the configurations shown in Fig. 4.24*b* and 4.24*c*. Here, dielectric waveguide structures are fabricated in the lateral direction. The variations in the real refractive index of the various materials in these structures control the lateral modes in the laser. Thus, these devices are called *index-guided lasers*. If a particular index-guided laser supports only the fundamental transverse mode and the fundamental longitudinal mode, it is known as a *single-mode laser*. Such a device emits a single, well-collimated beam of light that has an intensity profile that is a bell-shaped gaussian curve.

Index-guided lasers can have either positive-index or negative-index wave-confining structures. In a *positive-index waveguide,* the central region has a higher refractive index than the outer regions. Thus all of the guided light is reflected at the dielectric boundary, just as it is at the core-cladding interface in an optical fiber. By proper choice of the change in refractive index and the width of the higher-index region, one can make a device that supports only the fundamental lateral mode.

In a *negative-index waveguide,* the central region of the active layer has a lower refractive index than the outer regions. At the dielectric boundaries, part of the light is reflected and the rest is refracted into the surrounding material and is thus lost. This radiation loss appears in the far-field radiation pattern as narrow side lobes to the main beam, as shown in Fig. 4.24*c*. Since the fundamental mode in this device has less radiation loss than any other mode, it is the first to lase. The positive-index laser is the more popular of these two structures.

Index-guided lasers can be made using any one of four fundamental structures. These are the buried heterostructure, a selectively diffused construction, a varying-thickness structure, and a bent-layer configuration. To make the *buried heterostructure* (BH) laser shown in Fig. 4.25, one etches a narrow mesa stripe (1–2 μm wide) in double-heterostructure material. The mesa is then embedded in high-resistivity lattice-matched *n*-type material with an appropriate bandgap and low refractive index. This material is
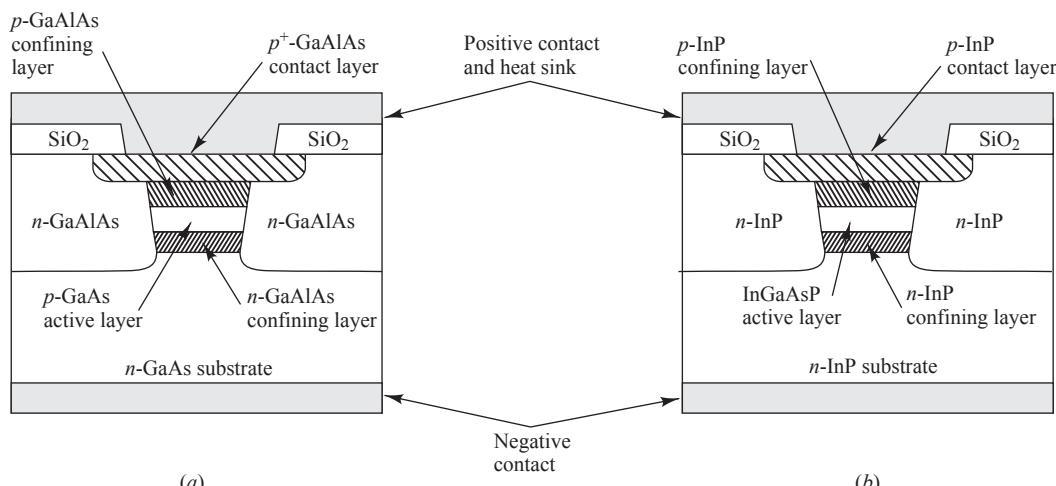

Fig. 4.25 *(a) Short-wavelength (800–900 nm) GaAlAs and (b) long-wavelength (1300–1600 nm) InGaAsP buried heterostructure laser diodes*
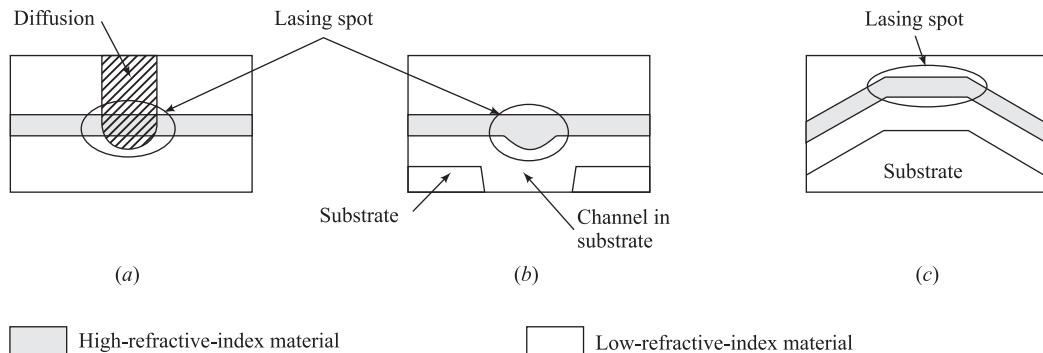
Fig. 4.26 *Positive-index optical-wave-confining structure of the (a) selectively diffused, (b) varying-thickness, and (c) bent-layer types. (Adapted with permission from Botez,[30] © 1985, IEEE.)*

GaAlAs in 800-to-900-nm lasers with a GaAs active layer, and is InP for 1300-to-1600-nm lasers with an InGaAsP active layer. This configuration thus strongly traps generated light in a lateral waveguide. A number of variations of this fundamental structure have been used to fabricate high-performing laser diodes.[3]

The *selectively diffused construction* is shown in Fig. 4.26a. Here, a chemical dopant, such as zinc for GaAlAs lasers and cadmium for InGaAsP lasers, is diffused into the active layer immediately below the metallic contact stripe. The dopant changes the refractive index of the active layer to form a lateral waveguide channel. In the *varying-thickness structure* shown in Fig. 4.26b, a channel (or other topological configuration, such as a mesa or terrace) is etched into the substrate. Layers of crystal are then regrown into the channel using liquid-phase epitaxy. This process fills in the depressions and partially dissolves the protrusions, thereby creating variations in the thicknesses of the active and confining layers. When an optical wave encounters a local increase in the thickness, the thicker area acts as a positive-index waveguide of higher-index material. In the *bent-layer structure,* a mesa is etched into the substrate as shown in Fig. 4.26c. Semiconductor material layers are grown onto this structure using vapor-phase epitaxy to exactly replicate the mesa configuration. The active layer has a constant thickness with lateral bends. As an optical wave travels along the flat top of the mesa in the active area, the lower-index material outside of the bends confines the light along this lateral channel.

In addition to confining the optical wave to a narrow lateral stripe to achieve continuous high optical output power, one also needs to restrict the drive current tightly to the active layer so more than 60 percent of the current contributes to lasing. Figure 4.27 shows the four basic *current-confinement methods.* In each method, the device architecture blocks current on both sides of the lasing region. This is achieved either by high-resistivity regions or by reverse-biased *pn* junctions, which prevent the current from flowing while the device is forward-biased under normal conditions. For structures with a continuous active layer, the current can be confined either above or below the lasing region. The diodes are forward-biased so that current travels from the *p*-type to the *n*-type regions. In the *preferential-dopant diffusion* method, partially diffusing a *p*-type dopant (Zn or Cd) through an *n*-type capping layer establishes a narrow path for the current, since back-biased *pn* junctions block the current outside the diffused region. The *proton implantation* method creates regions of high resistivity, thus restricting the current to a narrow path between these regions. The *inner-stripe confinement* technique grows the lasing structure above a
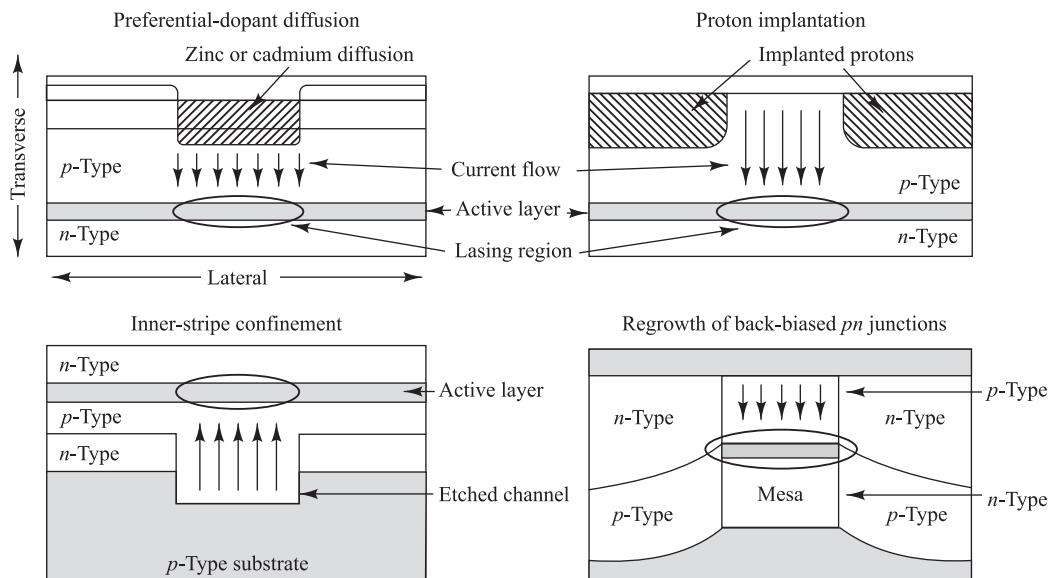
Fig. 4.27 *Four basic methods for achieving current confinement in laser diodes: (a) preferential-dopant diffusion, (b) proton implantation, (c) inner-stripe confinement, and (d) regrowth of back-biased pn junctions. (Adapted with permission from Botez,[30] © 1985, IEEE.)*

channel etched into planar material. Back-biased *pn* junctions restrict the current on both sides of the channel. When the active layer is discontinuous, as in a buried heterostructure, current can be blocked on both sides of the mesa by growing *pn* junctions that are reverse-biased when the device is operating. A laser diode can use more than one current-confining technique.

In a double-heterojunction laser, the highest-order transverse mode that can be excited depends on the waveguide thickness and on the refractive-index differentials at the waveguide boundaries.[1] If the refractive-index differentials are kept at approximately 0.08, then only the fundamental transverse mode will propagate if the active area is thinner than 1 μm.

When designing the width and thickness of the optical cavity, a tradeoff must be made between current density and output beam width. As either the width or the thickness of the active region is increased, a narrowing occurs of the lateral or transverse beam widths, respectively, but at the expense of an increase in the threshold current density. Most positive-index waveguide devices have a lasing spot 3 μm wide by 0.6 μm high. This is significantly greater than the active-layer thickness, since about half the light travels in the confining layers. Such lasers can operate reliably only up to continuous-wave (CW) output powers of 3–5 mW. Here, the transverse and lateral half-power beam widths shown in Fig. 4.18 are about $\theta_\perp \simeq 30$–$50°$ and $\theta_{||} \simeq 5$–$10°$, respectively.

Although the active layer in a standard double-heterostructure laser is thin enough (1–3 μm) to confine electrons and the optical field, the electronic and optical properties remain the same as in the bulk material. This limits the achievable threshold current density, modulation speed, and linewidth of the device. *Quantum-well lasers* overcome these limitations by having an active-layer thickness around 10 nm.[31–32] This changes the electronic and optical properties dramatically, because the dimensionality of
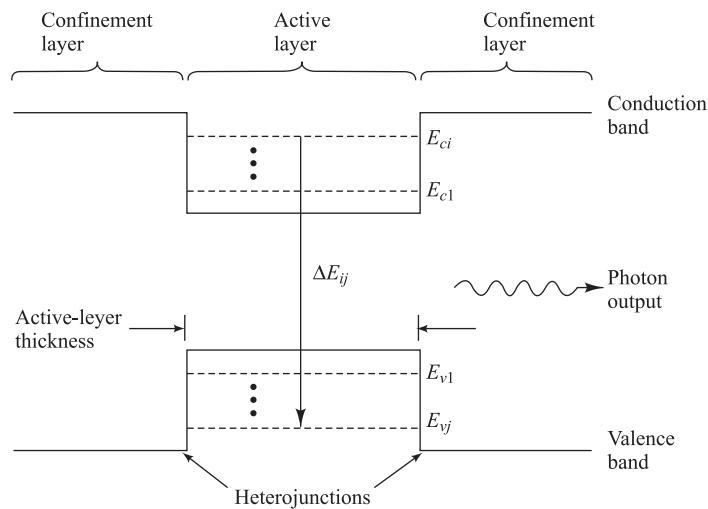
Fig. 4.28 *Energy-band diagram for a quantum layer in a multiple quantum-well (MQW) laser. The parameter ΔE_{ij} represents the allowed energy-level transitions.*

the free-electron motion is reduced from three to two dimensions. As shown in Fig. 4.28, the restriction of the carrier motion normal to the active layer results in a quantization of the energy levels. The possible energy-level transitions that lead to photon emission are designated by ΔE_{ij} (see Prob. 4.16). Both single quantum-well (SQW) and multiple quantum-well (MQW) lasers have been fabricated. These structures contain single and multiple active regions, respectively. The layers separating the active regions are called *barrier layers*. The MQW lasers have a better optical-mode confinement, which results in a lower threshold current density. The wavelength of the output light can be changed by adjusting the layer thickness d. For example, in an InGaAs quantum-well laser, the peak output wavelength moves from 1550 nm when $d = 10$ nm to 1500 nm when $d = 8$ nm.

4.3.6 Single-Mode Lasers

For high-speed long-distance communications one needs single-mode lasers, which must contain only a single longitudinal mode and a single transverse mode. Consequently, the spectral width of the optical emission is very narrow.

One way of restricting a laser to have only one longitudinal mode is to reduce the length L of the lasing cavity to the point where the frequency separation $\Delta\nu$ of the adjacent modes given in Eq. (4.45) is larger than the laser transition line width; that is, only a single longitudinal mode falls within the gain bandwidth of the device. For example, for a Fabry-Perot cavity, all longitudinal modes have nearly equal losses and are spaced by about 1 nm in a 250-μm-long cavity at 1300 nm. By reducing L from 250 μm to 25 μm, the mode spacing increases from 1 nm to 10 nm. However, these lengths make the device hard to handle, and they are limited to optical output powers of only a few milliwatts.[33]

Alternative devices were thus developed. Among these are vertical-cavity surface-emitting lasers, structures that have a built-in frequency-selective grating, and tunable lasers. Here, we look at the first two structures. Tunable lasers are discussed in Chapter 10 in relation to their use in multiple-wavelength optical links. The special feature of a vertical-cavity surface-emitting laser (VCSEL)[6,34–36] is that the light
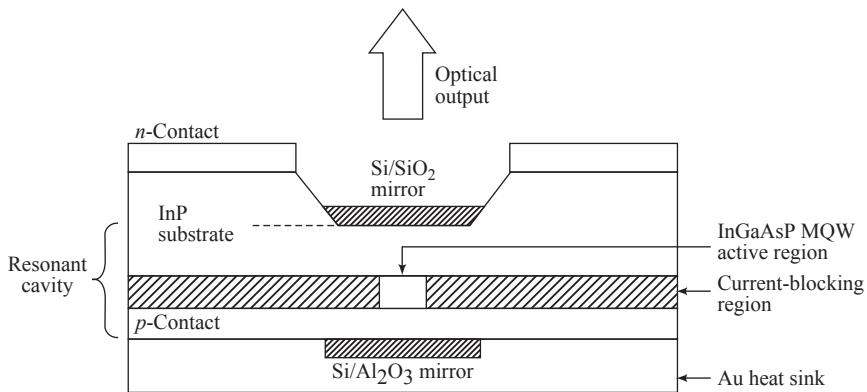
Fig. 4.29 *Basic architecture of a vertical-cavity surface-emitting laser (VCSEL)*

emission is perpendicular to the semiconductor surface, as shown in Fig. 4.29. This feature facilitates the integration of multiple lasers onto a single chip in one- or two-dimensional arrays, which makes them attractive for wavelength-division-multiplexing applications. The active-region volume of these devices is very small, which leads to very low threshold currents ($< 100\,\mu$A). In addition, for an equivalent output power compared to edge-emitting lasers, the modulation bandwidths are much greater, since the higher photon densities reduce radiative lifetimes. The mirror system used in VCSELs to form the resonant cavity is of critical importance, since maximum reflectivity is needed for efficient operation. Figure 4.29 shows one mirror system that consists of a semiconductor material, such as Si/SiO_2, as one material and an oxide layer, such as Si/Al_2O_3, as the other material.

Three types of laser configurations using a built-in *frequency-selective reflector* are shown in Fig. 4.30. In each case, the frequency-selective reflector is a corrugated grating that is a passive waveguide layer adjacent to the active region. The optical wave propagates parallel to this grating. The operation of these types of lasers is based on the distributed Bragg phase-grating reflector.[30] A phase grating is essentially a region of periodically varying refractive index that causes two counter propagating traveling waves to couple. The coupling is at a maximum for wavelengths close to the Bragg wavelength λ_B, which is related to the period Λ of the corrugations by

$$\lambda_B = \frac{2 n_e \Lambda}{k} \tag{4.47}$$

where n_e is the effective refractive index of the mode and k is the order of the grating. First-order gratings ($k = 1$) provide the strongest coupling, but sometimes second-order gratings are used because their larger corrugation period makes fabrication easier. Lasers based on this architecture exhibit good single-mode longitudinal operation with low sensitivity to drive-current and temperature variations.

In the *distributed-feedback* (DFB) laser,[1,28,29] the grating for the wavelength selector is formed over the entire active region. As shown in Fig. 4.31, in an ideal DFB laser the longitudinal modes are spaced symmetrically around λ_B at wavelengths given by

$$\lambda = \lambda_B \pm \frac{\lambda_B^2}{2 n_e L_e}\left(m + \frac{1}{2}\right) \tag{4.48}$$
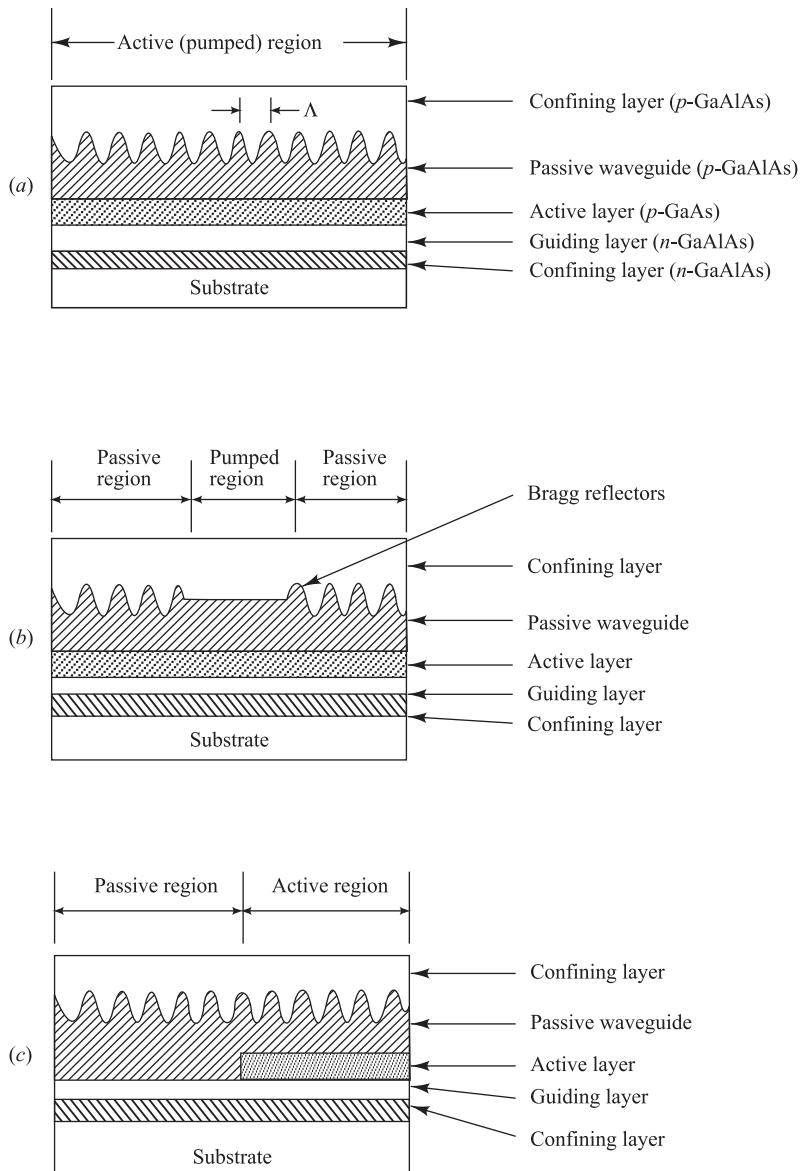
Fig. 4.30 *Three types of laser structures using built-in frequency-selective resonator gratings: (a) distributed-feedback (DFB) laser, (b) distributed-Bragg-reflector (DBR) laser, and (c) distributed-reflector (DR) laser*

where $m = 0, 1, 2, ...$ is the mode order and L_e is the effective grating length. The amplitudes of successively higher-order lasing modes are greatly reduced from the zero-order amplitude; for example, the first-order mode ($m = 1$) is usually more than 30 dB down from the zero-order amplitude ($m = 0$).
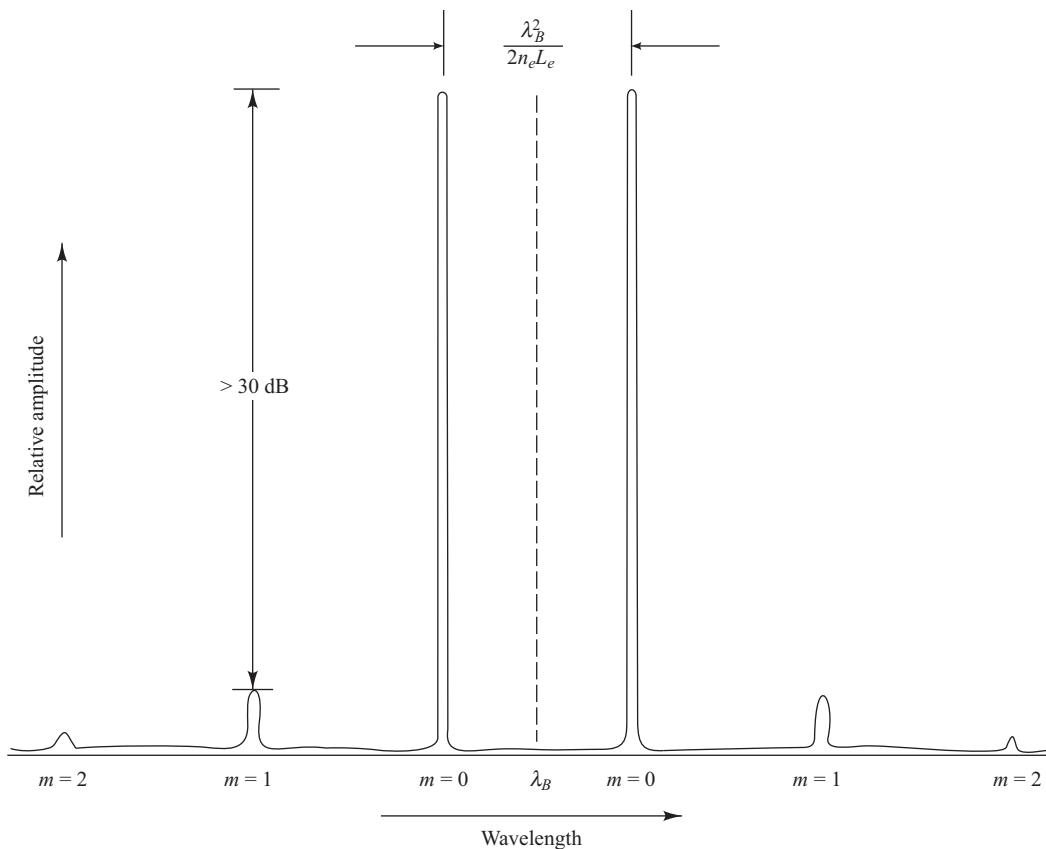
Fig. 4.31   *Output spectrum symmetrically distributed around λ_B in an idealized distributed-feedback (DFB) laser diode*

Theoretically, in a DFB laser that has both ends antireflection-coated, the two zero-order modes on either side of the Bragg wavelength should experience the same lowest threshold gain and would lase simultaneously in an idealized symmetrical structure. However, in practice, the randomness of the cleaving process lifts the degeneracy in the modal gain and results in single-mode operation. This facet asymmetry can be further increased by putting a high-reflection coating on one end and a low-reflection coating on the other; for example, around 2 percent on the front facet and 30 percent on the rear facet. Variations on the DFB design have been the introduction of a $\pi/2$ optical phase shift (i.e., a quarter wavelength) in the corrugation at the center of the optical cavity to make the laser oscillate near the Bragg wavelength, since reflections occur most effectively at this wavelength.

For the *distributed-Bragg-reflector* (DBR) laser,[29,37] the gratings are located at the ends of the normal active layer of the laser to replace the cleaved end mirrors used in the Fabry-Perot optical resonator (Fig. 4.30*b*). The *distributed-reflector* laser consists of active and passive distributed reflectors (Fig. 4.30*c*). This structure improves the lasing properties of conventional DFB and DBR lasers, and has a high efficiency and high output capability.

4.3.7 Modulation of Laser Diodes

The process of putting information onto a lightwave is called *modulation*. For data rates of less than approximately 10 Gb/s (typically 2.5 Gb/s), the process of imposing information on a laser-emitted light stream can be realized by *direct modulation*. This involves directly varying the laser drive current with the electrically formatted information stream to produce a correspondingly varying optical output power. For higher data rates one needs to use a device called an *external modulator* to temporally modify a steady optical power level emitted by the laser (see Sec. 4.3.8). A variety of external modulators are available commercially either as a separate device or as an integral part of the laser transmitter package.

The basic limitation on the direct modulation rate of laser diodes depends on the spontaneous and stimulated emission carrier lifetimes and on the photon lifetime. The *spontaneous carrier lifetime* τ_{sp} is a function of the semiconductor band structure and the carrier concentration. At room temperature this lifetime is about 1 ns in GaAs-based materials for dopant concentrations on the order of 10^{19} cm^{-3}. The *stimulated carrier lifetime* τ_{st} depends on the optical density in the lasing cavity and is on the order of 10 ps. The *photon lifetime* τ_{ph} is the average time that the photon resides in the lasing cavity before being lost either by absorption or by emission through the facets. In a Fabry-Perot cavity, the photon lifetime is[1]

$$\tau_{ph}^{-1} = \frac{c}{n}\left(\bar{\alpha} + \frac{1}{2L}\ln\frac{1}{R_1 R_2}\right) = \frac{c}{n}g_{th} \tag{4.49}$$

For a typical value of $g_{th} = 50$ cm^{-1} and a refractive index in the lasing material of $n = 3.5$, the photon lifetime is approximately $\tau_{ph} = 2$ ps. This value sets the upper limit to the direct modulation capability of the laser diode.

A laser diode can readily be pulse-modulated because the photon lifetime is much smaller than the carrier lifetime. If the laser is completely turned off after each pulse, the spontaneous carrier lifetime will limit the modulation rate. This is because, at the onset of a current pulse of amplitude I_p, a period of time t_d given by (see Prob. 4.19)

$$t_d = \tau \ln\frac{I_p}{I_p + (I_B - I_{th})} \tag{4.50}$$

is needed to achieve the population inversion necessary to produce a gain that is sufficient to overcome the optical losses in the lasing cavity. In Eq. (4.50) the parameter I_B is the bias current, which is a fixed dc current applied to the laser. The parameter τ is the average lifetime of the carriers in the combination region when the total current $I = I_p + I_B$ is close to the threshold current I_{th}. Equation (4.50) shows that the delay time can be eliminated by dc-biasing the diode at the lasing threshold current. Pulse modulation is then carried out by modulating the laser only in the operating region above threshold. In this region, the carrier lifetime is now shortened to the stimulated emission lifetime, so that high modulation rates are possible.

When using a directly modulated laser diode for high-speed transmission systems, the modulation frequency can be no larger than the frequency of the relaxation oscillations of the laser field. The relaxation oscillation depends on both the spontaneous lifetime and the photon lifetime. Theoretically, assuming a linear dependence of the optical gain on carrier density, the relaxation oscillation occurs
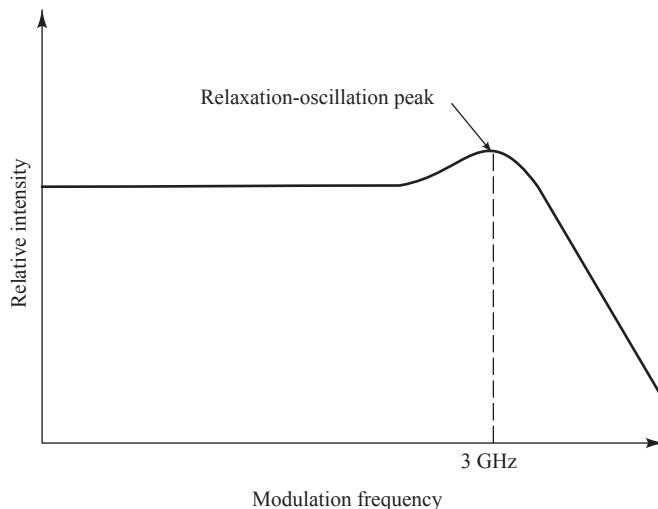
Fig. 4.32   *Example of the relaxation-oscillation peak of a laser diode*

approximately at[1]

$$f = \frac{1}{2\pi} \frac{1}{(\tau_{sp}\, \tau_{ph})^{1/2}} \left(\frac{I}{I_{th}} - 1 \right)^{1/2} \tag{4.51}$$

Since τ_{sp} is about 1 ns and τ_{ph} is on the order of 2 ps for a 300-μm-long laser, then when the injection current is about twice the threshold current, the maximum modulation frequency is a few gigahertz. An example of a laser that has a relaxation-oscillation peak at 3 GHz is shown in Fig. 4.32.

   Analog modulation of laser diodes is carried out by making the drive current above threshold proportional to the baseband information signal. A requirement for this modulation scheme is that a linear relation exist between the light output and the current input. However, signal degradation resulting from nonlinearities that are a consequence of the transient response characteristics of laser diodes make the implementation of analog intensity modulation susceptible to both intermodulation and cross-modulation effects. The use of pulse code modulation or special compensation techniques can alleviate these nonlinear effects.

4.3.8   Laser Linewidth

In non-semiconductor lasers, for example, in solid state lasers, it can be shown that noise arising from spontaneous emission effects results in a finite spectral width or *linewidth* $\Delta\nu$ for the lasing output. However, a semiconductor laser has a significantly higher linewidth than what is predicted by this simple theory. In a semiconductor material both the optical gain and the refractive index depend on the actual carrier density in the medium. This relationship leads to an index-gain coupling mechanism; that is, it gives rise in an interaction between phase noise and the light intensity. The theoretically calculated result is[38]

$$\Delta\nu = \frac{R_{sp}}{4\pi I}(1 + \alpha^2) \tag{4.52}$$

Here I is the average number of photons in the lasing cavity, R_{sp} is the spontaneous emission rate [see Eq. (4.30)], and the parameter α is the *linewidth enhancement factor*. Basically this shows that in semiconductor lasers the linewidth is increased by a factor $(1 + \alpha^2)$.

The linewidth expression in Eq. (4.52) can be rewritten in terms of the optical output power P_{out} as[39]

$$\Delta \nu = \frac{V_g^2 \, h\nu g_{th} n_{sp} \alpha_t}{8\pi P_{out}} (1 + \alpha^2)$$

(4.53)

where V_g is the group velocity of light, $h\nu$ is the photon energy, g_{th} is the threshold gain, α_t is the cavity loss [see Eq. (4.28)], and n_{sp} is the *spontaneous emission factor* (the ratio of spontaneous emission coupled into the lasing mode to the total spontaneous emission).

Equation (4.53) shows that a number of variables influence the magnitude of the laser linewidth. For example, typically $\Delta \nu$ decreases as the laser output power increases. The value of the α-factor also impacts the linewidth. Common values of the dimensionless α-factor range from 2.0 to 6.0 with calculated numbers being in good agreement with experimental measurements. In addition, the laser construction can influence the linewidth since the α-factor values are different depending on the material type and the laser diode structure. For example, the α-factor is smaller in MQW laser structures than in bulk material, and even smaller values are exhibited in devices such as quantum-dot lasers.[39,40] For DFB lasers the linewidth ranges from 5 to 10 MHz (or, equivalently, around 10^{-4} nm).

The spectral width of a laser also can increase significantly when direct modulation is used to vary the light output level. This line broadening is referred to as a *chirping effect*, which is explained in more detail in Sec. 8.3.3.

4.3.9 External Modulation

When direct modulation is used in a laser transmitter, the process of turning the laser on and off with an electrical drive current produces a widening of the laser linewidth. This phenomenon is referred to as *chirp* and makes directly modulated lasers undesirable for operation at data rates greater than about 2.5 Gb/s. For these higher-rate applications it is preferable to use an external modulator, as shown in Fig. 4.33. In such a configuration, the optical source emits a constant-amplitude light signal, which enters the external modulator. In this case, instead of varying the amplitude of the light coming out of the laser, the electrical driving signal dynamically changes the optical power level that exits the external modulator. This process thus produces a time-varying optical signal. The external modulator either can be integrated physically
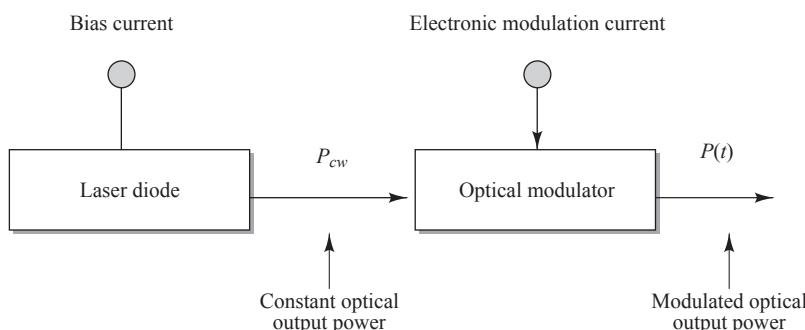
Fig. 4.33 *Operational concept of a generic external modulator*
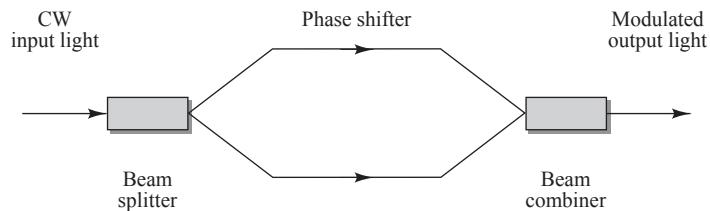
Fig. 4.34 *Operational concept of an electro-optical lithium niobate external modulator*

in the same package with the light source or it can be a separate device. The two main device types are the electro-optical phase modulator and the electro-absorption modulator.[41,42]

The *electro-optical* (EO) phase modulator (also called a *Mach-Zehnder Modulator* or MZM) typically is made of lithium niobate (LbNiO$_3$). In an EO modulator the light beam is split in half and then sent through two separate paths, as shown in Fig. 4.34. A high-speed electric signal then changes the phase of the light signal in one of the paths. This is done in such a manner that when the two halves of the signal meet again at the device output, they will recombine either constructively or destructively. The constructive recombination produces a bright signal and corresponds to a 1 pulse. On the other hand, destructive recombination results in the two signal halves canceling each other so there is no light at the output of the beam combiner. This corresponds to a 0 pulse. LbNiO$_3$ modulators are separately packaged devices and can be up to 12 cm (about 5 inches) long.
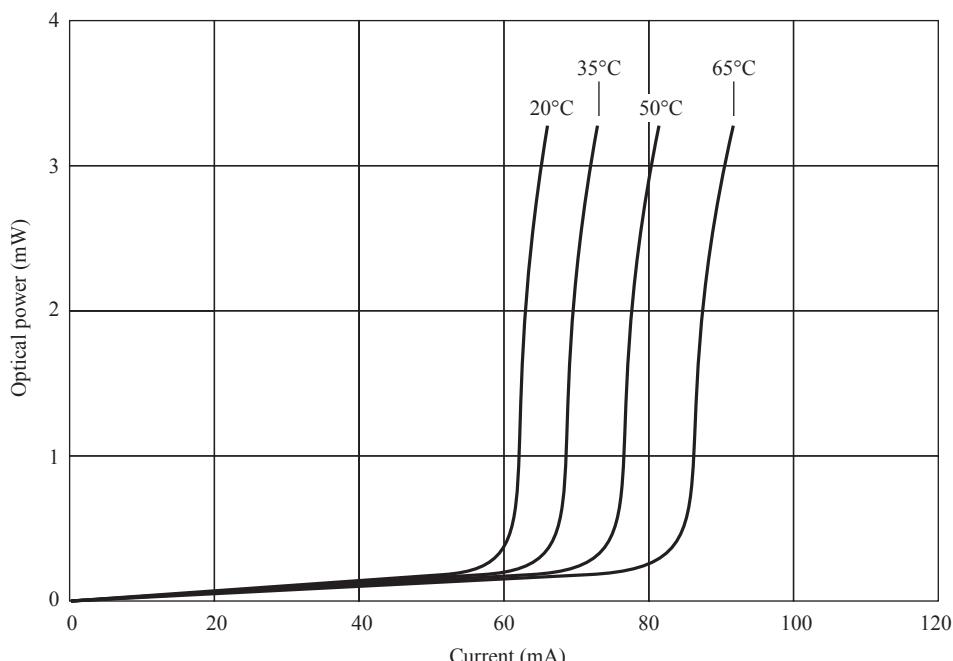
Fig. 4.35 *Temperature-dependent behavior of the optical output power as a function of the bias current for a particular laser diode with $T_0 = 135°C$ and $I_z = 52$ mA.*

The *electro-absorption modulator* (EAM) typically is constructed from *indium phosphide* (InP). It operates by having an electric signal change the transmission properties of the material in the light path to make it either transparent during a 1 pulse or opaque during a 0 pulse. Since InP is used as the material for an EAM, it can be integrated onto the same substrate as a DFB laser diode chip. The complete laser plus modulator module then can be put in a standard butterfly package, thereby reducing drive voltage, power, and space requirements compared to having separate laser and $LbNiO_3$ modulator packages.

4.3.10 Temperature Effects

An important factor to consider in the application of laser diodes is the temperature dependence of the threshold current $I_{th}(T)$. This parameter increases with temperature in all types of semiconductor lasers because of various temperature-dependent factors.[3] The complexity of these factors prevents the formulation of a single equation that holds for all devices and temperature ranges. However, the temperature variation of I_{th} can be approximated by the empirical expression

$$I_{th}(T) = I_z e^{T/T_0} \qquad (4.54)$$

where T_0 is a measure of the threshold temperature coefficient and I_z is a constant. For a conventional stripe-geometry GaAlAs laser diode, T_0 is typically 120–165°C in the vicinity of room temperature. An example of a laser diode with $T_0 = 135$°C and $I_z = 52$ mA is shown in Fig. 4.35. The variation in I_{th} with temperature is 0.8 percent/°C, as is shown in Fig. 4.36. Smaller dependences of I_{th} on temperature have been demonstrated for GaAlAs quantum-well heterostructure lasers. For these lasers, T_0 can be as high as 437°C. The temperature dependence of I_{th} for this device is also shown in Fig. 4.36. The threshold variation for this particular laser type is 0.23 percent/°C.

For the laser diode shown in Fig. 4.35, the threshold current increases by a factor of about 1.4 between 20 and 60°C. In addition, the lasing threshold can change as the laser ages. Consequently, if a constant optical output power level is to be maintained as the temperature of the laser changes or as the laser ages, it is necessary to adjust the dc-bias current level. One possible method for achieving this automatically is an optical feedback scheme.

Optical feedback can be carried out by using a photodetector either to sense the variation in optical power emitted from the rear facet of the laser or to tap off and monitor a small portion of the fiber-coupled power emitted from the front facet. The photodetector compares the optical power output with a reference level and adjusts the dc-bias current level automatically to maintain a constant peak light output relative to the reference. The photodetector used must have a stable long-term
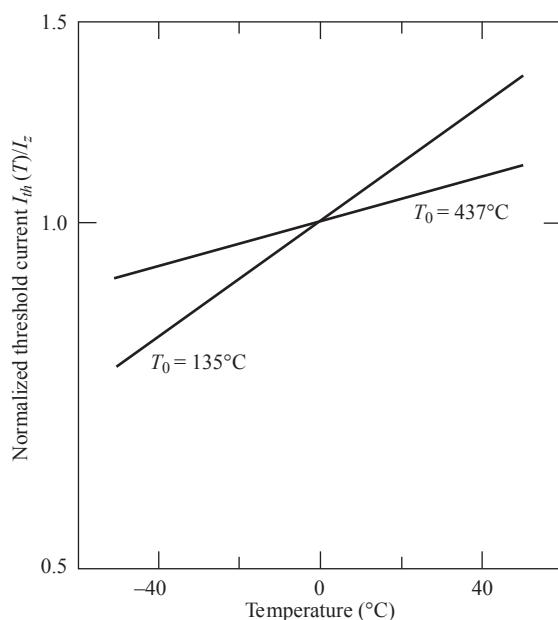
Fig. 4.36 *Variation with temperature of the threshold current I_{th} for two types of laser diodes*

Example 4.15 An engineer has a GaAlAs laser that has a threshold temperature coefficient $T_0 = 135°C$ and an InGaAsP laser with $T_0 = 55°C$. Compare the percent change in the threshold current for each of these lasers when the temperature increases from 20°C to 65°C.

Solution: (*a*) Letting $T_1 = 20°C$ and $T_2 = 65°C$, then from Eq. (4.54) for the GaAlAs laser we have that the threshold current increases by

$$\frac{I_{th}(65°C)}{I_{th}(20°C)} = e^{(\tau_2 - \tau_1)/\tau_0} = e^{(65-20)/135} = 1.40$$

$$= 140 \text{ percent}$$

(*b*) Similarly, letting $T_1 = 20°C$ and $T_2 = 65°C$, then from Eq. (4.54) for the InGaAsP laser we have that the threshold current increases by

$$\frac{I_{th}(65°C)}{I_{th}(20°C)} = e^{(65-20)/55} = 2.27 = 227 \text{ percent}$$

responsivity that remains constant over a wide temperature range. For operation in the 800-to-900-nm region, a silicon *pin* photodiode generally exhibits these characteristics (see Chapter 6).

Another standard method of stabilizing the optical output of a laser diode is to use a miniature thermoelectric cooler.[43] This device maintains the laser at a constant temperature and thus stabilizes the output level. Normally, a thermoelectric cooler is used in conjunction with a rear-facet detector feedback loop, as is shown in Fig. 4.37.
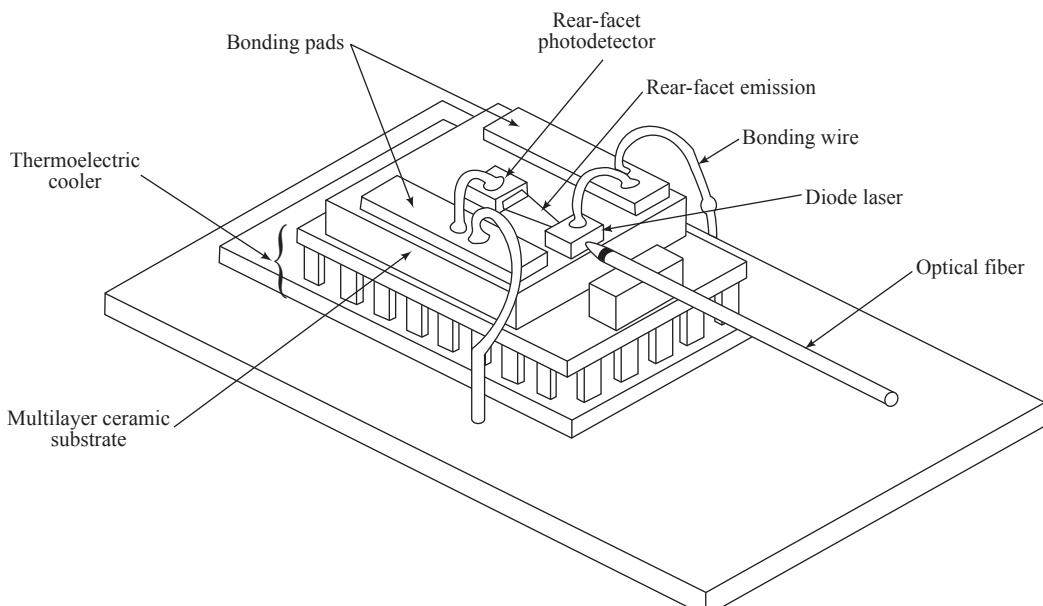
Fig. 4.37 *Construction of a laser transmitter that uses a rear-facet photodiode for output monitoring and a thermoelectric cooler for temperature stabilization*

4.4 Line Coding

In designing a communication link for transporting digitized information, a significant consideration is the format of the transmitted digital signal.[44–49] The signal format that is sent out from the transmitter is important because the receiver must be able to extract precise *timing information* from the incoming signal. The three main purposes of *timing* are:

- To allow the signal to be sampled by the receiver at the time the signal-to-noise ratio is a maximum
- To maintain a proper spacing between pulses
- To indicate the start and end of each timing interval

In addition, it may be desirable for the signal to have an inherent error-detecting capability, as well as an error-correction mechanism, if it is needed or is practical. These timing and error-minimizing features can be incorporated into the data stream by restructuring or *encoding* the digital signal. This process is called *channel coding* or *line coding*. This section examines the basic binary line codes that are used in optical fiber communication systems. Chapter 7 presents further details on specific signal formats in the discussions related to techniques for interpreting received optical signals with a high degree of fidelity.

One of the principal functions of a line code is to minimize errors in the bit stream that might arise from noise or other interference effects. Generally one does this by introducing extra bits into the raw data stream at the transmitter, arranging them in a specific pattern, and extracting the redundant bits at the receiver to recover the original signal. Depending on the amount of redundancy that is introduced into the data stream, various degrees of error reduction in the data can be achieved, provided that the data rate is less than the channel capacity. This is the result of the Shannon channel-capacity theory. Chapter 7 gives more details on error rates in digital fiber optic communication links.

4.4.1 NRZ and RZ Signal Formats

The simplest method for encoding data is the unipolar *nonreturn-to-zero* (NRZ) code. *Unipolar* means that a logic 1 is represented by a voltage or light pulse that fills an entire bit period, whereas for a logic 0 no pulse is transmitted, as shown in Fig. 4.38 for the data sequence 1010110. Since this process turns
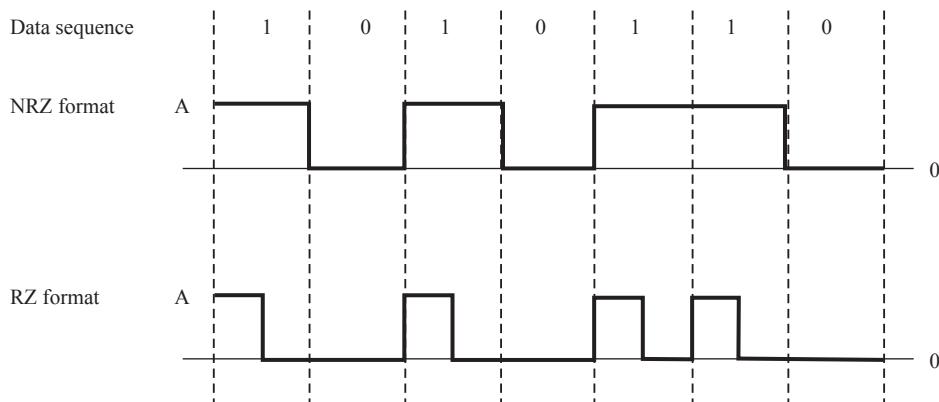
Fig. 4.38 *NRZ and RZ code patterns for the data sequence 1010110*

the light signal on and off, it is known as *amplitude shift keying* (ASK) or *on-off keying* (OOK). If 1 and 0 pulses occur with equal probability, and if the amplitude of the voltage pulse is A, then the average transmitted power for this code is $A^2/2$. In optical systems one typically describes a pulse in terms of its optical power level. In this case the average power for an equal number of 1 and 0 pulses is $P/2$, where P is the peak power in a 1 pulse.

The NRZ code needs the minimum bandwidth and is simple to generate and decode. However, the lack of timing capabilities in an NRZ code can lead to misinterpretations of the bit stream at the receiver. For example, since there are no level transitions from which to extract timing information in a long sequence of NRZ ones or zeros, a long string of N identical bits could be interpreted as either $N + 1$ or $N - 1$ bits, unless highly stable (and expensive) timing clocks are used. Two common techniques for restricting the longest time interval in which no level transitions occur are the use of block codes (see Sec. 4.4.2) and scrambling. *Scrambling* produces a random data pattern by the modulo-2 addition of a known bit sequence to the data stream. At the receiver the same known bit sequence is again modulo-2 added to the received data, which results in the recovery of the original bit sequence.

If an adequate bandwidth margin exists, the timing problem associated with NRZ encoding can be alleviated with a *return-to-zero* (RZ) code. As shown in the bottom half of Fig. 4.38, the RZ code has an amplitude transition at the beginning of each bit interval when a binary 1 is transmitted and no transition for a binary 0. Thus for a RZ pulse a 1 bit occupies only part of the bit interval and returns to zero in the remainder of the bit interval. No pulse is used for a 0 bit.

Although the RZ pulse nominally occupies exactly half a bit period in electronic digital transmission systems, in an optical communication link the RZ pulse might occupy only a fraction of a bit period. A variety of RZ formats are used for links that send data at rates of 10 Gb/s and higher.

4.4.2 Block Codes

Introducing *redundant bits* into a data stream can be used to provide adequate timing and to have error-monitoring features. A popular and efficient encoding method for this is the class of mBnB *block codes*. In this class of codes, blocks of m binary bits are converted to longer blocks of $n > m$ binary bits. As a result of the additional redundant bits, the required bandwidth increases by the ratio n/m. For example, in an mBnB code with $m = 1$ and $n = 2$, a binary 1 is mapped into the binary pair 10, and a binary 0 becomes 01. The overhead for such a code is 50 percent.

Suitable mBnB codes for high data rates are the 3B4B, 4B5B, 5B6B, and 8B10B codes. If simplicity of the encoder and decoder circuits is the main criterion, then the 3B4B format is the most convenient code. The 5B6B code is the most advantageous if bandwidth reduction is the major concern. Various versions of Ethernet use the 3B4B, 4B5B, or 8B10B formats.

4.5 Light Source Linearity

High-radiance LEDs and laser diodes are well-suited optical sources for wideband analog applications provided a method is implemented to compensate for any nonlinearities of these devices. In an analog system, a time-varying electric analog signal $s(t)$ is used to modulate an optical source directly about a bias current point I_B, as shown in Fig. 4.39. With no signal input, the optical power output is P_t. When the signal $s(t)$ is applied, the time-varying (analog) optical output power $P(t)$ is

$$P(t) = P_t[1 + ms(t)] \qquad (4.55)$$
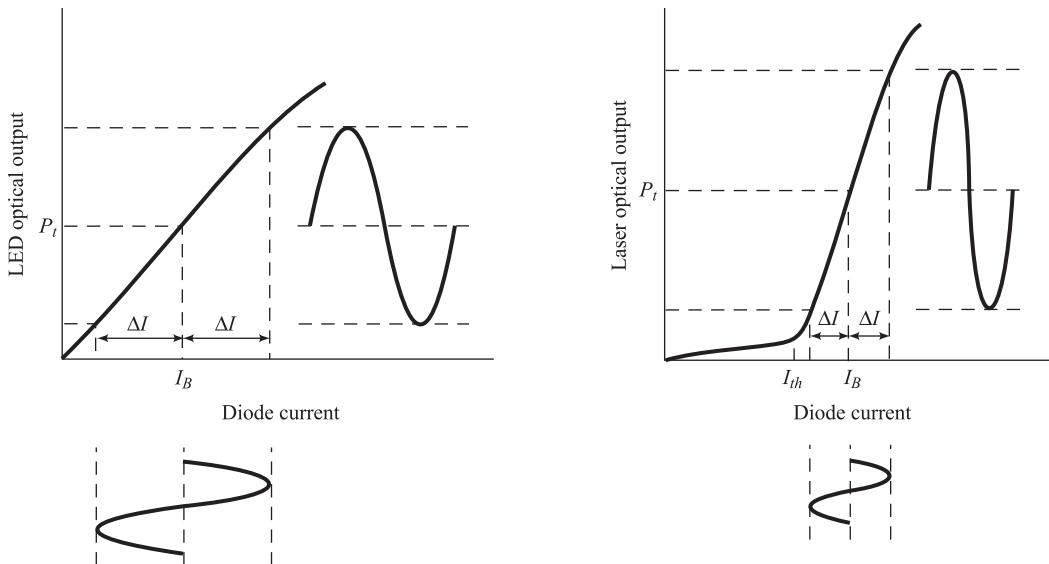
Fig. 4.39   *Bias point and amplitude modulation range for analog applications of LEDs (left) and laser diodes (right)*

Here, *m* is the *modulation index* (or *modulation depth*) defined as

$$m = \frac{\Delta I}{I'_B} \tag{4.56}$$

where $I'_B = I_B$ for LEDs and $I'_B = I_B - I_{th}$ for laser diodes. The parameter ΔI is the variation in current about the bias point. To prevent distortions in the output signal, the modulation must be confined to the linear region of the curve for optical output versus drive current. Furthermore, if ΔI is greater than I'_B (i.e., *m* is greater than 100 percent), the lower portion of the signal gets cut off and severe distortion will result. Typical *m* values for analog applications range from 0.25 to 0.50.

In analog applications, any device nonlinearities will create frequency components in the output signal that were not present in the input signal. Two important nonlinear effects are harmonic and intermodulation distortions. If the signal input to a nonlinear device is a simple cosine wave $x(t) = A \cos \omega t$, the output will be

$$y(t) = A_0 + A_1 \cos \omega t + A_2 \cos 2 \omega t + A_3 \cos 3 \omega t + ... \tag{4.57}$$

That is, the output signal will consist of a component at the input frequency ω plus spurious components at zero frequency, at the second harmonic frequency 2ω, at the third harmonic frequency 3ω, and so on. This effect is known as *harmonic distortion*. The amount of *n*th-order distortion in decibels is given by

$$n\text{th-order harmonic distortion} = 20 \log \frac{A_n}{A_1} \tag{4.58}$$

To determine *intermodulation distortion*, the modulating signal of a nonlinear device is taken to be the sum of two cosine waves $x(t) = A_1 \cos \omega_1 t + A_2 \cos \omega_2 t$. The output signal will then be of the form

$$y(t) = \sum_{m,n} B_{mn} \cos(m\omega_1 + n\omega_2) \tag{4.59}$$

where m and $n = 0, \pm 1, \pm 2, \pm 3, \ldots$. This signal includes all the harmonics of ω_1 and ω_2 plus cross-product terms such as $\omega_2 - \omega_1$, $\omega_2 + \omega_1$, $\omega_2 - 2\omega_1$, $\omega_2 + 2\omega_1$, and so on. The sum and difference frequencies give rise to the intermodulation distortion. The sum of the absolute values of the coefficients m and n determines the order of the intermodulation distortion. For example, the second-order intermodulation products are at $\omega_1 \pm \omega_2$ with amplitude B_{11}, the third-order intermodulation products are at $\omega_1 \pm 2\omega_2$ and $2\omega_1 \pm \omega_2$ with amplitudes B_{12} and B_{21}, and so on. (Harmonic distortions are also present wherever either $m \neq 0$ and $n = 0$ or when $m = 0$ and $n \neq 0$. The corresponding amplitudes are B_{m0} and B_{0n}, respectively.) In general, the odd-order intermodulation products having $m = n \pm 1$ (such as $2\omega_1 - \omega_2$, $2\omega_2 - \omega_1$, $3\omega_1 - 2\omega_2$, etc.) are the most troublesome because they may fall within the bandwidth of the channel. Of these, usually only the third-order terms are important, since the amplitudes of higher-order terms tend to be significantly smaller. If the operating frequency band is less than an octave, all other intermodulation products will fall outside the passband and can be eliminated with appropriate filters in the receiver.

4.6 Reliability Considerations

The operating lifetimes of light-emitting diodes and laser diodes are affected by both operating conditions and fabrication techniques. Thus it is important to understand the relationships between light source operation characteristics, degradation mechanisms, and system reliability requirements.[50–56]

Lifetime tests of optical sources are carried out either at room temperature or at elevated temperatures to accelerate the degradation process. A commonly used elevated temperature is 70°C. The two most popular techniques for determining the lifetime of an optical source either maintain a constant light output by increasing the bias current automatically or keep the current constant and monitor the optical output level. In the first case, the end of life of the device is assumed to be reached when the source can no longer put out a specified power at the maximum current value for CW (continuous-wave) operation. In the second case, the lifetime is determined by the time taken for the optical output power to decrease by 3 dB.

Degradation of light sources can be divided into three basic categories: internal damage and ohmic contact degradation, which hold for both lasers and LEDs, and damage to the facets of laser diodes.

The limiting factor on LED and laser diode lifetime is internal degradation. This effect arises from the migration of crystal defects into the active region of the light source. These defects decrease the internal quantum efficiency and increase the optical absorption. Fabrication steps that can be taken to minimize internal degradation include the use of substrates with low surface dislocation densities (less than 2×10^3 dislocation/cm^2), keeping work-damaged edges out of the diode current path, and minimizing stresses in the active region (to less than 10^8 dyn/cm^2).

For high-quality sources having lifetimes that follow a slow internal-degradation mode, the optical power P decreases with time according to the exponential relationship

$$P(t) = P_0 e^{-t/\tau_m} \tag{4.60}$$

Example 4.16 An optical source is selected from a batch of devices characterized as having lifetimes that follow a slow internal degradation mode. The –3-dB mean time to failure of these devices at room temperature is specified as 1×10^4 h. If the device initially emits 1 mW at room temperature, what is the expected optical output power after (*a*) 1 month of operation, (*b*) after 1 year, and (*c*) after 3 years?

Solution: From Eq. (4.60), the optical power decrease with time is $P = P_0 \exp(-t/\tau_m)$, where $P_0 = 1$ mW and $\tau_m = 2(1 \times 10^4)$ hrs $= 2 \times 10^4$ hrs.

(*a*) At 1 month = 720 hours, we have:

$P(\text{1 month}) = (1 \text{ mW}) \exp(-720/2 \times 10^4) = 0.965 \text{ mW}$

(*b*) At 1 year = 8760 hours, we have:

$P(\text{1 year}) = (1 \text{ mW}) \exp(-8760/2 \times 10^4) = 0.645 \text{ mW}$

(*c*) At 3 years = 3×8760 hours = 26280 hours, we have:

$P(\text{3 years}) = (1 \text{ mW}) \exp(-26280/2 \times 10^4) = 0.269 \text{ mW}$

Here, P_0 is the initial optical power at time $t = 0$, and τ_m is a time constant for the degradation process, which is approximately twice the –3-dB mean time to failure. Since the operating lifetime depends on both the current density J and the junction temperature T, internal degradation can be accelerated by increasing either one of these parameters.

The operating lifetime τ_s has been found experimentally to depend on the current density J through the relation

$$\tau_s \, \alpha \, J^{-n} \tag{4.61}$$

where $1.5 \leq n \leq 2.0$. For example, by doubling the current density, the lifetime decreases by a factor of 3–4. Since the degradation rate of optical sources increases with temperature, an Arrhenius relationship of the form

$$\tau_s = K \, e^{E_A/k_B T} \tag{4.62}$$

has been sought. Here, E_A is an activation energy characterizing the lifetime τ_s, k_B is Boltzmann's constant, T is the absolute temperature at which τ_s was evaluated, and K is a constant. The problem in establishing such an expression is that several competing factors are likely to contribute to the degradation, thereby making it difficult to estimate the activation energy E_A. Activation energies for laser degradation reported in the literature have ranged from 0.3 to 1.0 eV. For practical calculations, a value of 0.7 eV is generally used.

Equations (4.61) and (4.62) indicate that, to increase the light source lifetime, it is advantageous to operate these devices at as low a current and temperature as is practicable. Examples[50] of the luminescent output of InGaAsP LEDs as a function of time for different temperatures are shown in Fig. 4.40. At temperatures below 120°C the output power remains almost constant over the entire measured 15,000 h (1.7-year) operating time. At higher temperatures the power output drops as a function of time. For example, at 230°C the optical power has dropped to one-half its initial value (a 3-dB decrease) after approximately 3000 h (4.1 months) of operation. The activation energy of these lasers is about 1.0 eV.

A second fabrication-related degradation mechanism is ohmic contact deterioration. In LEDs and laser diodes the thermal resistance of the contact between the light source chip and the device heat sink occasionally increases with time. This effect is a function of the solder used to bond the chip to the heat sink, the current density through the contact, and the contact temperature. An increase in the thermal resistance results in a rise in the junction temperature for a fixed operating current. This, in turn, leads
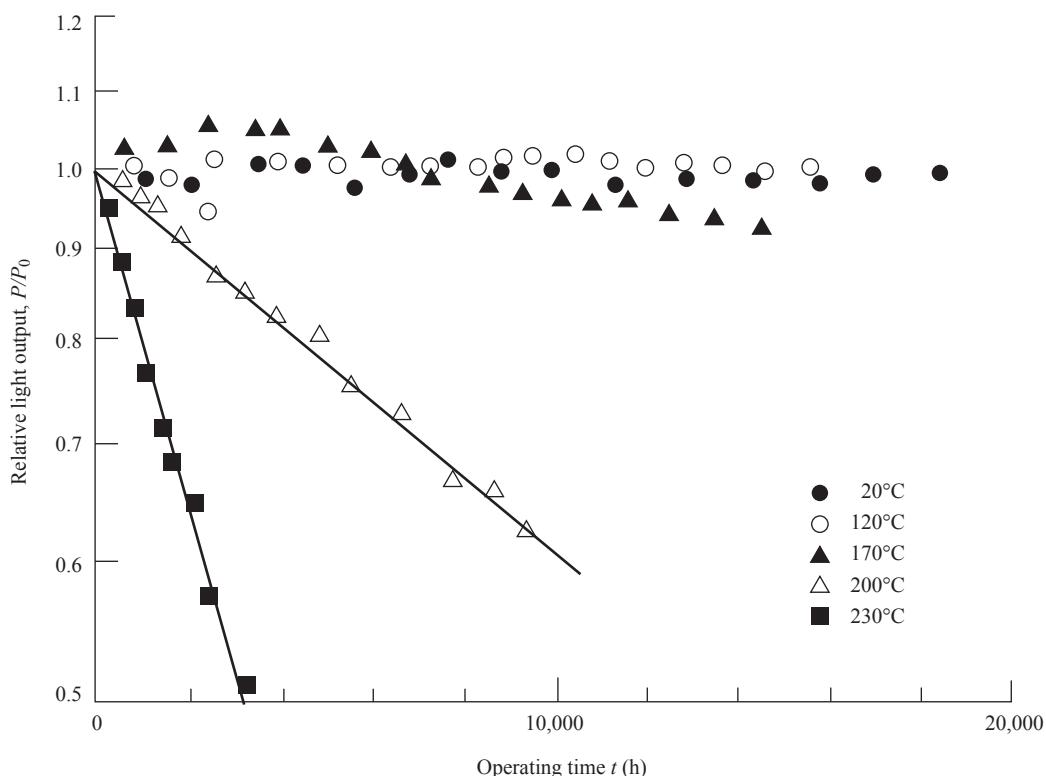
Fig. 4.40 *Normalized output power as a function of operating time for five ambient temperatures. P_0 is the initial optical output power. (Reproduced with permission from Yamakoshi et al.[50])*

to a decrease in the optical output power. However, careful designs and implementation of high-quality bonding procedures have minimized effects resulting from contact degradation.

Facet damage is a degradation problem that exists for laser diodes. This degradation reduces the laser mirror reflectivity and increases the nonradiative carrier recombination at the laser facets. The two types of facet damage that can occur are generally referred to as *catastrophic facet degradation* and *facet erosion*. Catastrophic facet degradation is mechanical damage of the facets that may arise after short operating times of laser diodes at high optical power densities. This damage tends to reduce greatly the facet reflectivity, thereby increasing the threshold current and decreasing the external quantum efficiency. The catastrophic facet degradation has been observed to be a function of the optical power density and the pulse length.

Facet erosion is a gradual degradation occurring over a longer period of time than catastrophic facet damage. The decrease in mirror reflectivity and the increase in nonradiative recombination at the facets owing to facet erosion lower the internal quantum efficiency of the laser and increase the threshold current. In GaAlAs lasers, facet erosion arises from oxidation of the mirror surface. It is speculated that the oxidation process is stimulated by the optical radiation emitted from the laser. Facet erosion is minimized by depositing a half-wavelength-thick Al_2O_3 film on the facet. This type of coating acts as a moisture barrier and does not affect the mirror reflectivity or the lasing threshold current.
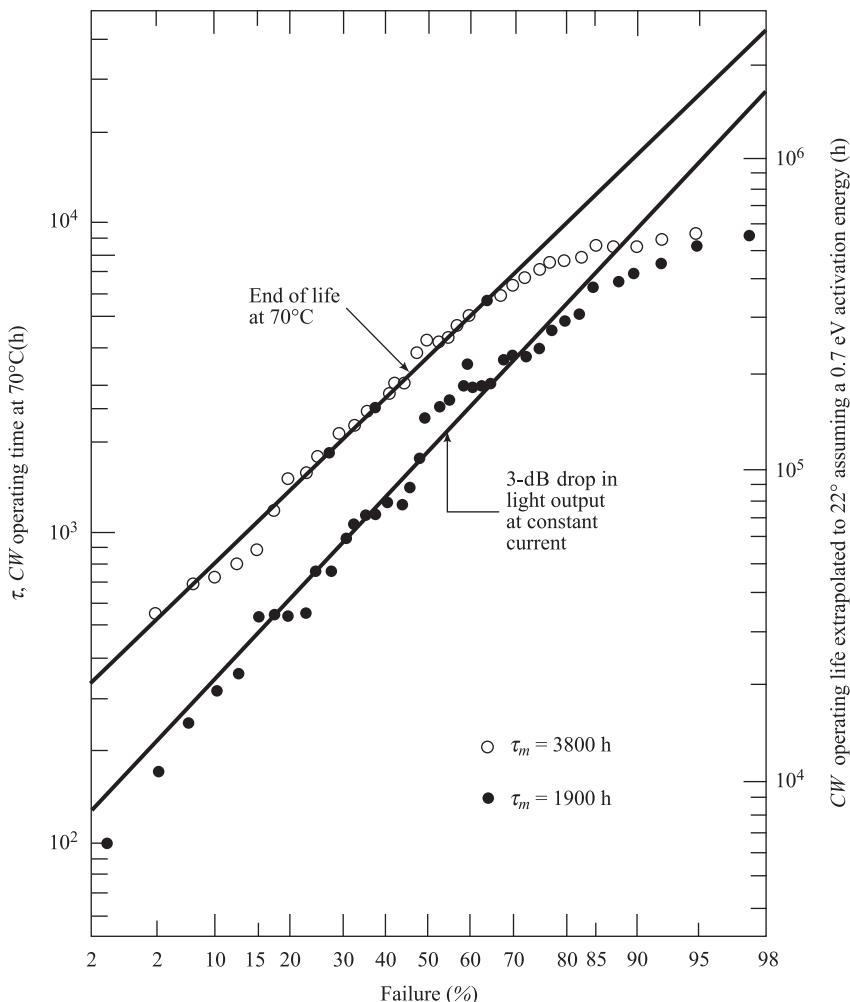
Fig. 4.41 *Time-to-failure plot on log-normal coordinates for 40 low-threshold (≈50 mA) oxide-defined stripe GaAlAs lasers at a 70°C heat-sink temperature. τ_m is the time it took for 50 percent of the lasers to fail. (Reproduced with permission from Ettenberg and Kressel,[51] © 1980, IEEE.)*

A comparison[51] of two definitions of failure for laser diodes operating at 70°C is shown in Fig. 4.41. The lower trace shows the time required for the laser output to drop to one-half its initial value when a constant current passes through the device. This is the "3-dB life."

The "end-of-life" failure is given by the top trace in Fig. 4.41. This condition is defined as the time at which the device can no longer emit a fixed power level (1.25 mW in this case) at the 70°C heat-sink temperature. The mean operating times (time for 50 percent of the lasers to fail) are 3800 h and 1900 h for the end-of-life and 3-dB-life conditions, respectively. The right-hand ordinate of Fig. 4.41 gives an estimate of the operating time at 22°C, assuming an activation energy of 0.7 eV.

4.7 Transmitter Packages

Manufacturers have developed a variety of transmitter packages for different applications. One popular transmitter configuration is shown in Fig. 4.42. This is called a *butterfly package,* which has an attached fiber flylead and contains the components illustrated in Fig. 4.37 (such as the diode laser, a monitoring photodiode, and a thermoelectric cooler). The term *flylead* or *pigtail* refers to a short length of optical fiber that is attached to a device in the factory. An optical connector can be attached easily to such a flylead for subsequent interfacing to an optical cable.

In a number of device configurations both the light source and the photodetector together with an optical fiber receptacle are integrated into single industry-standard packages called *transceivers*. This packaging is done for simplicity of device use, to save space on circuit boards, and to enable a high density of optical input and output ports. More complex integrated packages also contain electronics for carrying out various levels of digital link diagnostics.

Figure 4.43 illustrates and describes a variety of standard transceiver packages. Popular configurations are the *small form factor* (SFF) and the *SFF pluggable* (SFP) packages. Note that they have an optical fiber connector receptacle integrated into the package. This receptacle allows easy attachment of an optical connector for coupling to a transmission fiber. A key feature of a number of transceivers, such as the SFP modules, is that they are *hot-pluggable*. This means that one can insert and remove the modules from the circuit card without turning off the power to the device. This feature ensures that the equipment never stops running (known as *zero downtime interruption*) when performing online maintenance. The fact that the SFP modules can easily be exchanged also makes it simple to carry out system upgrades to higher speeds or increased capacity, it enables rapid link reconfigurations, and it helps reduce inventory costs as equipment circuit cards can be stocked without having specific optical transceivers on them.

Fig. 4.42 *An industry-standard laser diode mounted in a butterfly package. The fiber flylead is terminated with an LC optical connector. (Photo courtesy of Finisar Corporation; www.finisar.com.)*

Connector type	Features
SFP	• Short and long wavelength WDM use • Datacom applications: Fast/Gigabit Ethernet and 1x,2x,4x Fibre Channel • Telecom applications using OC-3/STM-1, OC-12/STM-4, and OC-48/STM-16 across all distances • Distances from very short links up to 100 km
SFF	• Short and long wavelength use • Datacom applications for Gigabit Ethernet and 1x,2x,4x Fibre Channel • Telecom applications using OC-3/STM-1, OC-12/STM-4, and OC-48/STM-16 across all distances • Distances from very short links up to 80 km
XFP	• Short and long wavelength DWDM use • Datacom applications using 10G Ethernet and 10x Fibre Channel • Telecom applications using OC-192/STM-64 • Distances up to 80 km • Supports bit rates up to 11.3 Gb/s
XPAK	• Short and long wavelength use • Datacom applications using 10G Ethernet and 10x Fibre Channel • Distances from short links up to 10 km • Supports bit rates up to 10.5 Gb/s • 1.2 V, 3.3 V, and 5 V operating voltage
XENPAK	• Short and long wavelength use • Datacom applications using 10G Ethernet • Distances from short links up to 10 km • Supports bit rates up to 10.3 Gb/s • 1.2 V, 3.3 V, and 5 V operating voltage
PON	• Long wavelength use • Telecom access network applications for GPON and GE-PON • Distances up to 20 km, Class B+, PX-20+ • 3.3 V operating voltage • Attached fiber flylead with optical connector
GBIC	• Short and long wavelength WDM use • Datacom applications using 1G Ethernet and 1x/2x Fibre Channel • Telecom applications using OC-3/STM-1, OC-12/STM-4, and OC-48/STM-16 across all distances • Distances from short links up to 160 km

Fig. 4.43 *Seven standard optical fiber transceiver packages and their features. (Photos courtesy of Finisar Corporation; www.finisar.com.)*

PROBLEMS

4.1 Measurements show that the bandgap energy E_g for GaAs varies with temperature according to the empirical formula

$$E_g(T) \approx 1.55 - 4.3 \times 10^{-4}\ T$$

where E_g is given in electron volts (eV).

(*a*) Using this expression, show that the temperature dependence of the intrinsic electron concentration n_i is

$$n_i = 5 \times 10^{15}\ T^{3/2}\ e^{-8991/T}$$

(*b*) Using a computer, plot the values of n_i as a function of temperature over the range $273\,\text{K} \le T \le 373\,\text{K}$ ($0°\text{C} \le T \le 100°\text{C}$).

4.2 Repeat the steps given in Example 4.2 for a *p*-type semiconductor. In particular, show that when the net acceptor concentration is much greater than n_i, we have $p_p = N_A$ and $n_p = n_i^2/N_A$.

4.3 An engineer has two $Ga_{1-x}Al_xAs$ LEDs: one has a bandgap energy of 1.540 eV and the other has x = 0.015.

(*a*) Find the aluminum mole fraction x and the emission wavelength for the first LED.

(*b*) Find the bandgap energy and the emission wavelength of the other LED.

4.4 The lattice spacing of $In_{1-x}Ga_xAs_yP_{1-y}$ has been shown to obey Vegard's law.[21] This states that for quaternary alloys of the form $A_{1-x}B_xC_yD_{1-y}$, where A and B are group III elements (e.g., Al, In, and Ga) and C and D are group V elements (e.g., As, P, and Sb), the lattice spacing $a(x, y)$ of the quaternary alloy can be approximated by

$$a(x, y) = xya(BC) + x(1 - y)\,a(BD)$$
$$+ (1 - x)\,ya(AC) + (1 - x)\,(1 - y)\,a(AD)$$

where the $a(IJ)$ are the lattice spacings of the binary compounds IJ.

(*a*) Show that for $In_{1-x}Ga_xAs_yP_{1-y}$ with

$$a(\text{GaAs}) = 5.6536\ \text{Å}$$
$$a(\text{GaP}) = 5.4512\ \text{Å}$$
$$a(\text{InAs}) = 6.0590\ \text{Å}$$
$$a(\text{InP}) = 5.8696\ \text{Å}$$

the quaternary lattice spacing becomes

$$a(x, y) = 0.1894y - 0.4184x$$
$$+ 0.0130xy + 5.8696\ \text{Å}$$

(*b*) For quaternary alloys that are lattice-matched to InP, the relation between x and y can be determined by letting $a(x, y) = a(\text{InP})$. Show that since $0 \le x \le 0.47$, the resulting expression can be approximated by $y \simeq 2.20x$.

(*c*) A simple empirical relation that gives the bandgap energy in terms of x and y is[21] $E_g(x, y) = 1.35 + 0.668x - 1.17y + 0.758x^2 + 0.18y^2 - 0.069xy - 0.322x^2y + 0.33xy^2$ eV. Find the bandgap energy and the peak emission wavelength of $In_{0.74}Ga_{0.26}As_{0.56}P_{0.44}$.

4.5 Using the expression $E = hc/\lambda$, show why the FWHM power spectral width of LEDs becomes wider at longer wavelengths.

4.6 A double-heterojunction InGaAsP LED emitting at a peak wavelength of 1310 nm has radiative and nonradiative recombination times of 25 and 90 ns, respectively. The drive current is 35 mA.

(*a*) Find the internal quantum efficiency and the internal power level.

(*b*) If the refractive index of the light source material is $n = 3.5$, find the power emitted from the device.

4.7 Assume the injected minority carrier lifetime of an LED is 5 ns and that the device has an optical output of 0.30 mW when a constant dc drive current is applied. Plot the optical output power when the LED is modulated at frequencies ranging from 20 to 100 MHz. Note what happens to the LED output power at higher modulation frequencies.

4.8 Consider an LED having a minority carrier lifetime of 5 ns. Find the 3-dB optical bandwidth and the 3-dB electrical bandwidth.

4.9 (*a*) A GaAlAs laser diode has a 500-μm cavity length, which has an effective absorption coefficient of 10 cm^{-1}. For uncoated facets the reflectivities are 0.32 at each end. What is the optical gain at the lasing threshold?

(*b*) If one end of the laser is coated with a dielectric reflector so that its reflectivity

is now 90 percent, what is the optical gain at the lasing threshold?

(c) If the internal quantum efficiency is 0.65, what is the external quantum efficiency in cases (a) and (b)?

4.10 Find the external quantum efficiency for a $Ga_{1-x}Al_xAs$ laser diode (with $x = 0.03$) that has an optical-power-versus-drive-current relationship of 0.5 mW/mA (e.g., as shown in Fig. 4.35).

4.11 Approximate expressions for the transverse and lateral optical-field confinement factors Γ_T and Γ_L, respectively, in a Fabry-Perot lasing cavity are

$$\Gamma_T = \frac{D^2}{2 + D^2} \text{ with } D = \frac{2\pi d}{\lambda}\left(n_1^2 - n_2^2\right)^{1/2}$$

and

$$\Gamma_L = \frac{W^2}{2 + W^2} \text{ with } W = \frac{2\pi w}{\lambda}\left(n_{\text{eff}}^2 - n_2^2\right)^{1/2}$$

where

$$n_{\text{eff}}^2 = n_2^2 + \Gamma_T\left(n_1^2 - n_2^2\right)$$

Here w and d are the width and thickness, respectively, of the active layer, and n_1 and n_2 are the refractive indices inside and outside the cavity, respectively.

(a) Consider a 1300-nm InGaAsP laser diode in which the active region is $0.1\ \mu m$ thick, $1.0\ \mu m$ wide, and $250\ \mu m$ long with refractive indices $n_1 = 3.55$ and $n_2 = 3.20$. What are the transverse and lateral optical-field confinement factors?

(b) Given that the total confinement factor is $\Gamma = \Gamma_T\Gamma_L$, what is the gain threshold if the effective absorption coefficient is $\overline{\alpha} = 30\ cm^{-1}$ and the facet reflectivities are $R_1 = R_2 = 0.31$?

4.12 A GaAs laser emitting at 800 nm has a $400\text{-}\mu m$ cavity length with a refractive index $n = 3.6$. If the gain g exceeds the total loss α_t throughout the range 750 nm $< \lambda <$ 850 nm, how many modes will exist in the laser?

4.13 A laser emitting at $\lambda_0 = 850$ nm has a gain-spectral width of $\sigma = 32$ nm and a peak gain of $g(0) = 50\ cm^{-1}$. Plot $g(\lambda)$ from Eq. 4.41. If $\alpha_t = 32.2\ cm^{-1}$, show the region where lasing takes place. If the laser is $400\ \mu m$ long and $n = 3.6$, how many modes will be excited in this laser?

4.14 The derivation of Eq. (4.46) assumes that the refractive index n is independent of wavelength.

(a) Show that when n depends on λ, we have

$$\Delta\lambda = \frac{\lambda^2}{2L(n - \lambda\, dn/d\lambda)}$$

(b) If the group refractive index $(n - \lambda\, dn/d\lambda)$ is 4.5 for GaAs at 850 nm, what is the mode spacing for a $400\text{-}\mu m$-long laser.

4.15 For laser structures that have strong carrier confinement, the threshold current density for stimulated emission J_{th} can to a good approximation be related to the lasing-threshold optical gain g_{th} by $g_{\text{th}} = \overline{\beta}J_{\text{th}}$ where $\overline{\beta}$ is a constant that depends on the specific device construction. Consider a GaAs laser with an optical cavity of length $250\ \mu m$ and width $100\ \mu m$. At the normal operating temperature, the gain factor $\overline{\beta} = 21 \times 10^{-3}$ A/cm^3 and the effective absorption coefficient $\overline{\alpha} = 10\ cm^{-1}$.

(a) If the refractive index is 3.6, find the threshold current density and the threshold current I_{th}. Assume the laser end faces are uncoated and the current is restricted to the optical cavity.

(b) What is the threshold current if the laser cavity width is reduced to $10\ \mu m$?

4.16 From quantum mechanics, the energy levels for electrons and holes in the quantum-well laser structure shown in Fig. 4.28 are given by

$$E_{ci} = E_c + \frac{h^2}{8d^2}\frac{i^2}{m_e} \text{ with } i = 1, 2, 3, \ldots \text{ for electrons}$$

and

$$E_{vj} = E_v - \frac{h^2}{8d^2}\frac{j^2}{m_h} \text{ with } j = 1, 2, 3, \ldots \text{ for holes}$$

where E_c and E_v are the conduction- and valence-band energies (see Fig. 4.1), d is the active layer thickness, h is Planck's constant, and m_e and m_h are the electron and hole masses as defined in Example 4.1. The possible energy-level transitions that lead to photon emission are given by

$$\Delta E_{ij} = E_{ci} - E_{vj} = E_g + \frac{h^2}{8d^2}\left(\frac{i^2}{m_e} + \frac{j^2}{m_v}\right)$$

If $E_g = 1.43$ eV for GaAs, what is the emission wavelength between the $i = j = 1$ states if the active layer thickness is $d = 5$ nm?

4.17 In a multiple quantum-well laser the temperature dependence of the differential or external quantum efficiency can be described by[57]

$$\eta_{\text{ext}}(T) = \eta_i(T)\frac{\alpha_{\text{end}}}{N_w[\alpha_w + \gamma(T - T_{\text{th}})] + \alpha_{\text{end}}}$$

where $\eta_i(T)$ is the internal quantum efficiency, α_{end} is the mirror loss of the lasing cavity as given in Eq. (4.28), N_w is the number of quantum wells, T_{th} is the threshold temperature, α_w is the internal loss of the wells at $T = T_{\text{th}}$ and γ is a temperature-dependent internal-loss parameter. Consider a six-well, 350-μm-long MQW laser having the following characteristics: $\alpha_w = 1.25$ cm^{-1}, $\gamma = 0.025$ cm^{-1}/K, and $T_{\text{th}} = 303$ K. The lasing cavity has a standard uncoated facet on the front ($R_1 = 0.31$) and a high-reflection coating on the near facet ($R_2 = 0.96$).

(*a*) Assuming that the internal quantum efficiency is constant, use a computer to plot the external quantum efficiency as a function of temperature over the range 303 K $\leq T \leq 375$ K. Let $\eta_{\text{ext}}(T) = 0.8$ at $T = 303$ K.

(*b*) Given that the optical output power at $T = 303$ K is 30 mW at a drive current of $I_d = 50$ mA, plot the power output as a function of temperature over the range 303 K $\leq T \leq 375$ K at this fixed bias current.

4.18 A distributed-feedback laser has a Bragg wavelength of 1570 nm, a second-order grating with $\Lambda = 460$ nm, and a 300-μm cavity length. Assuming a perfectly symmetrical DFB laser, find the zeroth-, first-, and second-order lasing wavelengths to a tenth of a nanometer. Draw a relative amplitude-versus-wavelength plot.

4.19 When a current pulse is applied to a laser diode, the injected carrier pair density n within the recombination region of width d changes with time according to the relationship

$$\frac{\partial n}{\partial t} = \frac{J}{qd} - \frac{n}{\tau}$$

(*a*) Assume τ is the average carrier lifetime in the recombination region when the injected carrier pair density is n_{th} near the threshold current density J_{th}. That is, in the steady state we have $\partial n/\partial t = 0$, so that

$$n_{\text{th}} = \frac{J_{\text{th}}\tau}{qd}$$

If a current pulse of amplitude I_p is applied to an unbiased laser diode, show that the time needed for the onset of stimulated emission is

$$t_d = \tau \ln\frac{I_p}{I_p - I_{\text{th}}}$$

Assume the drive current $I = JA$, where J is the current density and A is the area of the active region.

(*b*) If the laser is now prebiased to a current density $J_B = I_B/A$, so that the initial excess carrier pair density is $n_B = J_B\tau/qd$, then the current density in the active region during a current pulse I_p is $J = J_B + J_p$. Show that in this case Eq. (4.50) results.

4.20 A 3B4B code converts blocks of 3 bits to blocks of 4 bits according to the rules given in Table 4.3. When there are two or more consecutive blocks of three zeros, the coded binary blocks 0010 and 1101 are used alternately.

Table 4.3 *Conversion rules for a 3B4B code*

Original code word	3B4B code word	
	Mode 1	Mode 2
000	0010	1101
001	0011	
010	0101	
011	0110	
100	1001	
101	1010	
110	1100	
111	1011	0100

Similarly, the coded blocks 1011 and 0100 are used alternately for consecutive blocks of three ones.

(*a*) Using these translation rules, find the coded bit stream for the data input 01000 111111101000000001111110.

(*b*) What is the maximum number of consecutive identical bits in the coded pattern?

4.21 A 4B5B code has $2^4 = 16$ four-bit data characters. The code maps these characters into the 5-bit sequences listed in Table 4.4. Using this information, encode the following bit stream

010111010010111010100111

4.22 A laser diode has a maximum average output of 1 mW (0 dBm). The laser is to be amplitude-modulated with a signal $x(t)$ that has a dc component of 0.2 and a periodic component of ±2.56. If the current-input to optical-output relationship is $P(t) = i(t)/10$, find the values of I_0 and m if the modulating current is $i(t) = I_0 [1 + mx(t)]$.

4.23 Consider the following Taylor series expansion of the optical-power-versus-drive-current relationship of an optical source about a given bias point:

$$y(t) = a_1 x(t) + a_2 x^2(t) + a_3 x^3(t) + a_4 x^4(t)$$

Let the modulating signal $x(t)$ be the sum of two sinusoidal tones at frequencies ω_1 and ω_2

Table 4.4 *Data sequences used in 4B5B code conversion*

Data sequence	Encoded sequence
0000	11110
0001	01001
0010	10100
0011	10101
0100	01010
0101	01011
0110	01110
0111	01111
1000	10010
1001	10011
1010	10110
1011	10111
1100	11010
1101	11011
1110	11100
1111	11101

given by

$$x(t) = b_1 \cos \omega_1 t + b_2 \cos \omega_2 t$$

(*a*) Find the second-, third-, and fourth-order intermodulation distortion coefficients B_{mn} (where m and $n = \pm 1, \pm 2, \pm 3$, and ± 4) in terms of b_1, b_2, and the a_i.

(*b*) Find the second-, third-, and fourth-order harmonic distortion coefficients A_2, A_3 and A_4 in terms of b_1, b_2, and the a_i.

4.24 An optical source is selected from a batch characterized as having lifetimes that follow a slow internal degradation mode. The –3-dB mean time to failure of these devices at room temperature is specified as 5×10^4 h. If the device emits 1 mW at room temperature, what is the expected optical output power after 1 month of operation, after 1 year, and after 5 years?

4.25 A group of optical sources has operating lifetimes of 4×10^4 h at 60°C and 6500 h at 90°C. What is the expected lifetime at 20°C if the device lifetime follows an Arrhenius-type relationship?

REFERENCES

1. H. Kressel and J. K. Butler, *Semiconductor Lasers and Heterojunction LEDs*, Academic, New York, 1977.

2. T. P. Lee, C. A. Burrus Jr., and R. H. Saul, "Light emitting diodes for telecommunications," in S. E. Miller and I. P. Kaminow, eds., *Optical Fiber Telecommunications–II*, Academic, New York, 1988.

3. J. E. Bowers and M. A. Pollack, "Semiconductor lasers for telecommunications," in S. E. Miller and I. P. Kaminow, eds., *Optical Fiber Telecommunications–II*, Academic, New York, 1988.

4. L. A. Coldren and S. W. Corzine, *Diode Lasers and Photonic Integrated Circuits*, Wiley, New York, 1995.

5. M. Fukuda, *Optical Semiconductor Devices*, Wiley, Hoboken, NJ, 1999.

6. S. F. Yu, *Analysis and Design of Vertical Cavity Surface Emitting Lasers*, Wiley, Hoboken, NJ, 2003.

7. D. Meschede, *Optics, Light and Lasers: The Practical Approach to Modern Aspects of Photonics and Laser Physics*, Wiley, Hoboken, NJ, 2nd ed., 2007.

8. J. Buus, M. C. Amann, and D. J. Blumenthal, *Tunable Laser Diodes and Related Optical Sources*, Wiley-IEEE Press, Hoboken, NJ, 2nd ed., 2005.

9. R. S. Quimby, *Photonics and Lasers: An Introduction*, Wiley, Hoboken, NJ, 2006.

10. L. A. Coldren, G. A. Fish, Y. Akulova, J. S. Barton, L. Johansson, and C. W. Coldren, "Tunable semiconductor lasers," *J. Lightwave Tech.*, vol. 22, pp. 193–202, Jan. 2004.

11. J. Kani and K. Iwatsuki, "A wavelength-tunable optical transmitter using semiconductor optical amplifiers and an optical tunable filter for metro/access DWDM applications," *J. Lightwave Tech.,* vol. 23, pp. 1164–1169, Mar. 2005.

12. J. Buus and E. J. Murphy, "Tunable lasers in optical networks," *J. Lightwave Tech.*, vol. 24, pp. 5–11, Jan. 2006.

13. G. J. Pendock and D. D. Sampson, "Transmission performance of high bit rate spectrum-sliced WDM systems," *J. Lightwave Tech.*, vol. 14, pp. 2141–2148, Oct. 1996.

14. V. Arya and I. Jacobs, "Optical preamplifier receiver for spectrum-sliced WDM," *J. Lightwave Tech.*, vol. 15, pp. 576–583, Apr. 1997.

15. R. D. Feldman, "Crosstalk and loss in WDM systems employing spectral slicing," *J. Lightwave Tech.*, vol. 15, pp. 1823–1831, Nov. 1997.

16. B. L. Anderson and R. L. Anderson, *Fundamentals of Semiconductor Devices*, McGraw-Hill, New York, 2005.

17. D. A. Neamen, *An Introduction to Semiconductor Devices*, McGraw-Hill, New York, 2006.

18. S. O. Kasap, *Principles of Electronic Materials and Devices*, McGraw-Hill, New York, 2006.

19. O. Manasreh, *Semiconductor Heterojunctions and Nanostructures*, McGraw-Hill, New York, 2005.

20. C. A. Burrus and B. I. Miller, "Small-area double heterostructure AlGaAs electroluminescent diode sources for optical fiber transmission lines," *Opt. Commun.*, vol. 4, pp. 307–309, Dec. 1971.

21. R. E. Nahory, M. A. Pollack, W. D. Johnston Jr., and R. L. Barns, "Band gap versus composition and demonstration of Vegard's law for InGaAsP lattice matched to InP," *Appl. Phys. Lett.*, vol. 33, pp. 659–661, Oct. 1978.

22. B.E.A. Saleh and M. Teich, *Fundamentals of Photonics*, Wiley, Hoboken, NJ, 2nd ed., 2007.

23. T. P. Lee and A. J. Dentai, "Power and modulation bandwidth of GaAs-A1GaAs high radiance LEDs for optical communication systems," *IEEE J. Quantum Electron.*, vol. QE-14, pp. 150–159, Mar. 1978.

24. H. Namizaki, M. Nagano, and S. Nakahara, "Frequency response of GaA1As light emitting diodes," *IEEE Trans. Electron. Devices*, vol. ED-21, pp. 688–691, Nov. 1974.

25. Y. S. Liu and D. A. Smith, "The frequency response of an amplitude modulated GaAs luminescent diode," *Proc. IEEE*, vol. 63, pp. 542–544, Mar. 1975.

26. T. P. Lee, "Effects of junction capacitance on the rise time of LEDs and the turn-on delay of injection lasers," *Bell Sys. Tech. J.*, vol. 54, pp. 53–68, Jan. 1975.

27. I. Hino and K. Iwamoto, "LED pulse response analysis," *IEEE Trans. Electron. Devices*, vol. ED-26, pp. 1238–1242, Aug. 1979.

28. G. Morthier and P. Vankwikelberge, *Handbook of Distributed Feedback Lasers*, Artech House, Boston, 1997.

29. H. Ghafouri-Shiraz, *Distributed Feedback Laser Diodes and Optical Tunable Filters*, Wiley, Hoboken, NJ, 2004.

30. D. Botez, "Laser diodes are power-packed," *IEEE Spectrum*, vol. 22, pp. 43–53, June 1985.

31. C. Y. Tsai, F.-P. Shih, T.-L. Sung, T.-Y. Wu, C. H. Chen., and C. Y. Tsai, "A small-signal analysis of the modulation response of high-speed quantum-well lasers," *IEEE J. Quantum Electron.*, vol. 33, pp. 2084–2096, Nov. 1997.

32. A. Thränhardt, S. W. Koch, J. Hader, and J. V. Moloney, "Carrier dynamics in quantum well lasers," *Optical and Quantum Electronics*, vol. 38, pp. 361–368, Mar. 2006.

33. T. P. Lee, C. A. Burrus, R. A. Linke, and R. J. Nelson, "Short-cavity, single-frequency InGaAsP buried heterostructure lasers," *Electron. Lett.*, vol. 19, pp. 82–84, Feb. 1983.

34. Hongling Rao, M. J. Steel, R. Scarmozzino, and R. M. Osgood Jr., "VCSEL design using the bidirectional beam-propagation method," *IEEE J. Quantum Electron.*, vol. 37, pp. 1435–1440, Nov. 2001.

35. C. J. Chang-Hasnain, "Tunable VCSEL," *IEEE J. Sel. Topics Quantum Electron.*, vol. 6, pp. 978–987, Nov.–Dec. 2000.

36. H. C. Kuo, Y. H. Chang, Y.-A. Chang, F.-I. Lai, J. T. Chu, M. Y. Tsai, and S. C. Wang, "Single-mode 1.27-μm InGaAs:Sb-GaAs-GaAsP quantum well vertical cavity surface emitting lasers," *IEEE J. Sel. Topics .Quantum Electron.*, vol. 11, pp. 121–126, Jan.–Feb. 2005.

37. G. M. Smith, J. S. Hughes, R. M. Lammert, M. L. Osowski, G. C. Papen, J. T. Verdeyen, and J. J. Coleman, "Very narrow linewidth asymmetric cladding InGaAs-GaAs ridge waveguide distributed Bragg reflector lasers," *IEEE Photonics Tech. Lett.*, vol. 8, pp. 476–478, Apr. 1996.

38. C. H. Henry, "Theory of the linewidth of semiconductor lasers," *IEEE J. Quantum Electron.*, vol. QE-18, pp. 259–264, Feb. 1982.

39. K. Kojima, K. Kyuma, and T. Nakayama, "Analysis of the spectral linewidth of distributed feedback laser diodes," *J. Lightwave Tech.*, vol. LT-3, pp. 1048–1055, Oct. 1985.

40. J. Kim, Hui Su, S. Minin, and S.-L. Chuang, "Comparison of linewidth enhancement factor between p-doped and undoped quantum-dot lasers," *IEEE Photonics Tech. Lett.*, vol. 18, pp. 1022–1024, May 2006.

41. E. L. Wooten, K. M. Kissa, A. Yi-Yan, E. J. Murphy, D. A. Lafaw, P. F. Hallemeier, D. Maack, D. V. Attanasio, D. J. Fritz, G. J. McBrien, and D. E. Bossi, "A review of lithium niobate modulators for fiber-optic communication systems," *IEEE J. Sel. Topics Quantum Electron.*, vol. 6, pp. 69–82, Jan./Feb. 2000.

42. G. L. Li and P.K.L. Yu, "Optical intensity modulators for digital and analog applications," *J. Lightwave Tech.*, vol. 21, pp. 2010–2030, Sept. 2003.

43. M. Labudovic and J. Li, "Modeling of TE cooling of pump lasers," *IEEE Trans. Compon. Packag. Technol.*, vol. 27, pp. 724–730, Dec. 2004.

44. A. Leon-Garcia and I. Widjaja, *Communication Networks*, McGraw-Hill, Burr Ridge, IL, 2nd ed., 2004.

45. A. J. Jerri, "The Shannon sampling theory—its various extensions and applications: A tutorial review," *Proc. IEEE*, vol. 65, pp. 1565–1596, Nov. 1977.

46. G. Keiser, *Local Area Networks*, McGraw-Hill, New York, 2nd ed., 2002.

47. E. Forestieri and G. Prati, "Novel optical line codes tolerant to fiber chromatic dispersion," *J. Lightwave Tech.*, vol. 19, pp. 1675–1684, Nov. 2001.

48. H. Bourdoucen, "Effect of line codes and WDM wavelengths on the performance of transmission systems based on the Mach-Zehnder modulator," *J. Opt. Networking*, vol. 3, pp. 92–99, Feb. 2004.

49. A. W. Moore, L. B. James, M. Glick, A. Wonfor, R. G. Plumb, I. H. White, D. McAuley, and R. V. Penty, "Optical network packet error rate due to physical layer coding," *J. Lightwave Tech.*, vol. 23, pp. 3056–3065, Oct. 2005.

50. S. Yamakoshi, O. Hasegawa, H. Hamaguchi, M. Abe, and T. Yamaoka, "Degradation of high-radiance $Ga_{1-x}Al_xAs$ LEDs," *Appl. Phys. Lett.*, vol. 31, pp. 627–629, Nov. 1977.

51. M. Ettenberg and H. Kressel, "The reliability of (AlGa)As CW laser diodes," *IEEE J. Quantum Electron.*, vol. QE-16, pp. 186–196, Feb. 1980.

52. S. L. Chuang, A. Ishibashi, S. Kijima, N. Nakajama, M. Ukita, and S. Taniguchi, "Kinetic model for degradation of light-emitting diodes," *IEEE J. Quantum Electron.*, vol. 33, pp. 970–979, June 1997.

53. J.-H. Han and S. W. Park, "Theoretical and experimental study on junction temperature of packaged Fabry-Perot laser diode," *IEEE Trans. Device Materials Reliability*, vol. 4, pp. 292–294, June 2004.

54. M. Funabashi, H. Nasu, T. Mukaihara, T. Kimoto, T. Shinagawa, T. Kise, K. Takaki, T. Takagi, M. Oike, T. Nomura, and A. Kasukawa, "Recent advances in DFB lasers for ultradense WDM applications," *IEEE J. Selected Topics Quantum Electron.*, vol. 10, pp. 312–320, Mar./Apr. 2004.

55. R. Sobiestianskas, J. G. Simmons, G. Letal, and R. E. Mallard, "Experimental study on the intrinsic response, optical and electrical parameters of 1.55 μm DFB BH laser diodes during aging tests," *IEEE Trans. Device Materials Reliability*, vol. 5, pp. 659–664, Dec. 2005.

56. L. A. Johnson, "Laser diode burn-in and reliability testing," *IEEE Commun. Mag.*, vol. 44, pp. 4–7, Feb. 2006.

57. S. Seki and K. Yokahama, "Power penalty in 1.3-μm InP-based strained-layer multiple-quantum-well lasers at elevated temperatures," *IEEE Photonics Tech. Lett.*, vol. 9, pp. 1205–1207, Sept. 1997.

CHAPTER 5

Power Launching and Coupling

In implementing an optical fiber link, two of the major system questions are how to launch optical power into a particular fiber from some type of luminescent source and how to couple optical power from one fiber into another. Launching optical power from a source into a fiber entails considerations such as the numerical aperture, core size, refractive-index profile, and core-cladding index difference of the fiber, plus the size, radiance, and angular power distribution of the optical source.

A measure of the amount of optical power emitted from a source that can be coupled into a fiber is usually given by the *coupling efficiency η* defined as

$$\eta = \frac{P_F}{P_S}$$

Here, P_F is the power coupled into the fiber and P_S is the power emitted from the light source. The launching or coupling efficiency depends on the type of fiber that is attached to the source and on the coupling process; for example, whether or not lenses or other coupling improvement schemes are used.

In practice, many source suppliers offer devices with a short length of optical fiber (1 m or less) already attached in an optimum power-coupling configuration. This section of fiber is generally referred to as a *flylead* or a *pigtail*. The power-launching problem for these pigtailed sources thus reduces to a simpler one of coupling optical power from one fiber into another. The effects to be considered in this case include fiber misalignments; different core sizes, numerical apertures, and core refractive-index profiles; plus the need for clean and smooth fiber end faces that are either perpendicular to the axis or that are polished at a slight angle to prevent back reflections.

An alternate arrangement consists of light sources and optical fiber receptacles that are integrated within a transceiver package. To achieve fiber-to-fiber coupling in this case, the fiber connector from a cable is simply mated to the built-in connector in the transceiver package. Among the various commercially available configurations are the popular *small-form-factor* (SFF) and the SFF pluggable (SFP) devices.

5.1 | Source-to-Fiber Power Launching

A convenient and useful measure of the optical output of a luminescent source is its radiance (or brightness) *B* at a given diode drive current. *Radiance* is the optical power radiated into a unit solid angle per unit emitting surface area and is generally specified in terms of watts per square centimeter per steradian. Since the optical power that can be coupled into a fiber depends on the radiance (i.e., on the spatial distribution of the optical power), the radiance of an optical source rather than the total output power is the important parameter when considering source-to-fiber coupling efficiencies.

5.1.1 Source Output Pattern

To determine the optical power-accepting capability of a fiber, the spatial radiation pattern of the source must first be known. This pattern can be fairly complex. Consider Fig. 5.1, which shows a spherical coordinate system characterized by R, θ, and ϕ, with the normal to the emitting surface being the polar axis. The radiance may be a function of both θ and ϕ, and can also vary from point to point on the emitting surface. A reasonable assumption for simplicity of analysis is to take the emission to be uniform across the source area.

Surface-emitting LEDs are characterized by their lambertian output pattern, which means the source is equally bright when viewed from any direction. The power delivered at an angle θ, measured relative to a normal to the emitting surface, varies as $\cos\theta$ because the projected area of the emitting surface varies as $\cos\theta$ with viewing direction. The emission pattern for a lambertian source thus follows the relationship

$$B(\theta, \phi) = B_0 \cos\theta \qquad (5.1)$$

where B_0 is the radiance along the normal to the radiating surface. The radiance pattern for this source is shown in Fig. 5.2.

Edge-emitting LEDs and laser diodes have a more complex emission pattern. These devices have different
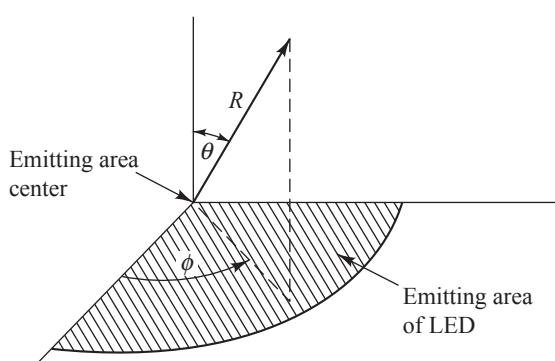

Fig. 5.1 *Spherical coordinate system for characterizing the emission pattern from an optical source*
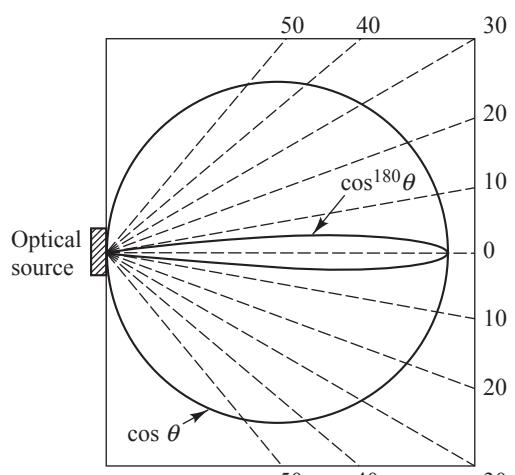

Fig. 5.2 *Radiance patterns for a lambertian source and the lateral output of a highly directional laser diode. Both sources have B_0 normalized to unity.*

radiances $B(\theta, 0°)$ and $B(\theta, 90°)$ in the planes parallel and normal, respectively, to the emitting-junction plane of the device. These radiances can be approximated by the general form[1,2]

$$\frac{1}{B(\theta, \phi)} = \frac{\sin^2 \phi}{B_0 \cos^T \theta} + \frac{\cos^2 \phi}{B_0 \cos^L \theta}$$ (5.2)

The integers T and L are the transverse and lateral power distribution coefficients, respectively. In general, for edge emitters, $L = 1$ (which is a lambertian distribution with a 120° half-power beam width) and T is significantly larger. For laser diodes, L can take on values over 100.

Example 5.1   Figure 5.2 compares a lambertian pattern with a laser diode that has a lateral ($\phi = 0°$) half-power beam width of $2\theta = 10°$. What is the lateral power distribution coefficient?

Solution: From Eq. (5.2), we have

$$B(\theta = 5°, \phi = 0°) = B_0 (\cos 5°)^L = \frac{1}{2} B_0$$

Solving for L, we have

$$L = \frac{\log 0.5}{\log(\cos 5°)} = \frac{\log 0.5}{\log 0.9962} = 182$$

The much narrower output beam from a laser diode allows significantly more light to be coupled into an optical fiber.

5.1.2   Power-Coupling Calculation

To calculate the maximum optical power coupled into a fiber, consider first the case shown in Fig. 5.3 for a symmetric source of brightness $B(A_s, \Omega_s)$, where A_s and Ω_s are the area and solid emission angle of the source, respectively. Here, the fiber end face is centered over the emitting surface of the source and is positioned as close to it as possible. The coupled power can be found using the relationship

$$P = \int_{A_f} dA_s \int_{\Omega_f} d\Omega_s\, B(A_s, \Omega_s)$$

$$= \int_0^{r_m} \int_0^{2\pi} \left[\int_0^{2\pi} \int_0^{\theta_A} B(\theta, \phi) \sin \theta\, d\theta\, d\phi \right] d\theta_s\, r\, dr$$ (5.3)
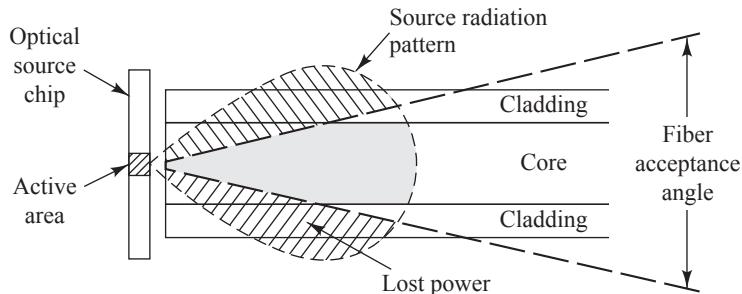

Fig. 5.3   *Schematic diagram of a light source coupled to an optical fiber. Light outside of the acceptance angle is lost.*

where the area A_f and solid acceptance angle Ω_f of the fiber define the limits of the integrals. In this expression, first the radiance $B(\theta, \phi)$ from an individual radiating point source on the emitting surface is integrated over the solid acceptance angle of the fiber. This is shown by the expression in square brackets, where θ_A is the acceptance angle of the fiber, which is related to the numerical aperture NA through Eq. (2.23). The total coupled power is then determined by summing up the contributions from each individual emitting-point source of incremental area $d\theta_s \, r \, dr$, that is, integrating over the emitting area. For simplicity, here the emitting surface is taken as being circular. If the source radius r_s is less than the fiber-core radius a, then the upper integration limit $r_m = r_s$; for source areas larger than the fiber-core area, $r_m = a$.

   As an example, assume a surface-emitting LED of radius r_s less than the fiber-core radius a. Since this is a lambertian emitter, Eq. (5.1) applies and Eq. (5.3) becomes

$$P = \int_0^{r_s} \int_0^{2\pi} \left(2\pi B_0 \int_0^{\theta_A} \cos\theta \sin\theta \, d\theta \right) d\theta_s \, r \, dr$$

$$= \pi B_0 \int_0^{r_s} \int_0^{2\pi} \sin^2\theta_A \, d\theta_s \, r \, dr$$

$$= \pi B_0 \int_0^{r_s} \int_0^{2\pi} NA^2 \, d\theta_s \, r \, dr \qquad (5.4)$$

where the numerical aperture NA is defined by Eq. (2.23). For step-index fibers the numerical aperture is independent of the positions θ_s and r on the fiber end face, so that Eq. (5.4) becomes (for $r_s < a$)

$$P_{\text{LED, step}} = \pi^2 r_s^2 B_0 (NA)^2 \simeq 2\pi^2 r_s^2 B_0 n_1^2 \Delta \qquad (5.5)$$

   Consider now the total optical power P_s that is emitted from the source of area A_s into a hemisphere (2π sr). This is given by

$$P_s = A_s \int_0^{2\pi} \int_0^{\pi/2} B(\theta, \phi) \sin\theta \, d\theta \, d\phi$$

$$= \pi r_s^2 \, 2\pi B_0 \int_0^{\pi/2} \cos\theta \sin\theta \, d\theta$$

$$= \pi^2 r_s^2 B_0 \qquad (5.6)$$

Equation (5.5) can, therefore, be expressed in terms of P_s:

$$P_{\text{LED, step}} = P_s (NA)^2 \quad \text{for } r_s \leq a \qquad (5.7)$$

When the radius of the emitting area is larger than the radius a of the fiber-core area, Eq. (5.7) becomes

$$P_{\text{LED, step}} = \left(\frac{a}{r_s} \right)^2 P_s (NA)^2 \quad \text{for } r_s > a \qquad (5.8)$$

Example 5.2 Consider an LED that has a circular emitting area of radius 35 μm and a lambertian emission pattern with 150 W/(cm^2·sr) axial radiance at a given drive current. Compare the optical powers coupled into two step-index fibers, one of which has a core radius of 25 μm with NA = 0.20 and the other which has a core radius of 50 μm with NA = 0.20.

Solution: For the larger core fiber, we use Eqs (5.6) and (5.7) to get

$$P_{\text{LED, step}} = P_s(\text{NA})^2 = \pi^2 r_s^2 B_0(\text{NA})^2$$

$$= \pi^2 (0.0035 \text{ cm})^2 \, [150 \text{ W/(cm}^2 \cdot \text{sr)}] \, (0.20)^2$$

$$= 0.725 \text{ mW}$$

For the case when the fiber end-face area is smaller than the emitting surface area, we use Eq. (5.8). Thus the coupled power is less than the above case by the ratio of the radii squared:

$$P_{\text{LED, step}} = \left(\frac{25 \, \mu\text{m}}{35 \, \mu\text{m}} \right)^2 P_s(\text{NA})^2$$

$$= \left(\frac{25 \, \mu\text{m}}{35 \, \mu\text{m}} \right)^2 (0.725 \text{ mW})$$

$$= 0.37 \text{ mW}$$

In the case of a graded-index fiber, the numerical aperture depends on the distance r from the fiber axis through the relationship defined by Eq. (2.80). Thus using Eqs (2.80a) and (2.80b), the power coupled from a surface-emitting LED into a graded-index fiber becomes (for $r_s < a$)

$$P_{\text{LED, graded}} = 2\pi^2 B_0 \int_0^{r_s} [n^2(r) - n_2^2] r \, dr$$

$$= 2\pi^2 r_s^2 B_0 \, n_1^2 \Delta \left[1 - \frac{2}{\alpha + 2} \left(\frac{r_s}{a} \right)^\alpha \right]$$

$$= 2P_s \, n_1^2 \Delta \left[1 - \frac{2}{\alpha + 2} \left(\frac{r_s}{a} \right)^\alpha \right] \tag{5.9}$$

where the last expression was obtained from Eq. (5.6).

These analyses assumed perfect coupling conditions between the source and the fiber. This can be achieved only if the refractive index of the medium separating the source and the fiber end matches the index n_1 of the fiber core. If the refractive index n of this medium is different from n_1, then, for perpendicular fiber end faces, the power coupled into the fiber reduces by the factor

$$R = \left(\frac{n_1 - n}{n_1 + n} \right)^2 \tag{5.10}$$

where R is the *Fresnel reflection* or the *reflectivity* at the fiber-core end face. The ratio $r = (n_1 - n)/(n_1 + n)$, which is known as the *reflection coefficient,* relates the amplitude of the reflected wave to the amplitude of the incident wave.

The calculation of power coupling for nonlambertian emitters following a cylindrical cos$^m \theta$ distribution is left as an exercise. The power launched into a fiber from an edge-emitting LED that has a noncylindrical distribution is rather complex.[3] Section 5.4 presents a simplified analysis of this coupling problem in the discussion on coupling LEDs to single-mode fibers.

Example 5.3 A GaAs optical source with a refractive index of 3.6 is coupled to a silica fiber that has a refractive index of 1.48. What is the power loss between the source and the fiber?

Solution: If the fiber end and the source are in close physical contact, then, from Eq. (5.10), the Fresnel reflection at the interface is

$$R = \left(\frac{n_1 - n}{n_1 + n} \right)^2 = \left(\frac{3.60 - 1.48}{3.60 + 1.48} \right)^2 = 0.174$$

This value of R corresponds to a reflection of 17.4 percent of the emitted optical power back into the source.

Given that

$$P_{\text{coupled}} = (1 - R)P_{\text{emitted}}$$

the power loss L in decibels is found from:

$$L = -10 \, \log \left(\frac{P_{\text{coupled}}}{P_{\text{emitted}}} \right) = -10 \log(1 - R)$$

$$= -10 \log(0.826) = 0.83 \text{ dB}$$

This number can be reduced by having an index-matching material between the source and the fiber end.

Example 5.4 An InGaAsP optical source that has a refractive index of 3.540 is closely coupled to a step-index fiber that has a core refractive index of 1.480. Assume that the source size is smaller than the fiber core and that the small gap between the source and the fiber is filled with a gel that has a refractive index of 1.520. (*a*) What is the power loss in decibels from the source into the fiber? (*b*) What is the power loss if no gel is used?

Solution: (*a*) Here we need to consider the reflectivity at two interfaces. First, using Eq. (5.10) we have that the reflectivity R_{sg} at the source-to-gel interface is

$$R_{sg} = \left(\frac{3.540 - 1.520}{3.540 + 1.520} \right)^2 = 0.159$$

Similarly, using Eq. (5.10) we have that the reflectivity R_{gf} at the gel-to-fiber interface is

$$R_{gf} = \left(\frac{1.480 - 1.520}{1.480 + 1.520} \right)^2 = 0.040$$

The total reflectivity then is $R = R_{sg} \times R_{gf} = (0.159) \times (0.040) = 0.0064$.

The power loss in decibels is (see Example 5.3)

$$L = -10 \log (1 - R) = -10 \log (0.994) = 0.0028 \text{ dB}$$

(*b*) If no index-matching gel is used, and if we assume there is no gap between the source and the fiber, then from Eq. (5.10) we have that the reflectivity is

$$R = \left(\frac{3.540 - 1.480}{3.540 + 1.480} \right)^2 = 0.168$$

In this case the power loss in decibels is

$$L = -10 \log (1 - R) = -10 \log (0.832) = 0.799 \text{ dB}$$

5.1.3 Power Launching versus Wavelength

It is of interest to note that the optical power launched into a fiber does not depend on the wavelength of the source but only on its brightness: that is, its radiance. To explore this concept a little further, Eq. (2.81) shows that the number of modes that can propagate in a multimode graded-index fiber of core size a and index profile α is

$$M = \frac{\alpha}{\alpha + 2} \left(\frac{2\pi a n_1}{\lambda} \right)^2 \Delta \tag{5.11}$$

Thus, for example, twice as many modes propagate in a given fiber at 900 nm than at 1300 nm.

The radiated power per mode, P_s/M, from a source at a particular wavelength is given by the radiance multiplied by the square of the nominal source wavelength,[4]

$$\frac{P_s}{M} = B_0 \lambda^2 \tag{5.12}$$

Thus twice as much power is launched into a given mode at 1300 nm than at 900 nm. Hence, two identically sized sources operating at different wavelengths but having identical radiances will launch equal amounts of optical power into the same fiber.

5.1.4 Equilibrium Numerical Aperture

As noted earlier, a light source may be supplied with a short (1- to 2-m) fiber flylead attached to it in order to facilitate coupling the source to a system fiber. To achieve a low coupling loss, this flylead should be connected to a system fiber that has a nominally identical NA and core diameter. A certain amount of optical power (ranging from 0.1 to 1 dB) is lost at this junction, the exact loss depending on the connecting mechanism and on the fiber type; this is discussed in Sec. 5.3. In addition to the coupling loss, an excess power loss will occur in the first few tens of meters of a multimode system fiber. This excess loss is a result of nonpropagating modes scattering out of the fiber as the launched modes come to an equilibrium condition. This loss is of particular importance for surface-emitting LEDs, which tend to launch power into all modes of the fiber. Fiber-coupled lasers are less prone to this effect because they tend to excite fewer nonpropagating fiber modes.

The excess power loss must be analyzed carefully in any system design as it can be significantly higher for some types of fibers than for others.[5] An example of the excess power loss is shown in Fig. 5.4 in terms of the fiber numerical aperture. At the input end of the fiber, the light acceptance is described in terms of the launch numerical aperture NA_{in}. If the light-emitting area of the LED is less than the cross-sectional area of the fiber core, then, at this point, the power coupled into the fiber is given by Eq. (5.7), where $NA = NA_{in}$.

However, when the optical power is measured in long multimode fibers after the launched modes have come to equilibrium (which is often taken to occur at 50 m), the effect of the equilibrium numerical aperture NA_{eq} becomes apparent. At this point, the optical power in the fiber scales as
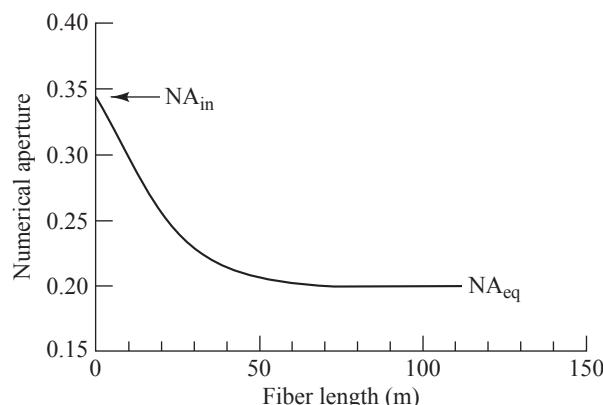

Fig. 5.4 *Example of the change in NA as a function of multimode fiber length*

$$P_{eq} = P_{50} \left(\frac{NA_{eq}}{NA_{in}} \right)^2 \qquad (5.13)$$

where P_{50} is the power expected in the fiber at the 50-m point based on the launch NA. The degree of mode coupling occurring in a fiber is primarily a function of the core-cladding index difference. It can thus vary significantly among different fiber types. Since most optical fibers attain 80–90 percent of their equilibrium NA after about 50 m, it is the value of NA_{eq} that is important when calculating launched optical power in multimode fibers.

5.2 Lensing Schemes for Coupling Improvement

The optical power-launching analysis given in Sec. 5.1 is based on centering a flat fiber end face directly over the light source as close to it as possible. If the source-emitting area is larger than the fiber-core area, then the resulting optical power coupled into the fiber is the maximum that can be achieved. This is a result of fundamental energy and radiance conservation principles[6] (also known as the *law of brightness*). However, if the emitting area of the source is smaller than the core area, a miniature lens may be placed between the source and the fiber to improve the power-coupling efficiency.

The function of the microlens is to magnify the emitting area of the source to match the core area of the fiber end face exactly. If the emitting area is increased by a magnification factor M, the solid angle within which optical power is coupled to the fiber from the source is increased by the same factor.

Several possible lensing schemes[1,2,7–12] are shown in Fig. 5.5. These include a rounded-end fiber, a small glass sphere (nonimaging microsphere) in contact with both the fiber and the source, a larger spherical lens used to image the source on the core area of the fiber end, a cylindrical lens generally formed from a short section of fiber, a system consisting of a spherical-surfaced LED and a spherical-ended fiber, and a taper-ended fiber.

Although these techniques can improve the source-to-fiber coupling efficiency, they also create additional complexities. One problem is that the lens size is similar to the source and fiber-core
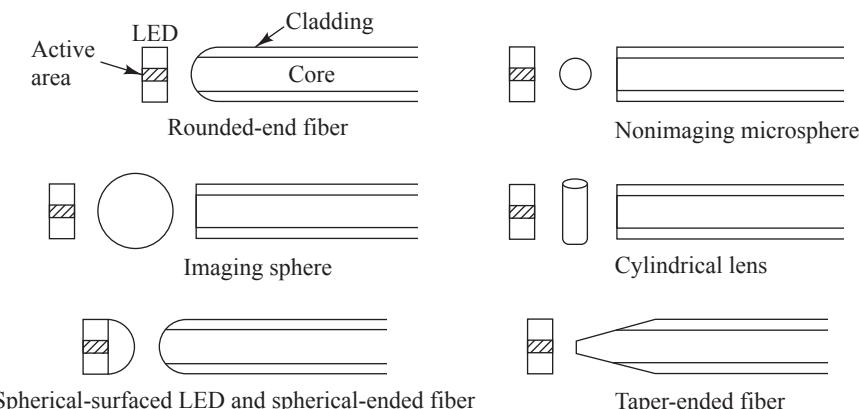

Fig. 5.5 *Examples of possible lensing schemes used to improve optical source-to-fiber coupling efficiency*

dimensions, which introduces fabrication and handling difficulties. In the case of the taper-ended fiber, the mechanical alignment must be carried out with greater precision since the coupling efficiency becomes a more sharply peaked function of the spatial alignment. However, alignment tolerances are increased for other types of lensing systems.
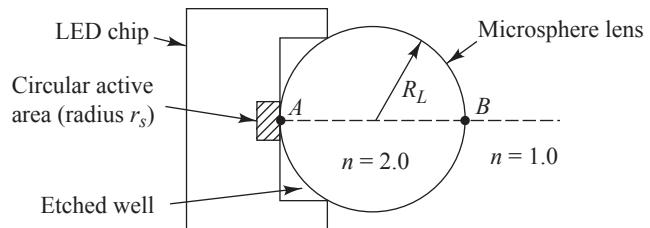

Fig. 5.6 *Schematic diagram of an LED emitter with a microsphere lens*

5.2.1 Nonimaging Microsphere

One of the most efficient lensing methods is the use of a nonimaging microsphere. Its use for a surface emitter is shown in Fig. 5.6. We first make the following practical assumptions: the spherical lens has a refractive index of about 2.0, the outside medium is air ($n = 1.0$), and the emitting area is circular. To collimate the output from the LED, the emitting surface should be located at the focal point of the lens. The focal point can be found from the gaussian lens formula[13]

$$\frac{n}{s} + \frac{n'}{q} = \frac{n' - n}{r} \tag{5.14}$$

where s and q are the object and image distances, respectively, as measured from the lens surface, n is the refractive index of the lens, n' is the refractive index of the outside medium, and r is the radius of curvature of the lens surface.

The following sign conventions are used with Eq. (5.14):
  1. Light travels from left to right.
  2. Object distances are measured as positive to the left of a vertex and negative to the right.
  3. Image distances are measured as positive to the right of a vertex and negative to the left.
  4. All convex surfaces encountered by the light have a positive radius of curvature, and concave surfaces have a negative radius.

Example 5.5 Using the sign conventions for Eq. (5.14), find the focal point for the right-hand surface shown in Fig. 5.6.

Solution: To find the focal point, we set $q = \infty$ and solve for s in Eq. (5.14), where s is measured from point B. With $n = 2.0$, $n' = 1.0$, $q = \infty$, and $r = -R_L$, Eq. (5.14) yields

$$S = f = 2R_L$$

Thus the focal point is located on the lens surface at point A.

Placing the LED close to the lens surface thus results in a magnification M of the emitting area. This is given by the ratio of the cross-sectional area of the lens to that of the emitting area:

$$M = \frac{\pi R_L^2}{\pi r_s^2} = \left(\frac{R_L}{r_s}\right)^2 \tag{5.15}$$

Using Eq. (5.4) one can show that, with the lens, the optical power P_L that can be coupled into a full aperture angle 2θ is given by

$$P_L = P_s \left(\frac{R_L}{r_s} \right)^2 \sin^2 \theta \qquad (5.16)$$

where P_s is the total output power from the LED without the lens.

The theoretical coupling efficiency that can be achieved is based on energy and radiance conservation principles.[14] This efficiency is usually determined by the size of the fiber. For a fiber of radius a and numerical aperture NA, the maximum coupling efficiency η_{max} for a lambertian source is given by

$$\eta_{max} = \begin{cases} \left(\dfrac{a}{r_s} \right)^2 (\text{NA})^2 & \text{for} \quad \dfrac{r_s}{a} > 1 \\[3mm] (\text{NA})^2 & \text{for} \quad \dfrac{r_s}{a} \leq 1 \end{cases} \qquad (5.17)$$

Thus when the radius of the emitting area is larger than the fiber radius, no improvement in coupling efficiency is possible with a lens. In this case, the best coupling efficiency is achieved by a direct-butt method.

Based on Eq. (5.17), the theoretical coupling efficiency as a function of the emitting diameter is shown in Fig. 5.7 for a fiber with a 50-μm core diameter.

Example 5.6 An optical source with a circular output pattern is closely coupled to a step-index fiber that has a numerical aperture of 0.22. If the source radius is $r_s = 50~\mu$m and the fiber core radius $a = 25~\mu$m, what is the maximum coupling efficiency from the source into the fiber?

Solution: Since the ratio $r_s/a > 1$, we can find the maximum coupling efficiency η_{max} from the top expression in Eq. (5.17):

$$\eta_{max} = \left(\frac{a}{r_s} \right)^2 (\text{NA})^2 = \left(\frac{25}{50} \right)^2 (.22)^2$$

$$= 0.25(.22)^2 = 0.012 = 1.2\%$$

Thus the coupling efficiency is reduced to 25 percent compared to the case in which the source and fiber radii are equal.

5.2.2 Laser Diode-to-Fiber Coupling

As we noted in Chapter 4, edge-emitting laser diodes have an emission pattern that nominally has a full width at half-maximum (FWHM) of 30–50° in the plane perpendicular to the active-area junction and an FWHM of 5–10° in the plane parallel to the junction. Since the angular output distribution of the laser is greater than the fiber acceptance angle, and since the laser emitting area is much smaller than the fiber core, spherical or cylindrical lenses or optical fiber tapers[15–19] also can be used to improve the coupling efficiency between edge-emitting laser diodes and optical fibers. This also works well for vertical-cavity surface-emitting lasers (VCSELs). Here, coupling efficiencies to multimode fibers of 35 percent result

for mass-produced connections of laser arrays to parallel optical fibers, and efficiencies of up to 90 percent are possible by direct (lensless) coupling from a single VCSEL source to a multimode fiber.[20]

The use of homogeneous glass microsphere lenses has been tested in a series of several hundred laser diode assemblies.[15] Spherical glass lenses with a refractive index of 1.9 and diameters ranging between 50 and 60 μm were epoxied to the ends of 50-μm core-diameter graded-index fibers having a numerical aperture of 0.2. The measured FWHM values of the laser output beams were as follows:

1. Between 3 and 9 μm for the near-field parallel to the junction.
2. Between 30 and 60° for the field perpendicular to the junction.
3. Between 15 and 55° for the field parallel to the junction.

Coupling efficiencies in these experiments ranged between 50 and 80 percent.
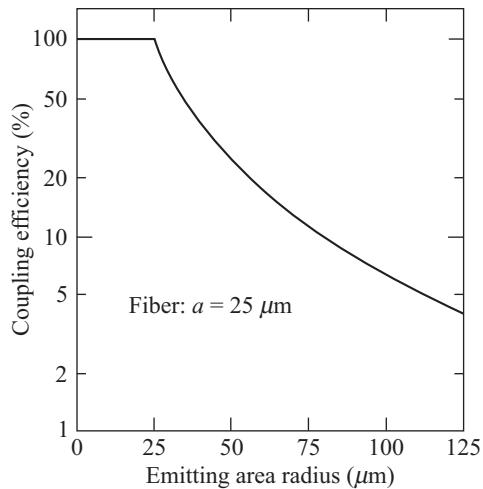
Fig. 5.7 *Theoretical coupling efficiency in units of (NA)² for a surface-emitting LED as a function of the light emitting diameter. Coupling is to a fiber with a core radius a = 25 μm.*

5.3 Fiber-to-Fiber Joints

A significant factor in any fiber optic system installation is the requirement to interconnect fibers in a low-loss manner. These interconnections occur at the optical source, at the photodetector, at intermediate points within a cable where two fibers are joined, and at intermediate points in a link where two cables are connected. The particular technique selected for joining the fibers depends on whether a permanent bond or an easily demountable connection is desired. A permanent bond is generally referred to as a *splice,* whereas a demountable joint is known as a *connector.*

Every joining technique is subject to certain conditions that can cause various amounts of optical power loss at the joint. The loss at a particular junction or through a component is called the *insertion loss.* These losses depend on parameters such as the input power distribution to the joint, the length of the fiber between the optical source and the joint, the geometrical and waveguide characteristics of the two fiber ends at the joint, and the fiber end-face qualities.

The optical power that can be coupled from one fiber to another is limited by the number of modes that can propagate in each fiber. For example, if a fiber in which 500 modes can propagate is connected to a fiber in which only 400 modes can propagate, then, at most, 80 percent of the optical power from the first fiber can be coupled into the second fiber (if we assume that all modes are equally excited). For a graded-index fiber with a core radius a and a cladding index n_2, and with $k = 2\pi/\lambda$, the total number of modes can be found from the expression (the derivation of this is complex)[6]

$$M = k^2 \int_0^a [n^2(r) - n_2^2] r \, dr \qquad (5.18)$$

where $n(r)$ defines the variation in the refractive-index profile of the core. This can be related to a general local numerical aperture $NA(r)$ through Eq. (2.80) to yield

$$M = k^2 \int_0^a NA^2(r) r \, dr$$

$$= k^2 NA^2(0) \int_0^a \left[1 - \left(\frac{r}{a} \right)^\alpha \right] r \, dr \tag{5.19}$$

In general, any two fibers that are to be joined will have varying degrees of differences in their radii a, axial numerical apertures $NA(0)$, and index profiles α. Thus the fraction of energy coupled from one fiber to another is proportional to the number of modes common to both fibers M_{comm} (if a uniform distribution of energy over the modes is assumed). The fiber-to-fiber coupling efficiency η_F is given by

$$\eta_F = \frac{M_{comm}}{M_E} \tag{5.20}$$

where M_E is the number of modes in the *emitting fiber* (the fiber that launches power into the next fiber).

The fiber-to-fiber coupling loss L_F is given in terms of η_F as

$$L_F = -10 \log \eta_F \tag{5.21}$$

An analytical estimate of the optical power loss at a joint between multimode fibers is difficult to make because the loss depends on the power distribution among the modes in the fiber.[21–23] For example, consider first the case where all modes in a fiber are equally excited, as shown in Fig. 5.8*a*. The emerging optical beam thus fills the entire exit numerical aperture of this emitting fiber. Suppose now that a second identical fiber, called the *receiving fiber,* is to be joined to the emitting fiber. For the receiving fiber to accept all the optical power emitted by the first fiber, there must be perfect mechanical alignment between the two optical waveguides, and their geometric and waveguide characteristics must match precisely.

On the other hand, if steady-state modal equilibrium has been established in the emitting fiber, most of the energy is concentrated in the lower-order fiber modes. This means that the optical power is concentrated near the center of the fiber core, as shown in Fig. 5.8*b*. The optical power emerging from the fiber then fills only the equilibrium numerical aperture (see Fig. 5.4). In this case, since the input NA of the receiving fiber is larger than the equilibrium NA of the emitting fiber, slight mechanical misalignments of the two joined fibers and small variations in their geometric characteristics do not contribute significantly to joint loss.

Steady-state modal equilibrium is generally established in long fiber lengths. Thus, when estimating joint losses between long fibers, calculations based on a uniform modal power distribution tend to lead to results that may be too pessimistic. However, if a steady-state equilibrium modal distribution is assumed, the estimate may be too optimistic because mechanical misalignments and fiber-to-fiber variations in operational characteristics cause a redistribution of power among the modes in the second fiber. As the power propagates along the second fiber, an additional loss will thus occur when a steady-state distribution is again established.

An exact calculation of coupling loss between different optical fibers, which takes into account non-uniform distribution of power among the modes and propagation effects in the second fiber, is lengthy and involved.[24] Here, we shall therefore make the assumption that all modes in the fiber are equally
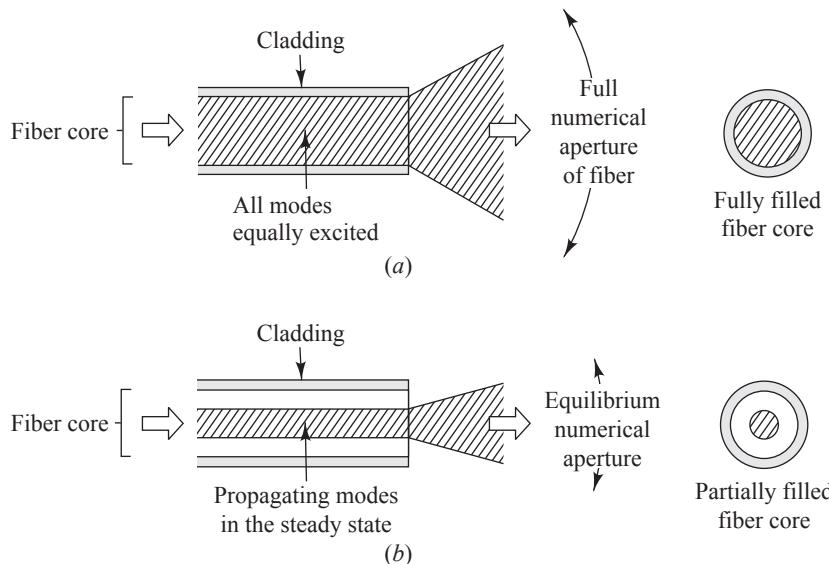
Fig. 5.8 *Different modal distributions of the optical beam emerging from a fiber result in different degrees of coupling loss. (a) When all modes are equally excited, the output beam fills the entire output NA; (b) for a steady-state modal distribution, only the equilibrium NA is filled by the output beam.*

excited. Although this gives a somewhat pessimistic prediction of joint loss, it will allow an estimate of the relative effects of losses resulting from mechanical misalignments, geometrical mismatches, and variations in the waveguide properties between two joined fibers.

5.3.1 Mechanical Misalignment

Mechanical alignment is a major problem when joining two fibers, owing to their microscopic size.[25–29] A standard multimode graded-index fiber core is 50–100 μm in diameter, which is roughly the thickness of a human hair, whereas single-mode fibers have core diameters on the order of 9 μm. Radiation losses result from mechanical misalignments because the radiation cone of the emitting fiber does not match the acceptance cone of the receiving fiber. The magnitude of the radiation loss depends on the degree of misalignment. The three fundamental types of misalignment between fibers are shown in Fig. 5.9.


(*a*) Lateral (axial)   (*b*) Longitudinal (end separation)   (*c*) Angular
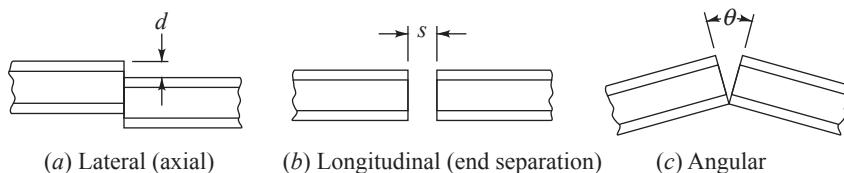
Fig. 5.9 *Three types of mechanical misalignments that can occur between two joined fibers*
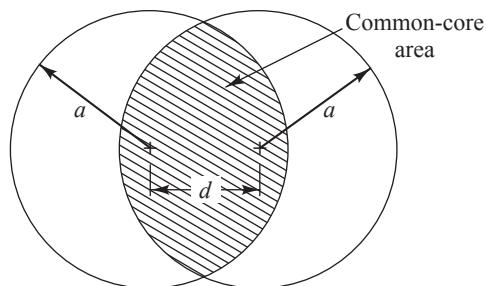
Fig. 5.10 *Axial offset reduces the shaded common core area of the two fiber end faces.*

Axial displacement (which is also often called *lateral displacement*) results when the axes of the two fibers are separated by a distance *d. Longitudinal separation* occurs when the fibers have the same axis but have a gap *s* between their end faces. *Angular misalignment* results when the two axes form an angle so that the fiber end faces are no longer parallel.

The most common misalignment occurring in practice, which also causes the greatest power loss, is axial displacement. This axial offset reduces the overlap area of the two fiber-core end faces, as illustrated in Fig. 5.10, and consequently reduces the amount of optical power that can be coupled from one fiber into the other.

To illustrate the effects of axial misalignment, first consider the simple case of two identical step-index fibers of radii *a*. Suppose that their axes are offset by a separation *d* as is shown in Fig. 5.10, and assume there is a uniform modal power distribution in the emitting fiber. Since the numerical aperture is constant across the end faces of the two fibers, the optical power coupled from one fiber to another is simply proportional to the common area A_{comm} of the two fiber cores. It is straightforward to show that this is (see Prob. 5.9)

$$A_{\text{comm}} = 2a^2 \arccos \frac{d}{2a} - d\left(a^2 - \frac{d^2}{4} \right)^{1/2} \tag{5.22}$$

For the step-index fiber, the coupling efficiency is simply the ratio of the common-core area to the core end-face area,

$$\eta_{F,\text{step}} = \frac{A_{\text{comm}}}{\pi a^2} = \frac{2}{\pi} \arccos \frac{d}{2a} - \frac{d}{\pi a}\left[1 - \left(\frac{d}{2a} \right)^2 \right]^{1/2} \tag{5.23}$$

Example 5.7 An engineer makes a joint between two identical step-index fibers. Each fiber has a core diameter of 50 μm. If the two fibers have an axial (lateral) misalignment of 5 μm, what is the insertion loss at the joint?

Solution: Using Eq. (5.23) we find that the coupling efficiency is

$$\eta_F = \frac{2}{\pi}\cos^{-1}\left(\frac{5}{50} \right) - \frac{5}{\pi(25)}\left[1 - \left(\frac{5}{50} \right)^2 \right]^{1/2} = 0.873$$

From Eq. (5.21) we find that the fiber-to-fiber insertion loss L_F is

$$L_F = -10 \log \eta_F = -10 \log 0.873 = -0.590 \text{ dB}$$

The calculation of power coupled from one graded-index fiber into another identical one is more involved, since the numerical aperture varies across the fiber end face. Because of this, the total power coupled into the receiving fiber at a given point in the common-core area is limited by the numerical aperture of the transmitting or receiving fiber, depending on which is smaller at that point.

If the end face of a graded-index fiber is uniformly illuminated, the optical power accepted by the core will be that power which falls within the numerical aperture of the fiber. The optical power density

$p(r)$ at a point r on the fiber end is proportional to the square of the local numerical aperture NA(r) at that point:[30]

$$p(r) = p(0)\frac{\text{NA}^2(r)}{\text{NA}^2(0)} \qquad (5.24)$$

where NA(r) and NA(0) are defined by Eqs (2.80a) and (2.80b), respectively. The parameter $p(0)$ is the power density at the core axis, which is related to the total power P in the fiber by

$$P = \int_0^{2\pi} \int_0^a p(r)r \, dr \, d\theta \qquad (5.25)$$

For an arbitrary index profile, the double integral in Eq. (5.25) must be evaluated numerically. However, an analytic expression can be found by using a fiber with a parabolic index profile ($\alpha = 2.0$). Using Eq. (2.80), the power density expression at a point r given by Eq. (5.24) becomes

$$p(r) = p(0)\left[1 - \left(\frac{r}{a}\right)^2\right] \qquad (5.26)$$

Using Eqs. (5.25) and (5.26), the relationship between the axial power density $p(0)$ and the total power P in the emitting fiber is

$$P = \frac{\pi a^2}{2}p(0) \qquad (5.27)$$

Let us now calculate the power transmitted across the butt joint of the two parabolic graded-index fibers with an axial offset d, as shown in Fig. 5.11. The overlap region must be considered separately for the areas A_1 and A_2. In area A_1 the numerical aperture is limited by that of the emitting fiber, whereas in area A_2 the numerical aperture of the receiving fiber is smaller than that of the emitting fiber. The vertical dashed line separating the two areas is the locus of points where the numerical apertures are equal.

To determine the power coupled into the receiving fiber, the power density given by Eq. (5.26) is integrated separately over areas A_1 and A_2. Since the numerical aperture of the emitting fiber is smaller than that of the receiving fiber in area A_1, all of the power emitted in this region will be accepted by the receiving fiber. The received power P_1 in area A_1 is thus

$$P_1 = 2\int_0^{\theta_1} \int_{r_1}^a p(r)r \, dr \, d\theta$$

$$= 2p(0) \int_0^{\theta_1} \int_{r_1}^a \left[1 - \left(\frac{r}{a}\right)^2\right]r \, dr \, d\theta$$

$$(5.28)$$
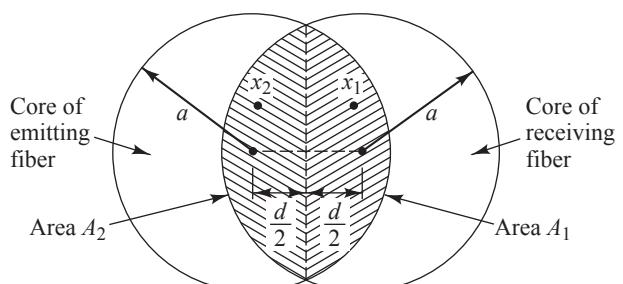
Fig. 5.11 *Core overlap region for two identical parabolic graded-index fibers with an axial separation d. Points x_1 and x_2 are arbitrary points of symmetry in areas A_1 and A_2.*
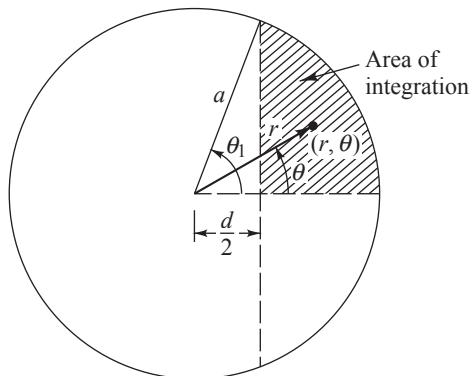
Fig. 5.12 *Area and limits of integration for the common core area of two parabolic graded-index fibers*

where the limits of integration, shown in Fig. 5.12, are

$$r_1 = \frac{d}{2\cos\theta}$$

and

$$\theta_1 = \arccos\frac{d}{2a}$$

Carrying out the integration yields

$$P_1 = \frac{a^2}{2} p(0) \left\{ \arccos\frac{d}{2a} - \left[1 - \left(\frac{d}{2a}\right)^2 \right]^{1/2} \frac{d}{6a}\left(5 - \frac{d^2}{2a^2} \right) \right\}$$

(5.29)

where $p(0)$ is given by Eq. (5.27). The derivation of Eq. (5.29) is left as an exercise.

In area A_2 the emitting fiber has a larger numerical aperture than the receiving fiber. This means that the receiving fiber will accept only that fraction of the emitted optical power that falls within its own numerical aperture. This power can be found from symmetry considerations.[31] The numerical aperture of the receiving fiber at a point x_2 in area A_2 is the same as the numerical aperture of the emitting fiber at the symmetrical point x_1 in area A_1. Thus the optical power accepted by the receiving fiber at any point x_2 in area A_2 is equal to that emitted from the symmetrical point x_1 in area A_1. The total power P_2 coupled across area A_2 is thus equal to the power P_1 coupled across area A_1. Combining these results, we have that the total power P_T accepted by the receiving fiber is

$$P_T = 2P_1$$

$$= \frac{2}{\pi} P \left\{ \arccos\frac{d}{2a} - \left[1 - \left(\frac{d}{2a}\right)^2 \right]^{1/2} \frac{d}{6a}\left(5 - \frac{d^2}{2a^2} \right) \right\}$$

(5.30)

Example 5.8 Suppose two identical graded-index fibers are misaligned with an axial offset of $d = 0.3a$. What is the power coupling loss between these two fibers?

Solution: From Eq. (5.30), the fraction of optical power coupled from the first fiber into the second fiber is

$$\frac{P_T}{P} = \frac{2}{\pi} \left\{ \cos^{-1}\left(\frac{0.3a}{2a}\right) - \left[1 - \left(\frac{0.3a}{2a}\right)^2 \right]^{1/2} \right.$$

$$\left. \times \left(\frac{0.3a}{6a}\right) \left[5 - \left(\frac{0.09}{2}\right) \right] \right\} = 0.748$$

Or, in decibels,

$$10\log\frac{P_T}{P} = -1.27 \text{ dB}$$
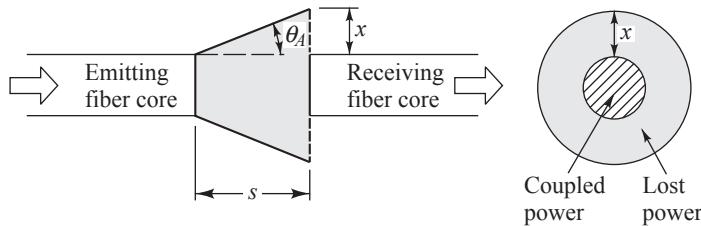
Fig. 5.13 *Optical power loss effect when fiber ends are separated longitudinally by a gap s*

When the axial misalignment d is small compared with the core radius a, Eq. (5.30) can be approximated by

$$P_T \simeq P\left(1 - \frac{8d}{3\pi a}\right) \tag{5.31}$$

This is accurate to within 1 percent for $d/a < 0.4$. The coupling loss for the offsets given by Eqs. (5.30) and (5.31) is

$$L_F = -10 \log \eta_F = -10 \log \frac{P_T}{P} \tag{5.32}$$

The effect of separating the two fiber ends longitudinally by a gap s is shown in Fig. 5.13. Not all the higher-mode optical power emitted in the ring of width x will be intercepted by the receiving fiber. The fraction of optical power coupled into the receiving fiber is given by the ratio of the cross-sectional area of the receiving fiber (πr^2) to the area $\pi(a + x)^2$ over which the emitted power is distributed at a distance s. From Fig. 5.13 we have $x = s \tan \theta_A$, where θ_A is the acceptance angle of the fibers, as defined in Eq. (2.2). From this ratio we find that the loss for an offset joint between two identical step-index fibers is

$$L_F = -10 \log\left(\frac{a}{a + x}\right)^2 = -10 \log\left(\frac{a}{a + s \tan \theta_A}\right)^2 = -10 \log\left[1 + \frac{s}{a}\sin^{-1}\left(\frac{\text{NA}}{n}\right)\right]^{-2} \tag{5.33}$$

where a is the fiber radius, NA is the numerical aperture of the fiber, and n is the refractive index of the material between the fiber ends (usually either air or an index matching gel).

Example 5.9 Two identical step-index fibers each have a 25-μm core radius and an acceptance angle of 14°. Assume the two fibers are perfectly aligned axially and angularly. What is the insertion loss for a longitudinal separation of 0.025 mm?

Solution: We can find the insertion loss due to a gap between fibers by using Eq. (5.23).

For a 0.025-mm = 25-μm gap

$$L_F = -10 \log\left(\frac{25}{25 + 25 \tan 14°}\right)^2 = 1.93 \text{ dB}$$

When the axes of two joined fibers are angularly misaligned at the joint, the optical power that leaves the emitting fiber outside of the solid acceptance angle of the receiving fiber will be lost. For two step-index fibers that have an angular misalignment θ, the optical power loss at the joint has been shown to be[32,33]

$$L_F = -10 \log \left(\cos\theta \left\{ \frac{1}{2} - \frac{1}{\pi} p(1-p^2)^{1/2} - \frac{1}{\pi}\arcsin p - q\left[\frac{1}{\pi} y(1-y^2)^{1/2} + \frac{1}{\pi}\arcsin y + \frac{1}{2} \right] \right\} \right)$$

(5.34)

where

$$p = \frac{\cos\theta_A (1-\cos\theta)}{\sin\theta_A \sin\theta}$$

$$q = \frac{\cos^3\theta_A}{(\cos^2\theta_A - \sin^2\theta)^{3/2}}$$

$$y = \frac{\cos^2\theta_A(1-\cos\theta) - \sin^2\theta}{\sin\theta_A \cos\theta_A \sin\theta}$$

The derivation of Eq. (5.34) again assumes that all modes are uniformly excited.

An experimental comparison[27] of the losses induced by the three types of mechanical misalignments is shown in Fig. 5.14. The measurements were based on two independent experiments using LED sources and graded-index fibers. The core diameters were 50 and 55 μm for the first and second experiments, respectively. A 1.83-m-long fiber was used in the first test and 20-m length in the second. In either case, the power output from the fibers was first optimized. The fibers were then cut at the center, so that the mechanical misalignment loss measurements were carried out on identical fibers. The axial offset and
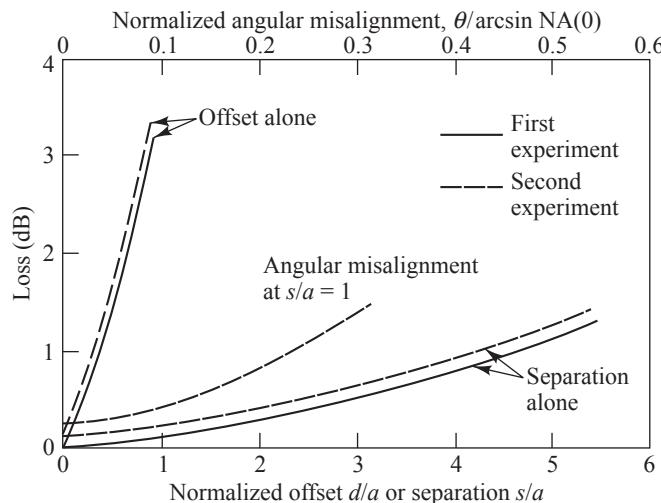

Fig. 5.14 *Experimental comparison of loss (in dB) as a function of mechanical misalignments. (Reproduced with permission from Chu and McCormick,[27] © 1978, AT&T.)*

longitudinal separation losses are plotted as functions of misalignment normalized to the core radius. A normalized angular misalignment of 0.1 corresponds to a 1° angular offset.

Figure 5.14 shows that, of the three mechanical misalignments, the dominant loss arises from lateral displacement. In practice, angular misalignments of less than 1° are readily achievable in splices and connectors. From the experimental data shown in Fig. 5.14, these misalignments result in losses of less than 0.5 dB.

For splices, the separation losses are normally negligible as the fibers should be in relatively close contact. In most connectors, the fiber ends are intentionally separated by a small gap. This prevents them from rubbing against each other and becoming damaged during connector engagement. Typical gaps in these applications range from 0.025 to 0.10 mm, which results in losses of less than 0.8 dB for a 50-μm-diameter fiber.

5.3.2 Fiber-Related Losses

In addition to mechanical misalignments, differences in the geometrical and waveguide characteristics of any two waveguides being joined can have a profound effect on fiber-to-fiber coupling loss. These include variations in core diameter, core-area ellipticity, numerical aperture, refractive-index profile, and core-cladding concentricity of each fiber. Since these are manufacturer-related variations, the user generally has little control over them. Theoretical and experimental studies[24,34–39] of the effects of these variations have shown that, for a given percentage mismatch, differences in core radii and numerical apertures have a significantly larger effect on joint loss than mismatches in the refractive-index profile or core ellipticity.

The joint losses resulting from core diameter, numerical aperture, and core refractive-index-profile mismatches can be found from Eqs. (5.19) and (5.20). For simplicity, let the subscripts E and R refer to the emitting and receiving fibers, respectively. If the radii a_E and a_R are not equal but the axial numerical apertures and the index profiles are equal [$\text{NA}_E(0) = \text{NA}_R(0)$ and $\alpha_E = \alpha_R$], then the coupling loss is

$$L_F(a) = \begin{cases} -10 \log\left(\dfrac{a_R}{a_E}\right)^2 & \text{for} \quad a_R < a_E \\ 0 & \text{for} \quad a_R \geq a_E \end{cases} \tag{5.35}$$

If the radii and the index profiles of the two coupled fibers are identical but their axial numerical apertures are different, then

$$L_F(\text{NA}) = \begin{cases} -10 \log\left[\dfrac{\text{NA}_R(0)}{\text{NA}_E(0)}\right]^2 & \text{for} \quad \text{NA}_R(0) < \text{NA}_E(0) \\ 0 & \text{for} \quad \text{NA}_R(0) \geq \text{NA}_E(0) \end{cases} \tag{5.36}$$

Example 5.10 Consider two joined step-index fibers that are perfectly aligned. What is the coupling loss if the numerical apertures are $\text{NA}_R = 0.20$ for the receiving fiber and $\text{NA}_E = 0.22$ for the emitting fiber?

Solution: From Eq. (5.36) we have

$$L_F(\text{NA}) = -10 \log\left(\frac{0.20}{0.22}\right)^2 = -10 \log 0.826$$
$$= -0.828 \text{ dB}$$

Finally, if the radii and the axial numerical apertures are the same but the core refractive-index profiles differ in two joined fibers, then the coupling loss is

$$
L_F(\alpha) = \begin{cases} -10 \log \dfrac{\alpha_R(\alpha_E + 2)}{\alpha_E(\alpha_R + 2)} & \text{for} \quad \alpha_R < \alpha_E \\ 0 & \text{for} \quad \alpha_R \geq \alpha_E \end{cases} \tag{5.37}
$$

This results because for $\alpha_R < \alpha_E$ the number of modes that can be supported by the receiving fiber is less than the number of modes in the emitting fiber. If $\alpha_R > \alpha_E$ then all modes in the emitting fiber can be captured by the receiving fiber. The derivations of Eqs. (5.35) to (5.37) are left as an exercise (see Probs. 5.13 through 5.15).

Example 5.11 Consider two joined graded-index fibers that are perfectly aligned. What is the coupling loss if the refractive index profiles are $\alpha_R = 1.98$ for the receiving fiber and $\alpha_E = 2.20$ for the emitting fiber?

Solution: From Eq. (5.37) we have

$$
L_F(\alpha) = -10 \log \frac{\alpha_R(\alpha_E + 2)}{\alpha_E(\alpha_R + 2)} = -10 \log 0.995
$$

$$
= -0.022 \text{ dB}
$$

5.3.3 Fiber End-Face Preparation

One of the first steps that must be followed before fibers are connected or spliced to each other is to prepare the fiber end faces properly. In order not to have light deflected or scattered at the joint, the fiber ends must be flat, perpendicular to the fiber axis, and smooth. End-preparation techniques that have been extensively used include sawing, grinding and polishing, controlled fracture, and laser cleaving.

Conventional grinding and polishing techniques can produce a very smooth surface that is perpendicular to the fiber axis. However, this method is quite time-consuming and requires a fair amount of operator skill. Although it is often implemented in a controlled environment such as a laboratory or a factory, it is not readily adaptable for field use. The procedure employed in the grinding and polishing technique is to use successively finer abrasives to polish the fiber end face. The end face is polished with each successive abrasive until the scratches created by the previous abrasive material are replaced by the finer scratches of the present abrasive. The number of abrasives used depends on the degree of smoothness that is desired.
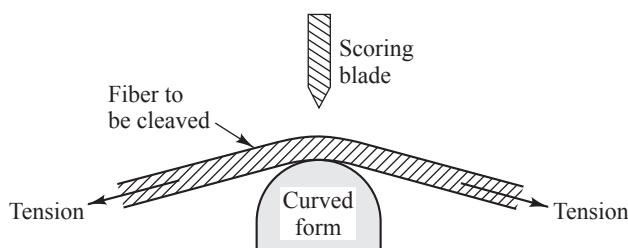

Fig. 5.15 *Controlled-fracture procedure for fiber end preparation*

Controlled-fracture techniques are based on score-and-break methods for cleaving fibers. In this operation, the fiber to be cleaved is first scratched to create a stress concentration at the surface. The fiber is then bent over a curved form while tension is simultaneously applied, as shown in Fig. 5.15. This action produces a stress distribution across the fiber. The maximum stress occurs at the scratch point so that a crack starts to propagate through the fiber.
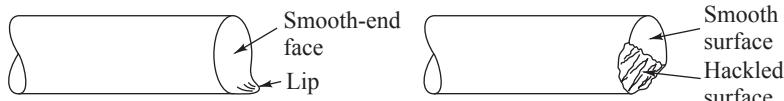
Fig. 5.16 *Two examples of improperly cleaved fiber ends*

One can produce a highly smooth and perpendicular end face in this way. A number of different tools based on the controlled-fracture technique have been developed and are being used both in the field and in factory environments. However, the controlled-fracture method requires careful control of the curvature of the fiber and of the amount of tension applied. If the stress distribution across the crack is not properly controlled, the fracture propagating across the fiber can fork into several cracks. This forking produces defects such as a lip or a hackled portion on the fiber end, as shown in Fig. 5.16. The EIA Fiber Optic Test Procedures (FOTP) 57 and 179 define these and other common end-face defects as follows:[40,41]

Lip. This is a sharp protrusion from the edge of a cleaved fiber that prevents the cores from coming in close contact. Excessive lip height can cause fiber damage.

Rolloff. This rounding-off of the edge of a fiber is the opposite condition to lipping. It is also known as *breakover* and can cause high insertion or splice loss.

Chip. A chip is a localized fracture or break at the end of a cleaved fiber.

Hackle. Figure 5.16 shows this as severe irregularities across a fiber end face.

Mist. This is similar to hackle but much less severe.

Spiral or step. These are abrupt changes in the end-face surface topology.

Shattering. This is the result of an uncontrolled fracture and has no definable cleavage or surface characteristics.

An alternative to a mechanical score-and-break method is the use of a laser to cleave fibers. As an example, by using an ultraviolet-excimer laser this technique can be applied to automated cleaving procedures for high-volume production of fiber optic devices.[42]

5.4 LED Coupling to Single-Mode Fibers

In the early years of optical fiber applications, LEDs were traditionally considered only for multimode-fiber systems. However, around 1985, researchers recognized that edge-emitting LEDs can launch sufficient optical power into a single-mode fiber for transmission at data rates up to 560 Mb/s over several kilometers.[43–49] The interest in this arose because of the cost and reliability advantages of LEDs over laser diodes. Edge-emitting LEDs are used for these applications because they have a laserlike output pattern in the direction perpendicular to the junction plane.

To rigorously evaluate the coupling between an LED and a single-mode fiber, we need to use the formalism of electromagnetic theory rather than geometrical optics because of the monomode nature of the fiber. However, coupling analyses of the output from an edge-emitting LED to a single-mode fiber can be carried out wherein the results of electromagnetic theory are interpreted from a geometrical point of view,[45–47] which involves defining a numerical aperture for the single-mode fiber. Agreement with experimental measurements and with a more exact theory are quite good.[46–49]

This section analyzes two cases: (1) direct coupling of an LED into a single-mode fiber, and (2) coupling into a single-mode fiber from a multimode flylead attached to the LED.[46] In general, edge-emitting LEDs have gaussian near-field output profiles with $1/e^2$ full widths of approximately 0.9 and 22 μm in the

directions perpendicular and parallel to the junction plane, respectively. The far-field patterns vary approximately as $\cos^7 \theta$ in the perpendicular direction and as $\cos \theta$ (lambertian) in the parallel direction.

For a source with a circularly asymmetric radiance $B(A_S, \Omega_S)$, Eq. (5.3) is, in general, not separable into contributions from the perpendicular and parallel directions. However, we can approximate the independent contributions by evaluating Eq. (5.3) as if each component were a circularly symmetric source, and then taking the geometric mean to find the total coupling efficiency. Calling these the x (parallel) and y (perpendicular) directions, and letting τ_x and τ_y be the x and y power transmissivities (directional coupling efficiencies), respectively, we can find the maximum LED-to-fiber coupling efficiency η from the relation

$$\eta = \frac{P_{\text{in}}}{P_s} = \tau_x \tau_y \tag{5.38}$$

where P_{in} is the optical power launched into the fiber and P_s is the total source output power.

Using a small-angle approximation, we first integrate over the effective solid acceptance angle of the fiber to get $\pi \text{NA}_{\text{SM}}^2$, where the geometrical-optics-based fiber numerical aperture $\text{NA}_{\text{SM}} = 0.11$. Assuming a gaussian output for the source, then for butt coupling of the LED to the single-mode fiber of radius a, the coupling efficiency in the y direction is

$$
\begin{aligned}
\tau_y &= \left(\frac{P_{\text{in}, y}}{P_s} \right)^{1/2} \\
&= \left[\frac{\displaystyle\int_0^{2\pi} \int_0^a B_0\, e^{-2r^2/\omega_y^2}\, r\, dr\, d\theta_s\, \pi \text{NA}_{\text{SM}}^2}{\displaystyle\int_0^{2\pi} \int_0^\infty B_0\, e^{-2r^2/\omega_y^2}\, y\, dy\, d\theta_s \int_0^{2\pi} \int_0^{\pi/2} \cos^7 \theta \sin \theta\, d\theta\, d\phi} \right]^{1/2}
\end{aligned}
\tag{5.39}
$$

where $P_{\text{in}, y}$ is the optical power coupled into the fiber from the y-direction source output, which has a $1/e^2$ LED intensity radius ω_y. One can write a similar set of integrals for τ_x. Letting $a = 4.5\ \mu\text{m}$, $\omega_x = 10.8\ \mu\text{m}$, and $\omega_y = 0.47\ \mu\text{m}$, calculations[46] give values of $\tau_x = -12.2$ dB and $\tau_y = -6.6$ dB to yield a total coupling efficiency $\eta = -18.8$ dB. Thus, for example, if the LED emits 200 μW (-7 dBm), then 2.6 μW (-25.8 dBm) gets coupled into the single-mode fiber.

When a 1- to 2-m multimode-fiber flylead is attached to an edge-emitting LED, the near-field profile of the multimode fiber has the same asymmetry as the LED. In this case, one can assume that the multimode-fiber optical output is a simple gaussian with different beam widths along the x and y directions. Using a similar coupling analysis with effective beam widths of $\omega_x = 19.6\ \mu$m and $\omega_y = 10.0\ \mu$m, the directional coupling efficiencies are $\tau_x = -7.8$ dB and $\tau_y = -5.2$ dB, yielding a total LED-to-fiber coupling efficiency $\eta = -13.0$ dB.

5.5 | Fiber Splicing

A fiber splice is a permanent or semipermanent joint between two fibers. These are typically used to create long optical links or in situations where frequent connection and disconnection are not needed. In making and evaluating such splices, one must take into account the geometrical differences in the two fibers, fiber misalignments at the joint, and the mechanical strength of the splice. This section first
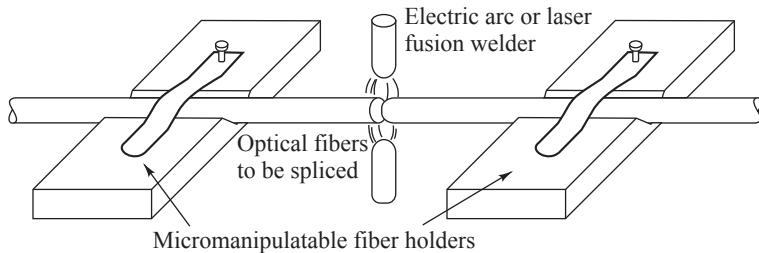
Fig. 5.17 *Fusion splicing of optical fibers*

addresses general splicing methods and then examines the factors contributing to loss when splicing single-mode fibers.

5.5.1 Splicing Techniques

Fiber splicing techniques include the fusion splice, the V-groove mechanical spice, and the elastic-tube splice.[50–59] The first technique yields a permanent joint, whereas the other two types of splices can be disassembled if necessary.

Fusion splices are made by thermally bonding together prepared fiber ends, as pictured in Fig. 5.17. In this method, the fiber ends are first prealigned and butted together. This is done either in a grooved fiber holder or under a microscope with micromanipulators. The butt joint is then heated with an electric arc or a laser pulse so that the fiber ends are momentarily melted and hence bonded together. This technique can produce very low splice losses, which typically are in the range of 0.05 to 0.10 dB for both single-mode and multimode fibers. However, care must be exercised in this technique, since bubbles arising from dirt or fiber coating particles that are trapped and gasified in the splice, incorrectly set fusion arc current, application of insufficient heat, and residual stresses induced near the joint as a result of changes in chemical composition arising from the material melting can produce a weak splice.[60,61]

In the V-groove splice technique, the prepared fiber ends are first butted together in a V-shaped groove, as shown in Fig. 5.18. They are then bonded together with an adhesive or are held in place by means of a cover plate. The V-shaped channel can be either a grooved silicon, plastic, ceramic, or metal substrate. The splice loss in this method depends strongly on the fiber size (outside dimensions and core-diameter variations) and eccentricity (the position of the core relative to the center of the fiber).

The elastic-tube splice shown cross-sectionally in Fig. 5.19 is a unique device that automatically performs lateral, longitudinal, and angular alignment. It splices multimode fibers to give losses in the same range as commercial fusion splices, but much less equipment and skill are needed. The splice mechanism is basically a tube made of an elastic material. The central hole diameter is slightly smaller than that of the fiber to be spliced and is tapered on each end for easy fiber insertion. When a fiber is inserted, it expands the hole diameter so that the elastic material exerts a symmetrical force on the fiber. This symmetry feature allows
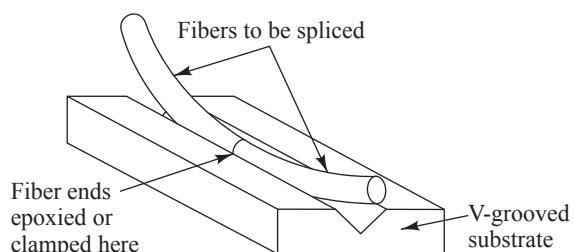

**Fig. 5.18** *V-groove optical fiber splicing technique*
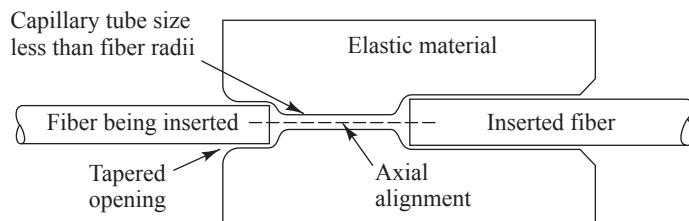
Fig. 5.19 *Alignment concept for an elastic-tube splice*

an accurate and automatic alignment of the axes of the two fibers to be joined. A wide range of fiber diameters can be inserted into the elastic tube. Thus, the fibers to be spliced do not have to be equal in diameter, since each fiber moves into position independently relative to the tube axis.

5.5.2 Splicing Single-Mode Fibers

As is the case in multimode fibers, in single-mode fibers the lateral (axial) offset misalignment presents the most serious loss. This loss depends on the shape of the propagating mode. For gaussian-shaped beams the loss between identical fibers is[62]

$$L_{SM, lat} = -10 \log \left\{ \exp \left[-\left(\frac{d}{w} \right)^2 \right] \right\} \tag{5.40}$$

where the spot size w is the mode-field radius defined in Eq. (2.74), and d is the lateral displacement shown in Fig. 5.9. Since the spot size is only a few micrometers in single-mode fibers, low-loss coupling requires a very high degree of mechanical precision in the axial dimension.

Example 5.12 A single-mode fiber has a normalized frequency $V = 2.20$, a core refractive index $n_1 = 1.47$, a cladding refractive index $n_2 = 1.465$, and a core diameter $2a = 9$ μm. What is the insertion loss of a fiber joint having a lateral offset of $d = 1$ μm? For the mode-field diameter use the expression (see Prob. 2.24) $w = a(0.65 + 1.619\ V^{-3/2} + 2.879\ V^{-6})$.

Solution: First, using the expression for the mode-field diameter from Prob. 2.24, we have

$$w = 4.5[0.65 + 1.619(2.20)^{-3/2} + 2.879(2.20)^{-6}]$$
$$= 5.27\ \mu m$$

Then, from Eq. (5.40), we have

$$L_{SM,lat} = -10 \log \{\exp[-(1/5.27)^2]\} = 0.156\ dB$$

For angular misalignment in single-mode fibers, the loss at a wavelength λ is[62]

$$L_{SM, ang} = -10 \log \left\{ \exp \left[-\left(\frac{\pi n_2 w \theta}{\lambda} \right)^2 \right] \right\} \tag{5.41}$$

where n_2 is the refractive index of the cladding, θ is the angular misalignment in radians shown in Fig. 5.9, and w is the mode-field radius.

Example 5.13 Consider the single-mode fiber described in Example 5.12. Find the loss at a joint having an angular misalignment of $1° = 0.0175$ radians at a 1300-nm wavelength.

Solution: From Eq. (5.41), we have

$$L_{SM, ang} = -10 \log \left\{ \exp \left[- \left(\frac{\pi(1.465)(5.27)(0.0175)}{1.3} \right)^2 \right] \right\}$$

$$= 0.46 \text{ dB}$$

For a gap s with a material of index n_3, and letting $G = s/kw^2$, the gap loss for identical single-mode fiber splices is

$$L_{SM, gap} = -10 \log \frac{64 n_1^2 n_3^2}{(n_1 + n_3)^4 (G^2 + 4)} \qquad (5.42)$$

See Eq. (5.43) for a more general equation for dissimilar fibers.

5.6 Optical Fiber Connectors

A wide variety of optical fiber connectors has evolved for numerous different applications. Their uses range from simple single-channel fiber-to-fiber connectors in a benign location to multichannel connectors used in harsh military field environments. Some of the principal requirements of a good connector design are as follows:

1. *Low coupling losses.* The connector assembly must maintain stringent alignment tolerances to assure low mating losses. These low losses must not change significantly during operation or after numerous connects and disconnects.
2. *Interchangeability.* Connectors of the same type must be compatible from one manufacturer to another.
3. *Ease of assembly.* A service technician should be able to install the connector easily in a field environment; that is, in a location other than the connector factory. The connector loss should also be fairly insensitive to the assembly skill of the technician.
4. *Low environmental sensitivity.* Conditions such as temperature, dust, and moisture should have a small effect on connector-loss variations.
5. *Low cost and reliable construction.* The connector must have a precision suitable to the application, but its cost must not be a major factor in the fiber system.
6. *Ease of connection.* Generally, one should be able to mate and demate the connector, simply, by hand.

5.6.1 Connector Types

Connectors are available in designs that screw on, twist on, or snap into place. The most commonly used connectors are the twist-on and snap-on design.[63–73] These include both single-channel and multichannel assemblies for cable-to-cable and for cable-to-circuit card connections. The basic coupling mechanisms used in these connectors belong to either the *butt-joint* or the *expanded-beam* classes.

The majority of connectors use a butt-joint coupling mechanism, as illustrated in Fig. 5.20. The elements shown in this figure are common to most connectors. The key components are a long, thin
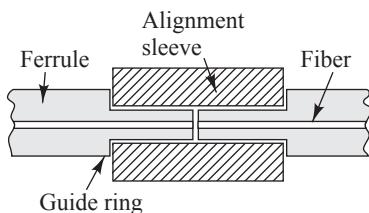
Fig. 5.20   *Example of an alignment scheme used in fiber-optic connectors*

stainless steel, glass, ceramic, or plastic cylinder known as a *ferrule* and a precision sleeve into which the ferrule fits. This sleeve is known variably as an *alignment sleeve*, an *adapter*, or a *coupling receptacle*. The center of the ferrule has a hole that precisely matches the size of the fiber cladding diameter. Typically the hole size is $125.0 \pm 1.0\ \mu$m. The fiber is secured in the hole with epoxy, and the end of the ferrule is polished flat to a smooth finish. The polished surface can be either perpendicular to the fiber axis or at a small angle (typically $8°$) relative to the axis. These two end-face orientations define a *polished connector* (PC) and an *angle-polished connector* (APC) configuration, respectively. The mechanical challenges in fabricating a ferrule include maintaining precisely both the dimensions of the hole and its position relative to the ferrule outer surface.

Normally a connector also has some type of strain-relief mechanism called a *boot*, which shields the junction of the connector body and the cable from bends and pulls. The TIA 568 color-code standard for connector bodies and boots is beige for multimode fiber, blue for single-mode fiber, and green for APC (angled end face) connectors.

Around 95 percent of the ferrules used in optical connectors are made of ceramic due to some of the desirable material properties. These properties include low insertion loss required for optical transmission, remarkable strength, small elasticity coefficient, easy control of product characteristics, and strong resistance to changes in environmental conditions such as temperature.

As shown in Fig. 5.20, since both ends of a cable have a plug type ferrule structure, an adapter mechanism is used to align the two fiber ends. In the connector the length of the sleeve and a guide ring on the ferrule determine the end separation of the two fibers. Note that an adapter also can be used to mate one type of connector to another. The only precaution in any case is to make sure that either the two fibers being joined have similar characteristics or that one is not trying to couple light from a multimode fiber into a single-mode fiber. However, it is not a problem to have light from a single-mode fiber couple into a multimode fiber. For example, this often is done at the receiving end of a link in which a multimode flylead on a photodetector receives light from a single-mode transmission fiber.

Figure 5.21 illustrates another coupling mechanism based on using an expanded-beam connector. This design employs lenses that either collimate the light emerging from the transmitting fiber or focus the expanded beam onto the core of the receiving fiber. The fiber-to-lens distance is equal to the focal length of the lens. The advantage of this scheme is that, since the beam is collimated, separation of the fiber ends may take place within the connector. Thus the connector is less dependent on lateral alignments.


Fig. 5.21   *Schematic representation of an expanded-beam fiber optic connector*

In addition, optical processing elements, such as beam splitters and switches, can easily be inserted into the expanded beam between the fiber ends.

Manufacturers have devised almost 100 different optical fiber connector styles and configurations. Many of these styles have now become obsolete owing to clever new designs and use of better materials, which resulted in smaller, lower-loss, and easier to use connectors. Table 5.1 lists six widely used connector types and gives their main features and applications.

Table 5.1 *Six popular fiber optic connectors with their features and applications (Photos courtesy of Mr. Frank Jaffer (SENKO GROUP) www.senko.com)*

Connector type	Features	Applications
ST	Uses a ceramic ferrule and a rugged metal housing. It is latched in place by twisting. Typical loss range is 0.20–0.50 dB.	Designed for distribution applications using either multimode or single-mode fibers.
SC	Designed by NTT for snap-in connection in tight spaces. Uses a ceramic ferrule in simplex or duplex plastic housings for either multimode or single-mode fibers. Typical loss range is 0.20–0.45 dB.	Widely used in Gigabit Ethernet, ATM, LAN, MAN, WAN, data communication, Fibre Channel, and telecommunication networks.
LC	SFF connector that uses a standard RJ-45 telephone plug housing and ceramic ferrules in simplex or duplex plastic housings. Typical loss range is 0.10–0.50 dB.	Available in simplex and duplex configurations for CATV, LAN, MAN, and WAN applications.
MU	SFF connector based on a 1.25-mm ceramic ferrule and a single free-floating ferrule. Typical loss range is 0.10–0.30 dB.	Used mainly in Japan. Suitable for board-mounted applications and for distribution-cable assemblies.
MT-RJ	SFF connector with two fibers in one molded plastic ferrule and an improved RJ-45 latch mechanism. Typical loss range is 0.25–0.75 dB.	Applications are for MANs and LANs, such as horizontal optical cabling to the desktop.
MPO/MTP	Can house up to twelve multimode or single-mode optical fibers in a single compact ferrule. Typical loss range is 0.25–1.00 dB.	Allows high-density connections between network equipment in telecom rooms.

Typically the different connector types are designated by combinations of two or three letters. The main ones are ST, SC, LC, MU, MT-RJ, MPO, and variations on MPO. Therefore, the vendors refer to these connectors as, for example, SC-type connectors or simply SC connectors. To understand the purpose of the letter designations, let us first take a brief look at their origins:

- ST is derived from the words *straight tip*, which refers to the ferrule configuration.
- The letters SC were coined by NTT to mean *subscriber connector* or *square connector*.
- Lucent developed a specific connector type, which they nicknamed the LC connector.
- The letters MU were selected by NTT to indicate a *miniature unit*.
- The designation MT-RJ is an acronym for *media termination–recommended jack*.
- The letters MPO were selected to indicate a *multiple-fiber push-on/pull-off* connecting function.

To improve on the screw-on mounting features of early connector designs, twist-on and snap-on single-fiber connectors such as the ST and SC styles were devised. Further emphasis on reducing connector sizes for higher packaging densities resulted in many concepts for *small-form-factor* (SFF) single-fiber connector types, such as the LC, MU, and MT-RJ designs. The biggest difference among the SFF connectors is whether they use ceramic or plastic ferrules. The SFF connectors are designed for fast termination in the field and also enable smaller fiber network interface cards (NICs) to be made for computer workstations and servers.

Another recent development was the introduction of an inexpensive, high-performance compact *multiple-fiber connector*. These connectors save space by providing up to at least twelve potential connections within a single ferrule that has the same physical size as a single-fiber SFF connector. This means that one such multiple-fiber connector can replace up to twelve single-fiber connectors. These components are known by various acronyms such as MPO, MTP, and MPX.

ST connector The ST connector is very popular for both data communication and telecommunication applications. It utilizes a precision zirconia (zirconium dioxide) ceramic ferrule. The ST connector employs a rugged metal bayonet coupling ring with radial ramps to allow easy engagement to the studs of the mating adapter. The connector is mated by pushing it into place and then twisting it to engage a spring-loaded bayonet socket. For multimode fibers this yields a typical insertion loss of 0.4 dB when using a manual polishing method or 0.2 dB when using an automated fiber polisher. Single-mode connectors typically achieve a 0.3-dB insertion loss and a 40-dB return loss using a simple manual polishing method.

SC connector The SC connector is mated by a simple snap-in method and can be disconnected by pushing in on a tab and then pulling the connector off. SC connectors are available in either simplex or duplex configurations. Analogous to the ST connector, the SC connector uses a ceramic ferrule and has similar loss characteristics. The mating and loss features allow the connector to be used in tight spaces where turning an outer connector housing, such as on an ST connector, is not practical. An example of such spaces is a patch panel where there is a high packing density of connections.

LC connector The LC connector was developed to meet the growing demand for small, high-density fiber optic connectivity on equipment bays, in distribution panels, and on wall plates. The LC design was based on the industry-standard RJ-45 telephone plug interface and ceramic ferrule technology. The advantage of the RJ-45 housing is that it provides a reliable and robust latching mechanism for the LC connector. The LC connector has a six-position tuning feature to achieve very low insertion loss performance by optimizing the alignment of the fiber cores. The LC connectors are available in both simplex and duplex configurations.

MU connector The MU connector is a SFF that was developed by NTT. Basically it can be considered as a smaller version of the SC type connector. It is based on a 1.25-mm ceramic ferrule and utilizes a single free-floating ferrule independent of the backbone. It has a plastic housing and uses a push-pull latching mechanism. It is available in simplex, duplex, and higher-count-channel styles. The MU connector is suitable both for board-mounted applications and for distribution-cable assemblies to allow connections in simplex networks.

MT-RJ connector The MT-RJ is a SFF connector with two multimode or single-mode fibers in one precision-molded plastic ferrule. It was designed to meet the desire for an interface technology that is significantly lower in cost and size than the SC duplex connector. The MT-RJ uses an improved version of an industry-standard RJ-45 type latch. Its principal application is for horizontal cabling needs to the desktop.

Multiple-fiber connectors The MPO is one of several variations of compact multiple-fiber connectors. They all use a simple push-pull latching mechanism for easy and intuitive insertion and removal. The end of the MPO connector may be polished flat or at an 8° angle. The MTO connector is the same size as the SC, but it can accommodate a maximum of twelve fibers and provides up to twelve times the density, thereby offering savings in circuit card and rack space.

5.6.2 Single-Mode Fiber Connectors

Because of the wide use of single-mode fiber optic links and because of the greater alignment precision required for these systems, this section addresses single-mode connector coupling losses. Based on the gaussian-beam model of single-mode fiber fields,[63] the following equation[74] gives the coupling loss (in decibels) between single-mode fibers that have unequal mode-field diameters (which is an intrinsic factor) and lateral, longitudinal, and angular offsets plus reflections (which are all extrinsic factors):

$$L_{\text{SM, ff}} = -10 \log \left[\frac{16 n_1^2 n_3^2}{(n_1 + n_3)^4} \frac{4\sigma}{q} \exp\left(-\frac{\rho u}{q} \right) \right] \tag{5.43}$$

where

$$\rho = (k w_1)^2$$
$$q = G^2 + (\sigma + 1)^2$$
$$u = (\sigma + 1) F^2 + 2\sigma FG \sin \theta + \sigma (G^2 + \sigma + 1) \sin^2 \theta$$
$$F = \frac{d}{k w_1^2}$$
$$G = \frac{s}{k w_1^2}$$
$$\sigma = (w_2 / w_1)^2$$
$$k = 2\pi n_3 / \lambda$$

n_1 = core refractive index of fibers
n_3 = refractive index of medium between fibers
λ = wavelength of source
d = lateral offset
s = longitudinal offset
θ = angular misalignment

w_1 = 1/e mode-field radius of transmitting fiber
w_2 = 1/e mode-field radius of receiving fiber

This general equation gives very good correlation with experimental investigations.[64]

5.6.3 Connector Return Loss

A connection point in an optical link can be categorized into four interface types. These consist of either a perpendicular or an angled end-face on the fiber, and either a direct physical contact between the fibers or a contact employing an index-matching material. Each of these methods has a basic application for which it is best suited. The physical-contact type connectors without index-matching material are traditionally used in situations where frequent reconnections are required, such as within a building or on localized premises. Index-matching connectors are standardly employed in outside cable plants where the reconnections are infrequent, but need to have a low loss.

This section gives some details on index-matched and direct physical contacts, and discusses angled interfaces. In each case, these connections require high return losses (low reflection levels) and low insertion losses (high optical-signal throughput levels). The low reflectance levels are desired since optical reflections provide a source of unwanted feedback into the laser cavity. This can affect the optical frequency response, the linewidth, and the internal noise of the laser, which results in degradation of system performance.

Figure 5.22 shows a model of an index-matched connection with perpendicular fiber end faces. In this figure and in the following analyses, offsets and angular misalignments are not taken into account. The connection model shows that the fiber end faces have a thin surface layer of thickness h having a high refractive index n_2 relative to the core index, which is a result of fiber polishing. The fiber core has an index n_0, and the gap width s between the end faces is filled with index-matching material having a refractive index n_1. The return loss RL_{IM} in decibels for the index-matched gap region is given by[75]

$$\mathrm{RL}_{IM} = -10 \log \left\{ 2R_1 \left[1 - \cos \left(\frac{4\pi n_1 s}{\lambda} \right) \right] \right\} \tag{5.44}$$

where

$$R_1 = \frac{r_1^2 + r_2^2 + 2r_1 r_2 \cos \delta}{1 + r_1^2 r_2^2 + 2r_1 r_2 \cos \delta} \tag{5.45}$$
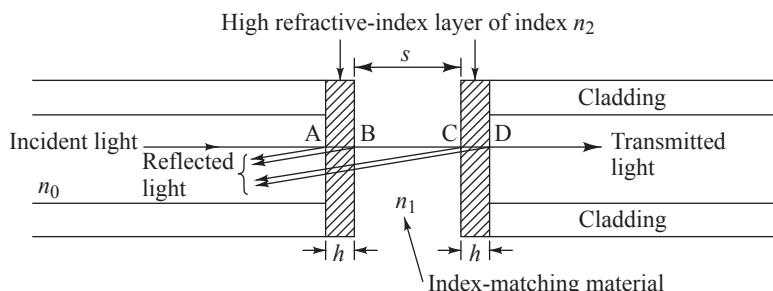

Fig. 5.22 *Model of an index-matched connection with perpendicular fiber end faces*

is the reflectivity at a single material-coated end face, and

$$r_1 = \frac{n_0 - n_2}{n_0 + n_2} \quad \text{and} \quad r_2 = \frac{n_2 - n_1}{n_2 + n_1} \tag{5.46}$$

are the reflection coefficients between the core and the high-index layer and between the high-index layer and the index-matching material, respectively. The parameter $\delta = (4\pi/\lambda)\, n_2 h$ is the phase difference in the high-index layer. The factor 2 in Eq. (5.44) accounts for reflections at both fiber end faces. The value of n_2 of the glass surface layer varies from 1.46 to 1.60, and the thickness h ranges from 0 to 0.15 μm.

When the perpendicular end faces are in direct physical contact, the return loss RL_{PC} in decibels is given by[75]

$$\text{RL}_{PC} = -10 \log \left\{ 2R_2 \left[1 - \cos\left(\frac{4\pi n_2}{\lambda} 2h \right) \right] \right\} \tag{5.47}$$

where

$$R_2 = \left(\frac{n_0 - n_2}{n_0 + n_2} \right)^2 \tag{5.48}$$

Here, R_2 is the reflectivity at the discontinuity between the refractive indices of the fiber core and the high-index surface layer. In this case, the return loss at a given wavelength depends on the value of the refractive index n_2 and the thickness h of the surface layer.

Connections with angled end faces are used in applications where an ultra-low reflection is required. Figure 5.23 shows a cross-sectional view of such a connection with a small gap of width s separating the fiber ends. The fiber core has an index n_0, and the material in the gap has a refractive index n_1. The end faces are polished at an angle θ_0 with respect to the plane perpendicular to the fiber axis. This angle is typically 8°. If I_i and I_t are the incident and throughput optical power intensities, respectively, then the transmitted efficiency T through the connector is[76]

$$T = \frac{I_t}{I_i} = \frac{(1 - R)^2}{(1 - R)^2 + 4R \sin^2(\beta/2)} \tag{5.49}$$
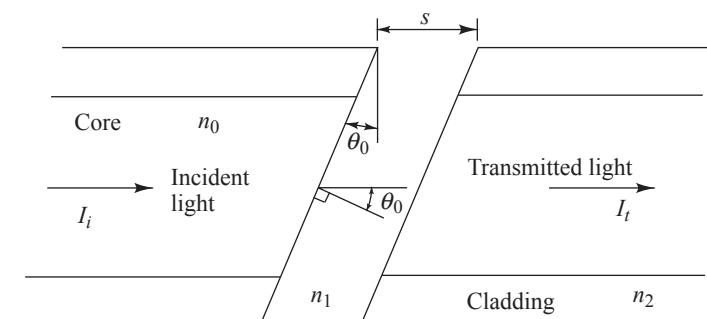
Fig. 5.23 *Connection with angled end faces having a small gap of width s separating the fiber ends*

where

$$\frac{\sin\theta_0}{\sin\theta} = \frac{n}{n_0}, \beta = \frac{4\pi n_1 s \cos\theta}{\lambda}, \quad \text{and} \quad R = \left(\frac{n_0 - n_1}{n_0 + n_1}\right)^2$$

The insertion loss for this type of connector with an 8° angle will vary from 0 dB for no gap to 0.6 dB for an air gap of width $s = 1.0$ μm. Note that when an index-matching material is used so that $n_0 = n_1$, then $R = 0$ and $T = 1$. When $n_0 \neq n_1$, the transmitted efficiency (and hence the connector loss) has an oscillatory behavior as a function of the wavelength and the end face angle.

■ PROBLEMS ■

5.1 Analogous to Fig. 5.2, use a computer to plot and compare the emission patterns from a lambertian source and a source with an emission pattern given by $B(\theta) = B_0 \cos^3 \theta$. Assume both sources have the same peak radiance B_0, which is normalized to unity in each case.

5.2 Consider light sources where the emission pattern is given by $B(\theta) = B_0 \cos^m \theta$. Use a computer to plot $B(\theta)$ as a function of m in the range $1 \leq m \leq 20$ at viewing angles of 10°, 20°, and 45°. Assume all sources have the same peak radiance B_0.

5.3 A laser diode has lateral ($\phi = 0°$) and transverse ($\phi = 90°$) half-power beam widths of $2\theta = 60°$ and 30°, respectively. What are the transverse and lateral power distribution coefficients for this device?

5.4 An LED with a circular emitting area of radius 20 μm has a lambertian emission pattern with a 100-W/(cm^2· sr) axial radiance at a 100-mA drive current. How much optical power can be coupled into a step-index fiber having a 100-μm core diameter and NA = 0.22? How much optical power can be coupled from this source into a 50-μm core-diameter graded-index fiber having $\alpha = 2.0$, $n_1 = 1.48$, and $\Delta = 0.01$?

5.5 A GaAs optical source that has a refractive index of 3.600 is closely coupled to a step-index fiber that has a core refractive index of 1.465. If the source size is smaller than the fiber core, and the small gap between the source and the fiber is filled with a gel that has a refractive index of 1.305, what is the power loss in decibels from the source into the fiber?

5.6 Use Eq. (5.3) to derive an expression for the power coupled into a step-index fiber from an LED that has a radiant distribution given by
$$B(\theta) = B_0 \cos^m \theta$$

5.7 On the same graph, plot the maximum coupling efficiencies as a function of the source radius r_s for the following fibers:
(a) Core radius of 25 μm and NA = 0.16.
(b) Core radius of 50 μm and NA = 0.20.
Let r_s range from 0 to 50 μm. In what regions can a lens improve the coupling efficiency?

5.8 The end faces of two optical fibers with core refractive indices of 1.485 are perfectly aligned and have a small gap between them. If this gap is filled with a gel that has a refractive index of 1.305, find the optical power in decibels reflected at one interface of this joint. If the gap is very small, what is the power loss in decibels through the joint when no index-matching material is used? Note that $n = 1.0$ for air.

5.9 Verify that Eq. (5.22) gives the common-core area of the two axially misaligned step-index fibers shown in Fig. 5.10. If $d = 0.1a$, what is the coupling efficiency?

5.10 Consider the three fibers having the properties listed in Table P5.10. Use Eq. (5.23) to complete this table for connector losses (in decibels) due to the indicated axial misalignments.

5.11 Show that when the axial misalignment of d is small compared with the core radius a, Eq. (5.30) can be approximated by Eq. (5.31). Compare

Table P5.10

	Coupling loss (dB) for given axial misalignment (μm)			
Fiber Size: core diameter (μm)/ clad diameter (μm)	1	3	5	10
50/125			0.590	
62.5/125				
100/140				

Table P5.16

Parameter	Fiber 1	Fiber 2
Core index n_1	1.46	1.48
Index difference Δ	0.010	0.015
Core radius a	50 μm	62.5 μm
Profile factor α	2.00	1.80

Eqs. (5.30) and (5.31) in terms of P_T/P as a function of d/a over the range $0 \leq d/a \leq 0.4$.

5.12 Consider an optical fiber that has a core refractive index $n_1 = 1.48$ and a numerical aperture NA = 0.20. Using Eqs. (5.32), (5.33), and (5.34), plot the three mechanical misalignment losses in decibels over the following ranges:
(*a*) $0 \leq d/a \leq 1.0$
(*b*) $0 \leq s/a \leq 3.0$
(*c*) $0 \leq \theta \leq 10°$

5.13 Using Eqs. (5.19) and (5.20), show that Eq. (5.35) gives the coupling loss for two fibers with unequal core radii. Plot the coupling loss in decibels as a function of a_R/a_E for $0.5 \leq a_R/a_E \leq 1.0$.

5.14 Using Eqs. (5.19) and (5.20), show that Eq. (5.36) gives the coupling loss for two fibers with unequal axial numerical apertures. Plot this coupling loss in decibels as a function of $NA_R(0)/NA_E(0)$ over the range $0.5 \leq NA_R(0)/NA_E(0) \leq 1.0$.

5.15 Show that Eq. (5.37) gives the coupling loss for two fibers with different core refractive-index profiles. Plot this coupling loss in decibels as a function of α_R/α_E over the range $0.75 \leq \alpha_R/\alpha_E \leq 1.0$. Take $\alpha_E = 2.0$.

5.16 Consider two multimode graded-index fibers that have the characteristics given in Table P5.16. If these two fibers are perfectly aligned with no gap between them, calculate the splice losses and the coupling efficiencies for the following cases:
(*a*) Light going from fiber 1 to fiber 2.
(*b*) Light going from fiber 2 to fiber 1.

5.17 Consider two identical single-mode optical fibers that have a core refractive index $n_1 = 1.48$ and a mode-field radius $w = 5$ μm

at 1300 nm. Assume the material between the fiber ends is air with an index of 1.0. Using Eq. (5.43), plot the following connector losses in decibels (in each case, vary only one alignment parameter, keeping the other two mechanical misalignments fixed at zero):
(*a*) Lateral offset over the range $0 \leq d \leq 4$ μm.
(*b*) Longitudinal offset over the range $0 \leq s \leq 40$ μm.
(*c*) Angular misalignment over the range $0 \leq \theta \leq 2°$.

5.18 Assuming that a single-mode connector has no losses due to extrinsic factors, show that a 10 percent mismatch in mode-field diameters yields a loss of 0.05 dB.

5.19 Consider two fibers that have core refractive indices $n_0 = 1.463$. Assume these fibers are separated by a gap $s = 0.22$ μm, which is filled with a material that has a refractive index $n_1 = 1.467$. Use Eq. (5.44) to plot the return loss as a function of the high-index-layer thickness h over the range $0 \leq h \leq 0.15$ μm for values of n_2 equal to 1.467, 1.500, and 1.600.

5.20 Consider a connector in which the fibers have angled end faces and core refractive indices $n_0 = 1.470$, as shown in Fig. 5.23. Assume the gap $s = 1$ μm and the face angle $\theta = 8°$. The connector experiences no loss when the gap is filled with an index-matching material that has $n_1 = 1.470$. Thus, use the relationship

$$L(\lambda) = 10 \log \left[\frac{T(\lambda, \, n_1 = 1.470, \, \theta = 8°)}{T(\lambda, \, n_1 = 1.00, \, \theta = 8°)} \right]$$

to plot the throughput loss with an air gap ($n_1 = 1.0$ for air) as a function of wavelength for 700 nm $\leq 1 \leq$ 1800 nm.

REFERENCES

1. Y. Uematsu, T. Ozeki, and Y. Unno, "Efficient power coupling between an MH LED and a taper-ended multimode fiber," *IEEE J. Quantum Electron.*, vol. 15, pp. 86–92, Feb. 1979.

2. H. Kuwahara, M. Sasaki, and N. Tokoyo, "Efficient coupling from semiconductor lasers into single-mode fibers with tapered hemispherical ends," *Appl. Opt.*, vol. 19, pp. 2578–2583, Aug. 1980.

3. D. Marcuse, "LED fundamentals: Comparison of front and edge-emitting diodes," *IEEE J. Quantum Electron.*, vol. 13, pp. 819–827, Oct. 1977.

4. B.E.A. Saleh and M. Teich, *Fundamentals of Photonics*, Wiley, Hoboken, NJ, 2nd ed., 2007.

5. TIA/EIA-455-54B, *Mode Scrambler Requirements for Overfilled Launching Conditions to Multimode Fibers*, Aug. 1998.

6. M. Born and E. Wolf, *Principles of Optics*, Cambridge University Press, Oxford, 7th ed., 1999.

7. K. Sakai, M. Kawano, H. Aruga, S.-I. Takagi, S.-I. Kaneko, J. Suzuki, M. Negishi, Y. Kondoh, and K.-I. Fukuda, "Photodiode packaging technique using ball lens and offset parabolic mirror," *J. Lightwave Technol.*, vol. 27, no. 17, pp. 3874–3879, Sept. 2009.

8. A. Nicia, "Lens coupling in fiber-optic devices: Efficiency limits," *Appl. Opt.*, vol. 20, pp. 3136–3145, Sept. 1981.

9. Z. L. Liau, D. Z. Tsang, and J. N. Walpole, "Simple compact diode-laser/microlens packaging," *IEEE J. Quantum Electron.*, vol. 33, pp. 457–461, Mar. 1997.

10. C. A. Edwards, H. M. Presby, and C. Dragone, "Ideal microlenses for laser to fiber coupling," *J. Lightwave Tech.*, vol. 11, pp. 252–257, Feb. 1993.

11. K. Keränen, J. T. Mäkinen, K. T. Kautio, J. Ollila, J. Petäjä, V. Heikkinen, J. Heilala, and P. Karioja, "Fiber pigtailed multimode laser module based on passive device alignment on an LTCC substrate," *IEEE Trans. Adv. Packag.*, vol. 29, pp. 463–472, Aug. 2006.

12. C. Tsou and Y. S. Huang, "Silicon-based packaging platform for light-emitting diode," *IEEE Trans. Adv. Packag.*, vol. 29, pp. 607–614, Aug. 2006.

13. See any general physics or introductory optics book; for example: (*a*) F. A. Jenkins and H. E. White, *Fundamentals of Optics*, McGraw-Hill, New York, 4th ed., 2001; (*b*) E. Hecht, *Optics*, Addison-Wesley, Boston, 4th ed., 2002; (*c*) R. Ditteon, *Modern Geometrical Optics*, Wiley, New York, 1997.

14. M. C. Hudson, "Calculation of the maximum optical coupling efficiency into multimode optical waveguides," *Appl. Opt.*, vol. 13, pp. 1029–1033, May 1974.

15. G. K. Khoe and G. Kuyt, "Realistic efficiency of coupling light from GaAs laser diodes into parabolic-index optical fibers," *Electron. Lett.*, vol. 14, pp. 667–669, Sept. 28, 1978.

16. S. M. Yeh, S. Y. Huang, and W. H. Cheng, "A new scheme of conical-wedge-shaped fiber endface for coupling between high-power laser diodes and single-mode fibers," *J. Lightwave Tech.*, vol. 23, pp. 1781–1786, Apr. 2005.

17. Y. Fu and N.K.A. Bryan, "A novel one step integration of edge-emitting laser diode with micro-elliptical lens using focused ion beam direct deposition," *IEEE Trans. Semiconductor Manufacturing*, vol. 15, pp. 2–8, Feb. 2002.

18. I. Moerman, P. P. Van Daele, and P. M. Demeester, "A review of fabrication technologies for the monolithic integration of tapers with III-V semiconductor devices," *IEEE J. Sel. Topics Quantum Electron.*, vol. 3, pp. 1308–1320, Dec. 1997.

19. A. Safaai-Jazi and V. Suppanitchakij, "A tapered graded-index lens: Analysis of transmission properties and applications in fiber-optic communication systems," *IEEE J. Quantum Electron.*, vol. 33, pp. 2159–2166, Dec. 1997.

20. G. Sialm, D. Lenz, D. Erni, G. L. Bona, C. Kromer, M. X. Jungo, T. Morf, F. Ellinger, and H. Jäckel, "Comparison of simulation and measurement of dynamic fiber-coupling effects for high-speed multimode VCSELs," *J. Lightwave Tech.*, vol. 23, pp. 2318–2330, July 2005.

21. D. H. Rice and G. E. Keiser, "Short-haul fiber-optic link connector loss," *Int. Wire & Cable Symp. Proc.*, Nov. 13–15, 1984, Reno, NV, pp. 190–192.

22. Y. Daido, E. Miyauchi, and T. Iwama, "Measuring fiber connection loss using steady-state power distribution: A method," *Appl. Opt.*, vol. 20, pp. 451–456, Feb. 1981.

23. M. J. Hackert, "Evolution of power distributions in fiber optic systems: Development of a measurement strategy," *Fiber & Integrated Optics*, vol. 8, pp. 163–167, 1989.

24. P. DiVita and U. Rossi, "Realistic evaluation of coupling loss between different optical fibers," *J. Opt. Common.*, vol. 1, pp. 26–32, Sept. 1980; "Evaluation of splice losses induced by mismatch in fiber parameters," *Opt. Quantum Electron.*, vol. 13, pp. 91–94, Jan. 1981.

25. M. J. Adams, D. N. Payne, and F.M.E. Staden, "Splicing tolerances in graded index fibers," *Appl. Phys. Lett.*, vol. 28, pp. 524–526, May 1976.

26. D. Gloge, "Offset and tilt loss in optical fiber splices," *Bell Sys. Tech. J.*, vol. 55, pp. 905–916, Sept. 1976.

27. T. C. Chu and A. R. McCormick, "Measurement of loss due to offset, end separation and angular misalignment in graded index fibers excited by an incoherent source," *Bell Sys. Tech. J.*, vol. 57, pp. 595–602, Mar. 1978.

28. P. DiVita and U. Rossi, "Theory of power coupling between multimode optical fibers," *Opt. Quantum Electron.*, vol. 10, pp. 107–117, Jan. 1978.

29. C. M. Miller, "Transmission vs. transverse offset for parabolic-profile fiber splices with unequal core diameters," *Bell Sys. Tech. J.*, vol. 55, pp. 917–927, Sept. 1976.

30. D. Gloge and E.A.J. Marcatili, "Multimode theory of graded-core fibers," *Bell Sys. Tech. J.*, vol. 52, pp. 1563–1578, Nov. 1973.

31. H. G. Unger, *Planar Optical Waveguides and Fibres*, Clarendon, Oxford, 1977.

32. F. L. Thiel and R. M. Hawk, "Optical waveguide cable connection," *Appl. Opt.*, vol. 15, pp. 2785–2791, Nov. 1976.

33. F. L. Thiel and D. H. Davis, "Contributions of optical-waveguide manufacturing variations to joint loss," *Electron. Lett.,* vol. 12, pp. 340–341, June 1976.

34. S. C. Mettler, "A general characterization of splice loss for multimode optical fibers," *Bell Sys. Tech. J.*, vol. 58, pp. 2163–2182, Dec. 1979.

35. D. J. Bond and P. Hensel, "The effects on joint losses of tolerances in some geometrical parameters of optical fibres," *Opt. Quantum Electron.*, vol. 13, pp. 11–18, Jan. 1981.

36. S. C. Mettler and C. M. Miller, "Optical fiber splicing," in S. E. Miller and I. P. Kaminow, eds., *Optical Fiber Telecommunications–II*, Academic, New York, 1988.

37. V.C.Y. So, R. P. Hughes, J. B. Lamont, and P. J. Vella, "Splice loss measurement using local launch and detect," *J. Lightwave Tech.*, vol. LT-5, pp. 1663–1666, Dec. 1987.

38. D. W. Peckham and C. R. Lovelace, "Multimode optical fiber splice loss: Relating system and laboratory measurements," *J. Lightwave Tech.*, vol. LT-5, pp. 1630–1636, Dec. 1987.

39. G. Cancellieri and U. Ravaioli, *Measurements of Optical Fibers and Devices*, Artech House, Dedham, MA, 1984.

40. TIA/EIA-455-57B (FOTP-57B), *Preparation and Examination of Optical Fiber Endface for Testing Purposes*, Feb. 1996.

41. TIA/EIA-455-179 (FOTP-179), *Inspection of Cleaved Fiber Endfaces by Interferometry*, May 1988.

42. G. Van Steenberge, P. Geerinck, S. Van Put, J. Watté, H. Ottevaere, H. Thienpont, and P. Van Daele, "Laser cleaving of glass fibers and glass fiber arrays," *J. Lightwave Tech.*, vol. 23, pp. 609–614, Feb. 2005.

43. D. M. Fye, R. Olshansky, J. LaCourse, W. Powazinik, and R. B. Lauer, "Low-current, 1.3-μm edge-emitting LED for single-mode subscriber loop applications," *Electron. Lett.*, vol. 22, pp. 87–88, Jan. 1986.

44. T. Tsubota, Y. Kashima, H. Takano, and Y. Hirose, "InGaAsP/InP long-wavelength high-efficiency edge-emitting LED for single-mode fiber optic communication," *Fiber Integr. Optics.*, vol. 7, no. 4, pp. 353–360, 1988.

45. D. N. Christodoulides, L. A. Reith, and M. A. Saifi, "Coupling efficiency and sensitivity of an LED to a single-mode fiber," *Electron. Lett.*, vol. 22, pp. 1110–1111, Oct. 1986.

46. L. A. Reith and P. A. Shumate, "Coupling sensitivity of an edge-emitting LED to a single-mode fiber," *J. Lightwave Tech.*, vol. LT-5, pp. 29–34, Jan. 1987.

47. B. Hillerich, "New analysis of LED to a single-mode fiber coupling," *Electron. Lett.*, vol. 22, pp. 1176–1177, Oct. 1986; "Efficiency and alignment tolerances of LED to a single-mode fiber coupling—theory and experiment," *Opt. Quantum Electron.*, vol. 19, no. 4, pp. 209–222, July 1987.

48. W. van Etten, "Coupling of LED light into a single-mode fiber," *J. Opt. Commun.*, vol. 9, no. 3, pp. 100–101, Sept. 1988.

49. D. N. Christodoulides, L. A. Reith, and M. A. Saifi, "Theory of LED coupling to single-mode fibers," *J. Lightwave Tech.*, vol. LT-5, pp. 1623–1629, Nov. 1987.

50. J. T. Krause, C. R. Kurkjian, and U. C. Paek, "Strength of fusion splices for fiber lightguides," *Electron. Lett.*, vol. 17, pp. 232–233, Mar. 1981.

51. T. Yamada, Y. Ohsato, M. Yoshinuma, T. Tanaka, and K.-I. Itoh, "Arc fusion splicer with profile alignment system for high-strength low-loss optical submarine cable," *J. Lightwave Tech.*, vol. 4, pp. 1204–1210, Aug. 1986.

52. F. El-Diasty, "Laser-scattering-based method for investigation of ultra-low-loss arc fusion-spliced single-mode optical fibers," *J. Lightwave Tech.*, vol. 22, pp. 1539–1542, June 2004.

53. A. D. Yablon and R. T. Bise, "Low-loss high-strength microstructured fiber fusion splices using GRIN fiber lenses," *IEEE Photonics Technol. Lett.*, vol. 17, pp. 118–120, Jan. 2005.

54. T. Katagiri, M. Tachikura, and Y. Murakami, "Evaluation of fiber slippage and core-axis offset distribution in optical fiber mechanical splice," *J. Lightwave Tech.*, vol. 19, pp. 195–204, Feb. 2001.

55. S. Fevrier, P. Viale, C. Kaczmarek, and P. Chanclou, "Low splice loss between 34-μm core diameter Bragg fibre and G-652 fibre by using micro-optics," *Electron. Lett.*, vol. 41, pp. 1166–1167, 13 Oct. 2005.

56. A. D. Yablon, *Optical Fiber Fusion Splicing*, Springer, New York, 2005.

57. Y. Yang, "Attenuation splice control in the manufacture of fiber optical communication system," *IEEE Trans. Control Sys. Technol.*, vol. 14, pp. 170–175, Jan. 2006.

58. P. Melman and W. J. Carlsen, "Elastic-tube splice performance with single-mode and multimode fibers," *Electron. Lett.*, vol. 18, pp. 320–321, Apr. 1982.

59. B. Bourliaguet, C. Paré, F. Émond, A. Croteau, A. Proulx, and R. Vallée, "Microstructured fusion splicing," *Optics Express*, vol. 11, pp. 3412–3417, Dec. 2003.

60. J. T. Krause, W. A. Reed, and K. L. Walker, "Splice loss of single-mode fiber as related to fusion time, temperature, and index profile alteration," *J. Lightwave Tech.*, vol. 4, pp. 837–840, July 1986.

61. E. Serafini, "Statistical approach to the optimization of optical fiber fusion splicing in the field," *J. Lightwave Tech.*, vol. 7, pp. 431–435, Feb. 1989.

62. D. Marcuse, D. Gloge, and E.A.J. Marcatili, "Guiding properties of fibers," in S. E. Miller and A. G. Chynoweth, eds., *Optical Fiber Telecommunications*, Academic, New York, 1979.

63. D. Marcuse, "Loss analysis of single-mode splices," *Bell Sys. Tech. J.*, vol. 56, pp. 703–718, May 1977.

64. W. C. Young and D. R. Frey, "Fiber connectors," in S. E. Miller and I. P. Kaminow, eds., *Optical Fiber Telecommunications–II*, Academic, New York, 1988.

65. P. Chanclou, H. Ramanitra, P. Gravey, and M. Thual, "Design and performance of expanded mode fiber using microoptics," *J. Lightwave Tech.*, vol. 20, pp. 836–842, May 2002.

66. M. He, J. Bu, B. H. Ong, and X. Yuan, "Two-microlens coupling scheme with revolved hyperboloid sol–gel microlens arrays for high-power-efficiency optical coupling," *J. Lightwave Tech.*, vol. 24, pp. 2940–2945, July 2006.

67. S. I. Chang and J. B. Yoon, "A 3-D planar microlens for an effective monolithic optical interconnection system," *IEEE Photonics Technol. Lett.*, vol. 18, pp. 814–816, Apr. 2006.

68. M. Uekawa, H. Sasaki, D. Shimura, K. Kotani, Y. Maeno, and T. Takamori, "Surface-mountable silicon microlens for low-cost laser modules,"

IEEE Photonics Technol. Lett., vol. 15, pp. 945–947, July 2003.

69. Y. Sunaga, R. Takahashi, T. Tokoro, and M. Kobayashi, "2 Gbit/s small form factor fiber-optic transceiver for single mode optical fiber," *IEEE Trans. Adv. Packag.*, vol. 23, pp. 176–181, May 2000.

70. M. Owen, "Agilent Technologies' singlemode small form factor (SFF) module incorporates micromachined silicon, automated passive alignment, and non-hermetic packaging to enable the next generation of low-cost fiber optic transceivers," *IEEE Trans. Adv. Packag.*, vol. 23, pp. 182–187, May 2000.

71. M. Labudovic and M. Burka, "Finite element analysis of post-weld shift during fiber pigtail of 980 nm pump lasers," *IEEE Trans. Adv. Packag.*, vol. 26, pp. 41–46, Feb. 2003.

72. R. Zhang and F. G. Shi, "Manufacturing of laser diode modules: Integration and automation of laser diode-fiber alignment and RIN characterization," *IEEE Trans. Adv. Packag.*, vol. 26, pp. 128–132, May 2003.

73. M. Takaya, "Design of enhanced MT-RJ type connector for physical contact," *IEEE Trans. Compon. Packag. Technol.*, vol. 27, pp. 283–290, June 2004.

74. S. Nemota and T. Makimoto, "Analysis of splice loss in single-mode fibers using a gaussian field approximation," *Optical Quantum Electron.*, vol. 11, no. 5, pp. 447–457, Sept. 1979.

75. M. Kihara, S. Nagasawa, and T. Tanifuji, "Return loss characteristics of optical fiber connectors," *J. Lightwave Tech.*, vol. 14, pp. 1986–1991, Sept. 1996.

76. M. Kihara, S. Nagasawa, and T. Tanifuji, "Design and performance of an angled physical contact type multifiber connector," *J. Lightwave Tech.*, vol. 14, pp. 542–548, Apr. 1996.

CHAPTER 6

Photodetectors

At the output end of an optical transmission line, there must be a receiving device that interprets the information contained in the optical signal. The first element of this receiver is a photodetector. The photodetector senses the luminescent power falling upon it and converts the variation of this optical power into a correspondingly varying electric current. Since the optical signal is generally weakened and distorted when it emerges from the end of the fiber, the photodetector must meet very high performance requirements. Among the foremost of these requirements are a high response or sensitivity in the emission wavelength range of the optical source being used, a minimum addition of noise to the system, and a fast response speed or sufficient bandwidth to handle the desired data rate. The photodetector should also be insensitive to variations in temperature, be compatible with the physical dimensions of the optical fiber, have a reasonable cost in relation to the other components of the system, and have a long operating life.

Several different types of photodetectors are in existence. Among these are photomultipliers, pyroelectric detectors, and semiconductor-based photoconductors, phototransistors, and photodiodes.[1–6] However, many of these detectors do not meet one or more of the foregoing requirements. Photomultipliers consisting of a photocathode and an electron multiplier packaged in a vacuum tube are capable of very high gain and very low noise. Unfortunately, their large size and high voltage requirements make them unsuitable for optical fiber systems. Pyroelectric photodetectors involve the conversion of photons to heat. Photon absorption results in a temperature change of the detector material. This gives rise to a variation in the dielectric constant, which is usually measured as a capacitance change. The response of this detector is quite flat over a broad spectral band, but its speed is limited by the detector cooling rate after it has been excited. Its principal use is for detecting high-speed laser pulses, and it is not well suited for optical fiber systems.

Of the semiconductor-based photodetectors, the photodiode is used almost exclusively for fiber optic systems because of its small size, suitable material, high sensitivity, and fast response time. The two types of photodiodes used are the *pin* photodetector and the avalanche photodiode (APD). We examine the fundamental characteristics of these two device types in the following sections. In describing these components, we make use of the elementary principles of semiconductor device physics given in Sec. 4.1. Basic discussions of photodetection processes can be found in various texts.[4–12]

6.1 Physical Principles of Photodiodes

6.1.1 The *pin* Photodetector

The most common semiconductor photodetector is the *pin* photodiode, shown schematically in Fig. 6.1. The device structure consists of *p* and *n* regions separated by a very lightly *n*-doped intrinsic (*i*) region. In normal operation a sufficiently large reverse-bias voltage is applied across the device so that the intrinsic region is fully depleted of carriers. That is, the intrinsic *n* and *p* carrier concentrations are negligibly small in comparison with the impurity concentration in this region.

As a photon flux Φ penetrates into a semiconductor, it will be absorbed as it progresses through the material. Suppose P_{in} is the optical power level falling on the photodetector at $x = 0$ and $P(x)$ is the power level at a distance x into the material. Then the incremental change $dP(x)$ in the optical power level as this photon flux passes through an incremental distance dx in the semiconductor is given by $dP(x) = -\alpha_s(\lambda) P(x) dx$, where $\alpha_s(\lambda)$ is the *photon absorption coefficient* at a wavelength λ. Integrating this relationship gives the power level at a distance x into the material as

$$P(x) = P_{in}\exp(-\alpha_s x) \tag{6.1}$$

Figure 6.1 gives an example of the power level as a function of the penetration depth into the intrinsic region, which has a width w. The width of the *p* region typically is very thin so that little radiation is absorbed there.

When an incident photon has an energy greater than or equal to the bandgap energy of the semiconductor material, the photon can give up its energy and excite an electron from the valence band to the conduction
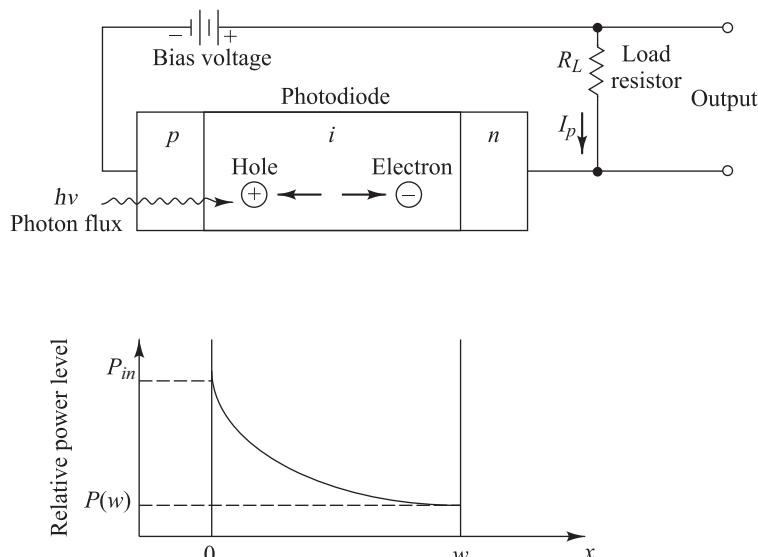
Fig. 6.1 *Representation of a pin photodiode circuit with an applied reverse bias. An incident optical power level decays exponentially inside the device.*

Example 6.1 If the absorption coefficient of $In_{0.53}Ga_{0.47}As$ is 0.8 μm^{-1} at 1550 nm, what is the penetration depth at which $P(x)/P_{in} = 1/e = 0.368$?

Solution: From Eq. (6.1),

$$\frac{P(x)}{P_{in}} = \exp(-a_s x) = \exp[(-0.8)x] = 0.368$$

Therefore

$$-0.8\,x = \ln 0.368 = -0.9997$$

which yields $x = 1.25\ \mu m$.

Example 6.2 A high-speed $In_{0.53}Ga_{0.47}As$ *pin* photodetector is made with a depletion layer thickness of 0.15 μm. What percent of incident photons are absorbed in this photodetector at 1310 nm if the absorption coefficient is 1.5 μm^{-1} at this wavelength?

Solution: From Eq. (6.1), the optical power level at $x = 0.15\ \mu m$ relative to the incident power level is

$$\frac{P(0.15)}{P_{in}} = \exp(-a_s x) = \exp[(-1.5)0.15] = 0.80$$

Therefore only 20 percent of the incident photons are absorbed.

band. This absorption process generates mobile electron–hole pairs, as Fig. 6.2 shows. These electrons and holes are known as *photocarriers*, since they are photon-generated charge carriers that are available to produce a current flow when a bias voltage is applied across the device. The number of charge carriers is controlled by the concentration level of impurity elements that are intentionally added to the material (see Sec. 4.1). The photodetector is normally designed so that these carriers are generated mainly in the depletion region (the depleted intrinsic region) where most of the incident light is absorbed. The high electric field present in the depletion region causes the carriers to separate and be collected across the reverse-biased junction. This gives rise to a current flow in an external circuit, with one electron flowing for every carrier pair generated. This current flow is known as the *photocurrent*.

As the charge carriers flow through the material, some electron–hole pairs will recombine and hence disappear. On the average, the charge carriers move a distance L_n or L_p for electrons and holes, respectively. This distance is known as the *diffusion length*. The time it takes for an electron or hole to recombine is known as the *carrier lifetime* and is represented by τ_n and τ_p, respectively. The lifetimes and the diffusion lengths are related by the expressions

$$L_n = (D_n \tau_n)^{1/2} \quad \text{and} \quad L_p = (D_p \tau_p)^{1/2}$$

where D_n and D_p are the electron and hole diffusion coefficients (or constants), respectively, which are expressed in units of centimeters squared per second.

The dependence of the optical absorption coefficient on wavelength is shown in Fig. 6.3 for several photodiode materials.[13] As the curves clearly show, α_s depends strongly on the wavelength. Thus a particular semiconductor material can be used only over a limited wavelength range. The upper wavelength cutoff
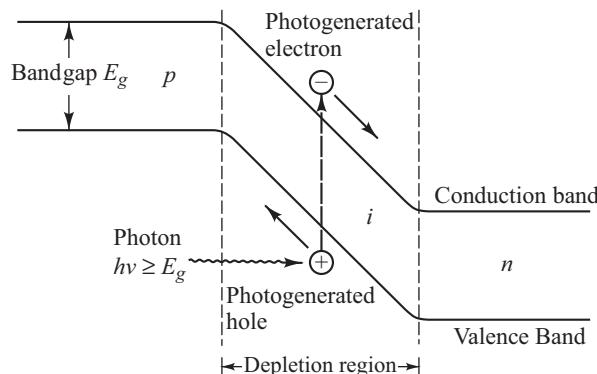

Fig. 6.2 *Simple energy-band diagram for a pin photodiode. Photons with energies greater than or equal to the bandgap energy E_g can generate free electron–hole pairs that act as photocurrent carriers.*
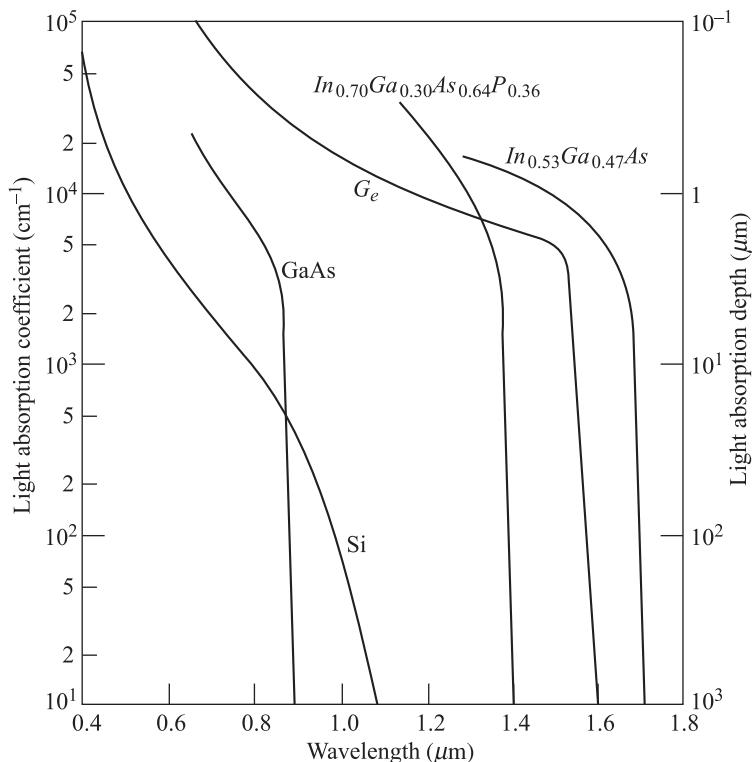
Fig. 6.3 *Optical absorption coefficient as a function of wavelength of several different photodetector materials. (Modified with permission from Miller, Marcatili, and Li,[13] © 1973, IEEE.)*

λ_c is determined by the bandgap energy E_g of the material. If E_g is expressed in units of electron volts (eV), then λ_c is given in units of micrometers (μm) by

$$\lambda_c(\mu\text{m}) = \frac{hc}{E_g} = \frac{1.24}{E_g\,(\text{eV})} \tag{6.2}$$

Example 6.3 A photodiode is constructed of GaAs, which has a bandgap energy of 1.43 eV at 300 K. What is the cutoff wavelength of this device?

Solution: From Eq. (6.2), the long-wavelength cutoff is

$$\lambda_c = \frac{hc}{E_g} = \frac{(6.625 \times 10^{-34}\,\text{J}\cdot\text{s})(3 \times 10^8\,\text{m}/\text{s})}{(1.43\,\text{eV})(1.6 \times 10^{-19}\,\text{J}/\text{ev})} = 869\,\text{nm}$$

This GaAs photodiode will not operate for photons of wavelength greater than 869 nm.

The cutoff wavelength is about 1.06 μm for Si and 1.6 μm for Ge. For longer wavelengths, the photon energy is not sufficient to excite an electron from the valence to the conduction band.

At the lower-wavelength end, the photoresponse cuts off as a result of the very large values of a_s at the shorter wavelengths. In this case, the photons are absorbed very close to the photodetector surface, where the recombination time of the generated electron–hole pairs is very short. The generated carriers thus recombine before they can be collected by the photodetector circuitry.

If the depletion region has a width w, then, from Eq. (6.1), the total power absorbed in the distance w is

$$P_{absorbed}(w) = \int_0^w \alpha_s P_{in} \exp(-\alpha_s x)\, dx = P_{in}(1 - e^{-\alpha_s w}) \tag{6.3}$$

If we take into account a reflectivity R_f at the entrance face of the photodiode, then the primary photocurrent I_p resulting from the power absorption of Eq. (6.3) is given by

$$I_p = \frac{q}{h\nu} P_{in}(1 - e^{-\alpha_s w})(1 - R_f) \tag{6.4}$$

where P_{in} is the optical power incident on the photodetector, q is the electron charge, and $h\nu$ is the photon energy.

Two important characteristics of a photodetector are its quantum efficiency and its response speed. These parameters depend on the material bandgap, the operating wavelength, and the doping and thickness of the $p, i,$ and n regions of the device. The *quantum efficiency* η is the number of the electron–hole carrier pairs generated per incident–absorbed photon of energy $h\nu$ and is given by

$$\eta = \frac{\text{number of electron–hole pairs generated}}{\text{number of incident–absorbed photons}} = \frac{I_p/q}{P_{in}/h\nu} \tag{6.5}$$

Here, I_p is the photocurrent generated by a steady-state optical power P_{in} incident on the photodetector.

In a practical photodiode, 100 photons will create between 30 and 95 electron–hole pairs, thus giving a detector quantum efficiency ranging from 30 to 95 percent. To achieve a high quantum efficiency, the depletion layer must be thick enough to permit a large fraction of the incident light to be absorbed. However, the thicker the depletion layer, the longer it takes for the photogenerated carriers to drift across the reverse-biased junction. Since the carrier drift time determines the response speed of the photodiode, a compromise has to be made between response speed and quantum efficiency. We discuss this further in Sec. 6.3.

The performance of a photodiode is often characterized by the *responsivity* $\mathscr{R}$. This is related to the quantum efficiency by

$$\mathscr{R} = \frac{I_p}{P_{in}} = \frac{\eta q}{h\nu} \tag{6.6}$$

This parameter is quite useful because it specifies the photocurrent generated per unit of optical power. Typical *pin* photodiode responsivities as a function of wavelength are shown in Fig. 6.4. Representative

---

Example 6.4 In a 100-ns pulse, 6×10^6 photons at a wavelength of 1300 nm fall on an InGaAs photodetector. On the average, 5.4×10^6 electron–hole (e–h) pairs are generated. The quantum efficiency is found from Eq. (6.5) as

$$\eta = \frac{\text{number of e–h pairs generated}}{\text{number of incident photons}} = \frac{5.4 \times 10^6}{6 \times 10^6} = 0.90$$

Thus here the quantum efficiency at 1300 nm is 90 percent.
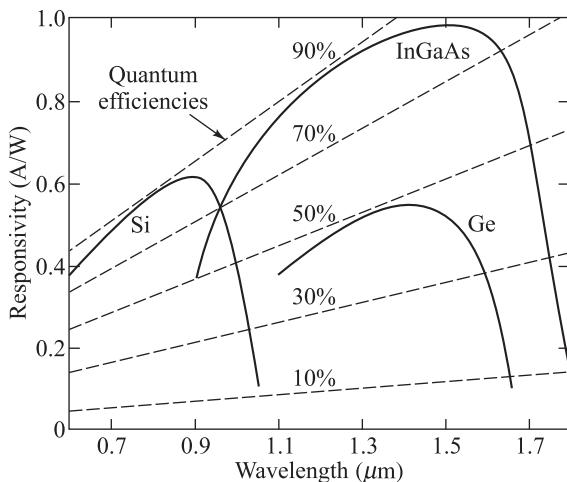
Fig. 6.4 *Comparison of the responsivity and quantum efficiency as a function of wavelength for pin photodiodes constructed of three different materials*

Example 6.5 Photons of energy 1.53×10^{-19} J are incident on a photodiode which has a responsivity of 0.65 A/W.

If the optical power level is 10 μW, then from Eq. (6.6) the photocurrent generated is

$$I_p = \mathscr{R}P_{in} = (0.65 \text{ A/W})(10 \text{ }\mu\text{W}) = 6.5 \text{ }\mu\text{A}$$

---

values are 0.65 A/W for silicon at 900 nm and 0.45 A/W for germanium at 1.3 μm. For InGaAs, typical values are 0.9 A/W at 1.3 μm and 1.0 A/W at 1.55 μm.

In most photodiodes the quantum efficiency is independent of the power level falling on the detector at a given photon energy. Thus the responsivity is a linear function of the optical power. That is, the photocurrent I_p is directly proportional to the optical power P_{in} incident upon the photodetector, so that the responsivity $\mathscr{R}$ is constant at a given wavelength (a given value of $h\nu$). Note, however, that the quantum efficiency is not a constant at all wavelengths because it varies according to the photon energy. Consequently, the responsivity is a function of the wavelength and of the photodiode material (since different materials have different bandgap energies). For a given material, as the wavelength of the incident photon becomes longer, the photon energy becomes less than that required to excite an electron from the valence band to the conduction band. The responsivity thus falls off rapidly beyond the cutoff wavelength, as can be seen in Fig. 6.4.

6.1.2 Avalanche Photodiodes

Avalanche photodiodes (APDs) internally multiply the primary signal photocurrent before it enters the input circuitry of the following amplifier. This increases receiver sensitivity, since the photocurrent is multiplied before encountering the thermal noise associated with the receiver circuit. In order for carrier multiplication to take place, the photogenerated carriers must traverse a region where a very high electric field is present. In this high-field region, a photogenerated electron or hole can gain enough energy so that it ionizes bound electrons in the valence band upon colliding with them. This carrier multiplication mechanism is known as *impact ionization*. The newly created carriers are also accelerated by the high

Example 6.6 As shown in Fig. 6.4, for the wavelength range 1300 nm < λ < 1600 nm, the quantum efficiency for InGaAs is around 90 percent. Thus in this wavelength range the responsivity is

$$\mathscr{R} = \frac{\eta q}{h\nu} = \frac{\eta q \lambda}{hc} = \frac{(0.90)(1.6 \times 10^{-19}\,\text{C})\lambda}{(6.625 \times 10^{-34}\,\text{J·s})(3 \times 10^8\,\text{m/s})}$$

$$= 7.25 \times 10^5 \lambda$$

For example, at 1300 nm we have

$$\mathscr{R} = [7.25 \times 10^5\,(\text{A/W})/\text{m}]\,(1.30 \times 10^{-6}\,\text{m})$$

$$= 0.92\,\text{A/W}$$

At wavelengths higher than 1600 nm, the photon energy is not sufficient to excite an electron from the valence band to the conduction band. For example, $In_{0.53}Ga_{0.47}As$ has an energy gap E_g = 0.73 eV, so that from Eq. (6.2) the cutoff wavelength is

$$\lambda_c = \frac{1.24}{E_g} = \frac{1.24}{0.73} = 1.7\,\mu\text{m}$$

At wavelengths less than 1100 nm, the photons are absorbed very close to the photodetector surface, where the recombination rate of the generated electron–hole pairs is very short. The responsivity thus decreases rapidly for smaller wavelengths, since many of the generated carriers do not contribute to the photocurrent.

electric field, thus gaining enough energy to cause further impact ionization. This phenomenon is the *avalanche effect*. Below the diode breakdown voltage a finite total number of carriers are created, whereas above breakdown the number can be infinite.

A commonly used structure for achieving carrier multiplication with very little excess noise is the *reach-through* construction[14–16] shown in Fig. 6.5. The reach-through avalanche photodiode (RAPD) is composed of a high-resistivity *p*-type material deposited as an epitaxial layer on a *p*$^+$ (heavily doped *p*-type) substrate. A *p*-type diffusion or ion implant is then made in the high-resistivity material, followed by the construction of an *n*$^+$ (heavily doped *n*-type) layer. For silicon, the dopants used to form these layers are normally boron and phosphorus, respectively. This configuration is referred to as *p*$^+\pi p n^+$ *reach-through* structure. The π layer is basically an intrinsic material that inadvertently has some *p* doping because of imperfect purification. Section 6.5 describes more complex structures used for InGaAs APDs.

The term "reach-through" arises from the photodiode operation. When a low reverse-bias voltage is applied, most of the potential drop is across the *pn*$^+$ junction. The depletion layer widens with increasing bias until a certain voltage is reached at which the peak electric field at the *pn*$^+$ junction is about 5–10 percent below that needed to cause avalanche breakdown. At this point, the depletion layer just "reaches through" to the nearly intrinsic π region.
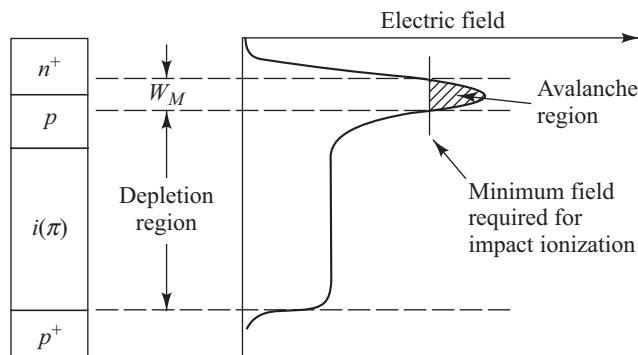

Fig. 6.5 *Reach-through avalanche photodiode structure and the electric fields in the depletion and multiplication regions*

In normal usage, the RAPD is operated in the fully depleted mode. Light enters the device through the p^+ region and is absorbed in the π material, which acts as the collection region for the photogenerated carriers. Upon being absorbed, the photon gives up its energy, thereby creating electron–hole pairs, which are then separated by the electric field in the π region. The photogenerated electrons drift through the π region in the pn^+ junction, where a high electric field exists. It is in this high-field region that carrier multiplication takes place.

The average number of electron–hole pairs created by a carrier per unit distance traveled is called the *ionization rate*. Most materials exhibit different *electron ionization rates* α and *hole ionization rates* β. Experimentally obtained values of α and β for five different semiconductor materials are shown in Fig. 6.6. The ratio $k = \beta/\alpha$ of the two ionization rates is a measure of the photodetector performance. As we shall see in Sec. 6.4, avalanche photodiodes constructed of materials in which one type of carrier largely dominates impact ionization exhibit low noise and large gain-bandwidth products. Of all the materials shown in Fig. 6.6, only silicon has a significant difference between electron and hole ionization rates.[17–23]
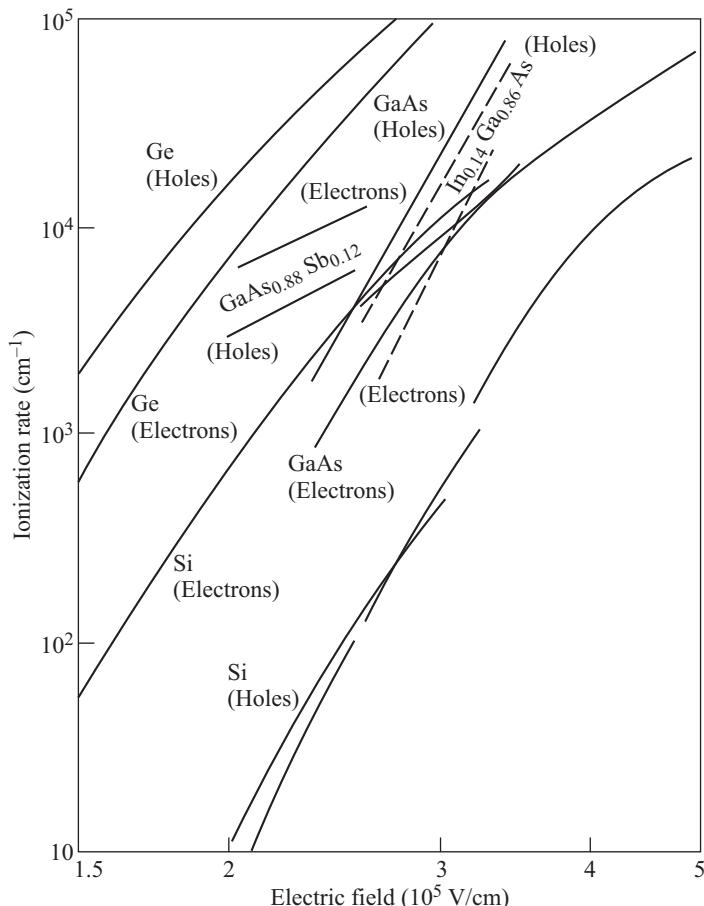
Fig. 6.6 *Carrier ionization rates obtained experimentally for silicon, germanium, gallium arsenide, gallium arsenide antimonide, and indium gallium arsenide. (Reproduced with permission from Melchior.[1])*

The multiplication M for all carriers generated in the photodiode is defined by

$$M = \frac{I_M}{I_p} \qquad (6.7)$$

where I_M is the average value of the total multiplied output current and I_p is the primary unmultiplied photocurrent defined in Eq. (6.4). In practice, the avalanche mechanism is a statistical process, since not every carrier pair generated in the diode experiences the same multiplication. Thus, the measured value of M is expressed as an average quantity.

Typical current gains for different wavelengths[15] as a function of bias voltage for a silicon reach-through avalanche photodiode are shown in Fig. 6.7. The dependence of the gain on the excitation wavelength is attributable to mixed initiation of the avalanche process by electrons and holes when most of the light is absorbed in the n^+p region close to the detector surface. This is especially noticeable at short wavelengths, where a larger portion of the optical power is absorbed closer to the surface than at longer wavelengths. In silicon, since the ionization coefficient for holes is smaller than that for electrons, the total current gain is reduced at the short wavelengths.

Analogous to the *pin* photodiode, the performance of an APD is characterized by its responsivity $\mathscr{R}_{APD}$, which is given by

$$\mathscr{R}_{APD} = \frac{nq}{h\nu} M = \mathscr{R}M \qquad (6.8)$$

where $\mathscr{R}$ is the unity gain responsivity.
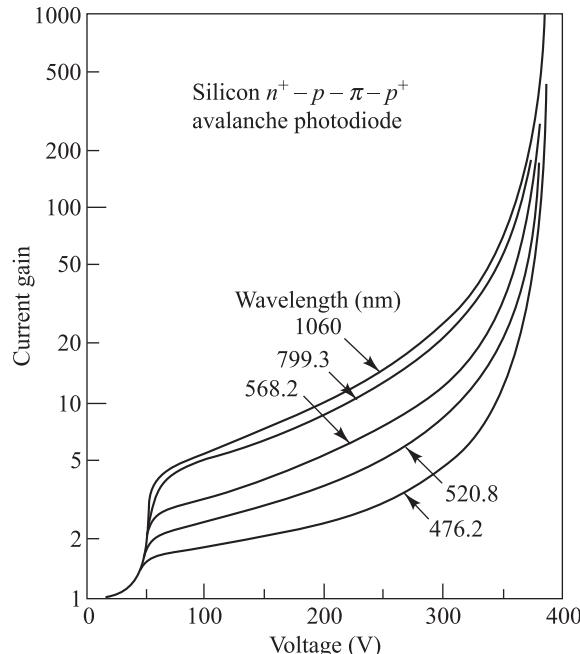
Fig. 6.7 *Typical room-temperature current gains of a silicon reach-through avalanche photodiode for different wavelengths as a function of bias voltage. (Reproduced with permission from Melchior, Hartman, Schinke, and Seidel,[15] © 1978, AT&T.)*

Example 6.7 A given silicon avalanche photodiode has a quantum efficiency of 65 percent at a wavelength of 900 nm. Suppose 0.5 μW of optical power produces a multiplied photocurrent of 10 μA. What is the multiplication M?

Solution: From Eq. (6.6), the primary photocurrent is

$$I_p = \mathscr{R}P_{in} = \frac{\eta q}{h\nu}P_{in} = \frac{\eta q\lambda}{hc}P_{in}$$

$$= \frac{(0.65)(1.6\times 10^{-19}\text{C})(9\times 10^{-7}\text{m})}{(6.625\times 10^{-34}\text{ J}\cdot\text{s})(3\times 10^{8}\text{ m/s})}5\times 10^{-7}\text{ W}$$

$$= 0.235\ \mu\text{A}$$

From Eq. (6.7), the multiplication is

$$M = \frac{I_M}{I_p} = \frac{10\,\mu\text{A}}{0.235\,\mu\text{A}} = 43$$

Thus the primary photocurrent is multiplied by a factor of 43.

6.2 Photodetector Noise

In fiber optic communication systems, the photodiode is generally required to detect very weak optical signals. Detection of the weakest possible optical signals requires that the photodetector and its following amplification circuitry be optimized so that a given signal-to-noise ratio is maintained. The power signal-to-noise ratio *S/N* (also designated by SNR) at the output of an optical receiver is defined by

$$SNR = \frac{S}{N} = \frac{\text{signal power from photocurrent}}{\text{photodetector noise power} + \text{amplifier noise power}} \tag{6.9}$$

The noise sources in the receiver arise from the photodetector noises resulting from the statistical nature of the photon-to-electron conversion process and the thermal noises associated with the amplifier circuitry.

To achieve a high signal-to-noise ratio, the following conditions should be met:
1. The photodetector must have a high quantum efficiency to generate a large signal power.
2. The photodetector and amplifier noises should be kept as low as possible.

In most applications, it is the noise currents that determine the minimum optical power level that can be detected, since the photodiode quantum efficiency is normally close to its maximum possible value.

The sensitivity of a photodetector in an optical fiber communication system is describable in terms of the *minimum detectable optical power*. This is the optical power necessary to produce a photocurrent of the same magnitude as the root mean square (rms) of the total noise current, or equivalently, a signal-to-noise ratio of 1. A thorough understanding of the source, characteristics, and interrelationships of the various noises in a photodetector is therefore necessary to make a reliable design and to evaluate optical receivers.

6.2.1 Noise Sources

To see the interrelationship of the different types of noises affecting the signal-to-noise ratio, let us examine the simple receiver model and its equivalent circuit shown in Fig. 6.8. The photodiode has a small series resistance R_s, a total capacitance C_d consisting of junction and packaging capacitances, and a bias (or load) resistor R_L. The amplifier following the photodiode has an input capacitance C_a and a resistance R_a. For practical purposes, R_s is much smaller than the load resistance R_L and can be neglected.
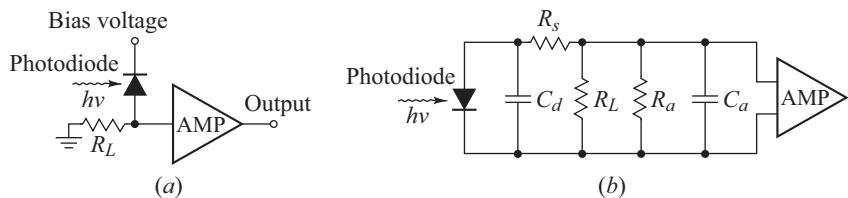
Fig. 6.8 *(a) Simple model of a photodetector receiver, and (b) its equivalent circuit*

If a modulated signal of optical power $P(t)$ falls on the detector, the primary photocurrent $i_{ph}(t)$ generated is

$$i_{ph}(t) = \frac{\eta q}{h\nu} P(t) \tag{6.10}$$

The primary current consists of a dc value I_p, which is the average photocurrent due to the signal power, and a signal component $i_p(t)$. For *pin* photodiodes the mean-square signal current $\langle i_s^2 \rangle$ is

$$\langle i_s^2 \rangle = \sigma_{s,\,pin}^2 = \langle i_p^2(t) \rangle \tag{6.11a}$$

where σ is the variance. For avalanche photodetectors,

$$\langle i_s^2 \rangle = \sigma_{s,\,APD}^2 = \langle i_p^2(t) \rangle M^2 \tag{6.11b}$$

where M is the average of the statistically varying avalanche gain as defined in Eq. (6.7). For a sinusoidally varying input signal of modulation index m, the signal component $\langle i_p^2 \rangle$ is of the form (see Prob. 6.5)

$$\langle i_p^2(t) \rangle = \sigma_p^2 = \frac{m^2}{2} I_p^2 \tag{6.12}$$

where m is defined in Eq. (4.54).

The principal noise sources associated with photodetectors that have no internal gain are quantum noise, dark-current noise generated in the bulk material of the photodiode, and surface leakage current noise. The *quantum* or *shot noise* arises from the statistical nature of the production and collection of photoelectrons when an optical signal is incident on a photodetector. It has been demonstrated[24] that these statistics follow a Poisson process. Since the fluctuations in the number of photocarriers created from the photoelectric effect are a fundamental property of the photodetection process, they set the lower limit on the receiver sensitivity when all other conditions are optimized. The shot noise current has a mean-square value in a receiver bandwidth B_e that is proportional to the average value of the photocurrent I_p.

$$\langle i_{shot}^2 \rangle = \sigma_{shot}^2 = 2q I_p B_e M^2 F(M) \tag{6.13}$$

where $F(M)$ is a noise figure associated with the random nature of the avalanche process. From experimental results, it has been found that to a reasonable approximation $F(M) \simeq M^x$, where x (with $0 \leq x \leq 1.0$) depends on the material. This is discussed in more detail in Sec. 6.4. For *pin* photodiodes, M and $F(M)$ are unity.

The photodiode dark current is the current that continues to flow through the bias circuit of the device when no light is incident on the photodiode. This is a combination of bulk and surface currents. The *bulk dark current i_{DB}* arises from electrons and/or holes that are thermally generated in the *pn* junction of the photodiode. In an APD, these liberated carriers also get accelerated by the high electric field present at the *pn* junction, and are therefore multiplied by the avalanche gain mechanism. The mean-square value of this current is given by

$$\langle i_{DB}^2 \rangle = \sigma_{DB}^2 = 2qI_D M^2 F(M)B_e \tag{6.14}$$

where I_D is the primary (unmultiplied) detector bulk dark current.

The *surface dark current* is also referred to as a *surface leakage current* or simply the leakage current. It is dependent on surface defects, cleanliness, bias voltage, and surface area. An effective way of reducing surface dark current is through the use of a guard ring structure that shunts surface leakage currents away from the load resistor. The mean-square value of the surface dark current is given by

$$\langle i_{DS}^2 \rangle = \sigma_{DS}^2 = 2qI_L B_e \tag{6.15}$$

where I_L is the surface leakage current. Note that since avalanche multiplication is a bulk effect, the surface dark current is not affected by the avalanche gain.

A comparison[20] of typical dark currents for Si, Ge, GaAs, and $In_xGa_{1-x}As$ photodiodes is given in Fig. 6.9 as a function of applied voltage normalized to the breakdown voltage V_B. Note that for $In_xGa_{1-x}As$ photodiodes the dark current increases with the composition x. Under a reverse bias, both dark currents also increase with the area. The surface dark current increases in proportion to the square root of the active area, and the bulk dark current is directly proportional to the area.

Since the dark currents and the signal current are uncorrelated, the total mean-square
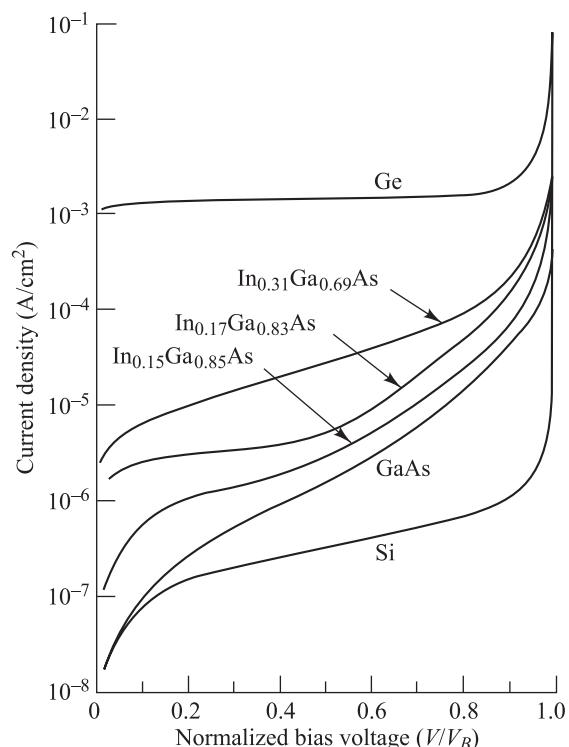

Fig. 6.9 *A comparison of typical dark currents for Si, Ge, GaAs, and InGaAs photodiodes as a function of normalized bias voltage. (Reproduced with permission from Susa, Yamauchi, and Kanbe,[20] © 1980, IEEE.)*

photodetector noise current $\langle i_N^2 \rangle$ can be written as

$$\langle i_N^2 \rangle = \sigma_N^2 = \langle i_{\text{shot}}^2 \rangle + \langle i_{DB}^2 \rangle + \langle i_{DS}^2 \rangle = \sigma_{\text{shot}}^2 + \sigma_{DB}^2 + \sigma_{DS}^2$$

$$= 2q(I_p + I_D)M^2 F(M)B_e + 2qI_L B_e \qquad (6.16)$$

To simplify the analysis of the receiver circuitry, we assume here that the amplifier input impedance is much greater than the load resistance, so that its thermal noise is much smaller than that of R_L. The photodetector load resistor contributes a mean-square thermal (Johnson) noise current

$$\langle i_T^2 \rangle = \sigma_T^2 = \frac{4k_B T}{R_L} B_e \qquad (6.17)$$

where k_B is Boltzmann's constant and T is the absolute temperature. This noise can be reduced by using a load resistor that is large but still consistent with the receiver bandwidth requirements.

Example 6.8   An InGaAs *pin* photodiode has the following parameters at a wavelength of 1300 nm: $I_D = 4$ nA, $\eta = 0.90$, $R_L = 1000\ \Omega$, and the surface leakage current is negligible. The incident optical power is 300 nW (−35 dBm), and the receiver bandwidth is 20 MHz. Find the various noise terms of the receiver.

Solution: (*a*) First, we need to find the primary photocurrent. From Eq. (6.6),

$$I_p = \mathscr{R}P_{in} = \frac{\eta q}{h\nu}P_{in} = \frac{\eta q\lambda}{hc}P_{in}$$

$$= \frac{(0.90)(1.6 \times 10^{-19}\ \text{C})(1.3 \times 10^{-6}\ \text{m})}{(6.625 \times 10^{-34}\ \text{J·s})(3 \times 10^{8}\ \text{m/s})}\, 3 \times 10^{-7}\ \text{W}$$

$$= 0.282\ \mu\text{A}$$

(*b*) From Eq. (6.13), the mean-square shot noise current for a *pin* photodiode is

$$\langle i_{\text{shot}}^2 \rangle = 2qI_p B_e$$

$$= 2(1.6 \times 10^{-19}\ \text{C})(0.282 \times 10^{-6}\ \text{A})(20 \times 10^{6}\ \text{Hz})$$

$$= 1.80 \times 10^{-18}\ \text{A}^2$$

or $\langle i_{\text{shot}}^2 \rangle^{1/2} = 1.34$ nA

(*c*) From Eq. (6.14), the mean-square dark current is

$$\langle i_{DB}^2 \rangle = 2qI_D B_e$$

$$= 2(1.6 \times 19^{-19}\ \text{C})(4 \times 10^{-9}\ \text{A})(20 \times 10^{6}\ \text{Hz})$$

$$= 2.56 \times 10^{-20}\ \text{A}^2$$

or

$$\langle i_{DB}^2 \rangle^{1/2} = 0.16\ \text{nA}$$

(*d*) The mean-square thermal noise current for the receiver is found from Eq. (6.17) as

$$\langle i_T^2 \rangle = \frac{4k_B T}{R_L} B_e = \frac{4(1.38 \times 10^{-23}\ \text{J/K})(293\,\text{K})}{1\ \text{k}\Omega} B_e$$

$$= 323 \times 10^{-18}\ \text{A}^2$$

or

$$\langle i_T^2 \rangle^{1/2} = 18\ \text{nA}$$

Thus for this receiver the rms thermal noise current is about 14 times greater than the rms shot noise current and about 100 times greater than the rms dark current.

6.2.2 Signal-to-Noise Ratio

Substituting Eqs. (6.11), (6.16), and (6.17) into Eq. (6.9) for the signal-to-noise ratio at the input of the amplifier, we have

$$\frac{S}{N} = \frac{\left\langle i_p^2 \right\rangle M^2}{2q(I_p + I_D) M^2 F(M) B_e + 2qI_L B_e + 4k_B TB_e / R_L} \tag{6.18a}$$

In general, in the expression given by Eq. (6.18a) one can ignore the negligible leakage current. Furthermore, the term involving I_D can be dropped when the average signal current is much larger than the dark current. The signal-to-nose ratio then becomes

$$\frac{S}{N} = \frac{\left\langle i_p^2 \right\rangle M^2}{2qI_p M^2 F(M) B_e + 4k_B TB_e / R_L} \tag{6.18b}$$

When the optical signal power is relatively high, then the shot noise power is much greater than the thermal noise power. In this case the SNR is called *shot-noise limited* or *quantum limited*. When the optical signal power is low, then thermal noise usually dominates over the shot noise. In this case the SNR is referred to as being *thermal-noise limited*.

> ***Example 6.9***   Consider the InGaAs *pin* photodiode described in Example 6.8. What is the SNR in decibels?
>
> ***Solution:***   Since the dark current noise is negligible compared to the shot noise and thermal noise, we can substitute the numerical results into Eq. (6.18b) to get
>
> $$\frac{S}{N} = \frac{(0.282 \times 10^{-6})^2}{1.80 \times 10^{-18} + 323 \times 10^{-18}} = 245$$
>
> In decibels the SNR is
>
> $$\frac{S}{N} = 10 \log 245 = 23.9$$

When *pin* photodiodes are used, the dominating noise currents are those of the detector load resistor (the thermal current i_T) and the active elements of the amplifier circuitry (i_{amp}). For avalanche photodiodes, the thermal noise is of lesser importance and the photodetector noises usually dominate.[25]

From Eq. (6.18a), it can be seen that the signal power is multiplied by M^2 and the shot noise plus bulk dark current is multiplied by $M^2 F(M)$. The surface-leakage current is not altered by the avalanche gain mechanism. Since the noise figure $F(M)$ increases with M, there always exists an optimum value of M that maximizes the signal-to-noise ratio. The optimum gain at the maximum signal-to-noise ratio can be found by differentiating Eq. (6.18a) with respect to M, setting the result equal to zero, and solving for M. Doing so for a sinusoidally modulated signal, with $m = 1$ and $F(M)$ approximated by M^x, yields

$$M_{\text{opt}}^{x+2} = \frac{2qI_L + 4k_B T / R_L}{xq(I_p + I_D)} \tag{6.19}$$

Example 6.10 Consider a Si APD operating at 300°K and with a load resistor $R_L = 1000 \, \Omega$. For this APD assume the responsivity $\mathcal{R} = 0.65$ A/W and let $x = 0.3$. (*a*) If dark current is neglected and 100 nW of optical power falls on the photodetector, what is the optimum avalanche gain? (*b*) What is the SNR if $B_e = 100$ MHz? (*c*) How does the SNR of this APD compare with the corresponding SNR of a Si *pin* photodiode? Assume the leakage current is negligible.

Solution: (*a*) Neglecting dark current and with $I_p = \mathcal{R}P$, we have

$$M_{opt} = \left(\frac{4k_B T}{xqR_L \mathcal{R}P} \right)^{1/(x+2)}$$

$$= \left[\frac{4(1.38 \times 10^{-23})(300)}{0.3(1.60 \times 10^{-19})(1000)(0.65)(100 \times 10^{-9})} \right]^{1/2.3} = 42$$

(*b*) Neglecting dark current and with $F(M) = M^x = (42)^{0.3}$, we have

$$\text{SNR} = \frac{(\mathcal{R}PM)^2}{\left[2q\mathcal{R}PM^{2.3} + \left(\frac{4k_B T}{R_L} \right) \right] B_e}$$

$$= \frac{[(0.65)(100 \times 10^{-9})(42)]^2}{\left[2(1.6 \times 10^{-19})(0.65)(100 \times 10^{-9})42^{2.3} + \left(\frac{4(1.38 \times 10^{-23})(300)}{(1000)} \right) \right](100 \times 10^6)} = 659$$

or in decibels, SNR = 10 log 659 = 28.2 dB

(*c*) For a *pin* photodiode with $M = 1$, the above equation yields SNR(*pin*) = 2.3 = 3.5 dB. Thus, compared to a *pin* photodiode, the APD improves the SNR by 24.7 dB.

6.2.3 Noise-Equivalent Power

The sensitivity of a photodetector in an optical fiber communication system is describable in terms of the *minimum detectable optical power*. This is the optical power necessary to produce a photocurrent of the same magnitude as the root mean-square (rms) of the total noise current, or equivalently, a signal-to-noise ratio of 1. This optical signal power is referred to as the *noise equivalent power* or NEP, which is designated in units of $\text{W}/\sqrt{\text{Hz}}$.

As an example, consider the thermal-noise limited case for a *pin* photodiode. Then Eq. (6.18*b*) becomes

$$\text{SNR} = \mathcal{R}^2 P^2 / (4k_B T B_e / R_L)$$

To find the NEP we set the SNR equal to 1 and solve for P to obtain

$$\text{NEP} = \frac{P_{\min}}{\sqrt{B_e}} = \sqrt{4k_B T / R_L} / \mathcal{R}$$

Example 6.11 For an InGaAs photodetector operating at 1550 nm, $\mathcal{R} = 0.90$ A/W. What is the NEP in the thermal-noise limited case if the load resistor is $R_L = 1000 \, \Omega$ and $T = 300°$K?

Solution: From the above expression for NEP we have

$$\text{NEP} = \sqrt{4(1.38 \times 10^{-23})(300)/1000}/0.90$$

$$= 4.52 \times 10^{-12} \, \text{W}/\sqrt{\text{Hz}}$$

6.3 Detector Response Time

6.3.1 Depletion Layer Photocurrent

To understand the frequency response of photodiodes, let us first consider the schematic representation of a reverse-biased *pin* photodiode shown in Fig. 6.10. Light enters the device through the *p* layer and produces electron–hole pairs as it is absorbed in the semiconductor material. Those electron–hole pairs that are generated in the depletion region or within a diffusion length of it will be separated by the reverse-bias-voltage-induced electric field, thereby leading to a current flow in the external circuit as the carriers drift across the depletion layer.

Under steady-state conditions, the total current density J_{tot} flowing through the reverse-biased depletion layer is

$$J_{\text{tot}} = J_{\text{dr}} + J_{\text{diff}} \tag{6.20}$$

Here, J_{dr} is the drift current density resulting from carriers generated inside the depletion region, and J_{diff} is the diffusion current density arising from the carriers that are produced outside of the depletion layer in the bulk of the semiconductor (i.e., in the *n* and *p* regions) and diffuse into the reverse-biased junction. The drift current density can be found from Eq. (6.4):

$$J_{\text{dr}} = \frac{I_p}{A} = q\Phi_0(1 - e^{-\alpha_s w}) \tag{6.21}$$

where A is the photodiode area and Φ_0 is the incident photon flux per unit area given by

$$\Phi_0 = \frac{P_{in}(1 - R_f)}{Ah\nu} \tag{6.22}$$

The surface *p* layer of a *pin* photodiode is normally very thin. The diffusion current is thus principally determined by hole diffusion from the bulk *n* region. The hole diffusion in this material can be determined by the one-dimensional diffusion equation[10]
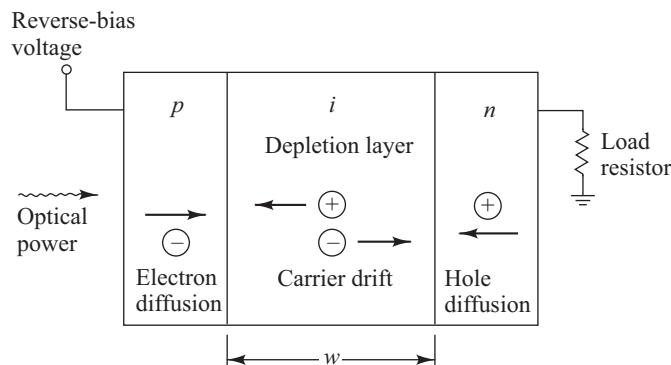
Fig. 6.10 *Schematic representation of a reverse-biased pin photodiode*

$$D_P \frac{\partial^2 p_n}{\partial x^2} - \frac{p_n - p_{n0}}{\tau_p} + G(x) = 0 \tag{6.23}$$

where D_p is the hole diffusion coefficient, p_n is the hole concentration in the *n*-type material, τ_p is the excess hole lifetime, p_{n0} is the equilibrium hole density, and $G(x)$ is the electron–hole generation rate given by

$$G(x) = \Phi_0 \alpha_s\, e^{-\alpha_s x} \tag{6.24}$$

From Eq. (6.23), the diffusion current density is found to be (see Prob. 6.10)

$$J_{\text{diff}} = q\Phi_0 \frac{\alpha_s L_P}{1 + \alpha_s L_p} e^{-\alpha_s w} + q p_{n0} \frac{D_p}{L_p} \tag{6.25}$$

Substituting Eqs. (6.21) and (6.25) into Eq. (6.20), we have that the total current density through the reverse-biased depletion layer is

$$J_{\text{tot}} = q\Phi_0 \left(1 - \frac{e^{-\alpha_s w}}{1 + \alpha_s L_p} \right) + q p_{n0} \frac{D_p}{L_p} \tag{6.26}$$

The term involving p_{n0} is normally small, so that the total photogenerated current is proportional to the photon flux Φ_0.

6.3.2 Response Time

The response time of a photodiode together with its output circuit (see Fig. 6.8) depends mainly on the following three factors:

1. The transit time of the photocarriers in the depletion region.
2. The diffusion time of the photocarriers generated outside the depletion region.
3. The *RC* time constant of the photodiode and its associated circuit.

The photodiode parameters responsible for these three factors are the absorption coefficient α_s, the depletion region width w, the photodiode junction and package capacitances, the amplifier capacitance, the detector load resistance, the amplifier input resistance, and the photodiode series resistance. The photodiode series resistance is generally only a few ohms and can be neglected in comparison with the large load resistance and the amplifier input resistance.

Let us first look at the transit time of the photocarriers in the depletion region. The response speed of a photodiode is fundamentally limited by the time it takes photogenerated carriers to travel across the depletion region. This transit time t_d depends on the carrier drift velocity v_d and the depletion layer width w, and is given by

$$t_d = \frac{w}{v_d} \tag{6.27}$$

In general, the electric field in the depletion region is large enough so that the carriers have reached their scattering-limited velocity. For silicon, the maximum velocities for electrons and holes are 8.4×10^6 and 4.4×10^6 cm/s, respectively, when the field strength is on the order of 2×10^4 V/cm. A typical high-speed silicon photodiode with a 10-μm depletion layer width thus has a response time limit of about 0.1 ns.
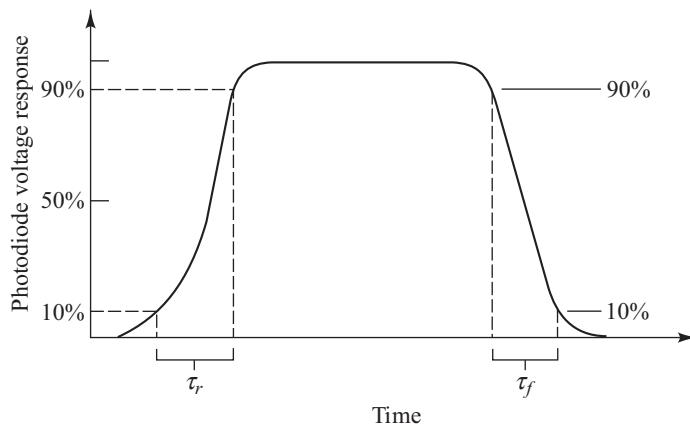
Fig. 6.11 *Photodiode response to an optical input pulse showing the 10- to 90-percent rise time and the 10- to 90-percent fall time*

The diffusion processes are slow compared with the drift of carriers in the high-field region. Therefore, to have a high-speed photodiode, the photocarriers should be generated in the depletion region or so close to it that the diffusion times are less than or equal to the carrier drift times. The effect of long diffusion times can be seen by considering the photodiode response time. This response time is described by the rise time and fall time of the detector output when the detector is illuminated by a step input of optical radiation. The rise time τ_r is typically measured from the 10- to 90-percent points of the leading edge of the output pulse, as is shown in Fig. 6.11. For fully depleted photodiodes the rise time τ_r and fall time τ_f are generally the same. However, they can be different at low bias levels where the photodiode is not fully depleted, since the photon collection time then starts to become a significant contributor to the rise time. In this case, charge carriers produced in the depletion region are separated and collected quickly. On the other hand, electron–hole pairs generated in the *n* and *p* regions must slowly diffuse to the depletion region before they can be separated and collected. A typical response time of a partially depleted photodiode is shown in Fig. 6.12. The fast carriers allow the device output to rise to 50 percent of its maximum value in approximately 1 ns, but the slow carriers cause a relatively long delay before the output reaches its maximum value.
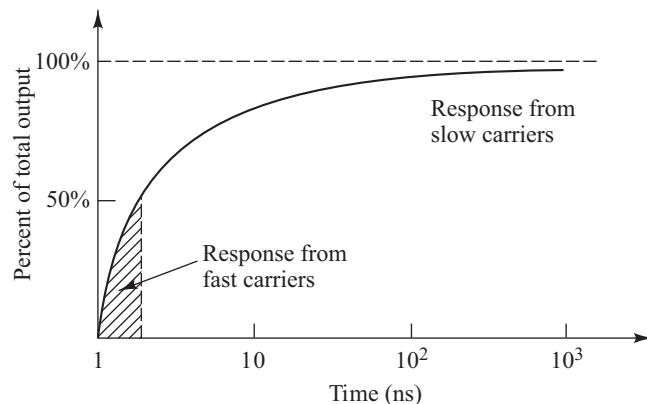
Fig. 6.12 *Typical response time of a photodiode that is not fully depleted*

To achieve a high quantum efficiency, the depletion layer width must be much larger than l/α_s (the inverse of the absorption coefficient), so that most of the light will be absorbed. The response to a rectangular input pulse of a low-capacitance photodiode having $w \gg 1/\alpha_s$ is shown in Fig. 6.13b. The rise and fall times of the photodiode follow the input pulse quite well. If the photodiode capacitance is larger, the response time becomes limited by the RC time constant of the load resistor R_L and the photodiode capacitance. The photodetector response then begins to appear as that shown in Fig. 6.13c.

If the depletion layer is too narrow, any carriers created in the undepleted material would have to diffuse back into the depletion region before they could be collected. Devices with very thin depletion regions thus tend to show distinct slow- and fast-response components, as shown in Fig. 6.13d. The fast component in the rise time is due to carriers generated in the depletion region, whereas the slow component arises from the diffusion of carriers that are created with a distance L_n from the edge of the depletion region. At the end of the optical pulse, the carriers in the depletion region are collected quickly, which results in the fast-detector-response component in the fall time. The diffusion of carriers that are within a distance L_n of the depletion region edge appears as the slowly decaying tail at the end of the pulse. Also, if w is too thin, the junction capacitance will become excessive. The junction capacitance C_j is

$$C_j = \frac{\epsilon_s A}{w} \tag{6.28}$$

where ϵ_s = the permittivity of the semiconductor material = $\epsilon_0 K_s$
         K_s = the semiconductor dielectric constant
         $\epsilon_0 = 8.8542 \times 10^{-12}$ F/m is the free-space permittivity
         A = the diffusion layer area

This excessiveness will then give rise to a large RC time constant, which limits the detector response time. A reasonable compromise between high-frequency response and high quantum efficiency is found for absorption region thicknesses between l/α_s and $2/\alpha_s$.

If R_T is the combination of the load and amplifier input resistances and C_T is the sum of the photodiode and amplifier capacitances, as shown in Fig. 6.8, the detector behaves approximately like a simple RC low-pass filter with a passband given by

$$B_c = \frac{1}{2\pi R_T C_T} \tag{6.29}$$

(a)   Rectangular input pulse

(b) $w \gg 1/\alpha_s$ Small C_j

(c) $w \gg 1/\alpha_s$ Large C_j

(d)   Slow component   Fast component   Diffusion component $w \leq 1/\alpha_s$ Small C_j
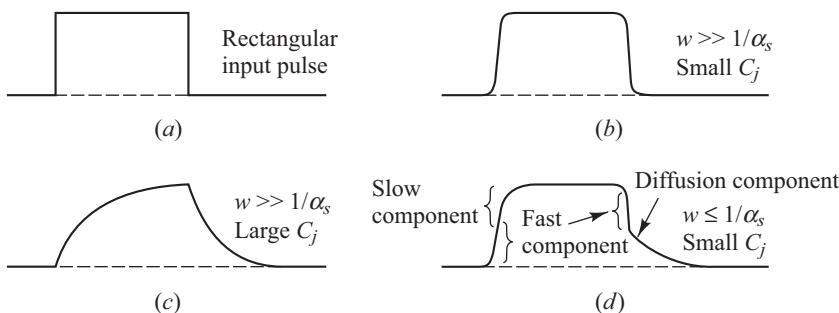
Fig. 6.13 *Photodiode pulse responses under various detector parameters*

Example 6.12 If the photodiode capacitance is 3 pF, the amplifier capacitance is 4 pF, the load resistor is 1 kΩ, and the amplifier input resistance is 1 MΩ, then $C_T = 7$ pF and $R_T = 1$ kΩ, so that the circuit bandwidth is

$$B_c = \frac{1}{2\pi R_T C_T} = 23\,\text{MHz} \qquad (6.30)$$

If we reduce the photodector load resistance to 50 Ω, then the circuit bandwidth becomes $B_c = 455$ MHz.

6.3.3 Double-Heterostructure Photodiodes

By using a double-heterostructure design similar to that employed in semiconductor lasers, one can significantly increase the performance of *pin* photodiodes. As shown in Fig. 6.14, in this design the central intrinsic layer (the depletion region) is sandwiched between different *p*- and *n*-type semiconductor layers. The bandgaps of these layers are chosen such that only the intrinsic region absorbs light. A popular *pin* photodiode structure for 1250-to-1650-nm applications uses In$_{1-x}$Ga$_x$As for the intrinsic layer with $x = 0.47$ and InP for the adjacent lattice-matched *p*- and *n*-type layers. The motivation

Fig. 6.14 *Double-heterostructure design of a pin photodiode*

for this structure is that the bandgap of InP is 1.35 eV, which from Eq. (6.2) means that it is transparent to light with a wavelength greater than 920 nm. When $x = 0.47$ the bandgap of the intrinsic region is 0.73 eV, which gives a cutoff wavelength of about 1700 nm in that material.

In the generic photodiode structure shown in Fig. 6.14, light enters the device from the top through the *n* layer. A common configuration is to have the top metallic contact be in the form of a ring, thus enabling light to enter through the area inside the ring.

6.4 Avalanche Multiplication Noise

The avalanche process is statistical in nature, since not every photogenerated carrier pair undergoes the same multiplication.[26–28] The probability distribution of possible gains that any particular electron–hole pair might experience is sufficiently wide so that the mean-square gain is greater than the average gain squared. That is, if *m* denotes the statistically varying gain, then

$$\langle m^2 \rangle > \langle m \rangle^2 = M^2 \qquad (6.31)$$

where the symbols $\langle\ \rangle$ denote an ensemble average and $\langle m \rangle = M$ is the average carrier gain defined in Eq. (6.7). Since the noise created by the avalanche process depends on the mean-square gain $\langle m^2 \rangle$, the noise in an avalanche photodiode can be relatively high. From experimental observations it has been found that, in general, $\langle m^2 \rangle$ can be approximated by

$$\langle m^2 \rangle \simeq M^{2+x} \qquad (6.32)$$

where the exponent *x* varies between 0 and 1.0 depending on the photodiode material and structure.

The ratio of the actual noise generated in an avalanche photodiode to the noise that would exist if all carrier pairs were multiplied by exactly M is called the *excess noise factor F* and is defined by

$$F = \frac{\langle m^2 \rangle}{\langle m \rangle^2} = \frac{\langle m^2 \rangle}{M^2} \tag{6.33}$$

This excess noise factor is a measure of the increase in detector noise resulting from the randomness of the multiplication process. It depends on the ratio of the electron and hole ionization rates and on the carrier multiplication.

The derivation of an expression for F is complex, since the electric field in the avalanche region (of width W_M, as shown in Fig. 6.5) is not uniform, and both holes and electrons produce impact ionization. For injected electrons and holes, the excess noise factors are[27]

$$F_e = \frac{k_2 - k_1^2}{1 - k_2} M_e + 2\left[1 - \frac{k_1(1 - k_1)}{1 - k_2}\right] - \frac{(1 - k_1)^2}{M_e(1 - k_2)} \tag{6.34}$$

$$F_h = \frac{k_2 - k_1^2}{k_1^2(1 - k_2)} M_h - 2\left[\frac{k_2(1 - k_1)}{k_1^2(1 - k_2)} - 1\right] + \frac{(1 - k_1)^2 k_2}{k_1^2(1 - k_2)M_h} \tag{6.35}$$

where the subscripts e and h refer to electrons and holes, respectively. The weighted ionization rate ratios k_1 and k_2 take into account the nonuniformity of the gain and the carrier ionization rates in the avalanche region. They are given by

$$k_1 = \frac{\int_0^{W_M} \beta(x) M(x)\, dx}{\int_0^{W_M} \alpha(x) M(x)\, dx} \tag{6.36}$$

$$k_2 = \frac{\int_0^{W_M} \beta(x) M^2(x)\, dx}{\int_0^{W_M} \alpha(x) M^2(x)\, dx} \tag{6.37}$$

where $\alpha(x)$ and $\beta(x)$ are the electron and hole ionization rates, respectively.

Normally, to a first approximation k_1 and k_2 do not change much with variations in gain and can be considered as constant and equal. Thus Eqs. (6.34) and (6.35) can be simplified as[14]

$$\begin{aligned} F_e &= M_e\left[1 - (1 - k_{\text{eff}})\left(1 - \frac{1}{M_e}\right)^2\right] \\ &= k_{\text{eff}} M_e + \left(2 - \frac{1}{M_e}\right)(1 - k_{\text{eff}}) \end{aligned} \tag{6.38}$$

for electron injection, and

$$F_h = M_h \left[1 - \left(1 - \frac{1}{k'_{\text{eff}}} \right) \left(1 - \frac{1}{M_h} \right)^2 \right]$$

$$= k'_{\text{eff}} M_h - \left(2 - \frac{1}{M_h} \right) (k'_{\text{eff}} - 1) \qquad (6.39)$$

for hole injection, where the effective ionization rate ratios are

$$k_{\text{eff}} = \frac{k_2 - k_1^2}{1 - k_2} \simeq k_2$$

$$k'_{\text{eff}} = \frac{k_{\text{eff}}}{k_1^2} \simeq \frac{k_2}{k_1^2} \qquad (6.40)$$

Figure 6.15 shows F_e as a function of the average electron gain M_e for various values of the effective ionization rate ratio k_{eff}. If the ionization rates are equal, the excess noise is at its maximum so that F_e is at its upper limit of M_e. As the ratio β/α decreases from unity, the electron ionization rate starts to be the dominant contributor to impact ionization, and the excess noise factor becomes smaller. If only electrons cause ionization, $\beta = 0$ and F_e reaches its lower limit of 2.

This shows that to keep the excess noise factor at a minimum, it is desirable to have small values of k_{eff}. Referring back to Fig. 6.6, we thus see the superiority of silicon over other materials for making avalanche photodiodes. The effective ionization rate ratio k_{eff} varies between 0.015 and 0.035 for silicon, between 0.3 and 0.5 for indium gallium arsenide, and between 0.6 and 1.0 for germanium.
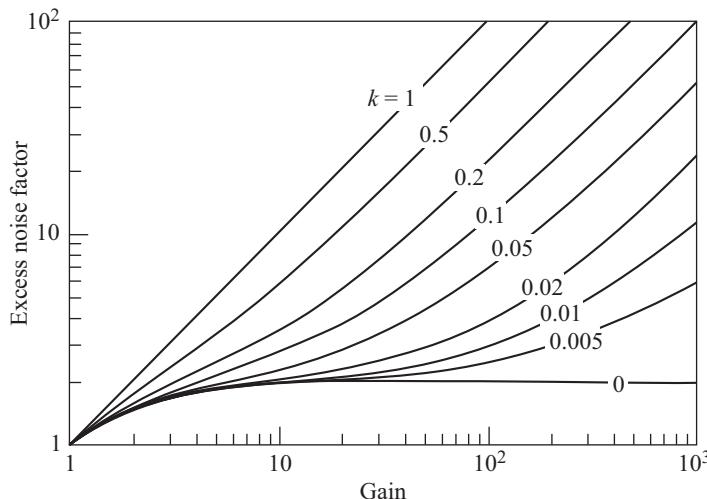
Fig. 6.15 *Variation of the electron excess noise factor F_e as a function of the electron gain for various values of the effective ionization rate ratio k_{eff}. (Webb, McIntyre, and Conradi.[14])*

From the empirical relationship for the mean-square gain given by Eq. (6.32), the excess noise factor can be approximated by

$$F = M^x \tag{6.41}$$

The parameter x takes on values of 0.3 for Si, 0.7 for InGaAs, and 1.0 for Ge avalanche photodiodes.

6.5 Structures for InGaAs APDs

To improve the performance of InGaAs APDs, various complex device architectures have been devised. One widely used structure is the *separate-absorption-and-multiplication* (SAM) APD configuration.[29–31] As Fig. 6.16 shows, this structure uses different materials in the absorption and multiplication regions, with each region being optimized for a particular function. Here, light enters the APD through the InP substrate. Since this material has a larger energy bandgap, it allows long-wavelength photons to pass through to the InGaAs absorption region where electron–hole pairs are generated. Following this is an InP layer that is used for the multiplication region because high electric fields needed for the gain mechanism can exist in InP without tunneling breakdown. This device structure gets its name SAM as a result of the separation of the absorption and multiplication regions.

Variations on the SAM structure include adding other layers to the device. These include

- Using a grading layer between the absorption and multiplication regions to increase the response time and bandwidth of the device.
- Adding a charge layer that provides better control of the electric field profile.
- Incorporating a resonant cavity that decouples the optical and electrical path lengths to achieve high quantum efficiencies and wide bandwidths simultaneously.

Another popular design for InGaAs APDs is the *superlattice structure*.[22] In these devices, the multiplication region is around 250 nm thick and consists of, for example, 13 layers of 9-nm-thick InAlGaAs quantum wells separated by 12-nm-thick InAlAs barrier layers. This structure improves the speed and sensitivity of InGaAs APDs, thereby allowing them to be used for applications such as 10-Gb/s long-distance systems (e.g., SONET OC-192/SDH STM-64 links).
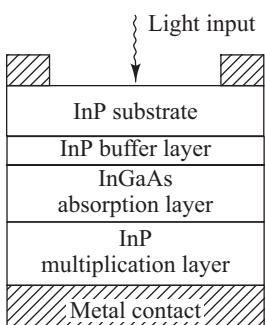
Fig. 6.16 *Simple diagram of a SAM APD structure (layers are not drawn to scale)*

6.6 Temperature Effect on Avalanche Gain

The gain mechanism of an avalanche photodiode is very temperature sensitive because of the temperature dependence of the electron and hole ionization rates. This temperature dependence is particularly critical at high bias voltages, where small changes in temperature can cause large variations in gain. An example of this is shown in Fig. 6.17 for a silicon avalanche photodiode. For example, if the operating temperature decreases and the applied bias voltage is kept constant, the ionization rates for electrons and holes will increase and so will the avalanche gain.
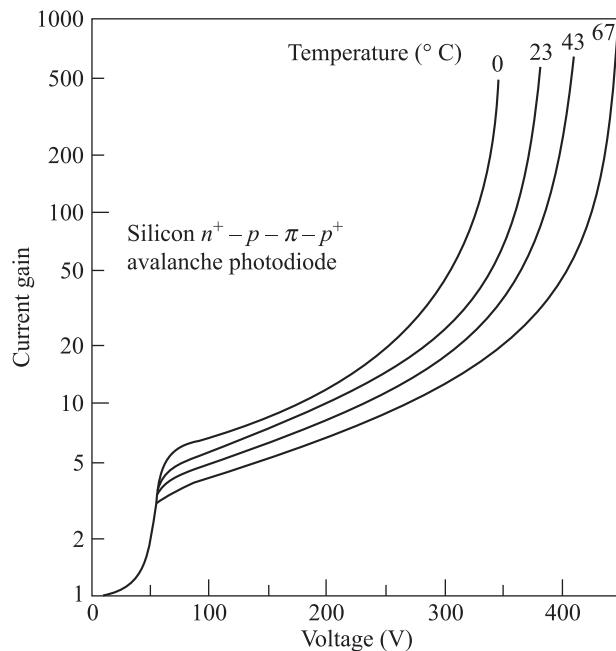
Fig. 6.17 *Example of how the gain mechanism of a silicon avalanche photodiode depends on temperature. The measurements for this device were carried out at 825 nm. (Reproduced with permission from Melchior, Hartman, Schinke, and Seidel,[15] © 1978, AT&T.)*

To maintain a constant gain as the temperature changes, the electric field in the multiplying region of the *pn* junction must also be changed. This requires that the receiver incorporate a compensation circuit that adjusts the applied bias voltage on the photodetector when the temperature changes.

A simple temperature-dependent expression for the gain can be obtained from the empirical relationship[32,33]

$$M = \frac{1}{1 - (V/V_B)^n} \tag{6.42}$$

where V_B is the breakdown voltage at which M goes to infinity; the parameter n varies between 2.5 and 7, depending on the material; and $V = V_a - I_M R_M$, with V_a being the reverse-bias voltage applied to the detector, I_M is the multiplied photocurrent, and R_M accounts for the photodiode series resistance and the detector load resistance. Since the breakdown voltage is known to vary with temperatures as[34]

$$V_B(T) = V_B(T_0)[1 + a(T - T_0)] \tag{6.43}$$

the temperature dependence of the avalanche gain can be approximated by substituting Eq. (6.43) into Eq. (6.42) together with the expression

$$n(T) = n(T_0)[1 + b(T - T_0)] \tag{6.44}$$

The constants a and b are positive for reach-through avalanche photodiodes and can be determined from experimental curves of gain versus temperature.

6.7 Comparisons of Photodetectors

This section summarizes some generic operating characteristics of Si, Ge, and InGaAs photodiodes. Tables 6.1 and 6.2 list the performance values for *pin* and avalanche photodiodes, respectively. The values were derived from various vendor data sheets and from performance numbers reported in the literature. They are given as guidelines for comparison purposes. Detailed values on specific devices for particular applications can be obtained from photodetector and receiver module suppliers.

For short-distance applications, Si devices operating around 850 nm provide relatively inexpensive solutions for most links. Longer links usually require operation in the 1300-nm and 1550-nm windows; here, one normally uses InGaAs-based devices.

Table 6.1 *Generic operating parameters of Si, Ge, and InGaAs pin photodiodes*

Parameter	Symbol	Unit	Si	Ge	InGaAs
Wavelength range	λ	nm	400–1100	800–1650	1100–1700
Responsivity	$\mathscr{R}$	A/W	0.4–0.6	0.4–0.5	0.75–0.95
Dark current	I_D	nA	1–10	50–500	0.5–2.0
Rise time	τ_r	ns	0.5–1	0.1–0.5	0.05–0.5
Modulation (bandwidth)	B_m	GHz	0.3–0.7	0.5–3	1–2
Bias voltage	V_B	V	5	5–10	5

TABLE 6.2 *Generic operating parameters of Si, Ge, and InGaAs avalanche photodiodes*

Parameter	Symbol	Unit	Si	Ge	InGaAs
Wavelength range	λ	nm	400–1100	800–1650	1100–1700
Avalanche gain	M	—	20–400	50–200	10–40
Dark current	I_D	nA	0.1–1	50–500	10–50
					@ $M = 10$
Rise time	τ_r	ns	0.1–2	0.5–0.8	0.1–0.5
Gain · bandwidth	$M \cdot B_m$	GHz	100–400	2–10	20–250
Bias voltage	V_B	V	150–400	20–40	20–30

PROBLEMS

6.1 Consider the absorption coefficient of silicon as a function of wavelength, as shown in Fig. 6.18. Ignoring reflections at the photodiode surface, plot the following two parameters over the wavelength range 600–1000 nm:

(*a*) The quantum efficiency for depletion layer widths of 1, 5, 10, 20, and 50 μm.

(*b*) The responsivity for a silicon *pin* photodiode having a 20-μm thick depletion layer.

6.2 If an optical power level P_{in} is incident on a photodiode, the electron–hole generation rate $G(x)$ is given by

$$G(x) = \Phi_0 \alpha_s \exp(-\alpha_s x)$$

Here Φ_0 is the incident photon flux per unit area given by

$$\Phi_0 = \frac{P_{in}(1 - R_f)}{Ah\nu}$$

10^4
Absorption coefficient (cm^{-1})
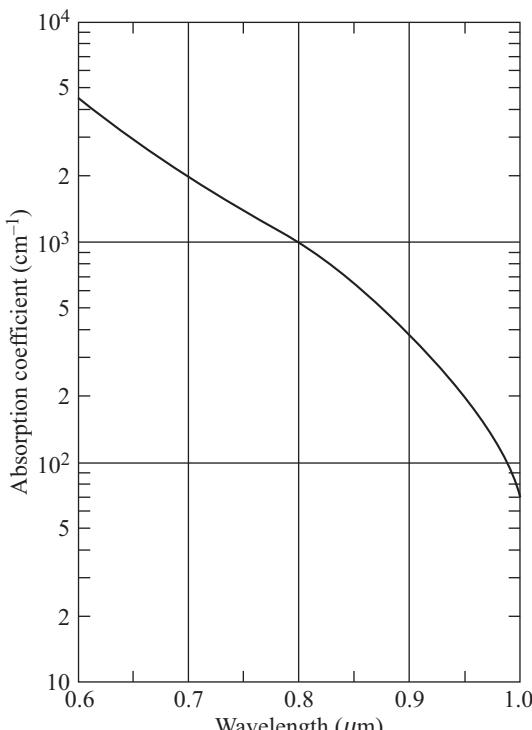
Fig. 6.18 *Absorption coefficient of Si as a function of wavelength*

where A is the detector area, R_f is the surface reflectance, and $h\nu$ is the photon energy. From this, show that the primary photocurrent in the depletion region of width w is given by Eq. (6.4).

6.3 If the absorption coefficient of silicon is $0.05\ \mu$m^{-1} at 860 nm, find the penetration depth at which $P(x)/P_{in} = 1/e = 0.368$.

6.4 The low-frequency gain M_0 of an avalanche photodiode depends on the carrier ionization rate and on the width of the multiplication region, both of which depend on the applied reverse-bias voltage V_a. This gain can be described by the empirical relationship[33]

$$M_0 = \frac{I_M}{I_p} = \frac{1}{1 - \left(\dfrac{V_a - I_M R_M}{V_B}\right)^n} \quad \text{(P6.4)}$$

where V_B is the breakdown voltage at which M_0 goes to infinity ($M_0 \rightarrow \infty$), I_M is the total multiplied current, and R_M accounts for the photodiode series resistance and the detector load resistance. The exponential factor n depends on the semiconductor material and its doping profile. Its value varies between about 2.5 and 7.

(*a*) Show that for applied voltages near the breakdown voltage, at which point $V_B \gg I_M R_M$, Eq. (P6.4) can be approximated by

$$M_0 = \frac{I_M}{I_p} \simeq \frac{V_B}{n(V_B - V_a + I_M R_M)}$$

$$\simeq \frac{V_B}{nI_M R_M}$$

(*b*) The maximum value of M_0 occurs when $V_a = V_B$. Show that, at this point,

$$M_{0,\text{max}} = \left(\frac{V_B}{nR_M I_p}\right)^{1/2}$$

6.5 Consider a sinusoidally modulated optical signal $P(t)$ of frequency ω, modulation index m, and average power P_0 given by

$$P(t) = P_0(1 + m \cos \omega t)^2$$

Show that when this optical signal falls on a photodetector, the mean-square signal current $\langle i_s^2 \rangle$ generated consists of a dc (average) component I_p and a signal current i_p given by

$$\langle i_s^2 \rangle = I_p^2 + \langle i_p^2 \rangle = (\mathscr{R} P_0)^2 + \frac{1}{2}(m\mathscr{R} P_0)^2$$

where the responsivity $\mathscr{R}$ is given by Eq. (6.6).

6.6 An InGaAs *pin* photodiode has the following parameters at 1550 nm: $I_D = 1.0$ nA, $\eta = 0.95$, $R_L = 500$ Ω, and the surface leakage current is negligible. The incident optical power is 500 nW (–33 dBm) and the receiver bandwidth is 150 MHz. Compare the noise currents given by Eqs. (6.14), (6.15), and (6.16).

6.7 Consider an avalanche photodiode receiver that has the following parameters: dark current $I_D = 1$ nA, leakage current $I_L = 1$ nA, quantum efficiency $\eta = 0.85$, gain $M = 100$, excess noise factor $F = M^{1/2}$, load resistor $R_L = 10^4$ Ω, and bandwidth $B_e = 10$ kHz. Suppose a sinusoidally varying 850-nm signal having a modulation index $m = 0.85$ falls on the photodiode, which is at room temperature ($T = 300$ K). To compare the contributions from the various noise terms to the signal-to-noise ratio for this particular set of parameters, plot the following terms in decibels [i.e., $10 \log(S/N)$] as a function of the average received optical power P_0. Let P_0 range from –70 to 0 dBm; that is, from 0.1 nW to 1.0 mW:

(a) $\left(\dfrac{S}{N}\right)_{\text{shot}} = \dfrac{\langle i_s^2 \rangle}{\langle i_{\text{shot}}^2 \rangle}$

(b) $\left(\dfrac{S}{N}\right)_{DB} = \dfrac{\langle i_s^2 \rangle}{\langle i_{DB}^2 \rangle}$

(c) $\left(\dfrac{S}{N}\right)_{DS} = \dfrac{\langle i_s^2 \rangle}{\langle i_{DS}^2 \rangle}$

(d) $\left(\dfrac{S}{N}\right)_{T} = \dfrac{\langle i_s^2 \rangle}{\langle i_{T}^2 \rangle}$

What happens to these curves if either the load resistor, the gain, the dark current, or the bandwidth is changed?

6.8 Suppose an avalanche photodiode has the following parameters: $I_L = 1$ nA, $I_D = 1$ nA, $\eta = 0.85$, $F = M^{1/2}$, $R_L = 10^3$ Ω, and $B_e = 1$ kHz. Consider a sinusoidally varying 850-nm signal, which has a modulation index $m = 0.85$ and an average power level $P_0 = -50$ dBm, to fall on the detector at room temperature. Plot the signal-to-noise ratio as a function of M for gains ranging from 20 to 100. At what value of M does the maximum signal-to-noise ratio occur?

6.9 Derive Eq. (6.19).

6.10 (a) Show that under the boundary conditions

$$p_n = p_{n0} \quad \text{for} \quad x = \infty$$

and

$$p_n = 0 \quad \text{for} \quad x = w$$

the solution to Eq. (6-23) is given by

$$p_n = p_{n0} - (p_{n0} + Be^{-\alpha_s w})e^{(w-x)/L_p} + Be^{-\alpha_s x}$$

where $L_p = (D_p \tau_p)^{1/2}$ is the diffusion length and

$$B = \left(\frac{\Phi_0}{D_p}\right)\frac{\alpha_s L_p^2}{1 - \alpha_s^2 L_p^2}$$

(b) Derive Eq. (6.25) using the relationship

$$J_{\text{diff}} = qD_p\left(\frac{\partial p_n}{\partial x}\right)_{x=w}$$

(c) Verify that J_{tot} is given by Eq. (6.26).

6.11 Consider a modulated photon flux density

$$\Phi = \Phi_0 e^{j\omega t} \text{ photons/(s . cm}^2)$$

to fall on a photodetector, where ω is the modulation frequency. The total current through the depletion region generated by this photon flux can be shown to be[35]

$$J_{\text{tot}} = \left(\frac{j\omega \epsilon_s V}{w} + q\Phi_0 \frac{1 - e^{j\omega t_d}}{j\omega t_d} \right) e^{j\omega t}$$

where ϵ_s is the material permittivity, V is the voltage across the depletion layer, and t_d is the transit time of carriers through the depletion region.

(*a*) From the short-circuit current density ($V = 0$), find the value of ωt_d at which the photocurrent amplitude is reduced by $\sqrt{2}$.

(*b*) If the depletion region thickness is assumed to be $1/\alpha_s$, what is the 3-dB modulation frequency in terms of α_s and v_d (the drift velocity)?

6.12 Suppose we have a silicon *pin* photodiode which has a depletion layer width $w = 20\ \mu$m, an area $A = 0.05$ mm^2, and a dielectric constant $K_s = 11.7$. If the photodiode is to operate with a 10-kΩ load resistor at 800 nm, where the absorption coefficient $\alpha_s = 10^3$ cm^{-1}, compare the *RC* time constant and the carrier drift time of this device. Is carrier diffusion time of importance in this photodiode?

6.13. Verify that, when the weighted ionization rate ratios k_1 and k_2 are assumed to be approximately equal, Eqs (6.34) and (6.35) can be simplified to yield Eqs (6.38) and (6.39).

6.14 Derive the limits of F_e given by Eq. (6.38) when (*a*) only electrons cause ionization; (*b*) the ionization rates α and β are equal.

REFERENCES

1. H. Melchior, "Detectors for lightwave communications," *Phys. Today,* vol. 30, pp. 32–39, Nov. 1977.

2. A. Beling and J.C. Campbell, "InP-based high-speed photodetectors: Tutorial," *J. Lightwave Tech.*, vol. 27, no. 3, pp. 343–355, Feb. 2009.

3. S. R. Forrest, "Optical detectors for lightwave communications," in S. E. Miller and I. P. Kaminow, eds., *Optical Fiber Telecommunications–II,* Academic, New York, 1988.

4. M Fukuda, *Optical Semiconductor Devices,* Wiley, Hoboken, NJ, 1998.

5. S. Donati, *Photodetectors: Devices, Circuits and Applications,* Prentice Hall, New York, 2000.

6. M. Johnson, *Photodetection and Measurement,* McGraw-Hill, New York, 2003.

7. H. Schneider and H. C. Liu, *Quantum Well Infrared Photodetectors,* Springer, New York, 2006.

8. B.E.A. Saleh and M. Teich, *Fundamentals of Photonics,* Wiley, Hoboken, NJ, 2nd ed., 2007.

9. B. L. Anderson and R. L. Anderson, *Fundamentals of Semiconductor Devices,* McGraw-Hill, New York, 2005.

10. D. A. Neamen, *An Introduction to Semiconductor Devices,* McGraw-Hill, New York, 2006.

11. S. O. Kasap, *Principles of Electronic Materials and Devices,* McGraw-Hill, New York, 2006.

12. O. Manasreh, *Semiconductor Heterojunctions and Nanostructures,* McGraw-Hill, New York. 2005.

13. S. E. Miller, E.A.J. Marcatili, and T. Li, "Research toward optical-fiber transmission systems," *Proc. IEEE*, vol. 61, pp. 1703–1751, Dec. 1973.

14. P. P. Webb, R. J. McIntyre, and J. Conradi, "Properties of avalanche photodiodes," *RCA Review,* vol. 35, pp. 234–278, June 1974.

15. H. Melchior, A. R. Hartman, D. P. Schinke, and T. E. Seidel, "Planar epitaxial silicon avalanche photodiode," *Bell Sys. Tech. J.,* vol. 57, pp. 1791–1807, July/Aug., 1978.

16. J. W. Hong, Y. W. Chen, W. L. Laih, Y. K. Fang, C. Y. Chang, and C. Gong, "The hydrogenated amorphous silicon reach-through avalanche photodiode," *IEEE J. Quantum Electron.,* vol. 26, pp. 280–284, Feb. 1990.

17. H. Sudo, Y. Nakano, and G. Iwanea, "Reliability of germanium avalanche photodiodes for optical transmission systems," *IEEE Trans. Electron Devices,* vol. 33, pp. 98–103, Jan. 1986.

18. F. Osaka, T. Mikawa, and T. Kaneda, "Impact ionization coefficients for electrons and holes in (100)-oriented GalnAsP," *IEEE J. Quantum Electron.,* vol. 21, pp. 1326–1338, Sept. 1985.

19. M. C. Brain and T. P. Lee, "Optical receivers for lightwave communication systems." *J. Lightwave Tech.*, vol. LT-3., pp. 1281–1300, Dec. 1985.

20. N. Susa, Y. Yamauchi, and H. Kanbe, "Vapor phase epitaxially grown InGaAs photodiodes," *IEEE Trans. Electron Devices,* vol. ED-27, pp. 92–98, Jan. 1980.

21. W. Wu, A. R. Hawkins, and J. E. Bowers, "Design of silicon hetero-interface photodetectors." *J. Lightwave Tech.*, vol. 15, pp. 1608–1615, Aug. 1997.

22. I. Watanabe, M. Tsuji, M. Hayashi, K. Makita, and K. Taguchi, "Design and performance of InAlGaAs/InAlAs superlattice avalanche photodiodes," *J. Lightwave Tech.*, vol. 15, pp. 1012–1019, June 1997.

23. J.T.K. Tang and K. B. Letaief, "The use of WMC distribution for performance evaluation of APD optical communication systems," *IEEE Trans. Commun.,* vol. 46, pp. 279–285, Feb. 1998.

24. B. M. Oliver, "Thermal and quantum noise," *IEEE Proc.,* vol. 53, pp. 436–454, May 1965.

25. W. M. Hubbard, "Utilization of optical-frequency carriers for low and moderate bandwidth channels." *Bell Sys. Tech. J.*, vol. 52, pp. 731–765, May/June, 1973.

26. R. S. Fyath and J. J. O'Reilly, "Performance degradation of APD-optical receivers due to dark current generated within the multiplication region," *J. Lightwave Tech.,* vol. 7, pp. 62–67, Jan. 1989.

27. R. J. McIntyre, "The distribution of gains in uniformly multiplying avalanche photodiodes: Theory," *IEEE Trans. Electron Devices,* vol. ED-19, pp. 703–713, June 1972.

28. J. Conradi, "The distribution of gains in uniformly multiplying avalanche photodiodes: Experimental," *IEEE Trans. Electron Devices,* vol. ED-19, pp. 713–718, June 1972.

29. D.S.G. Ong, J. S. Ng, M. M. Hayat, P. Sun, and J.P.R. David, "Optimization of InP APDs for high-speed lightwave systems," *J. Lightwave Tech.,* vol. 27, no. 15, pp. 3294–3302, Aug. 2009.

30. H. Nie, K. A. Anselm, C. Lenox, P. Yuan, C. Hu, G. Kinsey, B. G. Streetman, and J. C. Campbell, "Resonant-cavity separate absorption, charge and multiplication APDs with high-speed and high gain-bandwidth product," *IEEE Photonic Tech. Lett.*, vol. 10, pp. 409–411, Mar. 1998.

31. O. H. Kwon, M. M. Hayat, J. C. Campbell, B.E.A. Saleh, and M. C. Teich, "Gain-bandwidth product optimization of heterostructure avalanche photodiodes," *J. Lightwave Tech.,* vol. 23, pp. 1896–1906, May 2005.

32. J. Conradi, "Temperature effects in silicon avalanche photodiodes," *Solid State Electron.*, vol. 17, pp. 96–106, Jan. 1974.

33. S. L. Miller, "Avalanche breakdown in germanium," *Phys. Rev.*, vol. 99, pp. 1234–1241, Aug. 1955.

34. N. Susa, H. Nakagome, H. Ando, and H. Kanbe, "Characteristics in InGaAs/InP avalanche photodiodes with separated absorption and multiplication regions," *IEEE J. Quantum Electron.*, vol. QE-17, pp. 243–250, Feb. 1981.

35. W. W. Gaertner, "Depletion-layer photoeffects in semiconductors," *Phvs. Rev.,* vol. 116, pp. 84–87, Oct. 1959.

CHAPTER 7

Optical Receiver Operation

Having discussed the characteristics and operation of photodetectors in the previous chapter, we now turn our attention to the optical receiver. An optical receiver consists of a photodetector, an amplifier, and signal-processing circuitry. The receiver has the task of first converting the optical energy emerging from the end of a fiber into an electric signal, and then amplifying this signal to a large enough level so that it can be processed by the electronics following the receiver amplifier.

In these processes, various noises and distortions will unavoidably be introduced, which can lead to errors in the interpretation of the received signal. Depending on the magnitude of the received optical signal, the current generated by the photodetector could be very weak and is adversely affected by the random noises associated with the photodetection process. When this electric signal output from the photodiode is amplified, additional noises arising from the amplifier electronics will further corrupt the signal. Noise considerations are thus important in the design of optical receivers, since the noise sources operating in the receiver generally set the lowest limit for the signals that can be processed.

In designing a receiver, it is desirable to predict its performance based on mathematical models of the various receiver stages. These models must take into account the noises and distortions added to the signal by the components in each stage, and they must show the designer which components to choose so that the desired performance criteria of the receiver are met.

The average error probability is the most meaningful criterion for measuring the performance of a digital communication system. In an analog system the fidelity criterion normally is specified in terms of a peak signal-to-rms-noise ratio. The calculation of the error probability for a digital optical communication receiver differs from that of conventional electric systems. This is because of the discrete quantum nature of the photon arrival and detection processes and the random gain fluctuations when an avalanche photodiode is used. Researchers have used a variety of analytical methods to derive approximate predictions for receiver performance. In carrying out these analyses, one needs to make a tradeoff between calculation simplicity and accuracy of the approximation. General reviews and detailed concepts of optical receiver designs are given in the literature.[1–17]

In Sec. 7.1 we first present an overview of the fundamental operational characteristics of the various stages of an optical receiver. This consists of tracing the path of a digital signal through the receiver and showing what happens at each step along the way. Section 7.2 then outlines the fundamental methods for determining the *bit-error rate* or probability of error (the chance that a bit is corrupted and received in error) of a digital receiver based on signal-to-noise considerations. This section also discusses the

concept of *receiver sensitivity*, which is an important parameter for estimating the minimum received optical power that is needed to achieve a specific probability of error.

The *eye diagram* is a common measurement tool to determine the fidelity of the received signal. Eye diagrams have been used extensively for all types of communication links, including wire lines, wireless systems, and optical links. Section 7.3 describes the method for generating an eye diagram and how to interpret various signal fidelity parameters with it.

For *passive optical network* (PON) applications, the operational characteristics of an optical receiver located at the central telecommunications switching office differ significantly from receivers used in conventional point-to-point links. Section 7.4 gives a brief description of these devices, which are known as *burst-mode receivers*. Finally, Sec. 7.5 describes another optical receiver type that is used for analog links.

7.1 | Fundamental Receiver Operation

The design of an optical receiver is much more complicated than that of an optical transmitter because the receiver must be able to detect weak, distorted signals and make decisions on what type of data was sent based on an amplified and reshaped version of this distorted signal. To get an appreciation of the function of an optical receiver, we first examine what happens to a signal as it is sent through an optical fiber link. Since traditionally fiber optic communication links are intensity-modulated direct-detection (IM-DD) systems that use a binary *on-off keyed* (OOK) digital signal, as described in Chapter 4, first we analyze direct-detection receiver performance by using this OOK signal format.

7.1.1 Digital Signal Transmission

Figure 7.1 illustrates the shape of a digital signal at different points along an optical link. The transmitted signal is a two-level binary data stream consisting of either a 0 or a 1 in a time slot of duration T_b. This time slot is referred to as a *bit period*. Electrically there are many ways of sending a given digital message.[18–22] One of the simplest techniques for sending binary data is *amplitude-shift keying* (ASK) or *on-off keying* (OOK), wherein a voltage level is switched between two values, which are usually *on* or *off*. The resultant signal wave thus consists of a voltage pulse of amplitude V relative to the zero voltage level when a binary 1 occurs and a zero-voltage-level space when a binary 0 occurs. Depending on the coding scheme to be used, a binary 1 may or may not fill the time slot T_b. For simplicity, here we assume that when 1 is sent, a voltage pulse of duration T_b occurs, whereas for 0 the voltage remains at its zero level.

The function of the optical transmitter is to convert the electric signal to an optical signal. As shown in Sec. 4.3.7, one way of doing this is by directly modulating the light source drive current with the information stream to produce a varying optical output power $P(t)$. Thus in the optical signal emerging from the LED or laser transmitter, 1 is represented by a pulse of optical power (light) of duration T_b, whereas 0 is the absence of any light.

The optical signal that is coupled from the light source to the fiber becomes attenuated and distorted as it propagates along the fiber waveguide. Upon arriving at the end of a fiber, a receiver converts the optical signal back to an electrical format. Figure 7.2 shows the basic components of an optical receiver. The first element is either a *pin* or an avalanche photodiode, which produces an electric current that is proportional to the received power level. Since this electric current typically is very weak, a *front-end amplifier* boosts it to a level that can be used by the following electronics.

After the electric signal produced by the photodiode is amplified, it passes through a *low-pass filter* to reduce the noise that is outside of the signal bandwidth. This filter thus defines the receiver bandwidth.
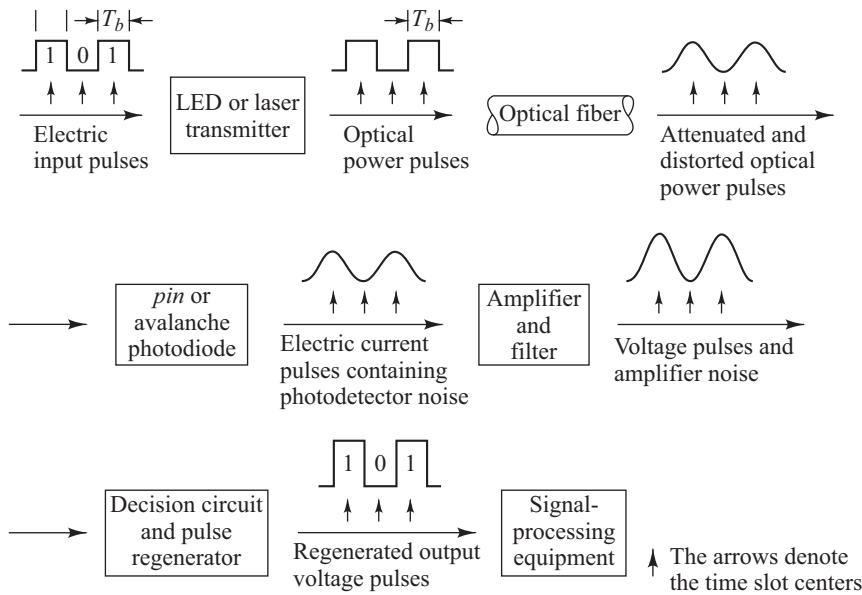
Fig. 7.1 *Signal path through an optical data link. (Adapted with permission from Personick et al.,[2] © 1977, IEEE.)*

In addition, to minimize the effects of intersymbol interference (ISI) the filter can reshape the pulses that have become distorted as they traveled through the fiber. This function is called *equalization* because it equalizes or cancels pulse-spreading effects.

In the final optical receiver module shown on the right in Fig. 7.2, a sampling and decision circuit samples the signal level at the midpoint of each time slot and compares it with a certain reference voltage known as the *threshold level*. If the received signal level is greater than the threshold level, 1 is said to have been received. If the voltage is below the threshold level, 0 is assumed to have been received. To accomplish this bit interpretation, the receiver must know where the bit boundaries are. This is done with the assistance of a periodic waveform called a *clock*, which has a periodicity equal to the bit interval. Thus this function is called *clock recovery* or *timing recovery*.

In some cases, an optical preamplifier is placed ahead of the photodiode to boost the optical signal level before photodetection takes place. This is done so that the signal-to-noise ratio degradation caused by thermal noise in the receiver electronics can be suppressed. Compared with other front-end devices, such as avalanche photodiodes or optical heterodyne detectors, an optical preamplifier provides a larger
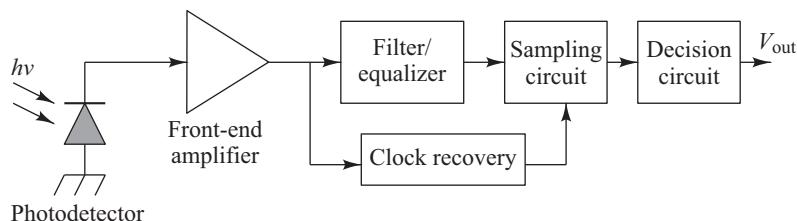
Fig. 7.2 *The basic sections of an optical receiver*

gain factor and a broader bandwidth. However, this process also introduces additional noise to the optical signal. Chapter 11 addresses optical amplifiers and their effects on system performance.

7.1.2 Error Sources

Errors in the detection mechanism can arise from various noises and disturbances associated with the signal detection system, as shown in Fig. 7.3. The term *noise* is used customarily to describe unwanted components of an electric signal that tend to disturb the transmission and processing of the signal in a physical system, and over which we have incomplete control. The noise sources can be either external to the system (e.g., electric power lines, motors, radio transmitters, lightning) or internal to the system (e.g., switch and power supply transients). Here, we shall be concerned mainly with internal noise, which is present in every communication system and represents a basic limitation on the transmission or detection of signals. This noise is caused by the spontaneous fluctuations of current or voltage in electric circuits. The two most common examples of these spontaneous fluctuations are shot noise and thermal noise.[19,21] Shot noise arises in electronic devices because of the discrete nature of current flow in the device. Thermal noise arises from the random motion of electrons in a conductor.

As discussed in Sec. 6.2.1, the random arrival rate of signal photons produces a quantum (or shot) noise at the photodetector. Since this noise depends on the signal level, it is of particular importance for *pin* receivers that have large optical input levels and for avalanche photodiode receivers. When using an avalanche photodiode, an additional shot noise arises from the statistical nature of the multiplication process. This noise level increases with larger avalanche gain M. Additional photodetector noises come from the dark current and leakage current. These are independent of the photodiode illumination and can generally be made very small in relation to other noise currents by a judicious choice of components.

Thermal noises arising from the detector load resistor and from the amplifier electronics tend to dominate in applications with low signal-to-noise ratio when a *pin* photodiode is used. When an avalanche photodiode is used in low-optical-signal-level applications, the optimum avalanche gain is determined by a design tradeoff between the thermal noise and the gain-dependent quantum noise.

Since the thermal noises are of a gaussian nature, they can be readily treated by standard techniques. The analysis of the noises and the resulting error probabilities associated with the primary photocurrent generation and the avalanche multiplication are complicated as neither of these processes is gaussian. The primary photocurrent generated by the photodiode is a time-varying Poisson process resulting from the random arrival of photons at the detector. If the detector is illuminated by an optical signal $P(t)$, then the average number of electron–hole pairs $\bar{N}$ generated in a time τ is

$$\bar{N} = \frac{\eta}{h\nu} \int_0^\tau P(t)\, dt = \frac{\eta E}{h\nu} \tag{7.1}$$
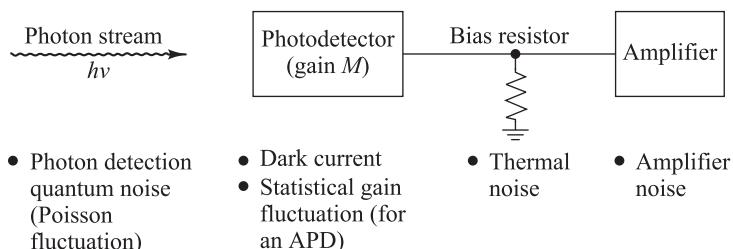
Fig. 7.3 *Noise sources and disturbances in an optical pulse detection mechanism*

Example 7.1 Consider the following two avalanche photodiodes: (*a*) a Si APD with an ionization rate ratio $k = 0.02$ and (*b*) an InGaAs APD with an ionization rate ratio $k = 0.50$. (*a*) What is the excess noise factor $F(M)$ for electron injection for each device if they both operate with a multiplication factor $M = 30$? (*b*) What is the estimate of $F(M)$ for the InGaAs APD if $M = 10$? (*c*) Compare the results in (*a*) and (*b*) with the empirical expression given in Eq. (7.4) if $x = 0.3$ for Si and 0.7 for InGaAs.

Solution: (*a*) From Eq. (7.3) we find that for the Si APD

$$F(M) = kM + \left(2 - \frac{1}{M}\right)(1 - k)$$
$$= 0.02(30) + (2 - 0.033)(1 - 0.02)$$
$$= 2.53 = 4.03 \text{ dB}$$

Similarly, for the InGaAs APD with $M = 30$

$$F(M) = kM + \left(2 - \frac{1}{M}\right)(1 - k)$$
$$= 0.50(30) + (2 - 0.033)(1 - 0.50)$$
$$= 16.0 = 12.0 \text{ dB}$$

(*b*) For the InGaAs APD with $M = 10$

$$F(M) = 0.50(10) + (2 - 0.10)(1 - 0.50) = 6.00 = 7.78 \text{ dB}$$

(*c*) For the Si APD with $M = 30$, Eq. (7.4) yields
$F(M) \approx 30^{0.3} = 2.77 = 4.42 \text{ dB}$
For the InGaAs APD with $M = 30$, Eq. (7.4) yields
$F(M) \approx 30^{0.7} = 10.8 = 10.3 \text{ dB}$
For the InGaAs APD with $M = 10$, Eq. (7.4) yields
$F(M) \approx 10^{0.7} = 5.01 = 7.00 \text{ dB}$

---

where η is the detector quantum efficiency, $h\nu$ is the photon energy, and E is the energy received in a time interval τ. The actual number of electron–hole pairs n that are generated fluctuates from the average according to the Poisson distribution

$$P_r(n) = \bar{N}^n \frac{e^{-\bar{N}}}{n!} \tag{7.2}$$

where $P_r(n)$ is the probability that n electrons are excited in an interval τ. The fact that it is not possible to predict exactly how many electron–hole pairs are generated by a known optical power incident on the detector is the origin of the type of shot noise called *quantum noise*. The random nature of the avalanche multiplication process gives rise to another type of shot noise. Recall from Sec. 6.4 that, for a detector with a mean avalanche gain M and an ionization rate ratio k, the excess noise factor $F(M)$ for electron injection is

$$F(M) = kM + \left(2 - \frac{1}{M}\right)(1 - k) \tag{7.3}$$

This equation is often approximated by the empirical expression

$$F(M) \simeq M^x \tag{7.4}$$

where the factor x ranges between 0 and 1.0 depending on the photodiode material.

A further error source is attributed to *intersymbol interference* (ISI), which results from pulse spreading in the optical fiber. When a pulse is transmitted in a given time slot, most of the pulse energy will arrive in the corresponding time slot at the receiver, as shown in Fig. 7.4. However, because of pulse spreading induced by the fiber, some of the transmitted energy
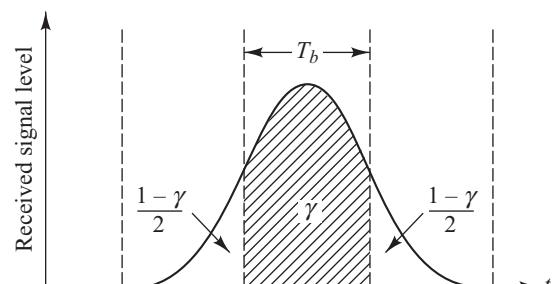
Fig. 7.4 *Pulse spreading in an optical signal that leads to intersymbol interference*
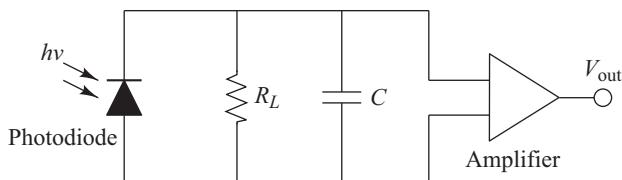
Fig. 7.5 *Generic structure of a high-impedance amplifier*

will progressively spread into neighboring time slots as the pulse propagates along the fiber. The presence of this energy in adjacent time slots results in an interfering signal, hence the term *intersymbol interference*. In Fig. 7.4 the parameter γ designates the fraction of energy remaining in the time slot T_b, so that $1 - \gamma$ is the fraction of energy that has spread into adjacent time slots.

7.1.3 Front-End Amplifier

Noise sources at the front end of a receiver dominate the sensitivity and bandwidth, so a major engineering emphasis has been on the design of a low-noise front-end amplifier. The goals generally are to maximize the receiver sensitivity while maintaining a suitable bandwidth. Front-end amplifiers used in optical fiber communication systems can be classified into two broad categories, which are the high-impedance and the transimpedance designs. These categories are not actually distinct because a continuum of intermediate designs is possible, but they serve to illustrate the design approaches.

A basic concern in front-end design is what load resistor R_L to choose. Recall from Sec. 6.2.1 that the thermal noise is inversely proportional to the load resistance. Thus R_L should be as large as possible to minimize thermal noise. For a *high-impedance amplifier* design shown in Fig. 7.5, a tradeoff must be made between noise and receiver bandwidth, since the bandwidth is inversely proportional to the resistance R_P seen by the photodiode. Since $R_P = R_L$ for a high-impedance front end, a high load resistance results in low noise but also gives a low receiver bandwidth. Although equalizers sometimes can be implemented to increase the bandwidth, if the bandwidth is much less than the bit rate, then such a front-end amplifier cannot be used.

The *transimpedance amplifier* design shown in Fig. 7.6 largely overcomes the drawbacks of the high-impedance amplifier. In this case, R_L is used as a negative feedback resistor around an inverting amplifier. Now R_L can be large since the negative feedback reduces the effective resistance seen by the photodiode by a factor of G, so that $R_P = R_L/(G + 1)$, where G is the gain of the amplifier. This means that compared to the high-impedance design the transimpedance bandwidth increases by a factor of $G + 1$ for the same load resistance. Although this does increase the thermal noise compared to a high-impedance amplifier, the increase usually is less than a factor of 2 and can easily be tolerated. Consequently, the transimpedance front-end design tends to be the amplifier of choice for optical fiber transmission links.

Note that in addition to the thermal noise differences resulting from selection of a particular load resistor, the electronic components in the front-end amplifier that follows the photodetector also add

Fig. 7.6 *Generic structure of a transimpedance amplifier*

further thermal noise. The magnitude of this additional noise depends on the design of the amplifier, for example, what type of bipolar or field-effect transistors are incorporated in the design.[8–15] This noise increase can be accounted for by introducing an *amplifier noise figure F_n* into the numerator of Eq. (6.17). This parameter is defined as the ratio of the input SNR to the output SNR of the amplifier. Typical values of the amplifier noise figure range from 3 to 5 dB (a factor of 2 to 3).

7.2 Digital Receiver Performance

Ideally, in a digital receiver the decision-circuit output signal voltage $v_{out}(t)$ would always exceed the threshold voltage when a 1 is present and would be less than the threshold when no pulse (a 0) was sent. In actual systems, deviations from the average value of $v_{out}(t)$ are caused by various noises, interference from adjacent pulses, and conditions wherein the light source is not completely extinguished during a zero pulse.

7.2.1 Probability of Error

In practice, there are several ways of measuring the rate of error occurences in a digital data stream.[23–27] Chapter 14 describes some of these methods. A simple approach is to divide the number N_e of errors occuring over a certain time interval t by the number N_t of pulses (ones and zeros) transmitted during this interval. This is called either the *error rate* or the *bit-error rate,* which is commonly abbreviated BER. Thus we have

$$\text{BER} = \frac{N_e}{N_t} = \frac{N_e}{Bt} \tag{7.5}$$

where $B = 1/T_b$ is the bit rate (i.e., the pulse transmission rate). The error rate is expressed by a number, such as 10^{-9}, for example, which states that, on the average, one error occurs for every billion pulses sent. Typical error rates for optical fiber telecommunication systems range from 10^{-9} to 10^{-12}. This error rate depends on the signal-to-noise ratio at the receiver (the ratio of signal power to noise power). The system error rate requirements and the receiver noise levels thus set a lower limit on the optical signal power level that is required at the photodetector.

To compute the bit-error rate at the receiver, we have to know the probability distribution[28–29] of the signal at the equalizer output. Knowing the signal probability distribution at this point is important because it is here that the decision is made as to whether a 0 or a 1 is sent. The shapes of two signal probability distributions are shown in Fig. 7.7. These are

$$P_1(v) = \int_{-\infty}^{v} p(y|1)\, dy \tag{7.6}$$

which is the probability that the equalizer output voltage is less than v when a logical 1 pulse is sent, and

$$P_0(v) = \int_{v}^{\infty} p(y|0)\, dy \tag{7.7}$$

which is the probability that the output voltage exceeds v when a logical 0 is transmitted. Note that the different shapes of the two probability distributions in Fig. 7.7 indicate that the noise power for a logical 0 is usually not the same as that for a logical 1. This occurs in optical systems because of signal distortion from transmission impairments (e.g., dispersion, optical amplifier noise, and distortion from nonlinear effects) and from noise and ISI contributions at the receiver. The functions $p(y|1)$ and $p(y|0)$ are the conditional probability distribution functions; that is, $p(y|x)$ is the probability that the output voltage is y, given that an x was transmitted.

If the threshold voltage is v_{th} then the error probability P_e is defined as

$$P_e = aP_1(v_{th}) + bP_0(v_{th}) \tag{7.8}$$
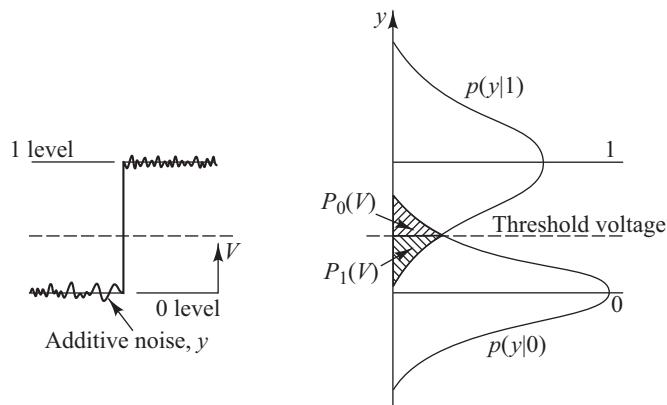
Fig. 7.7 *Probability distributions for received logic 0 and 1 signal pulses. The different widths of the two distributions are caused by various signal distortion effects.*

The weighting factors a and b are determined by the a priori distribution of the data. That is, a and b are the probabilities that either a 1 or a 0 occurs, respectively. For unbiased data with equal probability of 1 and 0 occurrences, $a = b = 0.5$. The problem to be solved now is to select the decision threshold at that point where P_e is minimum.

To calculate the error probability we require a knowledge of the mean-square noise voltage (v_N^2), which is superimposed on the signal voltage at the decision time. The statistics of the output voltage at the sampling time are very complicated, so that an exact calculation is rather tedious to perform. A number of different approximations[1–15] have therefore been used to calculate the performance of a binary optical fiber receiver. In applying these approximations, one has to make a tradeoff between computational simplicity and accuracy of the results. The simplest method is based on a gaussian approximation. In this method, it is assumed that, when the sequence of optical input pulses is known, the equalizer output voltage $v_{\text{out}}(t)$ is a gaussian random variable. Thus, to calculate the error probability, we need only to know the mean and standard deviation of $v_{\text{out}}(t)$. Other approximations that have been investigated are more involved[2–4,10,13] and will not be discussed here.

Thus, let us assume that a signal s (which can be either a noise disturbance or a desired information-bearing signal) has a gaussian probability distribution function with a mean value m. If we sample the signal voltage level $s(t)$ at any arbitrary time t_1, the probability that the measured sample $s(t_1)$ falls in the range s to $s + ds$ is given by

$$f(s)\,ds = \frac{1}{\sqrt{2\pi}\sigma} e^{-(s-m)^2/2\sigma^2}\,ds \qquad (7.9)$$

where $f(s)$ is the *probability density function,* σ^2 is the noise variance, and its square root σ is the *standard deviation,* which is a measure of the width of the probability distribution. By examining Eq. (7.9) we can see that the quantity $2\sqrt{2}\sigma$ measures the full width of the probability distribution at the point where the amplitude is $1/e$ of the maximum.

We can now use the probability density function to determine the probability of error for a data stream in which the 1 pulses are all of amplitude V. As shown in Fig. 7.8, the mean and variance of the gaussian output for a 1 pulse are b_{on} and σ_{on}^2, respectively, whereas for a 0 pulse they are b_{off} and σ_{off}^2, respectively. Let us first consider the case of a 0 pulse being sent, so that no pulse is present at the decoding time.
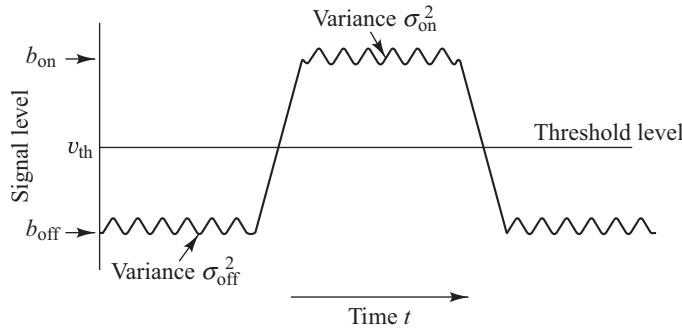
Fig. 7.8 *Gaussian noise statistics of a binary signal showing variances around the on and off signal levels*

The probability of error in this case is the probability that the noise will exceed the threshold voltage v_{th} and be mistaken for a 1 pulse. This probability of error $P_0(v)$ is the chance that the equalizer output voltage $v(t)$ will fall somewhere between v_{th} and ∞. Using Eqs. (7.7) and (7.9), we have

$$P_0(v_{th}) = \int_{v_{th}}^{\infty} p(y|0)\, dy = \int_{v_{th}}^{\infty} f_0(y)\, dy$$

$$= \frac{1}{\sqrt{2\pi}\, \sigma_{off}} \int_{v_{th}}^{\infty} \exp\left[-\frac{(v-b_{off})^2}{2\sigma_{off}^2} \right] dv \tag{7.10}$$

where the subscript 0 denotes the presence of a 0 bit.

Similarly, we can find the probability of error that a transmitted 1 is misinterpreted as a 0 by the decoder electronics following the equalizer. This probability of error is the likelihood that the sampled signal-plus-noise pulse falls below v_{th}. From Eqs. (7.6) and (7.9), this is given by

$$P_1(v_{th}) = \int_{-\infty}^{v_{th}} p(y|1)\, dy = \int_{-\infty}^{v_{th}} f_1(v)\, dv$$

$$= \frac{1}{\sqrt{2\pi}\, \sigma_{on}} \int_{-\infty}^{v_{th}} \exp\left[-\frac{(b_{on} - v)^2}{2\sigma_{on}^2} \right] dv \tag{7.11}$$

where the subscript 1 denotes the presence of a 1 bit.

If the probabilities of 0 and 1 pulses are equally likely [that is, $a = b = 0.5$ in Eq. (7.8)], then Eqs. (7.6) and (7.7) yield

$$P_0(v_{th}) = P_1(v_{th}) = \frac{1}{2} P_e \tag{7.12}$$

Thus, using Eqs (7.10) and (7.11), the bit error rate or the error probability P_e becomes

$$\text{BER} = P_e(Q) = \frac{1}{\sqrt{\pi}} \int_{Q/\sqrt{2}}^{\infty} e^{-x^2}\, dx$$

$$= \frac{1}{2}\left[1 - \text{erf}\left(\frac{Q}{\sqrt{2}} \right) \right] \approx \frac{1}{\sqrt{2\pi}} \frac{e^{-Q^2/2}}{Q} \tag{7.13}$$

Example 7.2 When there is little intersymbol interference, $\gamma - 1$ is small, so that $\sigma_{on}^2 \simeq \sigma_{off}^2$. Then, by letting $b_{off} = 0$, we have from Eq. (7.14) that

$$Q = \frac{b_{on}}{2\sigma_{on}} = \frac{1}{2}\frac{S}{N}$$

which is one-half the signal-to-noise ratio. In this case, $\upsilon_{th} = b_{on}/2$, so that the optimum decision threshold is midway between the 0 and 1 signal levels.

Example 7.3 For an error rate of 10^{-9} we have from Eq. (7.13) that

$$P_e(Q) = 10^{-9} = \frac{1}{2}\left[1 - \text{erf}\left(\frac{Q}{\sqrt{2}}\right)\right]$$

From Fig. 7.9 we have that $Q \simeq 6$ (an exact evaluation yields $Q = 5.99781$), which gives a signal-to-noise ratio of 12, or 10.8 dB [i.e., $10 \log(S/N) = 10 \log 12 = 10.8$ dB].

The approximation is obtained from the asymptotic expansion of erf (x). Here, the parameter Q is defined as

$$Q = \frac{\upsilon_{th} - b_{off}}{\sigma_{off}} = \frac{b_{on} - \upsilon_{th}}{\sigma_{on}} = \frac{b_{on} - b_{off}}{\sigma_{on} + \sigma_{off}} \tag{7.14}$$

and

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-y^2}\,dy \tag{7.15}$$

is the *error function*, which is tabulated in various mathematical handbooks.[30–31]

The factor Q is widely used to specify receiver performance, since it is related to the signal-to-noise ratio required to achieve a specific bit-error rate.[25–27] In particular, it takes into account that in optical fiber systems the variances in the noise powers generally are different for received logical 0 and 1 pulses. Figure 7.9 shows how the BER varies with Q. The approximation for P_e given in Eq. (7.13) and shown by the dashed line in Fig. 7.9 is accurate to 1 percent for $Q \approx 3$ and improves as Q increases. A commonly quoted Q value is 6, since this corresponds to a BER = 10^{-9}.

Let us consider the special case when $\sigma_{off} = \sigma_{on} = \sigma$ and $b_{off} = 0$, so that $b_{on} = V$. Then, from Eq. (7.14) we have that the threshold voltage $\upsilon_{th} = V/2$, so that $Q = V/2\sigma$. Since σ is usually called the *rms noise*, the ratio V/σ is the *peak signal-to-rms-noise ratio*. In this case, Eq. (7.13) becomes

$$P_e(\sigma_{on} = \sigma_{off}) = \frac{1}{2}\left[1 - \text{erf}\left(\frac{V}{2\sqrt{2}\,\sigma}\right)\right] \tag{7.16}$$

Example 7.4 demonstrates the exponential behavior of the probability of error as a function of the signal-to-noise ratio. By increasing V/σ by $\sqrt{2}$, that is, doubling S/N (a 3-dB power increase), the BER decreases by 10^4. Thus, there exists a narrow range of signal-to-noise ratios above which the error rate is tolerable and below which a highly unacceptable number of errors occur. The signal-to-noise ratio at which this transition occurs is called the *threshold level*. In general, a performance safety margin of 3–6 dB is included in the transmission link design to ensure that this BER threshold is not exceeded when system parameters such as transmitter output, line attenuation, or noise floor vary with time.
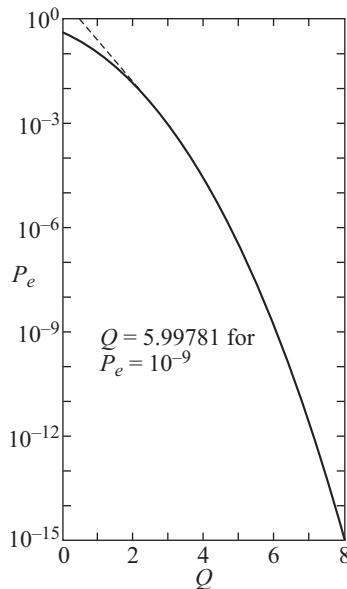
Fig. 7.9 *Plot of the BER (P_e) versus the factor Q. The approximation from Eq. (7.13) is shown by the dashed line.*

Example 7.4 Figure 7.10 shows a plot of the BER expression from Eq. (7.16) as a function of the signal-to-noise ratio. Let us look at two cases of transmission rates.

(*a*) For a signal-to-noise ratio of 8.5 (18.6 dB) we have $P_e = 10^{-5}$. If this is the received signal level for a standard DS1 telephone rate of 1.544 Mb/s, this BER results in a misinterpreted bit every 0.065 s, which is highly unsatisfactory. However, by increasing the signal strength so that $V/\sigma = 12.0$ (21.6 dB), the BER decreases to $P_e = 10^{-9}$. For the DS1 case, this means that a bit is misinterpreted every 650 s (or 11 min), which, in general, is tolerable.

(*b*) For high-speed SONET links, say the OC-12 rate that operates at 622 Mb/s, BERs of 10^{-11} or 10^{-12} are required. This means that we need to have at least $V/\sigma = 13.0$ (22.3 dB).

7.2.2 Receiver Sensitivity

Optical communication systems use a BER value to specify the performance requirements for a particular transmission link application. For example, SONET/SDH networks specify that the BER must be 10^{-10} or lower, whereas Gigabit Ethernet and Fibre Channel require a BER of no more than 10^{-12}. To achieve a desired BER at a given data rate, a specific minimum average optical power level must arrive at the photodetector. The value of this minimum power level is called the *receiver sensitivity*.

A common method of defining the receiver sensitivity is as an *average optical power* (P_{ave}) in dBm incident on the photodetector. Alternatively it may be defined as an *optical modulation amplitude* (OMA) given in terms of a peak-to-peak current at the photodetector output. The receiver sensitivity gives a measure of the minimum average power or OMA needed to maintain a maximum (worst-case) BER at a specific data rate. Many researchers have carried out numerous complex calculations of receiver sensitivity by taking into account different pulse-shape degradation factors.[1–14] This section presents a simplified analysis to illustrate the basis of receiver sensitivity.
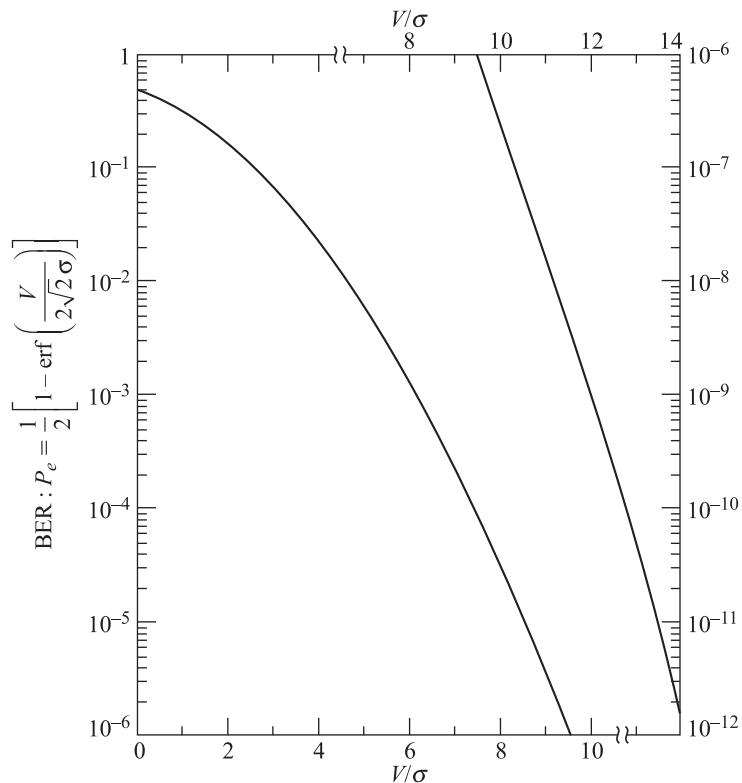
Fig. 7.10   *Bit-error rate as a function of signal-to-noise ratio when the standard deviations are equal ($\sigma_{on} = \sigma_{off}$) and when $b_{off} = 0$*

First, expressing Eq. (7.14) in terms of signal currents from 1 and 0 pulses (I_1 and I_0, respectively) and their corresponding noise current variations (σ_1 and σ_0, respectively), and assuming there is no optical power in a zero pulse, yields

$$Q = \frac{I_1 - I_0}{\sigma_1 + \sigma_0} \approx \frac{I_1}{\sigma_1 + \sigma_0} \tag{7.17}$$

Using Eqs. (6.6), (6.7), and (7.14), the receiver sensitivity $P_{\text{sensitivity}}$ is found from the average power contained in a bit period for the specified data rate as

$$P_{\text{sensitivity}} = P_1/2 = I_1/(2\,\mathscr{R}M) = Q\,(\sigma_1 + \sigma_0)/(2\,\mathscr{R}M) \tag{7.18}$$

where $\mathscr{R}$ is the unity-gain responsivity and M is the gain of the photodiode.

   If there is no optical amplifier in a fiber transmission link, then thermal noise and shot noise are the dominant noise effects in the receiver. As Sec. 6.2 describes, the thermal noise is independent of the incoming optical signal power, but the shot noise depends on the received power. Therefore, assuming there is no optical power in a received zero pulse, the noise variances for 0 and 1 pulses, respectively, are $\sigma_0^2 = \sigma_T^2$ and $\sigma_1^2 = \sigma_T^2 + \sigma_{\text{shot}}^2$. From Eqs. (6.6) and (6.13), and using the condition from Eq. (7.18),

the shot noise variance for a 1 pulse is

$$\sigma_{\text{shot}}^2 = 2q \, \mathcal{R} P_1 M^2 F(M) \, B_e = 4q \, \mathcal{R} P_{\text{sensitivity}} \, M^2 F(M) B/2 \tag{7.19}$$

where $F(M)$ is the photodiode noise figure and the electrical bandwidth B_e of the receiver is assumed to be half the bit rate B (i.e., $B_e = B/2$). Including the amplifier noise figure F_n in Eq. (6.17), the thermal noise current variance is

$$\sigma_T^2 = \frac{4k_B T}{R_L} F_n \frac{B}{2} \tag{7.20}$$

Substituting $\sigma_1 = \left(\sigma_{\text{shot}}^2 + \sigma_T^2 \right)^{1/2}$ and $\sigma_0 = \sigma_T$ into Eq. (7.18) and solving for $P_{\text{sensitivity}}$ then gives[4]

$$P_{\text{sensitivity}} = (1/\mathcal{R}) \frac{Q}{M} \left[\frac{qMF(M) \, BQ}{2} + \sigma_T \right] \tag{7.21}$$

Example 7.5 To see the behavior of the receiver sensitivity as a function of the BER, first consider the receiver to have a load resistor $R_L = 200 \, \Omega$ and let the temperature be $T = 300°K$. Letting the amplifier noise figure be $F_n = 3$ dB (a factor of 2), then from Eq. (7.20) the thermal noise current variance is $\sigma_T = 9.10 \times 10^{-12} B^{1/2}$. Next, select an InGaAs photodiode with a unity-gain responsivity $\mathcal{R} = 0.95$ A/W at 1550 nm and assume an operating BER $= 10^{-12}$ so that a value of $Q = 7$ is needed. If the photodiode gain is M, then the receiver sensitivity is

$$P_{\text{sensitivity}} =$$

$$\frac{7.37}{M} \left[5.6 \times 10^{-19} \, MF(M)B + 9.10 \times 10^{-12} \, B^{1/2} \right]$$

$$\tag{7.22}$$

Figure 7.11 shows the receiver sensitivity calculated from Eq. (7.22) as a function of data rate for typical InGaAs *pin* and avalanche photodiodes at 1550 nm for a 10^{-12} BER. In Fig. 7.11 the APD gain was taken to be $M = 10$ and $F(M) = 10^{0.7} = 5$. Note that the curves in Fig. 7.11 are for a BER given by $Q = 7$, a load resistor $R_L = 200 \, \Omega$, an amplifier noise figure $F_n = 3$ dB, and a 1550-nm wavelength. The sensitivity curves will change for different values of these parameters.

Example 7.6 Consider an InGaAs *pin* photodiode for which $M = 1$ and $F(M) = 1$. For the conditions in Eq. (7.22), what is the receiver sensitivity at a 1-Gb/s data rate for a 10^{-12} BER requirement?

Solution: From Eq. (7.22) we have

$$P_{\text{sensitivity}} = 7.37 \left[5.6 \times 10^{-19} (1 \times 10^9) \right.$$

$$\left. + 9.10 \times 10^{-12} (1 \times 10^9)^{\frac{1}{2}} \right]$$

$$= 2.12 \times 10^{-3} \text{ mW} = -26.7 \text{ dBm}$$

Example 7.7 Consider an InGaAs avalanche photodiode for which $M = 10$ and $F(M) = 5$. For the conditions in Eq. (7.22), what is the receiver sensitivity at a 1-Gb/s data rate for a 10^{-12} BER requirement?

Solution: From Eq. (7.22) we have

$$P_{\text{sensitivity}} = 0.737 \left[5.6 \times 10^{-19} (50)(1 \times 10^9) \right.$$

$$\left. + 9.10 \times 10^{-12} (1 \times 10^9)^{\frac{1}{2}} \right]$$

$$= 2.32 \times 10^{-4} \text{ mW} = -36.3 \text{ dBm}$$
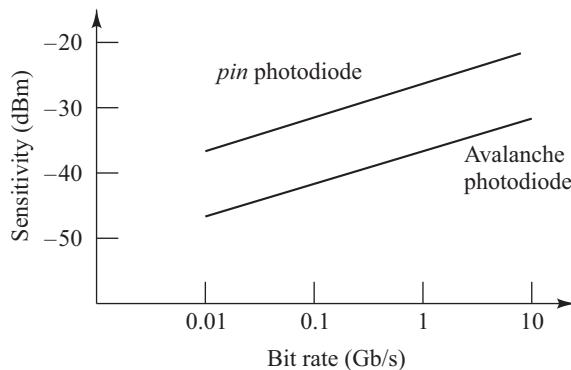
Fig. 7.11 *Sensitivities as a function of bit rate for generic pin and avalanche InGaAs photodiodes at 1550 nm for a 10^{-12} BER*

7.2.3 The Quantum Limit

In designing an optical system, it is useful to know what the fundamental physical bounds are on the system performance. Let us see what this bound is for the photodetection process. Suppose we have an ideal photodetector that has unity quantum efficiency and produces no dark current; that is, no electron–hole pairs are generated in the absence of an optical pulse. Given this condition, it is possible to find the minimum received optical power required for a specific bit-error rate performance in a digital system. This minimum received power level is known as the *quantum limit,* since all system parameters are assumed ideal and the performance is limited only by the photodetection statistics.

Assume that an optical pulse of energy E falls on the photodetector in a time interval τ. This can only be interpreted by the receiver as a 0 pulse if no electron–hole pairs are generated with the pulse present.

Example 7.8 A digital fiber optic link operating at 850 nm requires a maximum BER of 10^{-9}.

(a) Let us first find the quantum limit in terms of the quantum efficiency of the detector and the energy of the incident photon. From Eq. (7.23) the probability of error is

$$P_r(0) = e^{-\bar{N}} = 10^{-9}$$

Solving for $\bar{N}$, we have $\bar{N} = 9 \ln 10 = 20.7 \sim 21$. Hence, an average of 21 photons per pulse is required for this BER. Using Eq. (7.1) and solving for E, we get

$$E = 20.7 \frac{hv}{\eta}$$

(b) Now let us find the minimum incident optical power P_i that must fall on the photodetector to achieve a 10^{-9} BER at a data rate of 10 Mb/s for a simple binary-level

signaling scheme. If the detector quantum efficiency $\eta = 1$, then

$$E = P_i \tau = 20.7 hv = 20.7 \frac{hc}{\lambda}$$

where $1/\tau$ is one-half the data rate B; that is, $1/\tau = B/2$. (*Note:* This assumes an equal number of 0 and 1 pulses.) Solving for P_i,

$$P_i = 20.7 \frac{hcB}{2\lambda}$$

$$= \frac{20.7(6.626 \times 10^{-34}\ \text{J}\cdot\text{s})(3.0 \times 10^8\ \text{m/s})(10 \times 10^6\ \text{bits/s})}{2(0.85 \times 10^{-6}\ \text{m})}$$

$$= 24.2\ \text{pW}$$

or, when the reference power level is 1 mW,

$$P_i = -76.2\ \text{dBm}$$

From Eq. (7.2) the probability that $n = 0$ electrons are excited in a time interval t is

$$P_r(0) = e^{-\bar{N}} \tag{7.23}$$

where the average number of electron–hole pairs, $\bar{N}$, is given by Eq. (7.1). Thus, for a given error probability $P_r(0)$, we can find the minimum energy E required at a specific wavelength λ.

In practice, the sensitivity of most receivers is around 20 dB higher than the quantum limit because of various nonlinear distortions and noise effects in the transmission link. Furthermore, when specifying the quantum limit, one has to be careful to distinguish between average power and peak power. If one uses average power, the quantum limit given in Example 7.4 would be only 10 photons per bit for a 10^{-9} BER. Sometimes, the literature quotes the quantum limit based on these average powers. However, this can be misleading since the limitation on real components is based on peak and not average power.

7.3 | Eye Diagrams

The eye diagram is a powerful measurement tool for assessing the data-handling ability of a digital transmission system. This method has been used extensively for evaluating the performance of wire-line systems and also applies to optical fiber data links. Chapter 14 gives more details on BER test equipment and measurement methods.

7.3.1 Eye Pattern Features

The eye pattern measurements are made in the time domain and allow the effects of waveform distortion to be shown immediately on the display screen of standard BER test equipment. Figure 7.12 shows a typical display pattern, which is known as an *eye pattern* or an *eye diagram*. The basic upper and lower bounds are determined by the logic 1 and 0 levels, shown by b_{on} and b_{off}, respectively.
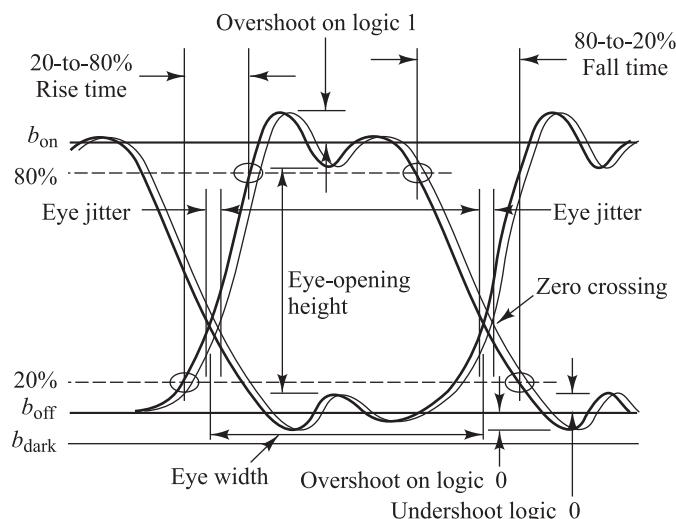
Fig. 7.12 *General configuration of an eye diagram showing the definitions of fundamental measurement parameters*
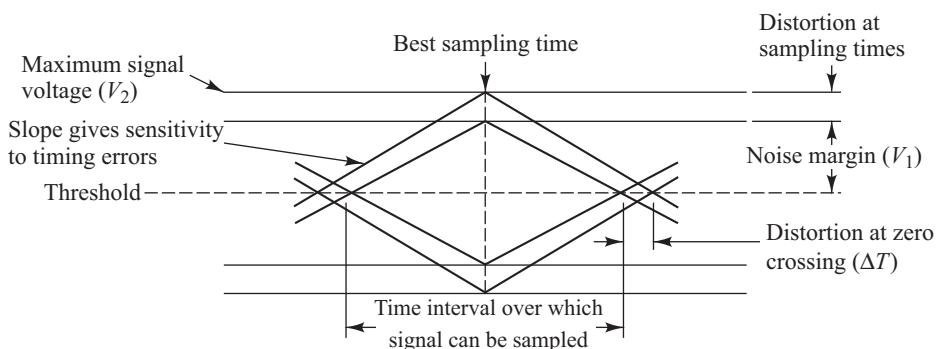
Fig. 7.13   *Simplified eye diagram showing the key performance parameters*

A great deal of system-performance information can be deduced from the eye pattern display. To interpret the eye pattern, consider Fig. 7.12 and the simplified drawing shown in Fig. 7.13. The following information regarding the signal amplitude distortion, timing jitter, and system rise time can be derived:

- The *width of the eye opening* defines the time interval over which the received signal can be sampled without error due to interference from adjacent pulses (known as *intersymbol interference*).
- The best time to sample the received waveform is when the *height of the eye opening* is largest. This height is reduced as a result of amplitude distortion in the data signal. The vertical distance between the top of the eye opening and the maximum signal level gives the degree of distortion. The more the eye closes, the more difficult it is to distinguish between ones and zeros in the signal.
- The height of the eye opening at the specified sampling time shows the noise margin or immunity to noise. *Noise margin* is the percentage ratio of the peak signal voltage V_1 for an alternating bit sequence (defined by the height of the eye opening) to the maximum signal voltage V_2 as measured from the threshold level, as shown in Fig. 7.13. That is

$$\text{Noise margin (percent)} = \frac{V_1}{V_2} \times 100 \text{ percent} \tag{7.24}$$

Example 7.9   Consider an eye diagram in which the center opening is about 90 percent due to intersymbol interference (ISI) degradation. What is the the ISI degradation in decibels?

Solution: The ISI degradation is given by

$$\text{ISI} = 20 \log \frac{V_1}{V_2} = 20 \log 0.90 = 0.915 \text{ dB}$$

- The rate at which the eye closes as the sampling time is varied (i.e., the slope of the eye pattern sides) determines the sensitivity of the system to *timing errors*. The possibility of timing errors increases as the slope becomes more horizontal.
- *Timing jitter* (also referred to as *edge jitter* or *phase distortion*) in an optical fiber system arises from noise in the receiver and pulse distortion in the optical fiber. Excessive jitter can result in bit errors, since such jitter can produce uncertainties in clock timing. This timing uncertainty will cause a receiver to lose synchronization with the incoming bit stream thereby incorrectly interpreting logic 1 and 0 pulses. If the signal is sampled in the middle of the time interval (i.e., midway between the times when the signal crosses the threshold level), then the amount of distortion ΔT

at the threshold level indicates the amount of jitter. Timing jitter is thus given by

$$\text{Timing jitter (percent)} = \frac{\Delta T}{T_b} \times 100 \text{ percent} \tag{7.25}$$

where T_b is one bit interval.

- Traditionally, the *rise time* is defined as the time interval between the points where the rising edge of the signal reaches 10 percent of its final amplitude to the time where it reaches 90 percent of its final amplitude. However, when measuring optical signals, these points are often obscured by noise and jitter effects. Thus, the more distinct values at the 20-percent and 80-percent threshold points normally are measured. To convert from the 20- to 80-percent rise time to a 10- to 90-percent rise time, one can use the approximate relationship

$$T_{10\text{-}90} = 1.25 \times T_{20\text{-}80} \tag{7.26}$$

A similar approach is used to determine the fall time.

- Any nonlinear effects in the channel transfer characteristics will create an asymmetry in the eye pattern. If a purely random data stream is passed through a purely linear system, all the eye openings will be identical and symmetrical.

Modern bit-error rate measurement instruments construct and display eye diagrams such as the example shown in Fig. 7.14. Ideally, if the signal impairments are small, the received pattern on the instrument display should exhibit sharp, clearly defined lines. However, time-varying signal impairments in the transmission path can lead to amplitude variations within the signal and to timing skews between the data signal and the associated clock signal. Note that a clock signal, which typically is encoded within a data signal, is used to help the receiver interpret the incoming data correctly. Thus in an actual link the received pattern will become wider or distorted on the sides and on the top and bottom, as shown in Fig. 7.14.

7.3.2   BER and Q-Factor Measurements

Since BER is a statistical parameter, its value depends on the measurement time and on the factors that cause the errors. If the errors are due to gaussian noise in a relatively stable transmission link, then a measurement time in which about 100 errors occur may be needed to ensure a statistically valid BER determination. Longer measurement times may be needed for systems in which bursts of errors can occur. For high-speed communications the required bit error rate typically needs to be 10^{-12} or lower. As an example, for a 10-Gb/s links a 10^{-12} BER means that one bit error occurs every 100 seconds. Such a level may be unacceptable, so even lower bit-error rates, such as 10^{-15}, may be required to assure customers of a high grade of service. Standards that define acceptable bit-error rates include the ITU-T 0.150 and 0.201 Recommendations and the ANSI T1.510 and T1.105 documents.[26,32,33]

Test times can be quite long. For example, to detect 100 errors for measuring a 10^{-12} BER in a 10-Gb/s link will require 2.8 hours. Thus test times on installed links could run
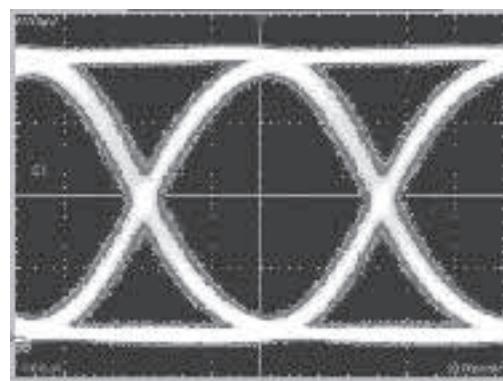
Fig. 7.14 *Typical eye diagram that shows relatively little signal distortion*
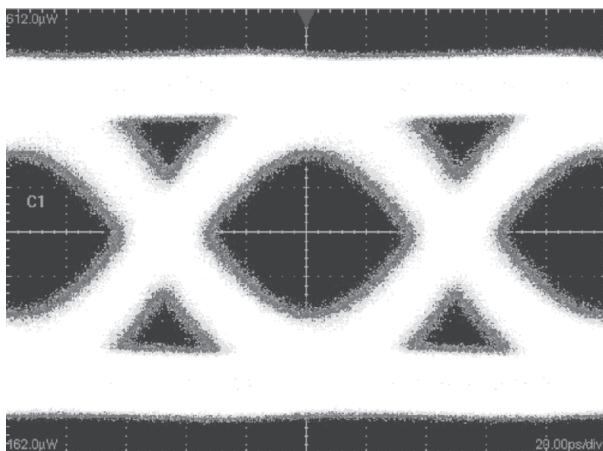
Fig. 7.15 *The inclusion of all possible signal distortion effects results in a stressed eye with only a small diamond-shaped opening.*

anywhere from 8 to 72 hours. To reduce such costly and time-consuming test periods, a *Q-factor technique* can be used. Although some accuracy is lost in this method, it reduces the test times to minutes instead of hours. In this method the receiver threshold is decreased, which increases the probability of errors and thus decreases test time.

A wide variety of sophisticated bit-error rate test equipment is available for both factory and field testing of optical communication equipment and transmission links. In addition to performing tests using standardized patterns or carrying out Q-factor-based measurements, more advanced equipment also measures performance by using a degraded signal that more closely represents what is seen in fielded links. This method is described in the IEEE 802.3ae specification for testing 10-Gigabit Ethernet (10-GbE) devices.[26,34] This *stressed-eye test* examines the worst-case performance condition by specifying a poor extinction ratio and adding multiple stresses, intersymbol interference (ISI) or vertical eye closure, sinusoidal interference, and sinusoidal jitter. The concept of this test is to assume that all different possible jitter and intersymbol interference impairments that might occur to a signal in a fielded link will close the eye down to a diamond shape, as shown in Fig. 7.15. If the eye opening of the optical receiver under test is greater than this diamond-shaped area of assured error-free operation, it is expected to operate properly in an actual fielded system. The stressed-eye template height typically is between 0.10 and 0.25 of the full pattern height. Chapter 14 gives more details on this stressed-eye test.

7.4 Burst-Mode Receivers

To address the continuously increasing demands by customers for higher-capacity connections to a central switching facility, network and service providers devised the concept of using a *passive optical network* (PON).[35–40] As described in more detail in Chapter 13, such network configurations have become known as *fiber-to-the-premises* (FTTP) networks. In a PON there are no active components between the central office and the customer premises. Instead, only passive optical components are placed in the network transmission path to guide the traffic signals contained within specific optical wavelengths to the user endpoints and back to the central office.

Figure 7.16 illustrates the architecture of a typical PON in which a fiber optic network connects switching equipment in a central office with a number of service subscribers. Examples of equipment in the central office include public telephone switches, video-on-demand servers, Internet protocol (IP) routers, Ethernet switches, and asynchronous transfer mode (ATM) switches. In the central office, data and digitized voice are combined and sent *downstream* to customers over an optical link by using a 1490-nm wavelength. The *upstream* (customer to central office) return path for the data and voice uses a 1310-nm wavelength. Video services are sent downstream with a 1550-nm wavelength. The transmission
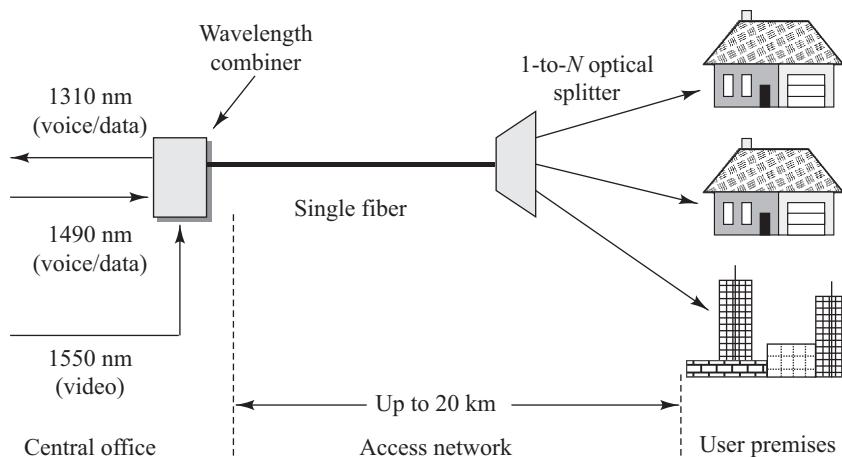
Fig. 7.16 *Architecture of a typical passive optical network*

equipment in the network consist of an *optical line termination* (OLT) situated at the central office and an *optical network termination* (ONT) at each customer premises.

Starting at the central office, one single-mode optical fiber strand runs to a passive *optical power splitter* near a housing complex, an office park, or some other campus environment. At this point a passive splitting device simply divides the optical power into *N* separate paths to the subscribers. The number of splitting directions can vary from 2 to 64, but normally there are 8, 16, or 32 paths. From the optical splitter, individual single-mode fibers then run to each building or serving equipment. The optical fiber transmission span from the central office to the user can be up to 20 km. For FTTP applications, at the central office the operational characteristics of an optical receiver in an OLT differ significantly from those of conventional point-to-point links.[35–38] This arises from the fact that the amplitude and phase of information packets received in successive time slots from different network user (customer) locations can vary widely from packet to packet, as is illustrated in Fig. 7.17. This is due to the 20-km possible distance variations of customers from the central office. In one case, suppose the closest and farthest customers attached to a common optical power splitter are 20 km apart and that the fiber attenuation is 0.5 dB/km. Then there is a 10-dB difference in the signal amplitudes that arrive at the OLT from these
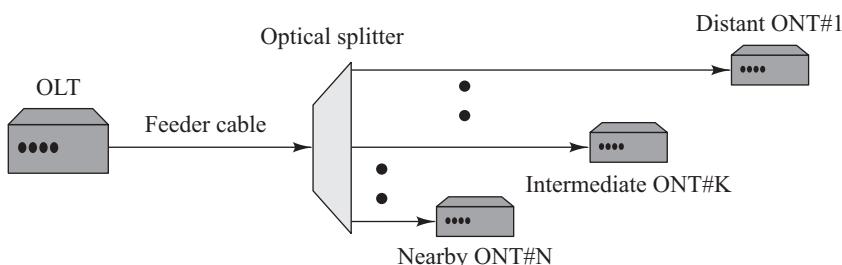
Fig. 7.17 *Large distance variations of customers from the central office result in different signal power losses across the PON.*
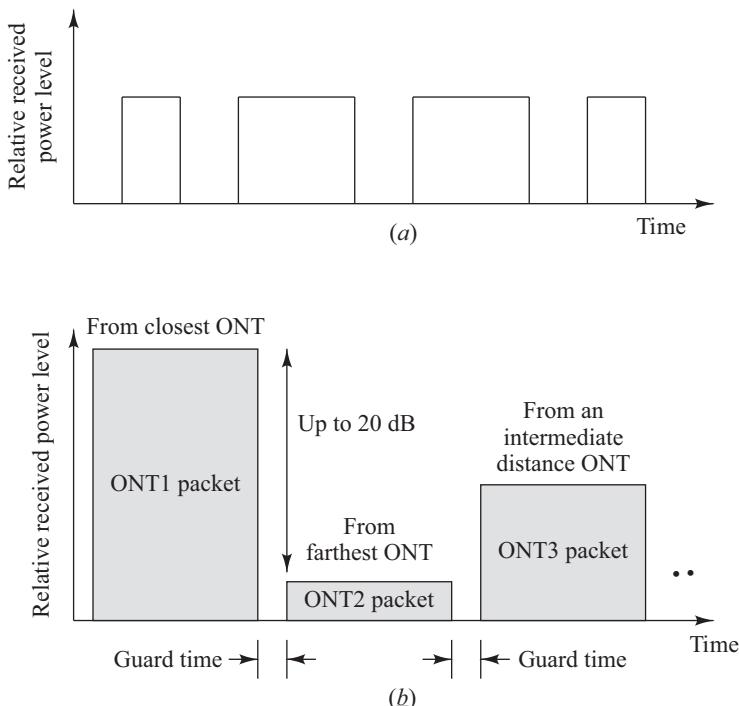
Fig. 7.18   *(a) Typical received data pattern in conventional point-to-point links; (b) Optical signal level variations in pulses that may arrive at an OLT*

two users if both have the same upstream laser output level. If there is an additional optical component in the transmission path going to one of the customer sites, then the difference in signal levels arriving at the OLT could vary up to 20 dB.

Figure 7.18 shows the consequence of this effect. The term ONT in this figure refers to the transceiver equipment at the customer location. The top part shows the type of data pattern that would be received in conventional point-to-point links, such as the signal levels arriving at a particular customer site from the central office. Here there is no amplitude variation in the received logic ones. The bottom part illustrates the optical signal pattern levels that may arrive at the OLT from various customers. In this case the signal amplitude changes from packet to packet depending on how far away each ONT is from the central office. The *guard time* shown in Fig. 7.18 provides a sufficient delay time to prevent collisions between successive packets that may come from different ONTs.

Since a conventional optical receiver is not capable of instantaneously handling rapidly changing differences in signal amplitude and clock phase alignment, a specially designed *burst-mode receiver* is needed. These receivers can quickly extract the decision threshold and determine the signal phase from a set of overhead bits placed at the beginning of each packet burst. However, this methodology results in a receiver sensitivity power penalty of up to 3 dB.

The key requirements of a burst-mode receiver are high sensitivity, wide dynamic range, and fast response time. The sensitivity is important in relation to the optical power budget, since, for example,

a 3-dB sensitivity improvement can double the size of the power splitter so that more customers can be attached to the PON. A wide dynamic range is essential for achieving a long network reach, that is, to be able to accommodate users located both close and far away from the central office.

The use of a conventional ac-coupling method is not possible in a burst-mode receiver, since the residual charge in a coupling capacitor following any particular data burst cannot dissipate fast enough in order not to affect the initial conditions of the next burst. The burst-mode receiver therefore requires additional circuitry to accommodate dc-coupled operation. Such receivers now are incorporated into standard commercially available OLT equipment.

7.5 Analog Receivers

In addition to the wide usage of fiber optics for the transmission of digital signals, there are many potential applications for analog links. These range from individual 4-kHz voice channels to microwave links operating in the multigigahertz region.[41–46] In the previous sections we discussed digital receiver performance in terms of error probability. For an analog receiver, the performance fidelity is measured in terms of a *signal-to-noise ratio*. This is defined as the ratio of the mean-square signal current to the mean-square noise current.

The simplest analog technique is to use amplitude modulation of the source. In this scheme, a time-varying electric signal $s(t)$ is used to modulate an optical source directly about some bias point defined by the bias current I_B, as shown in Fig. 7.19. The transmitted optical power $P(t)$ is thus of the form

$$P(t) = P_t\left[1 + ms(t)\right] \tag{7.27}$$
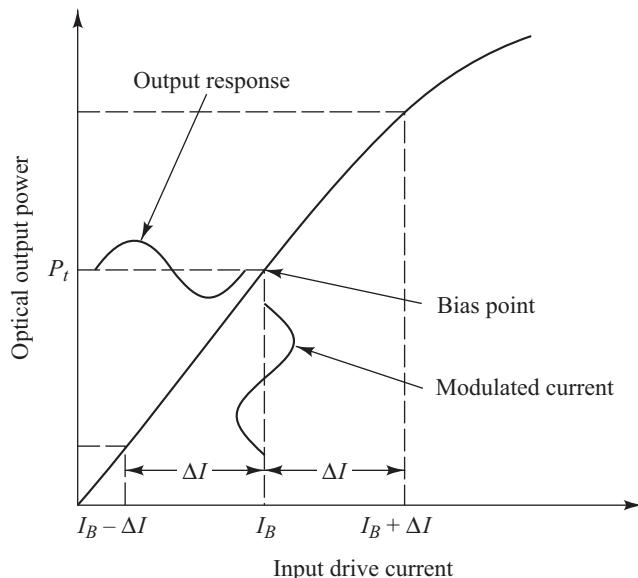
Fig. 7.19 *Direct analog modulation of an LED source*

where P_t is the average transmitted optical power, $s(t)$ is the analog modulation signal, and m is the modulation signal index defined by (see Sec. 4.5)

$$m = \frac{\Delta I}{I_B} \tag{7.28}$$

Here, ΔI is the variation in current about the bias point. In order not to introduce distortion into the optical signal, the modulation must be confined to the linear region of the light source output curve shown in Fig. 7.19. Also, if $\Delta I > I_B$, the lower portion of the signal gets cut off and severe distortion results.

At the receiver end, the photocurrent generated by the analog optical signal is

$$i_s(t) = \mathcal{R}MP_r[1 + ms(t)]$$

$$= I_pM[1 + ms(t)] \tag{7.29}$$

where $\mathcal{R}$ is the detector responsivitity, P_r is the average received optical power, $I_P = \mathcal{R}P_r$ is the primary photocurrent, and M is the photodetector gain. If $s(t)$ is a sinusoidally modulated signal, then the mean-square signal current at the photodetector output is (ignoring a dc term)

$$\left\langle i_s^2 \right\rangle = \frac{1}{2}\left(\mathcal{R}MmP_r\right)^2 = \frac{1}{2}\left(MmI_p\right)^2 \tag{7.30}$$

Recalling from Eq. (6.18) that the mean-square noise current for a photodiode receiver is the sum of the mean-square quantum noise current, the equivalent-resistance thermal noise current, the dark noise current, and the surface-leakage noise current, we have

$$\left\langle i_N^2 \right\rangle = 2q(I_p + I_D)M^2F(M)B_e + 2qI_LB_e + \frac{4k_BTB_e}{R_{eq}}F_t \tag{7.31}$$

where I_p = primary (unmultiplied) photocurrent = $\mathcal{R}P_r$
I_D = primary bulk dark current
I_L = surface-leakage current
$F(M)$ = excess photodiode noise factor $\simeq M^x(0 < x \leq 1)$
B_e = effective receiver noise bandwidth
R_{eq} = equivalent resistance of photodetector load and amplifier
F_t = noise figure of the baseband amplifier

By a suitable choice of the photodetector, the leakage current can be rendered negligible. With this assumption, the signal-to-noise ratio S/N is

$$\frac{S}{N} = \frac{\left\langle i_s^2 \right\rangle}{\left\langle i_N^2 \right\rangle} = \frac{\frac{1}{2}\left(\mathcal{R}MmP_r\right)^2}{2q\left(\mathcal{R}P_r + I_D\right)M^2F(M)B_e + \left(4k_BTB_e/R_{eq}\right)F_t}$$

$$= \frac{\frac{1}{2}\left(I_pMm\right)^2}{2q\left(I_p + I_D\right)M^2F(M)B_e + \left(4k_BTB_e/R_{eq}\right)F_t} \tag{7.32}$$

For a *pin* photodiode we have $M = 1$. When the optical power incident on the photodiode is small, the circuit (thermal) noise term dominates the noise current, so that

$$\frac{S}{N} \simeq \frac{\frac{1}{2} m^2 I_p^2}{\left(4 k_B T B_e / R_{\text{eq}}\right) F_t} = \frac{\frac{1}{2} m^2 \mathscr{R}^2 P_r^2}{\left(4 k_B T B_e / R_{\text{eq}}\right) F_t} \tag{7.33}$$

Here, the signal-to-noise ratio is directly proportional to the square of the photodiode output current and inversely proportional to the thermal noise of the circuit.

For large optical signals incident on a *pin* photodiode, the quantum (shot) noise associated with the signal detection process dominates, so that

$$\frac{S}{N} \simeq \frac{m^2 I_p}{4 q B_e} = \frac{m^2 \mathscr{R} P_r}{4 q B_e} \tag{7.34}$$

Since the signal-to-noise ratio in this case is independent of the circuit noise, it represents the fundamental or quantum limit for analog receiver sensitivity.

When an avalanche photodiode is employed at low signal levels and with low values of gain M, the circuit noise term dominates. At a fixed low signal level, as the gain is increased from a low value, the signal-to-noise ratio increases with gain until the quantum noise term becomes comparable to the circuit noise term. As the gain is increased further beyond this point, the signal-to-noise ratio *decreases* as $F(M)^{-1}$. Thus for a given set of operating conditions, there exists an optimum value of the avalanche gain for which the signal-to-noise ratio is a maximum. Since an avalanche photodiode increases the signal-to-noise ratio for small optical signal levels, it is the preferred photodetector for this situation.

For very large optical signal levels, the quantum noise term dominates the receiver noise. In this case, an avalanche photodiode serves no advantage, since the detector noise increases more rapidly with increasing gain M than the signal level. This is shown in Fig. 7.20, where we compare the signal-to-noise ratio for a *pin* and an avalanche photodiode receiver as a function of the received optical power. The signal-to-noise ratio for the avalanche photodetector is at the optimum gain (see Probs. 7.14 and 7.15). The parameter values chosen for this example are $B_e = 5$ MHz and 25 MHz, $x = 0.5$ for the avalanche photodiode and 0 for the *pin* diode, $m = 80$ percent, $\mathscr{R} = 0.5$ A/W, and $R_{\text{eq}}/F_t = 10^4$ Ω. We see that for low signal levels an avalanche
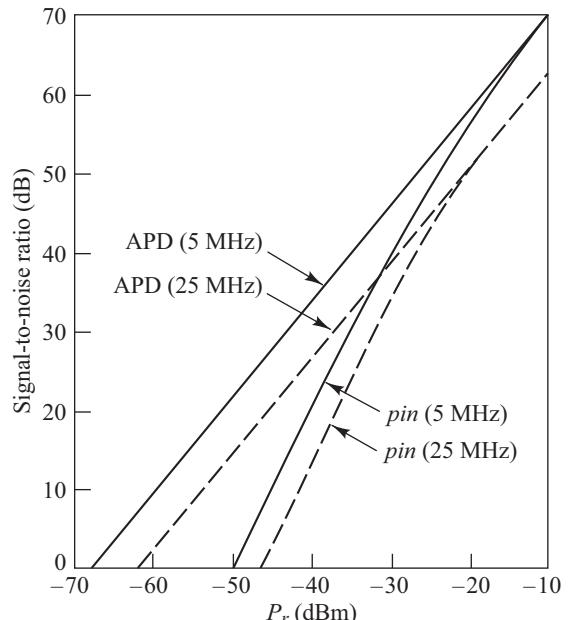
Fig. 7.20 *Comparison of the signal-to-noise ratio for pin and avalanche photodiodes as a function of received optical power for bandwidths of 5 and 25 MHz*

photodiode yields a higher signal-to-noise ratio, whereas at large received optical power levels a *pin* photodiode gives equal performance.

Example 7.10   Consider an analog optical fiber system operating at 1550 nm, which has an effective receiver noise bandwidth of 5 MHz. Assuming that the received signal is quantum noise limited, what is the incident optical power necessary to have a signal-to-noise ratio of 50 dB at the receiver? Assume the responsivity is 0.9 A/W and that m = 0.5.

Solution: First we note that a 50-dB SNR means that $S/N = 10^5$. Then, solving Eq. (7.34) for P_r yields

$$P_r = \frac{(S/N)4qB_e}{m^2 \mathcal{R}} = \frac{(1 \times 10^5)4(1.6 \times 10^{-19})(5 \times 10^6)}{(0.5)^2(0.90)}$$

$$= 1420 \text{ nW} = 1.42 \times 10^{-3} \text{ mW}$$

or in dBm

$$P_r(\text{dBm}) = 10 \log P_r = 10 \log 1.42 \times 10^{-3}$$

$$= -28.5 \text{ dBm}$$

PROBLEMS

7.1 In avalanche photodiodes the ionization ratio k is approximately 0.02 for silicon, 0.35 for indium gallium arsenide, and 1.0 for germanium. Show that, for gains up to 100 in Si and up to 25 in InGaAs and Ge, the excess noise factor $F(M)$ given by Eq. (7.3) can be approximated to within 10 percent by M^x, where x is 0.3 for Si, 0.7 for InGaAs, and 1.0 for Ge.

7.2 The equalizer in an optical receiver normally is a linear frequency-shaping filter used to mitigate the effects of signal distortion and intersymbol interference. To account for the fact that pulses arrive rounded and distorted at the receiver, the binary digital pulse train incident on the photodetector can be described by

$$P(t) = \sum_{n=-\infty}^{\infty} b_n h_p\left(t - nT_b\right)$$

Here $P(t)$ is the received optical power, T_b is the bit period, b_n represents the energy in the nth pulse ($b_n = b_0$ for a 0 pulse and b_1 for 1 pulse), and $h_p(t)$ is the received pulse shape.

   Show that the following pulse shapes satisfy the normalization condition

$$\int_{-\infty}^{\infty} h_p(t)\,dt = 1$$

(*a*) Rectangular pulse (α = constant)

$$h_p(t) = \begin{cases} \dfrac{1}{\alpha T_b} & \text{for} \quad \dfrac{-\alpha T_b}{2} < t < \dfrac{\alpha T_b}{2} \\ 0 & \text{otherwise} \end{cases}$$

(*b*) Gaussian pulse

$$h_p(t) = \frac{1}{\sqrt{2\pi}} \frac{1}{\alpha T_b} e^{-t^2/2(\alpha T_b)^2}$$

(*c*) Exponential pulse

$$h_p(t) = \begin{cases} \dfrac{1}{\alpha T_b} e^{-t/\alpha T_b} & \text{for } 0 \le t \le \infty \\ 0 & \text{otherwise} \end{cases}$$

7.3 Derive the error probability expression given by Eq. (7.16) from Eq. (7.8).

7.4 A transmission system sends out information at 200,000 b/s. During the transmission process, fluctuation noise is added to the signal so that at the decoder output the signal pulses are 1 V in amplitude and the rms noise voltage is 0.2 V.

(*a*) Assuming that ones and zeros are equally likely to be transmitted, what is the average time in which an error occurs?

(*b*) How is this time changed if the voltage amplitude is doubled with the rms noise voltage remaining the same?

7.5 Consider the probability distributions shown in Fig. 7.7, where the signal voltage for a binary 1 is V_1 and $v_{th} = V_1/2$.

(*a*) If $\sigma = 0.20\, V_1$ for $p(y|0)$ and $\sigma = 0.24\, V_1$ for $p(y|1)$, find the error probabilities $P_0(v_{th})$ and $P_1(v_{th})$.

(*b*) If $a = 0.65$ and $b = 0.35$, find P_e.

(*c*) If $a = b = 0.5$, find P_e.

7.6 Analogous to Fig. 7.11, plot the sensitivities at 1310 nm for a 10^{-9} BER of receivers that have InGaAs *pin* and avalanche photodiodes. Take the data rates to range from 10 Mb/s to 10 Gb/s. Let the temperature be 300°K and let the load resistor $R_L = 100\ \Omega$. Assume the InGaAs photodiodes have a unity-gain responsivity $\mathcal{R} = 0.90$ A/W at 1310 nm and that the gain of the APD is $M = 10$.

7.7 An LED operating at 1300 nm injects 25 μW of optical power into a fiber. If the attenuation between the LED and the photodetector is 40 dB and the photodetector quantum efficiency is 0.65, what is the probability that fewer than 5 electron–hole pairs will be generated at the detector in a 1-ns interval?

7.8 Show that, by using Eq. (7.14), the error-probability expressions given by Eq. (7.11) both reduce to Eq. (7.13).

7.9 A useful approximation to ½ (1 − erf x) for values of x greater than 3 is given by

$$\frac{1}{2}\left(1 - \operatorname{erf} x\right) \simeq \frac{\exp(-x^2)}{2\sqrt{\pi}x}$$

Using this approximation, consider an on-off binary system that transmits the signal levels 0 and A with equal probability in the presence of gaussian noise. Let the signal amplitude A be K multiplied by the standard deviation of the noise.

(*a*) Calculate the net probability of error if $K = 10$.

(*b*) Find the value of K required to give a net error probability of 10^{-5}.

7.10 If a bit-corrupting noise burst lasts for 2 ms, how many bits are affected at data rates of 10 Mb/s, 100 Mb/s, and 2.5 Gb/s?

7.11 Consider a thermal-noise-limited analog optical fiber system that uses a *pin* photodiode with a responsivity of 0.85 at 1310 nm. Assume the system uses a modulation index of 0.5 and operates in a 5-MHz bandwidth. Let the mean-square thermal noise current for the receiver be 2×10^{-23} A²/Hz. What is the peak-to-peak signal power to rms noise ratio at the receiver when the average incident optical power is −20 dBm?

7.12 Consider a quantum-noise-limited analog optical fiber system that uses a *pin* photodiode with a responsivity of 0.85 at 1310 nm. Assume the system uses a modulation index of 0.6 and operates in a 40-MHz bandwidth. If we neglect detector dark current, what is the signal-to-noise ratio when the incident optical power at the receiver is −15 dBm?

7. 13 Show that the signal-to-noise ratio given by Eq. (7.32) is a maximum when the gain is optimized at

$$M_{opt}^{2+x} = \frac{4k_B T F_t / R_{eq}}{q\left(I_P + I_D\right)x}$$

7.14 (*a*) Show that, when the gain M is given by the expression in Prob. 7.13, the signal-to-noise ratio given by Eq. (7.32) can be written as

$$\frac{S}{N} = \frac{xm^2}{2B(2+x)} \frac{I_p^2}{\left[q\left(I_p + I_D\right)x\right]^{2/(2+x)}} \left(\frac{R_{eq}}{4k_B T F_t}\right)^{x/(2+x)}$$

(*b*) Show that, when I_p is much larger than I_D, the foregoing expression becomes

$$\frac{S}{N} = \frac{m^2}{2Bx(2+x)} \left[\frac{\left(xI_p\right)^{2(1+x)}}{q^2\left(4k_B T F_t / R_{eq}\right)^x}\right]^{1/(2+x)}$$

7.15 Consider the signal-to-noise ratio expression given in Prob. 7.14a. Analogous to Fig. 7.20, plot S/N in dB [i.e., 10 log(S/N)] as a function of the received power level P_r in dBm when the dark current $I_D = 10$ nA and $x = 1.0$. Let $B_e = 5$ MHz, $m = 0.8$, $\mathcal{R} = 0.5$ A/W, $T = 300$ K, and $R_{eq}/F_t = 10^4\ \Omega$. Recall that $I_p = \mathcal{R}P_r$.

REFERENCES

1. S. D. Personick, "Receiver design for digital fiber optic communication systems," *Bell Sys. Tech. J.*, vol. 52, pp. 843–886, July/Aug. 1973.

2. S. D. Personick, P. Balaban, J. Bobsin, and P. Kumer, "A detailed comparison of four approaches to the calculation of the sensitivity of optical fiber receivers," *IEEE Trans. Commun.*, vol. COM-25, pp. 541–548, May 1977.

3. S. Bottacchi, *Noise and Signal Interference in Optical Fiber Transmission Systems,* Wiley, New York, 2009.

4. B. L. Kasper, O. Mizuhra, and Y.-K. Chen, "High bit-rate receivers, transmitters, and electronics," chap. 16, in I. P. Kaminow and T. Li, eds., *Optical Fiber Telecommunications–IVA*, Academic, New York, 2002.

5. K. Schneider and H. Zimmermann, *Highly Sensitive Optical Receivers,* Springer, New York, 2006.

6. J. Savoi and B. Razavi, *High-Speed CMOS Circuits for Optical Receivers,* Springer, New York, 2001.

7. T. V. Muoi, "Receiver design for high-speed optical-fiber systems," *J. Lightwave Tech.*, vol. LT-2, pp. 243–267, June 1984.

8. S. R. Forrest, "The sensitivity of photoconductive receivers for long-wavelength optical communications," *J. Lightwave Tech.*, vol. LT-3, pp. 347–360, April 1985.

9. M. Brain and T. P. Lee, "Optical receivers for lightwave communication systems," *J. Lightwave Tech.,* vol. LT-3, pp. 1281–1300, Dec. 1985.

10. B. L. Kasper, "Receiver design," in *Optical Fiber Telecommunications–II*, S. E. Miller and I. P. Kaminow, eds., Academic, New York, 1988.

11. G. F. Williams, "Lightwave receivers," in T. Li, ed., *Topics in Lightwave Transmission Systems*, Academic, New York, 1991.

12. Y. K. Park and S. W. Granlund, "Optical preamplifier receivers: Application to long-haul digital transmission," *Optical Fiber Tech.*, vol. 1, pp. 59–71, Oct. 1994.

13. D. A. Fishman and B. S. Jackson, "Transmitter and receiver design for amplified lightwave systems," chap. 3, pp. 69–114, in I. P. Kaminow and T. L. Koch, eds., *Optical Fiber Telecommunications–III*, vol. B, Academic, New York, 1997.

14. S. B. Alexander, *Optical Communication Receiver Design*, SPIE Optical Engineering Press, Bellingham, WA, 1997.

15. E. Säckinger, *Broadband Circuits for Optical Fiber Communication*, Wiley, Hoboken, NJ, 2005.

16. K. Schneider and H. K. Zimmermann, *Highly Sensitive Optical Receivers*, Springer, New York, 2006.

17. C. Hermans and M. Stayaert, *Broadband Opto-Electrical Receivers in Standard CMOS*, Springer, New York, 2007.

18. G. Keiser, *Local Area Networks*, McGraw-Hill, New York, 2nd ed., 2002.

19. R. E. Ziemer and W. H. Tranter, *Principles of Communications: Systems, Modulation, and Noise*, Wiley, Hoboken, NJ, 6th ed., 2009.

20. A. B. Carlson and P. Crilly, *Communication Systems*, McGraw-Hill, Burr Ridge, IL, 5th ed., 2010.

21. S. V. Vaseghi, *Advanced Digital Signal Processing and Noise Reduction*, Wiley, Hoboken, NJ, 4th ed., 2009.

22. L. W. Couch II, *Digital and Analog Communication Systems*, Prentice Hall, Upper Saddle River, NJ, 7th ed., 2007.

23. E. A. Newcombe and S. Pasupathy, "Error rate monitoring for digital communications," *Proc. IEEE*, vol. 70, pp. 805–828, Aug. 1982.

24. N. S. Bergano, F. W. Kerfoot, and C. R. Davidson, "Margin measurements in optical amplifier systems," *IEEE Photonics Tech. Lett.*, vol. 5, pp. 304–306, Aug. 1993.

25. S. W. Hinch and C. M. Miller, "Analysis of digital modulation on optical carriers," in D. Derickson, ed., *Fiber Optic Test and Measurement*, Prentice Hall, Upper Saddle River, NJ, 1998.

26. ITU-T Recommendation O.201, *Q-factor Test Equipment to Estimate the Transmission Performance of Optical Channels*, July 2003.

27. I. Shake, H. Takara, and S. Kawanishi, "Simple measurement of eye diagram and BER using high-speed asynchronous sampling," *J. Lightwave Tech.*, vol. 22, pp. 1296–1302, Jan. 2004.

28. A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, Burr Ridge, IL, 4th ed., 2002.

29. P. Z. Peebles Jr., *Probability, Random Variables, and Random Signal Principles*, McGraw-Hill, Burr Ridge, IL, 4th ed., 2001.

30. W. Navidi, *Principles of Statistics for Engineers and Scientists*, McGraw-Hill, New York, 2010.

31. D. Zwillinger, ed., *Standard Mathematical Tables and Formulae*, CRC Press, Boca Raton, FL, 31st ed., 2003.

32. American National Standards Institute, ANSI T1.510-1999, *Network Performance Parameters for Dedicated Digital Services for Rates Up to and Including DS3-Specifications*, 1999.

33. American National Standards Institute, ANSI T1.105.03-2003, *Synchronous Optical Network (SONET)—Jitter and Wander at Network and Equipment Interfaces*, 2003.

34. IEEE 802.3-2005, *LAN/MAN CSMA/CD Access Method and Physical Layer Specifications*, 2005.

35. G. Keiser, *FTTX Concepts and Applications*, Wiley, Hoboken, NJ, 2006.

36. P. E. Green, "Fiber to the home: The next big broadband thing," *IEEE Commun. Mag.,* vol. 42, pp. 100–106, Sept. 2004.

37. C. Su, L.-K. Chen, and K. W. Cheung, "Theory of burst-mode receiver and its applications in optical multiaccess networks," *J. Lightwave Tech.*, vol. 15, pp. 590–606, Apr. 1997.

38. K. Schneider and H. Zimmermann, "Fast transimpedance switching burst-mode CMOS optical receiver," *Inter. J. Circuit Theory App.,* vol. 35, no. 3, pp. 355–370, May 2007.

39. S. Nishihara, S. Kimura, T. Yoshida, M. Nakamura, J. Terada, K. Nishimura, K. Kishine, K. Kato, Y. Ohtomo, N. Yoshimoto, T. Imai, and M. Tsubokawa, "A burst-mode 3R receiver for 10-Gbit/s PON systems with high sensitivity, wide dynamic range, and fast response," *J. Lightwave Tech.,* vol. 26, no.1, pp. 99–107, Jan. 2008.

40. C. Mélange, B. Baekelandt, J. Bauwelinck, P. Ossieur, T. De Ridder, X.-Z. Qiu, and J. Vandewege, "Burst-mode CDR performance in long-reach high-split passive optical networks," *J. Lightwave Tech.,* vol. 27, no.17, pp. 3837–3844, Sept. 2008.

41. Special Issue on Microwave and Millimeter-Wave Photonics, *IEEE Microwave Theory Tech.,* vol. 49, part II, Oct. 2001.

42. Special Issue on Microwave Photonics, *IEEE Microwave Theory Tech.,* vol. 54, part II, Feb. 2006.

43. B. Razavi, *Design of Integrated Circuits for Optical Communications,* McGraw-Hill, New York, 2003.

44. S. Haykin and M. Moher, "An Introduction to Analog and Digital Communications," Wiley, Hoboken, NJ, 2nd ed., 2006.

45. C. Cox, *Analog Optical Links,* Cambridge University Press, Cambridge, U.K., 2004.

46. A. Brillant, *Digital and Analog Fiber Optic Communications for CATV and FTTx Applications,* Wiley, Hoboken, NJ, 2008.

CHAPTER 8

Digital Links

The preceding chapters have presented the fundamental characteristics of the individual building blocks of an optical fiber transmission link. These include the optical fiber transmission medium, the optical source, the photodetector and its associated receiver, and the connectors used to join individual fiber cables to each other and to the source and detector. Now we will examine how these individual parts can be put together to form a complete optical fiber transmission link. In particular, we will study basic digital links in this chapter, and analog links in Chapter 9. More complex transmission links are examined in Chapter 13.

The first discussion involves the simplest case of a point-to-point link. This includes examining the components that are available for a particular application and seeing how these components relate to the system performance criteria (such as dispersion and bit-error rate). For a given set of components and a given set of system requirements, we then carry out a power budget analysis to determine whether the fiber optic link meets the attenuation requirements or if amplifiers are needed within the link to boost the power level. The final step is to perform a system rise-time analysis to verify that the overall system performance requirements are met.

The analysis in Sec. 8.1 assumes that the optical power falling on the photodetector is a clearly defined function of time within the statistical nature of the quantum detection process. In reality, various signal impairments can degrade the link performance. These impairments can reduce the power of the optical signal arriving at the receiver from the ideal case, which is known as a *power penalty* for that effect. Section 8.2 describes the power penalties associated with some key impairments that may be observed in an optical link.

To control errors and to improve the reliability of a communication line, it is necessary to be able to detect the errors and then correct them or retransmit the information. Section 8.3 describes error detection and correction methods that are used in a variety of optical fiber communication links. Basic *coherent detection* schemes are addressed in Sec. 8.4. These schemes are used in place of direct-detection methods to improve receiver sensitivies, especially for high-speed links operating at 40 and 160 Gb/s. Chapter 13 describes additional signal modulation and detection formats, such as *differential phase-shift keying* (DPSK) and *differential quadrature phase-shift keying* (DQPSK).

Versatile and powerful modeling and simulation tools are commercially available for doing many of the tasks described in this chapter. These software-based tools can run on conventional personal computers and include functions such as estimations of the BER and power-penalty effects with different optical receiver models, calculations of link power budgets, and simulation of link performance when using

different components. The book website (see the URL given in the Preface) provides links to several of the vendor websites of such modeling and simulation tools. These websites contain a variety of interactive demonstration simulation modules related to material in the book that the reader can download. These modules illustrate various component and optical fiber link performances discussed in this and other chapters of the book. The reader has the ability to change various parameter values in these modules and see the effects in graphical representations.

8.1 Point-to-Point Links

The simplest transmission link is a point-to-point line that has a transmitter on one end and a receiver on the other, as is shown in Fig. 8.1. This type of link places the least demand on optical fiber technology and thus sets the basis for examining more complex system architectures.[1–10]

The design of an optical link involves many interrelated variables among the fiber, source, and photodetector operating characteristics, so that the actual link design and analysis may require several iterations before they are completed satisfactorily. Since performance and cost constraints are very important factors in fiber optic communication links, the designer must carefully choose the components to ensure that the desired performance level can be maintained over the expected system lifetime without overspecifying the component characteristics.

The following key system requirements are needed in analyzing a link:
1. The desired (or possible) transmission distance
2. The data rate or channel bandwidth
3. The bit-error rate (BER)

To fulfill these requirements, the designer has a choice of the following components and their associated characteristics:

1. Multimode or single-mode optical fiber
   (*a*) Core size
   (*b*) Core refractive-index profile
   (*c*) Bandwidth or dispersion
   (*d*) Attenuation
   (*e*) Numerical aperture or mode-field diameter

2. LED or laser diode optical source
   (*a*) Emission wavelength
   (*b*) Spectral line width
   (*c*) Output power
   (*d*) Effective radiating area
   (*e*) Emission pattern
   (*f*) Number of emitting modes
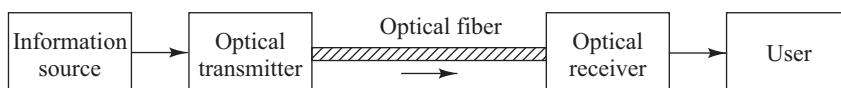
Fig. 8.1 *Simplex point-to-point link*

3. pin or avalanche photodiode

(*a*) Responsivity
(*b*) Operating wavelength
(*c*) Speed
(*d*) Sensitivity

Two analyses usually are carried out to ensure that the desired system performance can be met: these are the *link power budget* and the system *rise-time budget* analyses. In the link power budget analysis one first determines the power margin between the optical transmitter output and the minimum receiver sensitivity needed to establish a specified BER. This margin can then be allocated to connector, splice, and fiber losses, plus any additional margins required for other components, possible component degradations, transmission-line impairments, or temperature effects. If the choice of components did not allow the desired transmission distance to be achieved, the components might have to be changed or amplifiers might have to be incorporated into the link.

Once the link power budget has been established, the designer can perform a system rise-time analysis to ensure that the desired overall system performance has been met. We now examine these two analyses in more detail.

8.1.1   System Considerations

In carrying out a link power budget, we first decide at which wavelength to transmit and then choose components that operate in this region. If the distance over which the data are to be transmitted is not too far, we may decide to operate in the 770-to-910 nm region. On the other hand, if the transmission distance is relatively long, we may want to take advantage of the lower attenuation and dispersion that occurs in the O-band through U-band region.

Having decided on a wavelength, we next interrelate the system performances of the three major optical link building blocks; that is, the receiver, transmitter, and optical fiber. Normally, the designer chooses the characteristics of two of these elements and then computes those of the third to see if the system performance requirements are met. If the components have been over- or underspecified, a design iteration may be needed. The procedure we follow here is first to select the photodetector. We then choose an optical source and see how far data can be transmitted over a particular fiber before an amplifier is needed in the line to boost up the power level of the optical signal.

In choosing a particular photodetector, we mainly need to determine the minimum optical power that must fall on the photodetector to satisfy the bit-error rate (BER) requirement at the specified data rate. In making this choice, the designer also needs to take into account any design cost and complexity constraints. As noted in Chapters 6 and 7, a *pin* photodiode receiver is simpler, more stable with changes in temperature, and less expensive than an avalanche photodiode receiver. In addition, *pin* photodiode bias voltages are normally less than 5 V, whereas those of avalanche photodiodes range from 40 V to several hundred volts. However, the advantages of *pin* photodiodes may be overruled by the increased sensitivity of the avalanche photodiode if very low optical power levels are to be detected.

The system parameters involved in deciding between the use of an LED and a laser diode are signal dispersion, data rate, transmission distance, and cost. As shown in Chapter 4, the spectral width of the laser output is much narrower than that of an LED. This is of importance in the 770-to-910-nm region, where the spectral width of an LED and the dispersion characteristics of multimode silica fibers limit the data-rate-distance product to around 150 (Mb/s) · km. For higher values [up to 2500 (Mb/s) · km], a laser must be used at these wavelengths. At wavelengths around 1.3 μm, where signal dispersion is very low, bit-rate-distance products of at least 1500 (Mb/s) · km are achievable with LEDs in multimode fibers.

For InGaAsP lasers, distances of 150 m can be achieved at 100-Gb/s rates in OM4 multimode fiber at 1.3 μm (see Sec.13.4). A single-mode fiber can provide significantly higher rates over longer distances.

Since laser diodes typically couple from 10 to 15 dB more optical power into a fiber than an LED, greater repeaterless transmission distances are possible with a laser. This advantage and the lower dispersion capability of laser diodes may be offset by cost constraints. Not only is a laser diode itself more expensive than an LED, but also the laser transmitter circuitry is much more complex, since the lasing threshold has to be dynamically controlled as a function of temperature and device aging. However, a wide variety of cost-effective laser transmitters are on the market.

For the optical fiber, we have a choice between single-mode and multimode fiber, either of which could have a step- or a graded-index core. This choice depends on the type of light source used and on the amount of dispersion that can be tolerated. Light-emitting diodes (LEDs) tend to be used with multimode fibers. The optical power that can be coupled into a fiber from an LED depends on the core-cladding index difference Δ, which, in turn, is related to the numerical aperture of the fiber (for $\Delta = 0.01$, the numerical aperture NA $\simeq 0.21$). As Δ increases, the fiber-coupled power increases correspondingly. However, since dispersion also becomes greater with increasing Δ, a tradeoff must be made between the optical power that can be launched into the fiber and the maximum tolerable dispersion.

When choosing the attenuation characteristics of a cabled fiber, the excess loss that results from the cabling process must be considered in addition to the attenuation of the fiber itself. This must also include connector and splice losses as well as environmental-induced losses that could arise from temperature variations, radiation effects, and dust and moisture on the connectors.

8.1.2 Link Power Budget

An optical power loss model for a point-to-point link is shown in Fig. 8.2. The optical power received at the photodetector depends on the amount of light coupled into the fiber and the losses occurring in the fiber and at the connectors and splices. The link loss budget is derived from the sequential loss contributions of each element in the link. Each of these loss elements is expressed in decibels (dB) as

$$\text{loss} = 10 \log \frac{P_{\text{out}}}{P_{\text{in}}} \qquad (8.1)$$

where P_{in} and P_{out} are the optical powers entering and leaving the loss element, respectively. The loss value corresponding to a particular element generally is called the *insertion loss* for that element.

In addition to the link loss contributors shown in Fig. 8.2, a link power margin is normally provided in the analysis to allow for component aging, temperature fluctuations, and losses arising from components
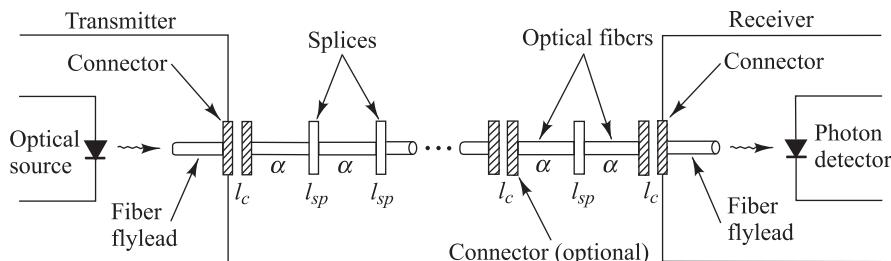
Fig. 8.2 *Optical power loss model for a point-to-point link. The losses occur at connectors (l_c), at splices (l_{sp}), and in the fiber (α).*

that might be added at future dates. A link margin of 3 to 6 dB is generally used for systems that are not expected to have additional components incorporated into the link in the future.

The link loss budget simply considers the total optical power loss P_T that is allowed between the light source and the photodetector, and allocates this loss to cable attenuation, connector loss, splice loss, and system margin. Thus, if P_S is the optical power emerging from the end of a fiber flylead attached to the light source or from a source-coupled connector, and if P_R is the receiver sensitivity, then

$$P_T = P_S - P_R$$

$$= 2l_c + \alpha L + \text{system margin} \qquad (8.2)$$

Example 8.1 To illustrate how a link loss budget is set up, let us carry out a specific design example. We begin by specifying a data rate of 20 Mb/s and a bit-error rate of 10^{-9} (i.e., at most one error can occur for every 10^9 bits sent). For the receiver, we choose a silicon *pin* photodiode operating at 850 nm. Figure 8.3 shows that the required receiver input signal is –42 dBm (42 dB below 1 mW). We next select a GaAlAs LED that can couple a 50-μW (–13-dBm) average optical power level into a fiber flylead with a 50-μm core diameter. We thus have a 29-dB allowable power loss. Assume further that a 1-dB loss occurs when the fiber flylead is connected to the cable and another 1-dB connector loss occurs at the cable-photodetector interface. Including a 6-dB system margin, the possible transmission distance for a cable with an attenuation α can be found from Eq. (8.2):

$$P_T = P_S - P_R = 29 \text{ dB}$$
$$= 2(1 \text{ dB}) + \alpha L + 6 \text{ dB}$$

If α = 3.5 dB/km, then a 6.0-km transmission path is possible.

The link power budget can be represented graphically as is shown in Fig. 8.4. The vertical axis represents the optical power loss allowed between the transmitter and the receiver. The horizontal axis gives the transmission distance. Here, we show a silicon *pin* receiver with a sensitivity of –42 dBm (at 20 Mb/s) and an LED with an output power of –13 dBm coupled into a fiber flylead. We subtract a 1-dB connector loss at each end, which leaves a total margin of 27 dB. Subtracting a 6-dB system safety margin leaves us with a tolerable loss of 21 dB that can be allocated to cable and splice loss. The slope of the line shown in Fig. 8.4 is the 3.5-dB/km cable (and splice, in this case) loss. This line starts at the –14-dBm point (which is the optical power coupled into the cabled fiber) and ends at the –35-dBm level (the receiver sensitivity minus a 1-dB connector loss and a 6-dB system margin). The intersection point D then defines the maximum possible transmission path length.

Fig. 8.3 *Receiver sensitivities as a function of bit rate. The Si pin, Si APD, and InGaAs pin curves are for a 10^{-9} BER. The InGaAs APD curve is for a 10^{-11} BER.*
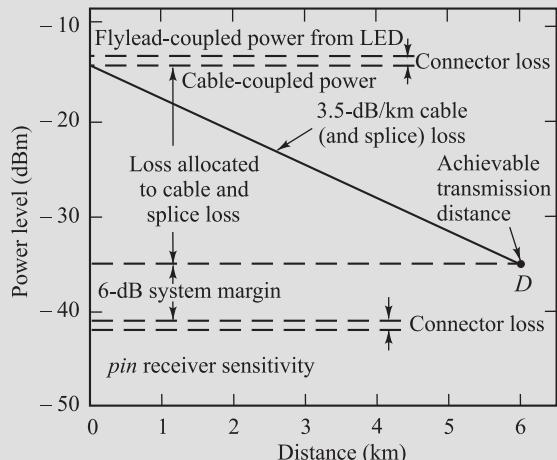
Fig. 8.4 *Graphical representation of a link loss budget for an 850-nm LED/pin system operating at 20 Mb/s*

Here, l_c is the connector loss, α is the fiber attenuation (dB/km), L is the transmission distance, and the system margin is nominally taken as 6 dB. Here, we assume that the cable of length L has connectors only on the ends and none in between. The splice loss is incorporated into the cable loss for simplicity.

A convenient procedure for calculating the power budget is to use a tabular or spreadsheet form. We illustrate this by way of an example for a 2.5-Gb/s link that may be used for SONET OC-48 or SDH STM-16.

Example 8.2 Consider a 1550-nm laser diode that launches a +3-dBm (2-mW) optical power level into a fiber flylead, an InGaAs APD with a –32-dBm sensitivity at 2.5 Gb/s, and a 60-km long optical cable with a 0.3-dB/km attenuation. Assume that here, because of the way the equipment is arranged, a 5-m optical jumper cable is needed at each end between the end of the transmission cable and the SONET equipment rack as shown in Fig. 8.5. Assume that each jumper cable introduces a loss of 3 dB. In addition, assume a 1-dB connector loss occurs at each fiber joint (two at each end because of the jumper cables).

Table 8.1 lists the components in column 1 and the associated optical output, sensitivity, or loss in column 2. Column 3 gives the power margin available after subtracting the component loss from the total optical power loss that is allowed between the light source and the photodetector, which, in this case, is 35 dB. Adding all the losses results in a final power margin of 7 dB.

Table 8.1 *Example of a spreadsheet for calculating an optical-link power budget*

Component/loss parameter	Output/sensitivity/loss	Power margin (dB)
Laser output	3 dBm	
APD sensitivity at 2.5 Gb/s	–32 dBm	
Allowed loss [3 – (–32)]		35
Source connector loss	1 dB	34
Jumper + connector loss	3 + 1 dB	30
Cable attenuation (60 km)	18 dB	12
Jumper + connector loss	3 + 1 dB	8
Receiver connector loss	1 dB	7 (final margin)
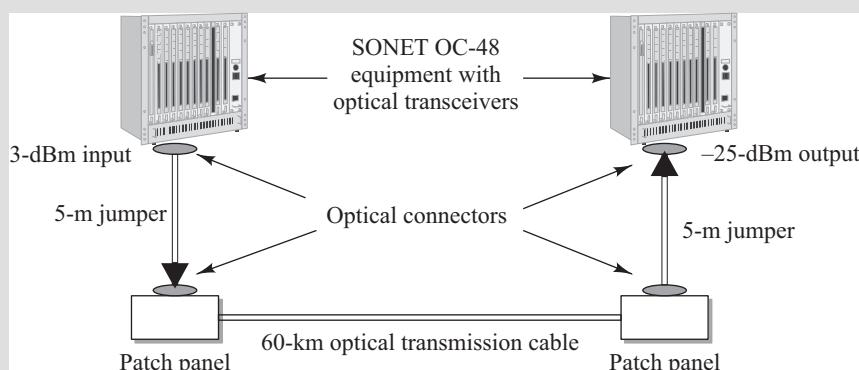
Fig. 8.5 *A 2.5-Gb/s 60-km optical fiber link with 5-m optical jumper cables at each end*

8.1.3 Rise-Time Budget

A rise-time budget analysis is a convenient method for determining the dispersion limitation of an optical fiber link. This is particularly useful for digital systems. In this approach, the total rise time t_{sys} of the link is the root sum square of the rise times from each contributor t_i to the pulse rise-time degradation:

$$t_{sys} = \left(\sum_{i=1}^{N} t_i^2 \right)^{1/2} \tag{8.3}$$

The four basic elements that may significantly limit system speed are the transmitter rise time t_{tx}, the group-velocity dispersion (GVD) rise time t_{GVD} of the fiber, the modal dispersion rise time t_{mod} of the fiber, and the receiver rise time t_{rx}. Single-mode fibers do not experience modal dispersion, so in these fibers the rise time is related only to GVD. Generally, the total transition-time degradation of a digital link should not exceed 70 percent of an NRZ (non-return-to-zero) bit period or 35 percent of a bit period for RZ (return-to-zero) data, where one bit period is defined as the reciprocal of the data rate (NRZ and RZ data formats are discussed in more detail in Sec. 4.4).

The rise times of transmitters and receivers are generally known to the designer. The transmitter rise time is attributable primarily to the light source and its drive circuitry. The receiver rise time results from the photodetector response and the 3-dB electrical bandwidth of the receiver front end. The response of the receiver front end can be modeled by a first-order lowpass filter having a step response[11–12]

$$g(t) = [1 - \exp(-2\pi B_e t)]u(t)$$

where B_e is the 3-dB electrical bandwidth of the receiver and $u(t)$ is the unit step function which is 1 for $t \geq 0$ and 0 for $t < 0$. The rise time t_{rx} of the receiver is usually defined as the time interval between $g(t) = 0.1$ and $g(t) = 0.9$. This is known as the *10- to 90-percent rise time.* Thus, if B_e is given in megahertz, then the receiver front-end rise time in nanoseconds is (see Prob. 8.3)

$$t_{rx} = \frac{350}{B_e} \tag{8.4}$$

In practice, an optical fiber link seldom consists of a uniform, continuous, jointless fiber. Instead, a transmission link nominally is formed from several concatenated (tandemly joined) fibers that may have different dispersion characteristics. This is especially true for dispersion-compensated links operating at 10 Gb/s and higher (see Chapter 13). In addition, multimode fibers experience modal distributions at fiber-to-fiber joints owing to misaligned joints, different core index profiles in each fiber, and/or different degrees of mode mixing in individual fibers. Determining the fiber rise times resulting from GVD and modal dispersion then becomes more complex than for the case of a single uniform fiber.

The fiber rise time t_{GVD} resulting from GVD over a length L can be approximated by Eq. (3.44) as

$$t_{GVD} \approx |D| L \sigma_\lambda \tag{8.5}$$

where σ_λ is the half-power spectral width of the source, and the dispersion D is given by Eq. (3.47) for a non-dispersion-shifted fiber and by Eq. (3.49) for a dispersion-shifted fiber. Since the dispersion value generally changes from fiber section to section in a long link, an average value should be used for D in Eq. (8.5).

The difficulty in predicting the bandwidth (and hence the modal rise time) of a series of concatenated multimode fibers arises from the observation that the total route bandwidth can be a function of the order in which fibers are joined. For example, instead of randomly joining together arbitrary (but very similar) fibers, an improved total link bandwidth can be obtained by selecting adjoining fibers with alternating over- and undercompensated refractive-index profiles to provide some modal delay equalization. Although the ultimate concatenated fiber bandwidth can be obtained by judiciously selecting adjoining fibers for optimum modal delay equalization, in practice this is unwieldy and time-consuming, particularly since the initial fiber in the link appears to control the final link characteristics.

A variety of empirical expressions for modal dispersion have thus been developed.[13–16] From practical field experience, it has been found that the bandwidth B_M in a link of length L can be expressed to a reasonable approximation by the empirical relation

$$B_M(L) = \frac{B_0}{L^q} \tag{8.6}$$

where the parameter q ranges between 0.5 and 1, and B_0 is the bandwidth of a 1-km length of cable. A value of $q = 0.5$ indicates that a steady-state modal equilibrium has been reached, whereas $q = 1$ indicates little mode mixing. Based on field experience, a reasonable estimate is $q = 0.7$.

Another expression that has been proposed for B_M based on curve fitting of experimental data, is

$$\frac{1}{B_M} = \left[\sum_{n=1}^{N} \left(\frac{1}{B_n} \right)^{1/q} \right]^q \tag{8.7}$$

where the parameter q ranges between 0.5 (quadrature addition) and 1.0 (linear addition), and B_n is the bandwidth of the nth fiber section. Alternatively, Eq. (8.7) can be written as

$$t_M(N) = \left[\sum_{n=1}^{N} (t_n)^{1/q} \right]^q \tag{8.8}$$

where $t_M(N)$ is the pulse broadening occurring over N cable sections in which the individual pulse broadenings are given by t_n.

We now need to find the relation between the fiber rise time and the 3-dB bandwidth. We assume that the optical power emerging from the fiber has a gaussian temporal response described by

$$g(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-t^2/2\sigma^2} \tag{8.9}$$

where σ is the rms pulse width.

The Fourier transform of this function is

$$G(\omega) = \frac{1}{\sqrt{2\pi}} e^{-\omega^2\sigma^2/2} \tag{8.10}$$

From Eq. (8.9) the time $t_{1/2}$ required for the pulse to reach its half-maximum value, that is, the time required to have

$$g(t_{1/2}) = 0.5\, g(0) \tag{8.11}$$

is given by

$$t_{1/2} = (2 \ln 2)^{1/2} \sigma \tag{8.12}$$

If we define the time t_{FWHM} as the full width of the pulse at its half-maximum value, then

$$t_{\text{FWHM}} = 2t_{1/2} = 2\sigma (2 \ln 2)^{1/2} \tag{8.13}$$

The 3-dB optical bandwidth $B_{3\text{dB}}$ is defined as the modulation frequency $f_{3\text{dB}}$ at which the received optical power has fallen to 0.5 of the zero frequency value. Thus, from Eqs. (8.10) and (8.13), we find that the relation between the full-width half-maximum rise time t_{FWHM} and the 3-dB optical bandwidth is

$$f_{3\text{dB}} = B_{3\text{dB}} = \frac{0.44}{t_{\text{FWHM}}} \tag{8.14}$$

Using Eq. (8.6) for the 3-dB optical bandwidth of the fiber link and letting t_{FWHM} be the rise time resulting from modal dispersion, then, from Eq. (8.14),

$$t_{\text{mod}} = \frac{0.44}{B_M} = \frac{0.44L^q}{B_0} \tag{8.15}$$

If t_{mod} is expressed in nanoseconds and B_{M} is given in megahertz, then

$$t_{\text{mod}} = \frac{440}{B_M} = \frac{440L^q}{B_0} \tag{8.16}$$

Substituting Eqs. (3.20), (8.4), and (8.16) into Eq. (8.3) gives a total system rise time of

$$\begin{aligned}
t_{\text{sys}} &= \left[t_{tx}^2 + t_{\text{mod}}^2 + t_{\text{GVD}}^2 + t_{rx}^2 \right]^{1/2} \\
&= \left[t_{tx}^2 + \left(\frac{440L^q}{B_0} \right)^2 + D^2 \sigma_\lambda^2 L^2 + \left(\frac{350}{B_e} \right)^2 \right]^{1/2}
\end{aligned} \tag{8.17}$$

Example 8.3   As an example of a rise-time budget for a multimode link, let us continue the analysis of the link we started to examine in Sec. 8.1.2. We assume that the LED together with its drive circuit has a rise time of 15 ns. Taking a typical LED spectral width of 40 nm, we have a material-dispersion-related rise-time degradation of 21 ns over the 6-km link. Assuming the receiver has a 25-MHz bandwidth, then from Eq. (8.4) the contribution to the rise-time degradation from the receiver is 14 ns. If the fiber we select has a 400-MHz · km bandwidth-distance product and with $q = 0.7$ in Eq. (8.6), then from Eq. (8.15)

the modal-dispersion-induced fiber rise time is 3.9 ns. Substituting all these values back into Eq. (8.17) results in a link rise time of

$$\begin{aligned}
t_{\text{sys}} &= \left(t_{tx}^2 + t_{\text{mat}}^2 + t_{\text{mod}}^2 + t_{rx}^2 \right)^{1/2} \\
&= [(15 \text{ ns})^2 + (21 \text{ ns})^2 + (3.9 \text{ ns})^2 + (14 \text{ ns})^2]^{1/2} \\
&= 30 \text{ ns}
\end{aligned}$$

This value falls below the maximum allowable 35-ns rise-time degradation for our 20-Mb/s NRZ data stream (0.70/bit rate). The choice of components was thus adequate to meet our system design criteria.

where all the times are given in nanoseconds, σ_λ is the half-power spectral width of the source, and the dispersion D [expressed in ns/(nm · km)] is given by Eq. (3.47) for a non-dispersion-shifted fiber and by Eq. (3.49) for a dispersion-shifted fiber. As indicated by the curves in Fig. 3.28 for G.652 single-mode fiber, the dispersion D is less than +3.5 ps/(nm · km) in the O-band and about +17 ps/(nm · km) at 1550 nm. For G.655 fiber the dispersion values range from −10 to −3 ps/(nm · km) across the O-band and from +5 to +10 ps/(nm-km) in the C-band.

Analogous to power budget calculations, a convenient procedure for keeping track of the various rise-time values in the rise-time budget is to use a tabular or spreadsheet form. We illustrate this by way of an example for the SONET OC-48 (2.5 Gb/s) link we looked at in Example 8.2.

Example 8.4 Assume that the laser diode together with its drive circuit has a rise time of 0.025 ns (25 ps). Taking a 1550-nm laser diode spectral width of 0.1 nm and an average dispersion of 2 ps/(nm · km) for the fiber, we have a GVD-related rise-time degradation of 12 ps (0.012 ns) over a 60-km long optical cable. Assuming the InGaAs-APD-based receiver has a 2.5-GHz bandwidth, then from Eq. (8.4) the receiver rise time is 0.14 ns. Using Eq. (8.17) to add up the various contributions, we have a total rise time of 0.14 ns.

Table 8.2 lists the components in column 1 and the associated rise times in column 2. Column 3 gives the allowed system rise-time budget of 0.28 ns for a 2.5-Gb/s NRZ data stream at the top. This is found from the expression $0.7/B_{\text{NRZ}}$ where B_{NRZ} is the bit rate for the NRZ signal. The calculated system rise time of 0.14 ns is shown at the bottom. The system rise time, in this case, is dominated by the receiver and is well within the required limits.

Table 8.2 *Example for a tabular form for keeping track of component contributions to an optical-link rise-time budget*

Component	Rise time	Rise-time budget
Allowed rise-time budget		$t_{\text{sys}} = 0.7/B_{\text{NRZ}} = 0.28$ ns
Laser transmitter	25 ps	
GVD in fiber	12 ps	
Receiver rise time	0.14 ns	
System rise time [Eq. (8.17)]		0.14 ns

8.1.4 Short-Wavelength Band

Figure 8.6 shows the attenuation and dispersion limitation on the repeaterless transmission distance as a function of data rate for the short-wavelength (770–910-nm) LED/*pin* combination. The BER was taken as 10^{-9} for all data rates. The fiber-coupled LED output power was assumed to be a constant −13 dBm for all data rates up to 200 Mb/s. The attenuation limit curve was then derived by using a fiber loss of 3.5 dB/km and the receiver sensitivities shown in Fig. 8.3. Since the minimum optical power required at the receiver for a given BER becomes higher for increasing data rates, the attenuation limit curve slopes downward to the right. We have also included a 1-dB connector-coupling loss at each end and a 6-dB system operating margin.

The dispersion limit depends on material and modal dispersion. Material dispersion at 800 nm is taken as 0.07 ns/(nm · km) or 3.5 ns/km for an LED with a 50-nm spectral width. The curve shown is the material dispersion limit in the absence of modal dispersion. This limit was taken to be the distance
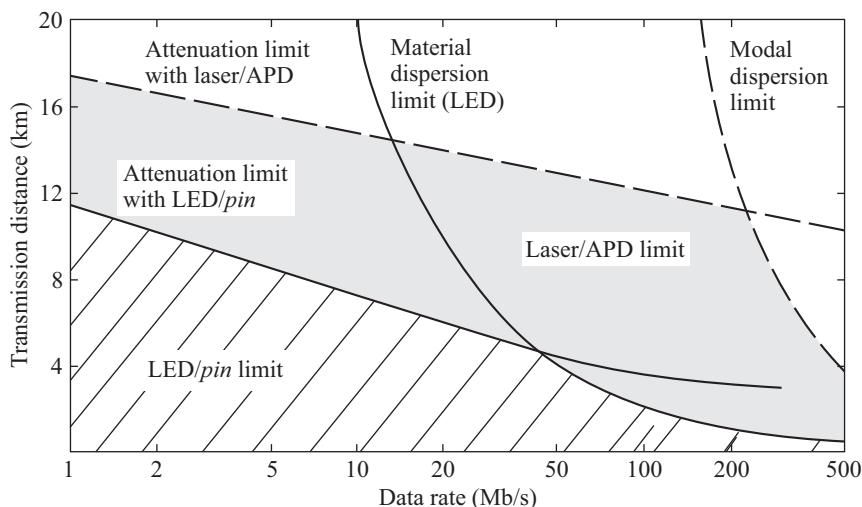
Fig. 8.6 *Transmission-distance limits as a function of data rate for an 800-MHz · km fiber, a combination of an 800-nm LED source with a Si pin photodiode, and an 850-nm laser diode with a Si APD*

at which t_{mat} is 70 percent of a bit period. The modal dispersion was derived from Eq. (8.15) for a fiber with an 800-MHz·km bandwidth–distance product and with $q = 0.7$. The modal dispersion limit was then taken to be the distance at which t_{mod} is 70 percent of a bit period. The achievable transmission distances are those that fall below the attenuation limit curve and to the left of the dispersion line, as indicated by the hatched area. The transmission distance is attenuation-limited up to about 40 Mb/s, after which it becomes material-dispersion-limited.

Greater transmission distances are possible when a laser diode is used in conjunction with an avalanche photodiode. Let us consider an AlGaAs laser emitting at 850 nm with a spectral width of 1 nm that couples 0 dBm (1 mW) into a fiber flylead. The receiver uses an APD with a sensitivity depicted in Fig. 8.3. The fiber is the same as described in Sec. 8.1.4. In this case, the material-dispersion-limit curve lies off the graph to the right of the modal-dispersion-limit curve, and the attenuation limit (with an 8-dB system margin) is as shown in Fig. 8.6. The achievable transmission distances now include those indicated by the shaded area.

8.1.5 Attenuation-Limited Distances for Single-Mode Links

For single-mode links there is no modal dispersion. In this case, in addition to attenuation factors, the repeaterless transmission distance is limited by dispersion arising from the source spectral width, from polarization mode dispersion, and from nonlinear effects in the fiber. In this section we examine the limits that signal attenuation imposes on repeaterless transmission distances. The transmission limits due to chromatic and polarization-mode dispersions are described in Sec. 8.2. In this chapter we will assume that the optical power launched into the fiber is no more than 0 dBm (1 mW), so that there are negligible nonlinear effects on the optical signals. Discussions on signal distortion arising from the nonlinear effects resulting from high optical powers in fibers are given in Chapter 12.

To illustrate the attenuation-limited repeaterless transmission limits, let us examine two single-mode links operating at 1550 nm based on using the *pin* and APD receivers described for Fig. 7.11.

The component and performance characteristics for the two links are taken to be as follows:
1. The optical source is a DFB laser that has a fiber-coupled output of 0 dBm at 1550 nm.
2. At 1550 nm the single-mode fiber has a 0.20-dB/km attenuation.
3. Consider the receiver to have a load resistor $R_L = 200$ Ω and let the temperature be 300°K.
4. The performances of the two links are measured at a 10^{-12} BER, so that a value of $Q = 7$ is needed.
5. The InGaAs *pin* and APD photodiodes have a responsivity of 0.95 A/W. The gain of the APD is $M = 10$ and the noise figure $F(M) = 5$ dB.

Example 8.5 What are the attenuation-limited repeaterless transmission distances?

(*a*) From the receiver sensitivity curves shown in Fig. 7.11, we can deduce that for an InGaAs *pin* photodiode operating at 1550 nm with a 10^{-12} BER, the receiver sensitivity can be approximated by the straight-line equation $P_R = 8 \log B - 28$ dBm, where B is the data rate in Gb/s. To find the attenuation-limited repeaterless transmission distance L_{pin}, we use Eq. (8.2) with a combined connector loss plus system margin of 3 dB, so that

$$L_{pin} = (P_S - P_R - 3 \text{ dB})/\alpha$$
$$= (0 \text{ dBm} - 8 \log B + 28 \text{ dB} - 3 \text{ dB})/\alpha$$
$$= (- 8 \log B + 25)/0.2$$

(*b*) Similarly, from the receiver sensitivity curves shown in Fig. 7.11, for the InGaAs APD the receiver sensitivity can be approximated by the straight-line equation $P_R = 5 \log B - 38$ dBm, where B is the data rate in Gb/s. Again, we use Eq. (8.2) with a combined connector loss plus system margin of 3 dB, so that attenuation-limited repeaterless transmission distance L_{pin} when using an APD is

$$L_{APD} = (P_S - P_R - 3 \text{ dB})/\alpha$$
$$= (0 \text{ dBm} - 5 \log B + 38 \text{ dB} - 3 \text{ dB})/\alpha$$
$$= (- 5 \log B + 35)/0.2$$

The results for the attenuation-limited repeaterless transmission distances L_{pin} and L_{APD} are plotted in Fig. 8.7.
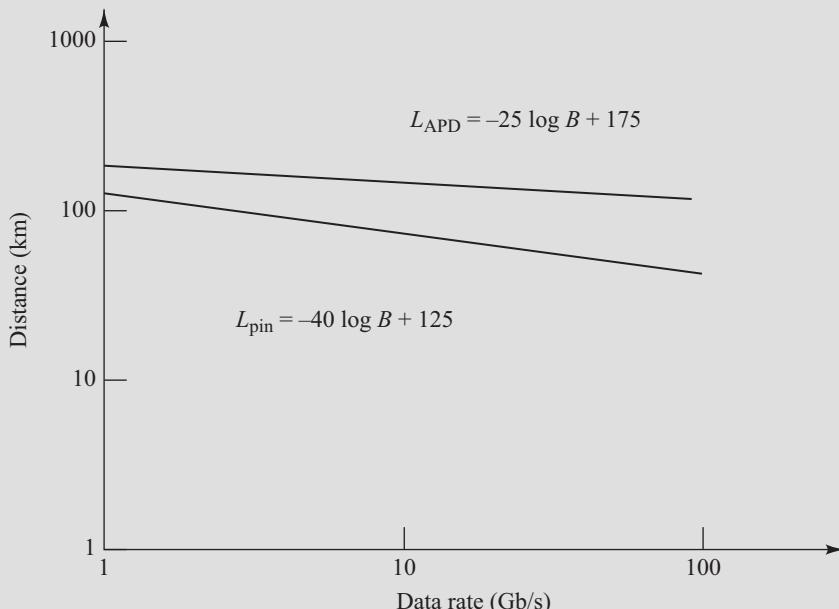
Fig. 8.7 *Transmission-distance limits due to attenuation as a function of data rate for 1550-nm laser diodes with 0-dBm fiber-coupled power, InGaAs pin and APD photodiodes, and a single-mode fiber with a 0.2-dB/km attenuation*

8.2 Power Penalties

The analysis in Sec. 8.1 assumed that the optical power falling on the photodetector is a clearly defined function of time within the statistical nature of the quantum detection process. In reality, a number of signal impairments that are inherent in optical fiber transmission systems can degrade the link performance.

When any signal impairments are present in a link, a lower optical power level arrives at the receiver compared to the ideal reception case. This lower power results in a reduced signal-to-noise ratio of the link compared to the case when there are no impairments. Since a reduced SNR leads to a higher bit-error rate (BER), a higher signal power is required at the receiver in order to maintain the same BER as in the ideal case. The ratio of the reduced received signal power to the ideal received power is known as the *power penalty* for that effect and generally is expressed in decibels. If P_{ideal} and P_{impair} are the received optical powers for the ideal and impaired cases, respectively, then the power penalty PP_x in decibels for impairment condition x is given by

$$PP_x = -10 \log \frac{P_{\text{impair}}}{P_{\text{ideal}}} \qquad (8.18)$$

In some cases one can increase the optical power level at the receiver to reduce the power penalty. For other situations, for example for some nonlinear effects described in Chapter 12, increasing the power level will have no effect on the power penalty. The main power penalties are due to chromatic and polarization-mode dispersions, modal or speckle noise, mode-partition noise, the extinction ratio, wavelength chirp, timing jitter, optical reflection noise, and nonlinear effects that arise when there is a high optical power level in a fiber link. Modal noise is present only in multimode links, but all the other effects can be serious in single-mode links. This section addresses all these performance impairments except nonlinear effects, which Chapter 12 describes. Additional power penalties due to optical amplifiers and WDM-channel crosstalk are given in Chapters 11 and 12, respectively.

8.2.1 Chromatic Dispersion Penalty

Chromatic dispersion originates from the fact that each wavelength travels at a slightly different velocity in a fiber, and thus they arrive at different times at the fiber end. Therefore, the range of arrival times at the fiber end of the spectrum of wavelengths will lead to pulse spreading. As noted in Sec. 3.3, chromatic dispersion is a fixed quantity at a specific wavelength and is measured in units of ps/(nm·km). Figure 3.28 shows the chromatic dispersion behavior as a function of wavelength for several different standard single-mode fiber types. For example, a G.652 single-mode fiber typically has a chromatic dispersion value of $D_{CD} = 18$ ps/(nm · km) at 1550 nm.

The accumulated chromatic dispersion increases with distance along a link. Therefore, either a transmission system has to be designed to tolerate the total dispersion, or some type of dispersion compensation method has to be employed.[17–26] A basic estimate of what limitation chromatic dispersion imposes on link performance can be made by specifying that the accumulated dispersion should be less than a fraction ε of the bit period $T_b = 1/B$, where B is the bit rate. This gives the relationship $|D_{CD}|L\,\sigma_\lambda < \varepsilon T_b$, or equivalently,

$$|D_{CD}|\, L\, B\, \sigma_\lambda < \varepsilon \qquad (8.19)$$

The ITU-T Recommendation G.957 for SDH and the *Telcordia Generic Requirement* GR-253 for SONET specify that for a 1-dB power penalty the accumulated dispersion should be less than 0.306 of a bit period.[23,24] For a 2-dB power penalty the requirement is $\varepsilon = 0.491$.

Example 8.6   What are the dispersion-limited repeaterless transmission distances L_{CD} at 1550 nm as a function of the bit rate in a G.652 single-mode fiber for the following three cases? Let the chromatic dispersion be $D_{CD} = 18$ ps/(nm · km) at 1550 nm.

(*a*) A directly modulated laser source with a $\sigma_\lambda = 1.0$-nm spectral width

(*b*) A directly modulated laser source with a $\sigma_\lambda = 0.2$-nm spectral width

(*c*) An externally modulated single-longitudinal-mode (SLM) DFB laser source with a spectral width that corresponds to the modulation bandwidth

Solution: Here we select an NRZ data format and choose the criterion that for the maximum allowed pulse dispersion with a 2-dB penalty the product $D_{CD}\sigma_\lambda L_{CD}$ is less than or equal to 0.491 of the bit period $1/B$. Thus we need to have the condition $D_{CD}BL_{CD}\sigma_\lambda \le 0.491$.

(*a*) Solving for the bit-rate distance product yields, where B is the data rate in Gb/s,

$$B \cdot L_{CD} \le \frac{0.491}{D_{CD}\sigma_\lambda} = \frac{0.491}{[(18 \text{ ps/nm} \cdot \text{km}) \times 1 \text{ nm}]}$$

$$= 27 \text{ Gb/s} \cdot \text{km}$$

This spectral width imposes a severe limitation on the transmission distance. The chromatic dispersion-limited repeaterless transmission distances L_{CD} are plotted as a function of the bit rate in Fig. 8.8.

(*b*) Solving for the bit-rate distance product yields, where B is the data rate in Gb/s,

$$B \cdot L_{CD} \le \frac{0.491}{D_{CD}\sigma_\lambda} = \frac{0.491}{[(18 \text{ ps/nm} \cdot \text{km}) \times 0.2 \text{ nm}]}$$

$$= 135 \text{ Gb/s} \cdot \text{km}$$

Narrowing the spectral width to 0.2 nm shows some improvement in the transmission distance, but it is still not adequate for high-speed long-distance optical communication systems. The chromatic dispersion-limited repeaterless transmission distances L_{CD} for this case also are plotted as a function of the bit rate in Fig. 8.8.

(*c*) When external modulation is used, the spectral width of the signal is proportional to the bit rate. For example, using a factor of $\Delta f = B$, a 10-Gb/s externally modulated signal would have a spectral width of $\Delta f = 10$ GHz. To view this particular spectral width in terms of wavelength, we differentiate the basic equation $c = f\lambda$ to get $\Delta\lambda = (c/f^2)\Delta f = (\lambda^2/c) \Delta f$. Substituting $\sigma_\lambda = \Delta\lambda = (\lambda^2/c)B$ into Eq. (8.19) yields (for a 2-dB power penalty)

$$D_{CD}B^2 L_{CD}\lambda^2/c \le 0.491$$

Using the parameter values $D_{CD} = 18$ ps/(nm · km) at $\lambda = 1550$ nm yields

$$B^2 L_{CD} \le 3406 \text{ (Gb/s)}^2 \cdot \text{km}$$

Thus, at $B = 2.5$ Gb/s the transmission distance limit is 545 km, whereas at $B = 10$ Gb/s the transmission distance limit is 34 km when using a G.652 fiber at 1550 nm. This condition for the chromatic dispersion-limited repeaterless transmission distances L_{CD} is plotted as a function of the bit rate in Fig. 8.8.

Shifting the operating wavelength to 1310 nm where $D_{CD} \approx 6$ ps/(nm·km) will increase the maximum transmission distance for a 10-Gb/s data rate to about 100 km. However, since the fiber attenuation is larger at 1310 nm than at 1550 nm, operation at 1310 nm may become attenuation limited.

Several methods have been examined to mitigate the effects of chromatic dispersion-induced intersymbol interference. First dispersion-shifted fibers were developed to reduce the value of D_{CD} at 1550 nm. Although this is useful for links carrying a single wavelength, these fibers are not suitable for WDM systems (see Chapters 10 and 13) owing to nonlinear crosstalk between different wavelength channels.

Fig. 8.8   *Chromatic dispersion limits for two different chromatic dispersion values and two different source spectral widths*

A more successful method to overcome the dispersion limit is by means of dispersion compensation. This is done via dispersion-compensating modules (DCM) in which the dispersion has the opposite sign of that in the transmission fiber.[17–19] Through a proper design of such modules, the overall accumulated dispersion in a lightwave transmission system can be reduced to an acceptable level.

8.2.2   Polarization-Mode Dispersion Penalty

As Sec. 3.2 describes, *polarization-mode dispersion* (PMD) results from the fact that light-signal energy at a given wavelength in a single-mode fiber actually occupies two orthogonal polarization states or modes. Figure 3.17 shows this condition. PMD arises because the two fundamental orthogonal polarization modes travel at slightly different speeds owing to fiber birefringence. The resulting difference in propagation times between the two orthogonal polarization modes will result in pulse spreading. This PMD effect cannot be mitigated easily and can be a very serious impediment for links operating at 10 Gb/s and higher.

PMD is not a fixed quantity but fluctuates with time due to factors such as temperature variations and stress changes on the fiber.[26–32] Since these external stresses vary slowly with time, the resulting PMD also fluctuates slowly. PMD varies as the square root of distance and thus is specified as a maximum value in units of $ps/\sqrt{km}$. A typical PMD value for a fiber is $D_{PMD} = 0.05\ ps/\sqrt{km}$, but the cabling process can increase this value. The PMD value does not fluctuate widely for cables that are enclosed in underground ducts or in buildings. However it can increase periodically to over $1 ps/\sqrt{km}$ for outside cables that are suspended on poles, since such cables are subject to wide variations in temperature, wind-induced stresses, and elongations caused by ice loading.

To have a power penalty of less than 1.0 dB, the pulse spreading $\Delta\tau_{\text{PMD}}$ resulting from polarization mode dispersion must on the average be less than 10 percent of a bit period T_b. Using Eq. (3.40) this condition is given by

$$\Delta\tau_{\text{PMD}} = D_{\text{PMD}} \sqrt{L} < 0.1\, T_b \tag{8.20}$$

Example 8.7 Consider a 100-km long fiber for which $D_{\text{PMD}} = 0.5 \text{ ps}/\sqrt{\text{km}}$. What is the maximum possible data rate for an NRZ-encoded signal if the pulse spread can be no more than 10 percent of a pulse width?

Solution: From Eq. (8.20) the pulse spread over the 100-km distance is $\Delta\tau_{\text{PMD}} = 5.0$ ps. Since this pulse spread can be no more than 10 percent of a pulse width, we have

$$\Delta\tau_{\text{PMD}} = 5.0 \text{ ps} \leq 0.1 T_b$$

Therefore the maximum NRZ bit rate is $1/T_b = 0.1/(5 \text{ ps}) = 20$ Gb/s.

8.2.3 Extinction Ratio Penalty

The *extinction ratio* r_e in a laser is defined as the ratio of the optical power level P_1 for a logic 1 to the power level P_0 for a logic 0, that is, $r_e = P_1/P_0$. Ideally one would like the extinction ratio to be infinite, so that there would be no power penalty from this condition. In this case, if P_{ave} is the average power, then $P_0 = 0$ and $P_1 = 2P_{\text{ave}} = P_{\text{ideal}}$. However, the extinction ratio must be finite in an actual system in order to reduce the rise time of laser pulses.

Letting $P_{1\text{-}ER}$ and $P_{0\text{-}ER}$ be the 1 and 0 power levels, respectively, with a nonzero extinction ratio, and defining $r_e = P_{1\text{-}ER}/P_{0\text{-}ER}$, the average power is

$$P_{\text{ave}} = \frac{P_{1-ER} + P_{0-ER}}{2} = P_{0-ER}\,\frac{r_e + 1}{2} = P_{1-ER}\,\frac{r_e + 1}{2r_e} \tag{8.21}$$

When receiver thermal noise dominates, then the 1 and 0 noise powers are equal and independent of the signal level. In this case, letting $P_0 = 0$ and $P_1 = 2P_{\text{ave}}$, the power penalty given by Eq. (8.18) becomes

$$PP_{ER} = -10\log\frac{P_{1-ER} - P_{0-ER}}{P_1} = -10\log\frac{r_e - 1}{r_e + 1} \tag{8.22}$$

In practice, optical transmitters have minimum extinction ratios ranging from 7 to 10 (8.5 to 10 dB), for which the power penalties range from 1.25 to 0.87 dB. A minimum extinction ratio of 18 is needed in order to have a power penalty of less than 0.5 dB. Note that the power penalty increases significantly for lower extinction ratios.

8.2.4 Modal Noise

When light from a coherent laser is launched into a multimode fiber, normally a number of propagating modes of the fiber are excited.[14,33–38] As long as these modes retain their relative phase coherence, the radiation pattern seen at the end of the fiber (or at any point along the fiber) takes on the form of a speckle pattern. This is the result of constructive and destructive interference between propagating modes at any given plane. An example of this is shown in Fig. 8.9. The number of speckles in the pattern approximates
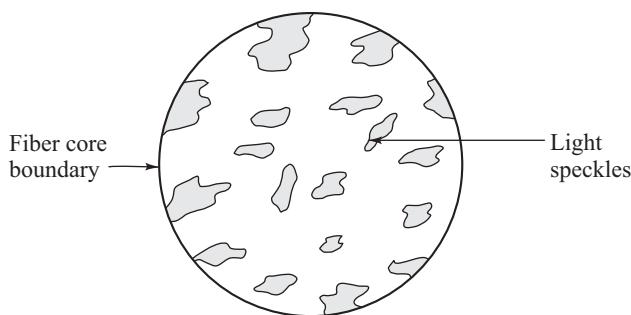
Fig. 8.9 *Example of a speckle pattern that is produced when coherent laser light is launched into a multimode fiber*

the number of propagating modes. As the light travels along the fiber, a combination of mode-dependent losses, changes in phase between modes, and fluctuations in the distribution of energy among the various fiber modes will change the modal interference and result in a different speckle pattern. *Modal* or *speckle noise* occurs when any losses that are speckle-pattern dependent are present in a link. Examples of such losses are splices, connectors, microbends, and photodetectors with nonuniform responsivity across the photosensitive area. Noise is generated when the speckle pattern *changes in time* so as to vary the optical power transmitted through the particular loss element. The continually changing speckle pattern that falls on the photodetector thus produces a time-varying noise in the received signal, which degrades receiver performance.

Fluctuations in the frequency of an optical source also can give rise to intermodal delays. A coherent source forms speckle patterns when its coherence time is greater than the intermodal dispersion time δT within the fiber. If the source has a frequency width δv, then its coherence time is $1/\delta v$. Modal noise occurs when the speckle pattern fluctuates; that is, when the source coherence time becomes much less than the intermodal dispersion time. The modal distortion resulting from interference between a single pair of modes will appear as a sinusoidal ripple of frequency

$$v = \delta T \frac{dv_{\text{source}}}{dt} \tag{8.23}$$

where dv_{source}/dt is the rate of change of optical frequency.

Figure 8.10 illustrates the error rates with the addition of modal noise to an avalanche-photodiode receiver system.[37] The analysis is for 280 Mb/s at 1200 nm with a gaussian-shaped received pulse. The factor M' in this figure is related to the number of speckles falling on the photodetector. For a very large number of speckles $(M' = 2910)$, the error-rate curve is very close to the case when there is no modal noise. As the number of speckles decreases, the performance degrades. When $M' = 50$, one needs an additional 1.0 dB of received optical power to maintain an error rate of 10^{-6}. When $M' = 20$, one must have 2.0 dB more power to achieve a 10^{-6} BER than in the case of no modal noise. This number becomes 4.9 dB when $M' = 4$.

The performance of a high-speed, laser-based multimode fiber link is difficult to predict because the degree of modal noise that can appear depends greatly on the particular installation. Thus the best policy is to take steps to avoid it. This can be done by the following measures:

1. Use LEDs (which are incoherent sources). This totally avoids modal noise.
2. Use a laser that has a large number of longitudinal modes (10 or more). This increases the graininess of the speckle pattern, thus reducing intensity fluctuations at mechanical disruptions in the link.
3. Use a fiber with a large numerical aperture because it supports a large number of modes and hence gives a greater number of speckles.
4. Use a single-mode fiber because it supports only one mode and thus has no modal interference.

Fig. 8.10 *Error-rate curves for a 280-Mb/s avalanche-photodiode-based system with the addition of modal noise. The factor M′ corresponds to the number of speckles. (Reproduced with permission from Chan and Tjhung,[37] ©1989, IEEE.)*

The last point needs some further explanation. If a connector or splice point couples some of the optical power from the fundamental mode into the first higher-order mode (the LP_{11} mode), then a significant amount of power could exist in the LP_{11} mode in a short section of fiber between two connectors or at a repair splice.[34,38] Figure 8.11 illustrates this effect. In a single-mode system, modal noise could occur in short connectorized patch-cords, in laser diode flyleads, or when two high-loss splices are a very short distance apart. To circumvent this problem, one should specify that the effective cutoff wavelength of short patch-cord and flylead fiber lengths is well below the system operating wavelength. Thus, mode coupling is not a problem in links that have long fiber lengths between connectors and splices, since the LP_{11} mode is usually sufficiently attenuated over the link length.
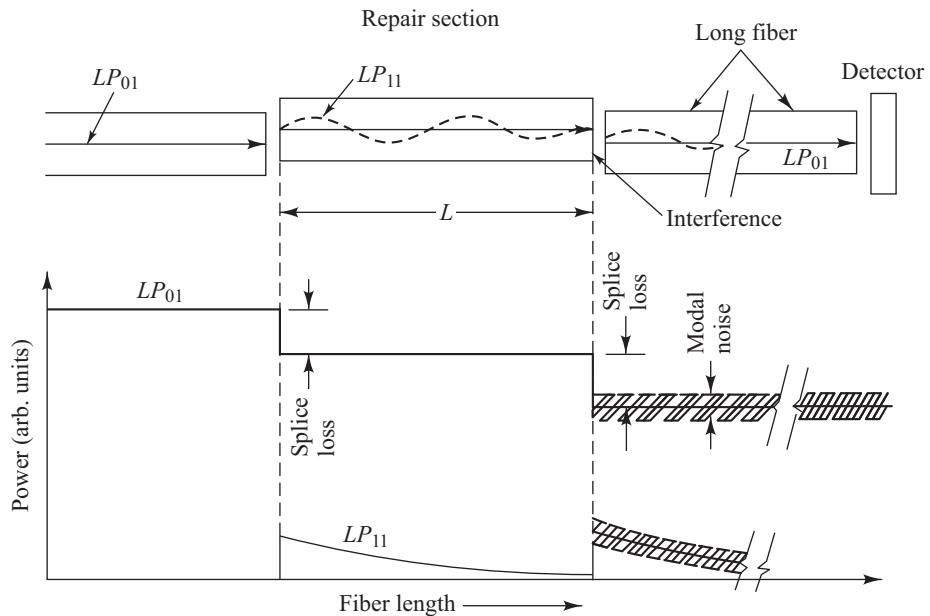
Fig. 8.11 *Repair sections can produce modal noise in a single-mode fiber link. This arises through the interchange of optical power between the LP_{01} and the LP_{11} modes at splice or connector joints. (Reproduced with permission from Sears, White, Kummer, and Stone,[34] © 1986, IEEE.)*

8.2.5 Mode-Partition Noise

Mode-partition noise is associated with intensity fluctuations in the longitudinal modes of a multimode laser diode;[39–46] that is, the side modes are not sufficiently suppressed. This is the dominant noise in single-mode fibers when using multimode devices, such as *FP* lasers. Intensity fluctuations can occur among the various modes in a multimode laser even when the total optical output is constant, as exhibited in Fig. 8.12. This power distribution can vary significantly both within a pulse and from pulse to pulse.

Because the output pattern of a laser diode is highly directional, the light from these fluctuating modes can be coupled into a single-mode fiber with high efficiency. Each of the longitudinal modes that is coupled into the fiber has a different attenuation and time delay because each is associated with a slightly different wavelength. Since the power fluctuations among the dominant modes can be quite large, significant variations in signal levels can occur at the receiver in systems with high fiber dispersion.

The signal-to-noise ratio due to mode-partition noise is independent of signal power, so that the overall system error rate cannot be improved beyond the limit set by this noise. This is an important difference from the degradation of receiver sensitivity normally associated with chromatic dispersion, which one can compensate for by increasing the signal power.

The power penalty in decibels caused by laser mode-partition noise can be approximated by[44]

$$PP_{mpn} = -5\frac{x+2}{x+1}\log\left[1 - \frac{k^2 Q^2}{2}\left(\pi BLD\sigma_\lambda\right)^4\right] \tag{8.24}$$

where x is the excess noise factor of an APD, Q is the signal-to-noise factor (see Fig. 7.9), B is the bit rate in Gb/s, L is the fiber length in km, D is the fiber chromatic dispersion in ps/(nm · km), σ_λ is the rms spectral width of the source in nm, and k is the mode-partition noise factor. The parameter k is difficult to quantify since it can vary from 0 to 1 depending on the laser. However, experimental values of k range from 0.6 to 0.8. To keep the power penalty less than 0.5 dB, a well-designed system should have the quantity $BLD\sigma_\lambda < 0.1$.

Mode-partition noise becomes more pronounced for higher bit rates. The errors due to mode-partition noise can be reduced and sometimes eliminated by setting the bias point of the laser above threshold. However, raising the bias power level reduces the available signal-pulse power, thereby reducing the achievable signal-to-thermal-noise ratio.

8.2.6 Chirping

A laser that oscillates in a single longitudinal mode under CW operation may experience dynamic line broadening when the injection current is directly modulated.[47–57] This line broadening is a frequency "chirp" associated with modulation-induced changes in the carrier density. Laser chirping can lead to significant dispersion effects

Fig. 8.12 *Time-resolved dynamic spectra of a laser diode. Different modes or groups of modes dominate the optical output at different times. The modes are approximately 1 nm apart.*

for intensity-modulated pulses when the laser emission wavelength is displaced from the zero-dispersion wavelength of the fiber. This is particularly true in systems operating at 1550 nm, where dispersion in G.652 non-dispersion-shifted fibers is much greater than at 1300 nm.

To a good approximation, the time-dependent frequency change $\Delta v(t)$ of the laser can be given in terms of the output optical power $P(t)$ as[51]

$$\Delta v(t) = \frac{-\alpha}{4\pi}\left[\frac{d}{dt}\ln P(t) + \kappa P(t)\right] \tag{8.25}$$

where α is the *linewidth enhancement factor*[54] and κ is a frequency-independent factor that depends on the laser structure.[51] The factor α ranges from −3.5 to −5.5 for AlGaAs lasers[55] and from −6 to −8 for InGaAsP lasers.[56]

One approach to minimize chirp is to increase the bias level of the laser so that the modulation current does not drive it below threshold where $\ln P$ and P change rapidly. However, this results in a lower extinction ratio (the ratio of on-state power to off-state power), which leads to an extinction-ratio power penalty at the receiver because of a reduced signal-to-background noise ratio. This penalty could be several decibels. Figure 8.13 gives examples of this for two types of laser structures. For higher extinction ratios (bias points progressively

Fig. 8.13 *Extinction-ratio, chirping, and total-system power penalties at 1550 nm for a 4-Gb/s 100-km-long single-mode link having a fiber dispersion D = 17 ps/(nm · km) and a DFB laser with an active layer width of 1.75 μm. (Reproduced with permission from Corvini and Koch,[53] © 1987, IEEE.)*

lower than threshold), the extinction-ratio power penalty decreases. However, the chirping-induced power penalty increases with lower bias levels.

When the effect of laser chirp is small, the eye closure Δ can be approximated by[50]

$$\Delta = \left(\tfrac{4}{3}\pi^2 - 8\right)t_{chirp}DLB^2\delta\lambda\left[1 + \tfrac{2}{3}(DL\delta\lambda - t_{chirp})\right] \tag{8.26}$$

where t_{chirp} is the chirp duration, B is the bit rate, D is the fiber chromatic dispersion, L is the fiber length, and $\delta\lambda$ is the chirp-induced wavelength excursion.

The power penalty for an APD system can be estimated from the signal-to-noise ratio degradation (in dB) due to the signal amplitude decrease as

$$PP_{chirp} = -10\frac{x+2}{x+1}\log\left(1 - \Delta\right) \tag{8.27}$$

where x is the excess noise factor of an APD.

Figure 8.14 illustrates the effects of chirping at a 5-Gb/s transmission rate in different single-mode fiber links.[47] Here, the laser side-mode suppression is greater than 30 dB, the back-reflected optical power is more than 30 dB below the transmitted signal, and the extinction ratio is about 8 dB. At 1536 nm the standard non-dispersion-shifted fiber has a dispersion $D = 17.3$ ps/(nm · km) and the dispersion-shifted
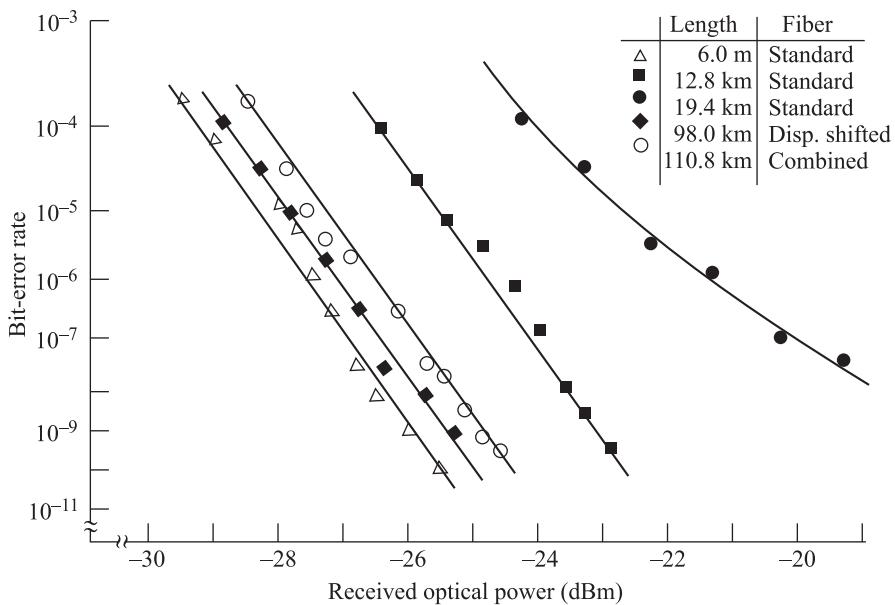
Fig. 8.14 *The effects of chirping at 5 Gb/s in different single-mode fiber links. The laser side-mode suppression is > 30 dB, the reflected power is more than 30 dB below the transmitted signal, and the extinction ratio is ≈ 8 dB. At 1536 nm the standard fiber has D= 17.3 ps/(nm · km) and the dispersion-shifted fiber has D=−1.0 ps/(nm · km). (Reproduced with permission from Heidemann,[51] © 1988, IEEE.)*

fiber has $D = -1.0 \, \text{ps/(nm} \cdot \text{km)}$. The combined-fiber link consists of concatenated standard positive-dispersion and negative-dispersion fibers. This leads to a spectral compression of the signal so that dispersion compensation occurs. Thus Fig. 8.14 shows the dramatic reduction in chirping penalty when using a dispersion-shifted fiber, or when combining fibers with positive and negative dispersion.

8.2.7 Reflection Noise

When light travels through a fiber link, some optical power gets reflected at refractive-index discontinuities such as in splices, couplers, and filters, or at air–glass interfaces in connectors. The reflected signals can degrade both transmitter and receiver performance.[58–61] In high-speed systems, this reflected power causes optical feedback, which can induce laser instabilities. These instabilities show up as either intensity noise (output power fluctuations), jitter (pulse distortion), or phase noise in the laser, and they can change its wavelength, linewidth, and threshold current. Since they reduce the signal-to-noise ratio, these effects cause two types of power penalties in receiver sensitivities. First, as shown in Fig. 8.15a, multiple reflection points set up an interferometric cavity that feeds power back into the laser cavity, thereby converting phase noise into intensity noise. A second effect created by multiple optical paths is the appearance of spurious signals arriving at the receiver with variable delays, thereby causing intersymbol interference. Figure 8.15b illustrates this.

Unfortunately, these effects are signal-dependent, so that increasing the transmitted or received optical power does not improve the bit-error rate performance. Thus one has to find ways to eliminate reflections.
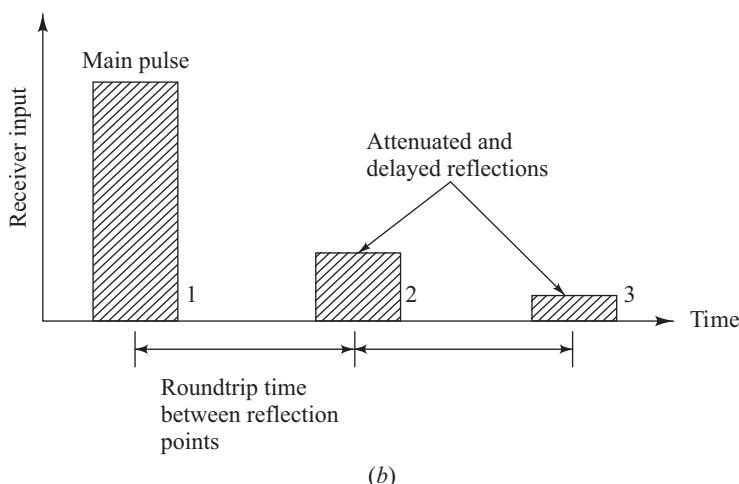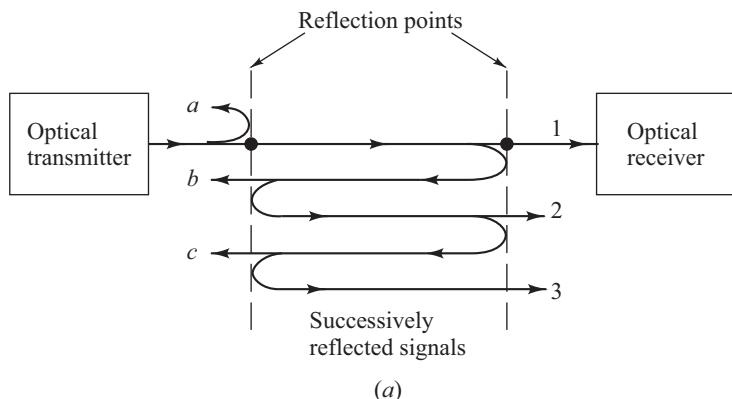
Fig. 8.15 (a) Refractive-index discontinuities can set up multiple reflections in a fiber link. (b) Each round trip of a light pulse reflection creates another attenuated and delayed pulse, which can cause intersymbol interference

Let us first look at their magnitudes. As we saw from Eq. (5.10), a cleaved silica-fiber end face in air typically will reflect about

$$R = \left(\frac{1.47 - 1.0}{1.47 + 1.0} \right)^2 = 3.6 \text{ percent}$$

This corresponds to an optical return loss of 14.4 dB down from the incident signal. Polishing the fiber ends can create a thin surface layer with an increased refractive index of about 1.6. This increases the reflectance to 5.3 percent (a 12.7-dB optical return loss). A further increase in the optical feedback level occurs when the distance between multiple reflection points equals an integral number of half-wavelengths of the transmitted wavelength. In this case, all roundtrip distances equal an integral number of in-phase wavelengths, so that

constructive interference arises. This quadruples the reflection to 14 percent or 8.5 dB for unpolished end faces and to over 22 percent (a 6.6-dB optical return loss) for polished end faces.

The power penalties can be reduced to a few tenths of a decibel by keeping the return losses below values ranging from −15 to −32 dB for data rates varying from 500 Mb/s to 4 Gb/s, respectively.[58] Techniques for reducing optical feedback include the following:

1. Prepare fiber end faces with a curved surface or an angle relative to the laser emitting facet. This directs reflected light away from the fiber axis, so it does not reenter the waveguide. Return losses of 45 dB or higher can be achieved with end-face angles of 5−15°. However, this increases both the insertion loss and the complexity of the connector.
2. Use index-matching oil or gel at air–glass interfaces. The return loss with this technique is usually greater than 30 dB. However, this may not be practical or recommended if connectors need to be remated often, since contaminants can collect on the interface.
3. Use connectors in which the end faces make physical contact (the so-called *PC connectors*). Return losses of 25–40 dB have been measured with these connectors.
4. Use optical isolators within the laser transmitter module. These devices easily achieve 25-dB return losses, but they also can introduce up to 1 dB of forward loss in the link.

8.3 | Error Control

In any digital transmission system, errors are likely to occur even when there is a sufficient signal-to-noise ratio to provide a low bit-error rate. The acceptance of a certain level of errors depends on the network user. For example, digitized speech or video can tolerate occasional high error rates. However, applications such as financial transactions require almost completely error-free transmission. In this case, the transport protocol of the network must compensate the difference between the desired and the actual bit-error rates.

To control errors and to improve the reliability of a communication line, first it is necessary to be able to detect the errors and then either to correct them or retransmit the information. Error detection methods encode the information stream to have a specific pattern. If segments in the received data stream violate this pattern, then errors have occurred. Sections 8.3.1 through 8.3.3 discuss the concept and several popular methods of error detection.

The two basic schemes for error correction are *automatic repeat request* (ARQ) and *forward error correction* (FEC).[62–68] ARQ schemes have been used for many years in applications such as computer communication links that use telephone lines and for data transmission over the Internet. As shown in Fig. 8.16, the ARQ technique uses a feedback channel between the receiver and the transmitter to request message retransmission in case errors are detected at the receiver. Since each such retransmission adds at least one roundtrip time of latency, ARQ may not be feasible for applications in which data must arrive within a certain time in order to be useful. Forward error correction avoids the shortcomings of ARQ for high-bandwidth optical networks requiring low delays. In FEC techniques, redundant information is
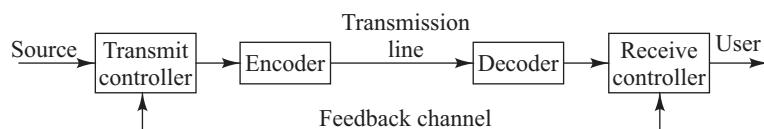
Fig. 8.16 *Basic setup for an automatic-repeat-request (ARQ) error-correction scheme*
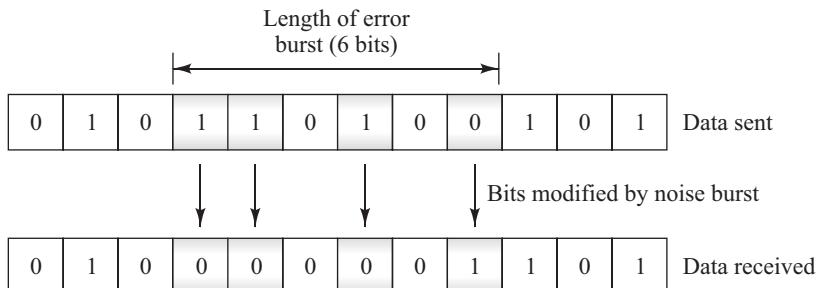
Fig. 8.17 *The length of an error burst is measured from the first to the last corrupted bit.*

transmitted along with the original information. If some of the original data is lost or received in error, the redundant information is used to reconstruct it. Section 8.3.4 gives an overview of Reed-Solomon codes used in FEC techniques.

8.3.1 Concept of Error Detection

An error in a data stream can be categorized as a single-bit error or a burst error. As its name implies, a *single-bit error* means that only one bit of a data unit (e.g., a byte, code word, a packet, or a frame) is changed from a 1 to a 0, or vice versa. Single-bit errors are not very common in a typical transmission system because most bit-corrupting noise effects last longer than a bit period.

A *burst error* refers to the fact that more than a single bit in a data unit has changed. This type of error happens most often in a typical transmission system because the duration of a noise burst lasts over several bit periods. A burst error does not necessarily change every bit in a data segment that contains errors. As shown in Fig. 8.17, the length of an error burst is measured from the first corrupted bit to the last corrupted bit. Not all the bits in this particular segment were damaged.

Example 8.8 The number of bits affected by a burst error depends on the data rate and the duration of the noise burst. If a bit-corrupting burst noise lasts for 1 ms, then 10 bits are affected for a 10-kb/s data rate, whereas a 10,000-bit segment is damaged for a 10-Mb/s rate.

The basic concept of error detection is straightforward. Prior to being inserted into a transmission channel, the information bit stream coming from a communication device is encoded so that it satisfies a certain *pattern* or a *specific set of code words*. At the destination the receiver checks the arriving information stream to verify that the pattern is satisfied. If the data stream contains segments (that is, invalid code words) that do not conform to the intended pattern, then an error has occurred in that segment.

8.3.2 Linear Error Detection Codes

The *single parity check code* is one of the simplest error detection methods. This code forms a *code word* from the combination of *k* information bits and a single added *check bit*. If the *k* information bits contain an odd number of 1 bits, then the check bit is set to 1; otherwise it is set to 0. This procedure ensures that the code word has an even number of ones, which is called having an *even parity*. Hence the check bit is called a *parity bit*. The single parity check code thus can detect when an odd number of errors has occurred in a code word. However, if the received code word contains an even number of errors, this

method will fail to detect the errors. The single parity check code is called a *linear code* because the parity bit b_{k+1} is calculated as the modulo 2 sum of the k information bits, that is,

$$b_{k+1} = b_1 + b_2 + \cdots + b_k \text{ modulo } 2 \tag{8.28}$$

where $b_1, b_2, \ldots, b_k$ are the information bits.

A more general linear code with stronger error detection capabilities is called a *binary linear code*. This linear code adds $n-k$ check bits to a group of k information bits, thereby forming a code word consisting of n bits. Such a code is designated by the notation (n, k). One example is the $(7, 4)$ linear Hamming code in which the first four bits of a code word are the information bits b_1, b_2, b_3, b_4 and the next three bits b_5, b_6, b_7 are check bits. Among the wide variety of *Hamming codes*, this particular one can detect all single and double bit errors, but fails to detect some triple errors.

8.3.3 Polynomial Codes

Polynomial codes are used widely for error detection because these codes are easy to implement using shift-register circuits. The term *polynomial code* comes from the fact that the information symbols, the code words, and the error vector are represented by polynomials with binary coefficients. Here, if a transmitted code word has n bits, then the error vector is defined by $(e_1, e_2, \ldots, e_n)$, where, $e_j = 1$ if an error has occurred in the jth transmitted bit and $e_j = 0$ otherwise. Since the encoding process generates check bits by means of a process called a *cyclic redundancy check* (CRC), a polynomial code also is known as a CRC code.

The *cyclic redundancy check* technique is based on a binary division process involving the data portion of a packet and a sequence of redundant bits. Figure 8.18 outlines the following basic CRC procedure:

- *Step 1.* At the sender end a string of n zeros is added to the data unit on which error detection will be performed. For example, this data unit may be a packet (a grouping of data plus routing and control bits). The characteristic of the redundant bits is such that the result (packet plus redundant bits) is exactly divisible by a second predetermined binary number.
- *Step 2.* The new enlarged data unit is divided by the predetermined divisor using binary division. If the number of bits added to the data unit is n, then the number of bits in the predetermined divisor is $n + 1$. The remainder that results from this division is called the *CRC remainder* or simply the CRC. The number of digits in this remainder is equal to n. For example, if $n = 3$ it may be the binary number 101. Note that the remainder also might be 000, if the two numbers are exactly divisible.
- *Step 3.* The n zeros that were added to the data unit in step 1 are replaced by the n-bit CRC. The composite data unit then is sent through the transmission channel.

Replace the n zeros with the CRC and send the composite data unit.

Accept data if remainder is zero; otherwise reject.
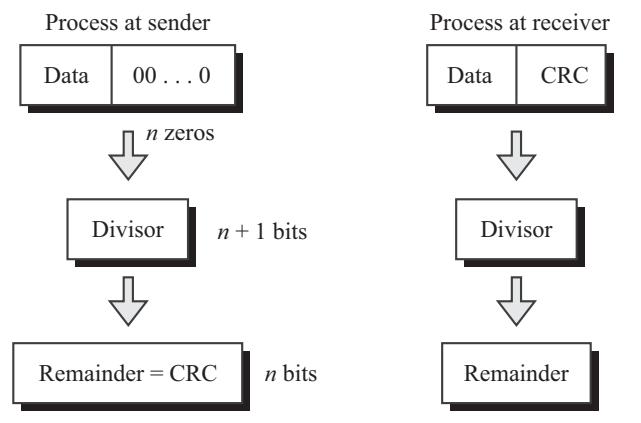
Fig. 8.18 *The basic procedure for the cyclic redundancy check (CRC) technique*

Table 8.3 *Commonly used polynomials and their binary equivalents for CRC generation*

CRC type	Generator polynomial	Binary equivalent
CRC-8	$x^8 + x^2 + x + 1$	100000111
CRC-10	$x^{10} + x^9 + x^5 + x^4 + x + 1$	11000110011
CRC-16	$x^{16} + x^{15} + x^2 + 1$	11000000000000101
CRC-32	$x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} +$ $x^{10} + x^8 + x^7 + x^5 + x^4 + x^2 + x + 1$	100000100110000010001110110110111

- *Step 4.* When the data unit plus the appended CRC arrives at the destination, the receiver divides this incoming composite unit by the same divisor that was used to generate the CRC.
- *Step 5.* If there is no remainder after this division occurs, then the assumption is that there are no errors in the data unit and it is accepted by the receiver. A remainder indicates that some bits became corrupted during the transmission process and therefore the data unit is rejected.

Instead of using a string of 1 and 0 bits, the CRC generator normally is represented by an algebraic polynomial with binary coefficients. The advantage of using a polynomial is that it is simple to visualize and perform the division mathematically. Table 8.3 shows examples of several commonly used polynomials and their binary equivalents for the CRC generation. These are designated as CRC-8, CRC-10, CRC-16, and CRC-32. The numbers 8,10,16, and 32, respectively, refer to the size of the CRC remainder. Thus the CRC divisors for these polynomials are 9,11,17, and 33 bits, respectively. The first two polynomials are used in ATM (asynchronous transfer mode) networks, whereas CRC-32 is used in IEEE-802 LANs. CRC-16 is used in bit-oriented protocols, such as the High-Level Data Link Control (HDLC) Standard, where frames are viewed as a collection of bits.

Example 8.9 The generator polynomial $x^7 + x^5 + x^2 + x + 1$ can be written as

$$1 \times x^7 + 0 \times x^6 + 1 \times x^5 + 0 \times x^4 + 0 \times x^3$$
$$+ 1 \times x^2 + 1 \times x^1 + 1 \times x^0$$

where the exponents on the variable x represent bit positions in a binary number and the coefficients correspond to the binary digits at these positions. Thus the generator polynomial given here corresponds to the 8-bit binary representation 10100111.

Example 8.10 The generator polynomial $x^3 + x + 1$ can be written in binary form as 1011. For the information unit 11110 the CRC can be found through either binary or algebraic division using steps 1 through 3 outlined earlier. Since there are 4 bits in the divisor, three 0s are added to the data for the binary arithmetic operation. Figure 8.19 shows the two different procedures using polynomial and binary arithmetic division. For the polynomial division process the remainder is $x^2 + 1$, which is equivalent to the remainder 101 found by the binary division method. The resulting composite data unit plus CRC that get transmitted is 11110101. Note that when following the binary division method, if the leftmost bit of a remainder is zero, one must use 0000 as the divisor instead of the original 1011 divisor.

A polynomial needs to have the following properties:
- It should not be divisible by *x*. This condition guarantees that the CRC can detect all burst errors that have a length less than or equal to the degree of the polynomial.
- It should be divisible by *x* + 1. This allows the CRC to detect all bursts that affect an odd number of bits.

(*a*) Polynomial division

(*b*) Binary division

Fig. 8.19 *Two different procedures for finding the CRC using polynomial and binary arithmetic divisions*

- Given these two rules, the CRC also can find with an error-detection probability

$$P_{ed} = 1 - 1/2^N \tag{8.29}$$

any burst errors that have a length greater than the degree N of the generator polynomial.

Example 8.11   The CRC-32 given in Table 8.3 has a degree of 32. Thus it will detect all burst errors affecting an odd number of bits, all burst errors with a length less than or equal to 32, and from Eq. (8.29) more than 99.99 percent of burst errors with a length of 32 or more.

8.3.4   Forward Error Correction

Error correction may be done by the use of *redundancy* in the data stream. With this method, extra bits are introduced into the raw data stream at the transmitter on a regular and logical basis and are extracted at the receiver. These digits themselves convey no information but allow the receiver to detect and correct a certain percentage of errors in the information-bearing bits. The degree of error-free transmission that

can be achieved depends on the amount of redundancy introduced. Note that the data rate that includes this redundancy must be less than or equal to the channel capacity.

The method of introducing redundant bits into the information stream at the transmitter for error-reducing purposes is called *forward error correction* (FEC). Typically the amount of added redundancy is small, so the FEC scheme does not use up much additional bandwidth and thus remains efficient. The most popular error-correcting codes are *cyclic codes*, such as *Reed-Solomon* (RS) *codes*. These codes add a redundant set of *r* symbols to blocks of *k* data symbols, with each symbol being *s* bits long, for example, $s = 8$. The codes are designated by the notation (*n, k*) where *n* equals the number of original information symbols *k* plus the number of redundant symbols *r*. For a given symbol size *s*, the maximum length of a Reed-Solomon code word is $n = 2^s - 1$.

Example 8.12 The (255,239) Reed-Solomon code with $s = 8$ (one byte) is used in high-speed undersea optical fiber links. This means that $r = n - k = 255 - 239 = 16$ redundant bytes are sent for every block of 239 information bytes. The code is quite efficient, since the 16 redundant bytes add less than 7 percent of overhead to the information stream.

A Reed-Solomon decoder can correct up to *t* symbol errors, where $2t = n - k$. For example, the (255,239) RS code can correct up to 8 errors in a block of 239 bytes. One symbol error occurs when one or more bits in a symbol are wrong. Thus the number of bits that are corrected depends on the distribution of the errors. If each incorrect byte contains only one bit error, then the (255,239) RS code will correct 8 bit errors. At the other extreme, if all the bits in each of the 8 incorrect bytes are corrupted, then the (255,239) code will correct $8 \times 8 = 64$ bit errors. Thus, a key feature of RS codes is their ability to correct burst errors, where a sequence of bytes is received incorrectly.

Another advantage of a Reed-Solomon code is that it allows transmission at a lower power level to achieve the same BER that would result without encoding. The resulting power saving is called the *coding gain*. The (255,239) RS code provides about a 6-dB coding gain. Concatenated Reed-Solomon codes (several codes used sequentially) can provide even higher coding gains.

Current terrestrial and undersea high-speed optical communication systems use a number of different FEC codes. For example, as part of the G.709 Digital Wrapper Recommendation, the ITU-T has selected the (255,239) and (255,223) Reed-Solomon codes.[69–71] The (255,223) code has a higher overhead (15 percent) compared to the (255,239) code, but is somewhat stronger since it is able to correct 16 errors in a block of 223 bits. The Digital Wrapper uses the same error-monitoring techniques as is employed in SDH and SONET. The performance metrics that are calculated include code violations in the incoming bit stream, the number of seconds in which at least one error occurs, the number of seconds in which multiple errors occur (called *severely errored seconds*), and the total number of seconds in which service is not available.

8.4 Coherent Detection

The basic receiver analysis in Chapter 7 considered a simple and cost-effective lightwave transmission scheme in which the light intensity of the optical source is modulated linearly with respect to the input electrical signal voltage. This scheme pays no attention to the frequency or phase of the optical carrier, since a photodetector at the receiving end only responds to changes in the power level (intensity) that falls directly on it. The photodetector then transforms the optical power level variations back to the original electrical signal format. This method is known as *intensity modulation with direct detection*

(IM/DD). In this respect a direct-detection optical system is analogous to the way a primitive crystal radio detects broadcast signals. Although these IM/DD methods offer system simplicity and relatively low cost, their sensitivities are limited by noise generated in the photodetector and the receiver preamplifier. These noises degrade the receiver sensitivities of square-law IM/DD transmission systems by 10 to 20 dB from the fundamental quantum noise limit.

Around 1978 component researchers had improved the spectral purity and frequency stability of semiconductor lasers to the point where alternative techniques using homodyne or heterodyne detection of the optical signal appeared to be feasible. Optical communication systems that use homodyne or heterodyne detection are called *coherent optical communication systems*, since their implementation depends on *phase coherence* of the optical carrier. In coherent detection techniques the light is treated as a carrier medium that can be amplitude-, frequency-, or phase-modulated similar to the methods used in microwave radio systems.[72–77]

Coherent systems were examined extensively during the 1980s and early 1990s as a method for increasing the transmission spans for long-haul links. However, interest in these methods declined when optical amplifiers were introduced because these devices offered dramatic increases in the transmission distances of multi-wavelength OOK-based systems. Fortunately research on coherent techniques continued, since a decade later there was renewed interest as data transmission speeds moved to 10 Gb/s and beyond. This interest was spurred by the fact that coherent detection techniques enable a higher spectral efficiency and greater tolerance to chromatic and polarization-mode dispersions than direct detection methods.[78–82] As noted in Chapter 13, such high-speed links have been tested at data rates up to 160 Gb/s over installed standard single-mode fibers.

8.4.1 Fundamental Concepts

Figure 8.20 illustrates the fundamental concept in coherent lightwave systems. The key principle of the coherent detection technique is to provide gain to the incoming optical signal by combining or mixing it with a locally generated *continuous-wave* (CW) optical field. The term *mixing* means that when two waves that have frequencies ω_1 and ω_2 are combined, the result will be other waves with frequencies equal to $2\omega_1$, $2\omega_2$, and $\omega_1 \pm \omega_2$. For coherent lightwave systems, all frequency components except $\omega_1 - \omega_2$ are filtered out at the receiver. The device used for creating the CW signal is a narrow-linewidth laser called a *local oscillator* (LO). The result of this mixing procedure is that the dominant noise in the receiver is the shot noise coming from the local oscillator. This means the receiver can achieve a shot noise limited sensitivity.
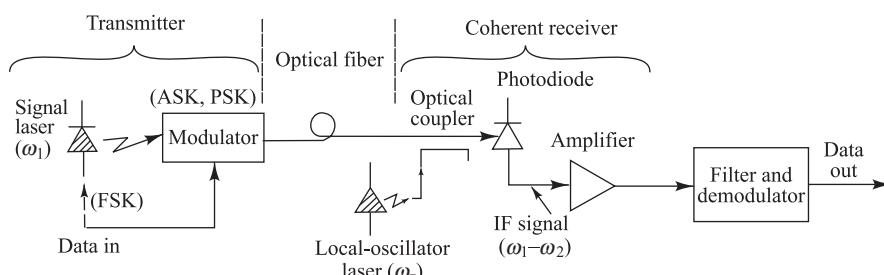
Fig. 8.20 *Fundamental concept of a coherent lightwave system. The three basic detection methods can have various modulation formats.*

For simplicity, to understand how this mixing can increase the coherent receiver performance, let us consider the electric field of the transmitted optical signal to be a plane wave having the form

$$E_s = A_s \cos\left[\omega_s t + \varphi_s(t)\right] \tag{8.30}$$

where A_s is the amplitude of the optical signal field, ω_s is the optical signal carrier frequency, and $\varphi_s(t)$ is the phase of the optical signal. To send information, one can modulate the amplitude, frequency, or phase of the optical carrier. Thus one of the following three modulation techniques can be implemented:

1. *Amplitude shift keying (ASK) or on-off keying (OOK).* In this case ϕ_s is constant and the signal amplitude A_s takes one of two values during each bit period, depending on whether a 0 or a 1 is being transmitted.
2. *Frequency shift keying (FSK).* For FSK modulation the amplitude A_s is constant and $\phi_s(t)$ is either $\omega_1 t$ or $\omega_2 t$, where the frequencies ω_1 and ω_2 represent binary signal values.
3. *Phase shift keying (PSK).* In the PSK method, information is conveyed by varying the phase with a sine wave $\phi_s(t) = \beta \sin \omega_m t$, where β is the modulation index and ω_m is the modulation frequency.

In a *direct-detection* system the electrical signal coming into the transmitter *amplitude-modulates* the optical power level of the light source. Thus the optical power is proportional to the signal current level. At the receiver the incoming optical signal is converted directly into a demodulated electrical output. This directly detected current is proportional to the intensity I_{DD} (the square of the electric field) of the optical signal, yielding

$$I_{DD} = E_s E_s^* = \frac{1}{2} A_s^2 \left[1 + \cos\left(2\omega_s t + 2\phi_s\right)\right] \tag{8.31}$$

The term involving $\cos(2\omega_s t + 2\phi_s)$ gets eliminated from the receiver, since its frequency, which is twice the optical carrier frequency, is beyond the response capability of the detector. Thus for direct detection Eq. (8.31) becomes

$$I_{DD} = E_s E_s^* = \frac{1}{2} A_s^2 \tag{8.32}$$

At the receiving end in coherent lightwave systems, the receiver first adds a locally generated optical wave to the incoming information-bearing signal and then detects the combination. There are four basic demodulation formats, depending on how the optical signal is mixed with the local oscillator (which gives heterodyne or homodyne detection) and how the electrical signal is detected (either synchronously or asynchronously). As we shall see in this section, for a given modulation format homodyne receivers are more sensitive than heterodyne receivers, and synchronous detection is more sensitive than asynchronous detection.

The mixing of the information-bearing and local-oscillator signals is done on the surface of the photodetector (before photodetection takes place). If the local-oscillator (LO) field has the form

$$E_{LO} = A_{LO} \cos\left[\omega_{LO} t + \phi_{LO}(t)\right] \tag{8.33}$$

where A_{LO} is the amplitude of the local oscillator field, and ω_{LO} and $\phi_{LO}(t)$ are the local-oscillator frequency and phase, respectively, then the detected current is proportional to the square of the total electric field of the signal falling on the photodetector. That is, the intensity $I_{coh}(t)$ is

$$\begin{aligned} I_{coh}(t) &= (E_s + E_{LO})^2 \\ &= \frac{1}{2} A_s^2 + \frac{1}{2} A_{LO}^2 + A_s A_{LO} \cos\left[\left(\omega_s - \omega_{LO}\right)t + \phi(t)\right]\cos\theta(t) \end{aligned} \tag{8.34}$$

where $\phi(t) = \phi_s(t) - \phi_{LO}(t)$ is the relative phase difference between the incoming information-bearing signal and the local-oscillator signal, and

$$\cos \theta(t) = \frac{E_s \cdot E_{LO}}{|E_s||E_{LO}|} \tag{8.35}$$

represents the polarization misalignment between the signal wave and the local-oscillator wave. Here again we have used the condition that the photodetector does not respond to terms oscillating near the frequency $2\omega_s$.

Since the optical power $P(t)$ is proportional to the intensity, at the photodetector we then have

$$P(t) = P_s + P_{LO} + 2\sqrt{P_s P_{LO}} \cos\left[\left(\omega_s - \omega_{LO}\right)t + \phi(t)\right]\cos\theta(t) \tag{8.36}$$

where P_s and P_{LO} are the signal and local-oscillator optical powers, respectively, with $P_{LO} \gg P_s$. Thus we see that the angular-frequency difference $\omega_{IF} = \omega_s - \omega_{LO}$ is an intermediate frequency, and the phase angle $\phi(t)$ gives the time-varying phase difference between the signal and local-oscillator levels. The frequency ω_{IF} is normally in the radio-frequency range of a few tens or hundreds of megahertz.

8.4.2 Homodyne Detection

When the signal-carrier and local-oscillator frequencies are equal, that is, when $\omega_{IF} = 0$, we have the special case of *homodyne detection.* Equation (8.36) then becomes

$$P(t) = P_s + P_{LO} + 2\sqrt{P_s P_{LO}} \cos\phi(t)\cos\theta(t) \tag{8.37}$$

Thus one can use either OOK [varying the signal level P_s while keeping $\phi(t)$ constant] or PSK [varying the phase $\phi_s(t)$ of the signal and keeping P_s constant] modulation schemes to transmit information. Note that since $P_{LO} \gg P_s$ and P_{LO} is constant, the last term on the right-hand side of Eq. (8.37) contains the transmitted information. Since this term increases with increasing laser power, the local oscillator effectively acts as a signal amplifier, thereby giving greater receiver sensitivity than direct detection.

As can be seen from Eq. (8.37), homodyne detection brings the signal directly to the baseband frequency, so that no further electrical demodulation is required. Homodyne receivers yield the most sensitive coherent systems. However, they are also the most difficult to build, since the local oscillator must be controlled by an optical phase-locked loop. In addition, the need for the signal and the local-oscillator lasers to have the same frequencies puts very stringent requirements on these two optical sources. These criteria include an extremely narrow spectral width (linewidth) and a high degree of wavelength tunability.

8.4.3 Heterodyne Detection

In *heterodyne detection* the intermediate frequency ω_{IF} is nonzero and an optical phase-locked loop is not needed. Consequently heterodyne receivers are much easier to implement than homodyne receivers. However, the price for this simplification is a 3-dB degradation in sensitivity compared to homodyne detection.[83]

Either OOK, FSK, or PSK modulation techniques can be used. Let us consider the output current at the receiver. Since $P_s \ll P_{LO}$, we can ignore the first term on the right-hand side of Eq. (8.36). The receiver output current then contains a dc term given by

$$i_{dc} = \frac{\eta q}{h\nu} P_{LO} \tag{8.38}$$

and a time-varying IF, term given by

$$i_{\text{IF}}(t) = \frac{2\eta q}{hv} \sqrt{P_s P_{\text{LO}}} \cos\left[\omega_{\text{IF}} + \phi(t)\right] \cos\theta(t) \qquad (8.39)$$

The dc-current is normally filtered out in the receiver, and the IF current gets amplified. One then recovers the information from the amplified current using conventional RF demodulation techniques.

8.4.4 BER Comparisons

Now we make a comparison between the various coherent detection techniques. Generally one characterizes the performance of a digital communication system in terms of the bit-error rate. The BER depends on the signal-to-noise ratio (SNR) and the probability density function (PDF) at the receiver output (at the input to the comparator). Since for high local oscillator powers the PDF is gaussian for both homodyne and heterodyne techniques, the BER depends only on the signal-to-noise ratio. Thus one can describe receiver sensitivity in terms of the SNR available at the receiver output, which is directly proportional to the received optical signal power. Since traditionally the receiver sensitivity for coherent detection techniques has been described in terms of the average number of photons required to achieve a 10^{-9} BER, we shall use that criterion here.

Direct-Detection OOK

Suppose we have an OOK system in which a sequence of 1 and 0 pulses occur with equal probability. Since the OOK data stream is in an on state only half of the time, the average number of photons required per bit $\bar{N}_p$ is half the number required per 1 pulse. Thus if $\bar{N}$ and 0 electron-hole pairs are created during 1 and 0 pulses, respectively, then the average number of photons per bit $\bar{N}_p$ for unity quantum efficiency ($\eta = 1$) is

$$\bar{N}_p = \frac{1}{2}\bar{N} + \frac{1}{2}(0) \qquad (8.40)$$

or $\bar{N} = 2\bar{N}_p$. From Eq. (7.23) we then have that the chance of making an error is

$$\frac{1}{2}P_r(0) = \frac{1}{2}e^{-2\bar{N}_p} \qquad (8.41)$$

Equation (8.41) implies that about 10 photons per bit are required to get a BER of 10^{-9} for a direct-detection OOK system.

In practice this fundamental quantum limit is very difficult to achieve for direct-detection receivers. The amplification electronics that follow the photodetector add both thermal noise and shot noise, so that the required received power level lies between 13 and 20 dB above the quantum limit.

OOK Homodyne System

As noted in Sec. 8.4.1, either homodyne or heterodyne type receivers can be used with OOK modulation. Let us first look at the homodyne case. When a 0 pulse of duration T is received, the average number $\bar{N}_0$ of electron–hole pairs created is simply the number generated by the local oscillator; that is,

$$\bar{N}_0 = A_{\text{LO}}^2 T \qquad (8.42)$$

For a 1 pulse, the average number of electron–hole pairs, $\bar{N}_1$, is

$$\bar{N}_1 = (A_{LO} + A_s)^2 T \simeq (A_{LO}^2 + 2A_{LO}A_s)T \tag{8.43}$$

where the approximation arises from the condition $A_{LO}^2 \gg A_S^2$. Since the local-oscillator output power is much higher than the received signal level, the voltage V seen by the decoder in the receiver during a 1 pulse is

$$V = \bar{N}_1 - \bar{N}_0 = 2A_{LO}A_s T \tag{8.44}$$

and the associated rms noise σ is

$$\sigma \simeq \sqrt{\bar{N}_1} \simeq \sqrt{\bar{N}_0} \tag{8.45}$$

Thus from Eq. (7.16) we have that the BER

$$P_e = \text{BER} = \frac{1}{2}\left[1 - \text{erf}\left(\frac{V}{2\sqrt{2}\,\sigma}\right)\right] = \frac{1}{2}\,\text{erfc}\left(\frac{V}{2\sqrt{2}\,\sigma}\right) = \frac{1}{2}\,\text{erfc}\left(\frac{A_s T^{1/2}}{\sqrt{2}}\right) \tag{8.46}$$

where $\text{erfc}(x) = 1 - \text{erf}(x)$ is the complementary error function.

From Example 7.8 we recall that to achieve a BER of 10^{-9} we need $V/\sigma = 12$. Using Eqs (8.44) and (8.45), this implies

$$A_s^2 T = 36 \tag{8.47}$$

which is the expected number of signal photons created per pulse. Thus for OOK homodyne detection, the average energy of each pulse must produce 36 electron–hole pairs. In the ideal case when the quantum efficiency is unity, a 10^{-9} BER is achieved with an average received optical energy of 36 photons per pulse. If we assume an OOK sequence of 1 and 0 pulses, which occur with equal probability, then the average number of received photons per bit of information, $\bar{N}_p$, is 18 (half the number required per pulse). Thus for OOK homodyne detection the BER is given by

$$\text{BER} = \frac{1}{2}\,\text{erfc}\left(\sqrt{\eta \bar{N}_p}\right) \tag{8.48}$$

To simplify this we note that a useful approximation to $\text{erfc}\left(\sqrt{x}\right)$ for $x \geq 5$ is

$$\text{erfc}\left(\sqrt{x}\right) \simeq \frac{e^{-x}}{\sqrt{\pi x}} \tag{8.49}$$

so that

$$\text{BER} \simeq \frac{e^{-\eta \bar{N}_p}}{\left(\pi \eta \bar{N}_p\right)^{1/2}} \tag{8.50}$$

for $\eta \bar{N}_p \geq 5$ in OOK homodyne detection.
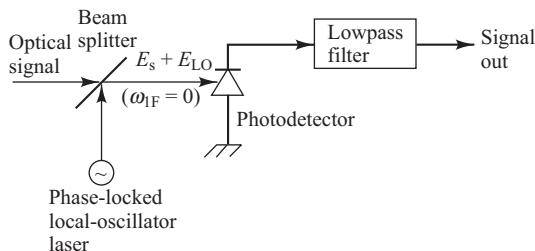
Fig. 8.21 *Fundamental setup of a homodyne receiver*

PSK Homodyne System

Homodyne detection of PSK modulation gives the best theoretical receiver sensitivity, but it is also the most difficult method to implement. Figure 8.21 shows the fundamental setup for a homodyne receiver. The incoming optical signal is first combined with a strong optical wave being emitted from the local oscillator. This is done using either a fiber directional coupler (see Chapter 11) or a partially reflecting plate called a *beam splitter*. When a beam splitter is used, it is made almost completely transparent, since the incoming signal is much weaker than the local-oscillator output.

As we saw in Eq. (8.37), the information is sent by changing the phase of the transmitted wave. For a 0 pulse the signal and local oscillator are out of phase, so that the resultant number of electron–hole pairs generated is

$$\bar{N}_0 = (A_{\mathrm{LO}} - A_s)^2 T \tag{8.51}$$

Similarly, for a 1 pulse the signals are in phase, so that

$$\bar{N}_1 = (A_{\mathrm{LO}} + A_s)^2 T \tag{8.52}$$

Consequently we have that the voltage seen by the decoder in the receiver is

$$V = \bar{N}_1 - \bar{N}_0 = (A_{\mathrm{LO}} + A_s)^2 T - (A_{\mathrm{LO}} - A_s)^2 T = 4 A_{\mathrm{LO}} A_s T \tag{8.53}$$

and the associated rms noise is

$$\sigma = \sqrt{A^2_{\mathrm{LO}} T} \tag{8.54}$$

Again, as in the case of homodyne OOK detection, the condition $V/\sigma = 12$ for a BER of 10^{-9} yields

$$A^2_{\mathrm{LO}} T = 9 \tag{8.55}$$

This says that for ideal PSK homodyne detection ($\eta = 1$), an average of 9 photons per bit is required to achieve a 10^{-9} BER. Note that here we do not need to consider the difference between photons per pulse and photons per bit as in the OOK case, since a PSK optical signal is on all the time.

Again using Eq. (7.16), we have that

$$\mathrm{BER} = \frac{1}{2} \, \mathrm{erfc} \sqrt{2 \eta \bar{N}_p} \tag{8.56}$$

for PSK homodyne detection.

Heterodyne Detection Schemes

The analysis for heterodyne receivers is more complicated than in the homodyne case because the photodetector output appears at an intermediate frequency ω_{IF}. The detailed derivations of the BER for various modulation schemes are given in the literature[72–77] so only the results are given here.

An attractive feature of heterodyne receivers is that they can employ either synchronous or asynchronous detection. Figure 8.22 shows the general receiver configuration. Let us look at this for PSK. In synchronous PSK detection (Fig. 8.22*a*) one uses a carrier-recovery circuit, which is usually a microwave phase-locked loop (PLL), to generate a local phase reference. The intermediate-frequency carrier is recovered by mixing the output of the PLL with the intermediate-frequency signal. One then uses a low-pass filter to recover the baseband signal. The BER for synchronous heterodyne PSK is given by

$$\text{BER} = \frac{1}{2} \, \text{erfc} \sqrt{\eta \overline{N}_p} \qquad (8.57)$$

In this case the ideal PSK receiver requires 18 photons per bit for a 10^{-9} BER. Note that this is the same as for OOK homodyne detection.

A simpler but robust technique that does not use a PLL is *asynchronous detection,* as illustrated in Fig. 8.22*b*. This technique is called *differential PSK* or DPSK.[74–76] Here the carrier-recovery circuit is replaced by a simple one-bit delay line. Since with a PSK method information is encoded by means of changes in the optical phase, the mixer will produce a positive or negative output depending on whether the phase of the received signal has changed from the previous bit. The transmitted information is thus
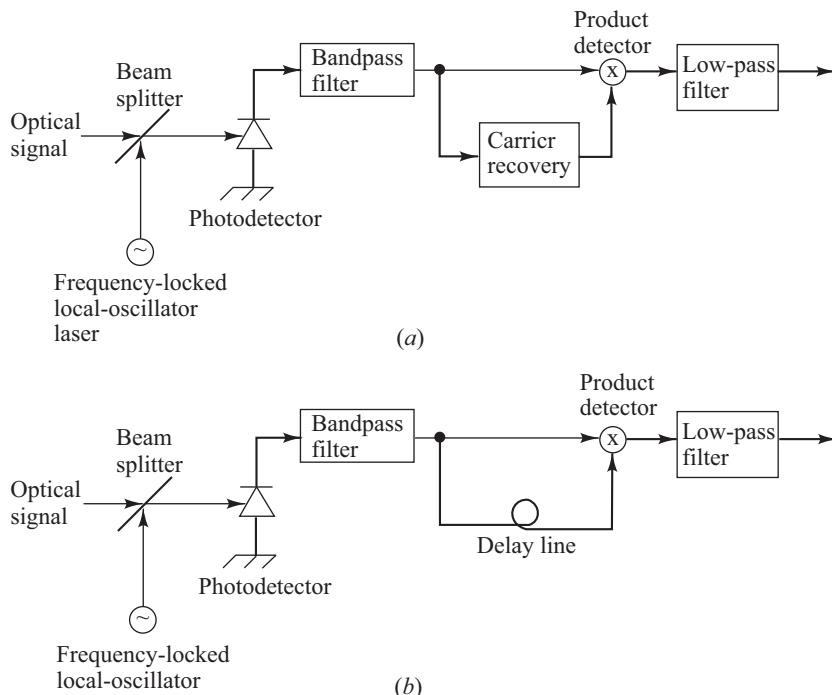
Fig. 8.22 *General heterodyne receiver configurations. (a) Synchronous detection uses a carrier-recovery circuit. (b) Asynchronous detection uses a one-bit delay line.*

recovered from this output. This DPSK technique has a sensitivity close to that of synchronous heterodyne detection of PSK, with a bit-error rate of

$$\text{BER} = \frac{1}{2} \exp\left(-\eta \bar{N}_p\right) \tag{8.58}$$

For a BER of 10^{-9} we thus require 20 photons per bit, which is a 0.5-dB penalty with respect to synchronous heterodyne detection of PSK.

Analogous to the PSK case, synchronous heterodyne OOK detection is 3 dB less sensitive than homodyne OOK. Thus the BER is given by

$$\text{BER} = \frac{1}{2}\text{erfc}\sqrt{\frac{1}{2}\eta \bar{N}_p} \tag{8.59}$$

Here one needs a minimum of 36 photons per bit for a 10^{-9} bit error rate. In the case of asynchronous heterodyne OOK detection, the bit error rate is given by

$$\text{BER} = \frac{1}{2}\exp\left(-\frac{1}{2}\eta \bar{N}_p\right) \tag{8.60}$$

Thus asynchronous heterodyne OOK detection requires 40 photons per bit for a 10^{-9} BER, which is 3 dB less sensitive than DPSK.

The receiver sensitivities for the various modulation techniques are summarized in Tables 8.4 and 8.5. Table 8.4 gives the probability of error as a function of the average number of received photons per

Table 8.4 *Summary of the probability of error as a function of the number of received photons per bit for coherent optical fiber systems*

| | | Probability of error | | |
| | | Heterodyne | | |
Modulation	*Homodyne*	*Synchronous detection*	*Asynchronous detection*	*Direct detection*
On-off keying (OOK)	$\frac{1}{2}\ \text{erfc}\left(\eta \bar{N}_p\right)^{1/2}$	$\frac{1}{2}\ \text{erfc}\left(\frac{1}{2}\eta \bar{N}_p\right)^{1/2}$	$\frac{1}{2}\ \exp\left(-\frac{1}{2}\eta \bar{N}_p\right)$	$\frac{1}{2}\exp\left(-2\eta \bar{N}_p\right)$
Phase-shift keying (PSK)	$\frac{1}{2}\ \text{erfc}\left(2\eta \bar{N}_p\right)^{1/2}$	$\frac{1}{2}\ \text{erfc}\left(\eta \bar{N}_p\right)^{1/2}$	$\frac{1}{2}\exp\left(-\eta \bar{N}_p\right)$	—
Frequency-shift keying (FSK)	—	$\frac{1}{2}\ \text{erfc}\left(\frac{1}{2}\eta \bar{N}_p\right)^{1/2}$	$\frac{1}{2}\exp\left(-\frac{1}{2}\eta \bar{N}_p\right)$	—

Table 8.5 *Summary of the number of photons required for a 10^{-9} BER by an ideal receiver having a photodetector with unify quantum efficiency*

| | | Number of photons | | |
| | | Heterodyne | | |
Modulation	Homodyne	Synchronous detection	Asynchronous detection	Direct detection
On-off keying (OOK)	18	36	40	10
Phase-shift keying (PSK)	9	18	20	—
Frequency-shift keying (FSK)	—	36	40	—

bit, $\bar{N}_p$, and Table 8.5 shows the number of photons required for a 10^{-9} BER by an ideal receiver having a photodetector with a quantum efficiency of $\eta = 1$.

A summary of the requirements of linewidth versus photons per bit for heterodyne PSK, FSK, and OOK systems is given in Fig. 8.23 for a 10^{-9} BER. PSK gives the best sensitivity for sources with very narrow linewidths. However, when the linewidth is greater than 0.2 percent of the bit rate, the sensitivity degrades quickly. In comparison, since FSK and OOK modulated signals can be detected using an optical power measurement that is not sensitive to phase noise, they maintain good performance with sensitivities below 60 photons per bit for linewidth-to-bit-rate ratios approaching unity.
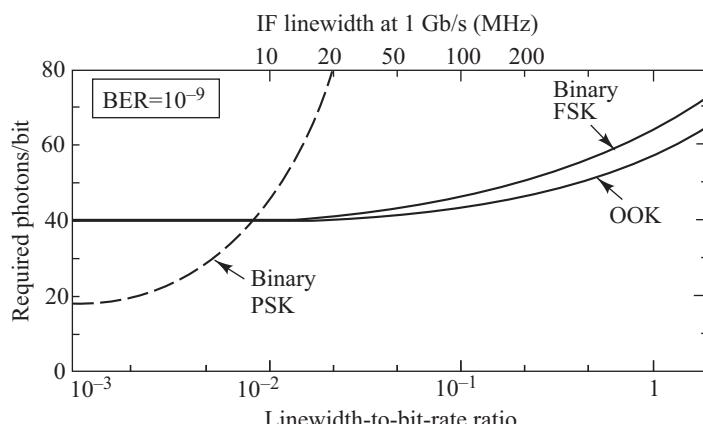
Fig. 8.23 *Calculated quantum-limited receiver sensitivity as a function of laser linewidth at 1 Gb/s or, equivalently, as a function of linewidth-to-bit-rate ratio. (Reproduced with permission from Linke,[83] © 1989, IEEE.)*

8.5 Differential Quadrature Phase-Shift Keying

Until about 2002, traffic was transmitted over most optical communication systems at data rates up to 2.5 Gb/s per wavelength using OOK signals in either NRZ or RZ formats. As the desire grew to transmit data at higher speeds, such as 10 and 40 Gb/s, the idea of using a multilevel modulation format received much attention.[84–92] Of particular interest for high-speed transmission is the use of the *differential quadrature phase-shift keying* (DQPSK) method. In a multilevel modulation format, more than one bit per symbol is transmitted. In the DQPSK method, information is encoded by means of the four phase shifts $\{0, +\pi/2, -\pi/2, \pi\}$. The set of bit pairs $\{00, 10, 01, 11\}$ can be assigned to each of the four phase shifts, respectively. For example, a phase shift of π means that the bit pair 11 was sent. Thus DQPSK transmits at a *symbol rate of half the aggregate bit rate*.

Commonly a DQPSK transmitter is implemented using two nested Mach-Zehnder modulators (see Fig. 4.34) as shown in Fig. 8.24. Here the output from a continuous-wave (CW) laser is inserted into an optical beam splitter that sends half of the laser power into each MZM. An optical $\pi/2$ phase shifter in one of the paths and an optical combiner at the MZM outputs then produce a single output signal with the four phase shifts $\{0, +\pi/2, -\pi/2, \pi\}$. The two MZMs are driven by incoming binary electrical signals. As illustrated in Fig. 8.24, an incoming bit stream is split into two parts with alternate bits being used to drive the top and bottom MZMs, respectively.

At the receiver the DQPSK optical signal is divided into two equal parts using an optical splitter (see Chapter 10). These parts are the inputs for two balanced receivers, which have differently biased delay interferometers (DI). The phase difference is set to $+\pi/4$ and $-\pi/4$ so that the two parallel receivers simultaneously demodulate the binary data streams contained in the DQPSK signal. In this setup the DI delay must be equal to the *symbol duration* for the DQPSK demodulation, which is *twice* the bit duration.

Since for a given data rate the symbol rate in DQPSK is reduced by a factor of two, then compared to a modulation scheme such as OOK, the spectral occupancy is reduced and the transmitter and receiver requirements are lowered. In addition, the chromatic dispersion and polarization-mode dispersion limits are extended. However, compared to DPSK, the SNR needed to achieve a specific BER is increased by a factor of 1 to 2 dB. Also, the design of the receiver becomes more complex because the tolerance to frequency drifts between the transmit laser and the delay interferometers is six times lower than for DPSK. For example, at a 40 Gb/s data rate and for a 1-dB penalty, DPSK tolerates a laser-to-DI mismatch of ±1.2 GHz, whereas for DQPSK only ±200 MHz is allowed.
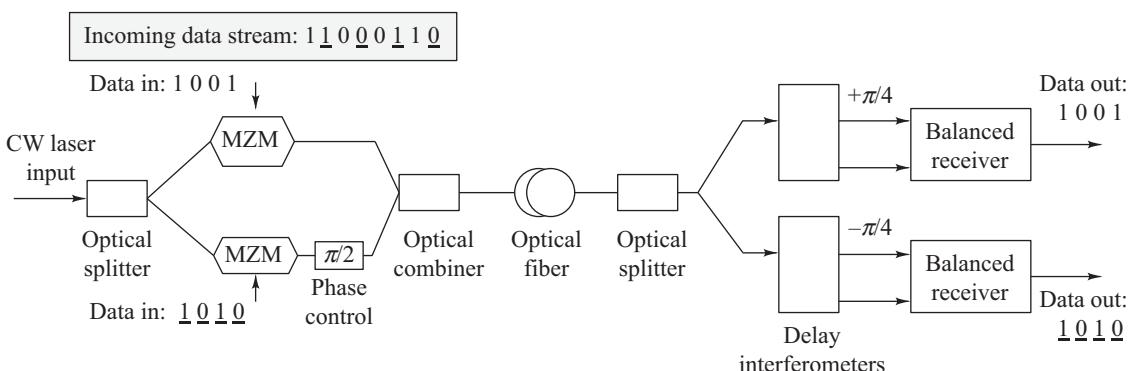
Fig. 8.24 *Common implementation of a DQPSK optical transmission system*

PROBLEMS

8.1 Make a graphical comparison, as in Fig. 8.4, and a spreadsheet calculation of the maximum attenuation-limited transmission distance of the following two systems operating at 100 Mb/s:

System 1 operating at 850 nm

(*a*) GaAlAs laser diode: 0-dBm (1-mW) fiber-coupled power.

(*b*) Silicon avalanche photodiode: −50-dBm sensitivity.

(*c*) Graded-index fiber: 3.5-dB/km attenuation at 850 nm.

(*d*) Connector loss: 1 dB/connector.

System 2 operating at 1300 nm

(*a*) InGaAsP LED: −13-dBm fiber-coupled power.

(*b*) InGaAs *pin* photodiode: −38-dBm sensitivity.

(*c*) Graded-index fiber: 1.5-dB/km attenuation at 1300 nm.

(*d*) Connector loss: 1 dB/connector.

Allow a 6-dB system operating margin in each case.

8.2 An engineer has the following components available:

(*a*) GaAlAs laser diode operating at 850 nm and capable of coupling 1 mW (0 dBm) into a fiber.

(*b*) Ten sections of cable each of which is 500 m long, has a 4-dB/km attenuation, and has connectors on both ends.

(*c*) Connector loss of 2 dB/connector.

(*d*) A *pin* photodiode receiver.

(*e*) An avalanche photodiode receiver.

Using these components, the engineer wishes to construct a 5-km link operating at 20 Mb/s. If the sensitivities of the *pin* and APD receivers are −45 and −56 dBm, respectively, which receiver should be used if a 6-dB system operating margin is required?

8.3 Using the step response $g(t) = [1-\exp(-2\pi B_e t)]u(t)$, show that the 10- to 90-percent receiver rise time is given by Eq. (8.4).

8.4 (*a*) Verify the steps leading from Eq. (8.11) to Eq. (8.12).

(*b*) Show that Eq. (8.14) follows from Eqs. (8.10) and (8.13).

8.5 Show that, if t_e is the full width of the gaussian pulse in Eq. (8.9) at the $1/e$ points, then the relationship between the 3-dB optical bandwidth and t_e is given by

$$f_{3dB} = \frac{0.53}{t_e}$$

8.6 A 90-Mb/s NRZ data transmission system that sends two DS3 (45-Mb/s) channels uses a GaAlAs laser diode that has a 1-nm spectral width. The rise time of the laser transmitter output is 2 ns. The transmission distance is 7 km over a graded-index fiber that has an 800-MHz · km bandwidth-distance product.

(*a*) If the receiver bandwidth is 90 MHz and the mode-mixing factor $q = 0.7$, what is the system rise time? Does this rise time meet the NRZ data requirement of being less than 70 percent of a pulse width?

(*b*) What is the system rise time if there is no mode mixing in the 7-km link; that is, $q = 1.0$?

8.7 Verify the plot in Fig. 8.6 of the transmission distance versus data rate of the following system. The transmitter is a GaAlAs laser diode operating at 850 nm. The laser power coupled into a fiber flylead is 0 dBm (1 mW), and the source spectral width is 1 nm. The fiber has a 3.5-dB/km attenuation at 850 nm and a bandwidth of 800 MHz · km. The receiver uses a silicon avalanche photodiode that has the sensitivity versus data rate shown in Fig. 8.3. For simplicity, the receiver sensitivity (in dBm) can be approximated from curve fitting by

$$P_R = 9 \log B - 68.5$$

where B is the data rate in Mb/s. For the data rate range of 1–1000 Mb/s, plot the attenuation-limited transmission distance (including a 1-dB connector loss at each end and a 6-dB system margin), the modal dispersion limit for full mode mixing

($q = 0.5$), the modal dispersion limit for no mode mixing ($q = 1.0$), and the material dispersion limit.

8.8 Make a plot analogous to Fig. 8.6 of the transmission distance versus data rate of the following system. The transmitter is an InGaAsP LED operating at 1300 nm. The fiber-coupled power from this source is −13 dBm (50 μW), and the source spectral width is 40 nm. The fiber has a 1.5-dB/km attenuation at 1300 nm and a bandwidth of 800 MHz · km. The receiver uses an InGaAs *pin* photodiode which has the sensitivity versus data rate shown in Fig. 8.3. For simplicity, the receiver sensitivity (in dBm) can be approximated from curve fitting by

$$P_R = 11.5 \log B - 60.5$$

where B is the data rate in Mb/s. For the data rate range of 10–1000 Mb/s, plot the attenuation-limited transmission distance (including a 1-dB connector loss at each end and a 6-dB system margin), the modal dispersion limit for no mode mixing ($q = 1.0$), and the modal dispersion limit for full mode mixing ($q = 0.5$). Note that the material dispersion is negligible in this case, as can be seen from Fig. 3.13.

8.9. A 1550-nm single-mode digital fiber optic link needs to operate at 622 Mb/s over 80 km without amplifiers. A single-mode InGaAsP laser launches an average optical power of 13 dBm into the fiber. The fiber has a loss of 0.35 dB/km, and there is a splice with a loss of 0.1 dB every kilometer. The coupling loss at the receiver is 0.5 dB, and the receiver uses an InGaAs APD with a sensitivity of −39 dBm. Excess-noise penalties are predicted to be 1.5 dB. Set up an optical power budget for this link and find the system margin. What is the system margin at 2.5 Gb/s with an APD sensitivity of −31 dBm?

8.10 Make a plot of the extinction ratio power penalty in decibels as a function of the extinction ratio ranging from 5 to 20 (from 7 to 13 dB).

8.11 Consider Eq. 8.24 for the power penalty caused by laser mode-partition noise.
 (*a*) Plot the power penalty (in dB) as a function of the factor $BLD\sigma_\lambda$ (ranging from 0 to 0.2) at a 10^{-9} BER (Q = 6) for mode-partition-noise factors $k = 0.4, 0.6, 0.8$, and 1, when using an InGaAs APD with $x = 0.7$.
 (*b*) Given that a multimode laser has a spectral width of 2.0 nm, what are the minimum dispersions required for 100-km spans operating at 155 Mb/s and 622 Mb/s with a 0.5-dB power penalty?

8.12 (*a*) Using Eq. (8.26) and assuming an excess noise factor $x = 0.7$, plot the chirp-induced power penalty in decibels as a function of the factor $DL\delta\lambda$ (product of the total dispersion and wavelength excursion) for the following parameter values (let $DL\delta\lambda$ range from 0 to 1.5 ns):
   (1) $t_{chirp} = 0.1$ ns and $B = 2.5$ Gb/s
   (2) $t_{chirp} = 0.1$ ns and $B = 622$ Mb/s
   (3) $t_{chirp} = 0.05$ ns and $B = 2.5$ Gb/s
   (4) $t_{chirp} = 0.05$ ns and $B = 622$ Mb/s
 (*b*) Find the distance limitation at 2.5 Gb/s if a 0.5-dB power penalty is allowed with $D = 1.0$ ps/(nm · km), $t_{chirp} = 0.05$ ns, and $\delta\lambda = 0.5$ nm.

8.13 In the (7,4) linear Hamming code the first four bits of a code word are the information bits b_1, b_2, b_3, b_4 and the next three bits b_5, b_6, b_7 are the check bits, which are given by

$b_5 = b_1 + b_3 + b_4$
$b_6 = b_1 + b_2 + b_4$
$b_7 = b_2 + b_3 + b_4$

Make a table listing the sixteen possible 4-bit information words, that is, 0000 through 1111, and the corresponding 7-bit code words.

8.14 (*a*) Find the binary equivalent of the polynomial $x^8 + x^7 + x^3 + x + 1$.
 (*b*) Find the polynomial equivalent of 10011011110110101

8.15 Consider the 10-bit data unit 1010011110 and the divisor 1011. Use both binary and algebraic division to find the CRC.

8.16 Consider the generator polynomial $x^3 + x + 1$.
 (*a*) Show that the CRC for the data unit 1001 is given by 110.

(*b*) If the resulting code word has an error in the first bit when it arrives at the destination, what is the CRC calculated by the receiver?

8.17 Why can the (255, 223) Reed-Solomon code correct up to 16 bytes, whereas the (255, 239) is limited to correcting 8 byte errors? What is the overhead for each of these two codes?

8.18 Verify the resulting expression in Eq. (8.34) for the intensity due to the combined signal and local oscillator fields.

8.19 A homodyne ASK receiver has a 100-MHz bandwidth and contains a 1310-nm *pin* photodiode with a responsivity of 0.6 A/W. It is shot noise limited and needs a signal-to-noise ratio of 12 to achieve a 10^{-9} BER. Find the photocurrent that is generated if the local oscillator power is -3 dBm and the phase error is $10°$. Assume that both the signal and the local oscillator have the same polarization.

8.20 For a bit error rate of 10^{-9} assume that the combined spectral width of the signal carrier wave and the local oscillator should be 1 percent of the transmitted bit rate.
(*a*) What spectral width is needed at 1310 nm for a 100-Mb/s data rate?
(*b*) What is the maximum allowed spectral width at 2.5 Gb/s?

8.21 (*a*) Verify that 10 photons per bit are required to get a bit error rate of 10^{-9} for a direct-detection OOK system.
(*b*) Show that for an ideal OOK homodyne system, one needs 36 photons per pulse to achieve a 10^{-9} BER.

8.22 Using the approximation given by Eq. (8.49) for erfc(*x*), find simplified expressions for PSK homodyne and PSK heterodyne detection. Letting $\eta = 1.0$, plot these expressions as a function of the number of received photons per bit for the range $5 < N_p < 20$.

REFERENCES

1. I. Jacobs, "Design considerations for long-haul lightwave systems," *IEEE J. Sel. Areas Commun.*, vol. 4, pp. 1389–1395, Dec. 1986.

2. P. S. Henry, R. A. Linke, and A. H. Gnauck, "Introduction to lightwave systems," in S. E. Miller and I. P. Kaminow, eds., *Optical Fiber Telecommunications–II,* Academic, New York, 1988.

3. D. H. Rice and G. E. Keiser, "Applications of fiber optics to tactical communication systems," *IEEE Commun. Mag.,* vol. 23, pp. 46–57, May 1985.

4. T. Kimura, "Factors affecting fiber-optic transmission quality," *J. Lightwave* Tech., vol. 6, pp. 611–619, May 1988.

5. A. Lowery, O. Lenzmann, I. Koltchanov, R. Moosburger, R. Freund, A. Richter, S. Georgi, D. Breuer, H. Hamster, "Multiple signal representation simulation of photonic devices, systems, and networks" *IEEE J. Selected Topics in Quantum Electronics,* vol. 6, no. 2, pp. 282–296, 2000.

6. A. J. Lowery, "Photonic simulation tools," in *Optical Fiber Telecommunications IV-B: Systems and Impairments*. I. Kaminow and T. Li, eds., Academic, San Diego, 2002.

7. P. Pepeljugoski, M. J. Hackert, J. S. Abbott, S. E. Swanson, S. E. Golowich, A. J. Ritger, P. Kolesar, Y. C. Chen, and P. Pleunis, "Deployment of system specification for laser-optimized 50-μm multimode fiber for multigigabit short-wavelength LANs," *J. Lightwave Tech.,* vol. 21, pp. 1256–1275, May 2003.

8. J. B. Schlager, M. J. Hackert, P. Pepeljugoski, and J. Gwinn, "Measurement for enhanced bandwidth performance over 62.6-μm multimode fiber in short-wavelength local area networks," *J. Lightwave Tech.,* vol. 21, pp. 1277–1285, May 2003.

9. I. B. Djordjevic, B. Vasic, M. Ivkovic, and I. Gabitov, "Achievable information rates for high-speed long-haul optical transmission," *J. Lightwave Tech.,* vol. 23, pp. 3755–3763, Nov. 2005.

10. ITU-T Recommendation G.959.1, *Optical Transport Network Physical Layer Interfaces,* March 2006.

11. A. B. Carlson and P. Crilly, *Communication Systems,* McGraw-Hill, Burr Ridge, IL, 5th ed., 2009.

12. B. Razavi, *Design of Integrated Circuits for Optical Communications,* McGraw-Hill, New York, 2003.

13. M. Eve, "Multipath time dispersion theory of an optical network," *Opt. Quantum Electron.,* vol. 10, pp. 45–51, Jan. 1978.

14. T. Kanada, "Evaluation of modal noise in multimode fiber-optic systems," *J. Lightwave Tech.,* vol. 2, pp. 11–18, Feb. 1984.

15. D. A. Nolan, R. M. Hawk, and D. B. Keck, "Multimode concatenation modal group analysis," *J. Lightwave Tech.,* vol. 5, pp. 1727–1732, Dec. 1987.

16. R. D. de la Iglesia and E. T. Azpitarte, "Dispersion statistics in concatenated single-mode fibers." *J. Lightwave Tech.,* vol. 5, pp. 1768–1772, Dec. 1987.

17. M. Suzuki and N. Edagawa, "Dispersion-managed high-capacity ultra-long-haul transmission," *J. Lightwave Tech.,* vol. 21, pp. 916–929, Apr. 2003.

18. B.-H. Choi, M. Attygalle, Y. J. Wen, and S. D. Dods, "Dispersion map optimization and dispersion slope mismatch issue on 40 channel × 10 Gbit/s transmission over 3000 km using standard SMF and EDFA amplification," *Optics Comm.,* vol. 242, pp. 525–532, Dec. 2004.

19. L. Grüner-Nielsen, M. Wandel, P. Kristensen, C. Jørgensen, L. V. Jørgensen, B. Edvold, B. Pálsdóttir, and D. Jakobsen, "Dispersion-compensating fibers," *J. Lightwave Tech.,* vol. 23, pp. 3566–3579, Nov. 2005.

20. H. Bülow, F. Buchali, and A. Klekamp, "Electronic dispersion compensation," *J. Lightwave Tech.,* vol. 26, pp. 158–167, Jan. 2008.

21. M. E. McCarthy, J. Zhao, A. D. Ellis, and P. Gunning, "Full-field electronic dispersion compensation of a 10 Gbit/s OOK signal over 4 124 km field-installed single-mode fibre," *J. Lightwave Tech.,* vol. 27, pp. 5327–5334, Dec. 2009.

22. N. Ohkawa and T. Takahashi, "Optimum bandwidth limitation method to overcome GVD-based effect in practical fibre-optic transmission systems,", *IEE Proc.-Commun.,* vol. 150, pp. 64–68, Feb. 2003.

23. ITU-T Recommendation G.957, *Optical Interfaces for Equipments and Systems Relating to the Synchronous Digital Hierarchy,* 1999.

24. Telcordia, *SONET Transport Systems—Common Generic Criteria GR-253,* Issue 4, Dec. 2005.

25. G. Bosco and P. Poggiolini, "On the joint effect of receiver impairments on direct-detection DQPSK systems," *J. Lightwave Tech.,* vol. 24, pp. 1323–1333, Mar. 2006.

26. J. Wang and J. M. Kahn, "Impact of chromatic and polarization-mode dispersions on DPSK systems using interferometric demodulation and direct detection," *J. Lightwave Tech.,* vol. 22, pp. 362–371, Feb. 2004.

27. C. D. Poole and J. Nagel, "Polarization effects in lightwave systems," in I. P. Kaminow and T. L Koch, eds., *Optical Fiber Telecommunications–III.,* vol. A, Academic, New York, 1997.

28. R. Khosravani and A. E. Willner, "System performance evaluation in terrestrial systems with high polarization mode dispersion and the effect of chirping," *IEEE Photonics Technol. Lett.,* vol. 13, pp. 296–298, Apr. 2001.

29. A. B. dos Santos, M. R. Jimenez, J. P. von der Weid, and A. Djupsjöbacka, "Statistical measurements of BER fluctuations due to PMD in 10-Gb/s optical transmissions," *IEEE Photonics Technol. Lett,* vol. 14, pp. 926–928, July 2002.

30. A. E. Willner, S.M.R. Motaghian Nezam, L. Yan, Z. Pan, and M. C. Hauer, "Monitoring and control of polarization-related impairments in optical fiber systems," *J. Lightwave Tech.,* vol. 22, pp. 106–125, Jan. 2004.

31. J. Cameron, L. Chen, X. Bao, and J. Stears, "Time evolution of polarization mode dispersion in optical fibers," *IEEE Photonics Tech. Lett.,* vol. 10, pp. 1265–1267, Sept. 1998.

32. D. S. Waddy, L. Chen, and X. Bao, "Polarization effects in aerial fibers," *Optical Fiber Tech.,* vol. 11, no. 1, pp. 1–19, Jan. 2005.

33. P. E. Couch and R. E. Epworth, "Reproducible modal-noise measurements in system design and analysis," *J. Lightwave Tech.,* vol. LT-1, pp. 591–595, Dec. 1983.

34. F. M. Sears, I. A. White, R. B. Kummer, and F. T. Stone, "Probability of modal noise in

single-mode lightguide systems," *J. Lightwave Tech.,* vol. LT-4, pp. 652–655, June 1986.

35. K. Petermann and G. Arnold, "Noise and distortion characteristics of semiconductor lasers in optical fiber communication systems," *IEEE J. Quantum Electron.,* vol. QE-18, pp. 543–554, Apr. 1982.

36. A. M. J. Koonen, "Bit-error-rate degradation in a multimode fiber optic transmission link due to modal noise," *IEEE J. Sel. Areas. Commun.,* vol. SAC-4, pp. 1515–1522, Dec. 1986.

37. P. Chan and T. T. Tjhung, "Bit-error-rate performance for optical fiber systems with modal noise," *J. Lightwave Tech.,* vol. 7, pp. 1285–1289, Sept. 1989.

38. P. M. Shankar, "Bit-error-rate degradation due to modal noise in single-mode fiber optic communication systems," *J. Opt. Commun.,* vol. 10, pp. 19–23, Mar. 1989.

39. N. H. Jensen, H. Olesen, and K. E. Stubkjaer, "Partition noise in semiconductor lasers under CW and pulsed operation," *IEEE J. Quantum Electron.,* vol. QE-18, pp. 71–80, Jan. 1987.

40. M. Ohtsu and Y. Teramachi, "Analyses of mode partition and mode hopping in semiconductor lasers," *IEEE J. Quantum Electron.,* vol. 25, pp. 31–38, Jan. 1989.

41. C. H. Henry, P. S. Henry, and M. Lax, "Partition fluctuations in nearly single longitudinal mode lasers," *J. Lightwave Tech.,* vol. LT-2, pp. 209–216, June 1984.

42. E. E. Basch, R. F. Kearns, and T. G. Brown, "The influence of mode partition fluctuations in nearly single-longitudinal-mode lasers on receiver sensitivity," *J. Lightwave Tech.,* vol. LT-4, pp. 516–519, May 1986.

43. J. C. Cartledge, "Performance implications of mode partition fluctuations in nearly single longitudinal mode lasers," *J. Lightwave Tech.,* vol. 6, pp. 626–635, May 1988.

44. K. Ogawa, "Analysis of mode partition noise in laser transmission systems," *IEEE J. Quantum Electron.,* vol. QE-18, pp. 849–855, May 1982.

45. N. A. Olsson, W. T. Tsang, H. Temkin, N. K. Dutta, and R. A. Logan, "Bit-error-rate saturation due to mode-partition noise induced by optical feedback in 1.5 μm single longitudinal-mode C^3 and DFB semiconductor lasers," *J. Lightwave Tech.,* vol. LT-3, pp. 215–218, Apr. 1985.

46. S. E. Miller, "On the injection laser contribution to mode partition noise in fiber telecommunication systems," *IEEE J. Quantum Electron.,* vol. 25, pp. 1771–1781, Aug. 1989.

47. S. Arahira, H. Yaegashi, K. Nakamura, and Y. Ogawa, "Chirp control and broadband wavelength-tuning of 40-GHz monolithic actively mode-locked laser diode module with an external CW light injection," *IEEE J. Sel. Topics Quantum Electron.,* vol. 11, pp. 1103–1111, Sept./Oct. 2005.

48. S. Chandrasekhar, C. R. Doerr, L. L. Buhl, Y. Matsui, D. Mahgerefteh, X. Zheng, K. McCallion, Z. Fan, and P. Tayebati, "Repeaterless transmission with negative penalty over 285 km at 10 Gb/s using a chirp managed laser," *IEEE Photonics Tech. Lett.,* vol. 17, pp. 2454–2456, Nov. 2005.

49. Y. Matsui, D. Mahgerefteh, X. Zheng, C. Liao, Z. F. Fan, K. McCallion, and P. Tayebati, "Chirp-managed directly modulated laser (CML)," *IEEE Photonics Tech. Lett.,* vol. 18, pp. 385–387, Jan. 2006.

50. H. Halbritter, C. Sydlo, B. Kögel, F. Riemenschneider, H. L. Hartnagel, and P. Meissner, "Linewidth and chirp of MEMS-VCSELs," *IEEE Photonics Tech. Lett.,* vol. 18, pp. 2180–2182, Oct. 2006.

51. R. Heidemann, "Investigation on the dominant dispersion penalties occurring in multigigabit direct detection systems," *J. Lightwave Tech.,* vol. 6, pp. 1693–1697, Nov. 1988.

52. S. Yamamoto, M. Kuwazuru, H. Wakabayashi, and Y. Iwamoto, "Analysis of chirp power penalty in 1.55-μm DFB-LD high-speed optical fiber transmission system," *J. Lightwave Tech.,* vol. LT-5, pp. 1518–1524, Oct. 1987.

53. P. J. Corvini and T. L. Koch, "Computer simulation of high-bit-rate optical fiber transmission using single-frequency laser," *J. Lightwave Tech.,* vol. LT-5, pp. 1591–1595, Nov. 1987.

54. J. C. Cartledge and G. S. Burley, "The effect of laser chirping on lightwave system performance," *J. Lightwave Tech.,* vol. 7, pp. 568–573, Mar. 1989.

55. C. H. Henry, "Theory of the linewidth of semiconductor lasers," *JEEE J. Quantum Electron.,* vol. QE-18, pp. 259–264, Feb. 1982.

56. C. H. Harder, K. Vahala, and A. Yariv, "Measurement of the linewidth enhancement factor of semiconductor lasers," *Appl. Phys. Lett.,* vol. 42, pp. 328–330, Apr. 1983.

57. R. Schimpe, J. E. Bowers, and T. L. Koch, "Characterization of frequency response of 1.5-μm InGaAsP DFB laser diode and InGaAs *PIN* photodiode by heterodyne measurement technique," *Electron. Lett.,* vol. 22, pp. 453–454, Apr. 24, 1986.

58. M. Shikada, S. Takano, S. Fujita, I. Mito, and K. Minemura, "Evaluation of power penalties caused by feedback noise of distributed feedback laser diodes," *J. Lightwave Tech.,* vol. 6, pp. 655–659, May 1988.

59. M. Nakazawa, "Rayleigh backscattering theory for single-mode fibers," *J. Opt. Soc. Amer.,* vol. 73, pp. 1175–1180, Sept. 1983.

60. P. Wan and J. Conradi, "Impact of double Rayleigh backscatter noise on digital and analog fiber systems," *J. Lightwave Tech.,* vol. 14, pp. 288–297, Mar. 1996.

61. R. A. Salvatore and R. T. Sahara, "Reduction in reflection-induced chirp from photonic integrated sources," *IEEE Photonics Tech. Lett.,* vol. 14, pp. 1662–1664, Dec. 2002.

62. E. W. Biersack, "Performance of forward error correction in an ATM environment," *IEEE J. Sel. Areas Commun.,* vol. 11, pp. 631–640, May 1993.

63. S.-M. Lei, "Forward error correction codes for MPEG2 over ATM," *IEEE Trans. Circuits Sys. for Video Tech.,* vol. 4, pp. 200–203, Apr. 1994.

64. M. Tomizawa, Y. Yamabayashi, K. Murata, T. Ono, Y. Kobayashi, and K. Hagimoto, "Forward error correcting codes in synchronous fiber optic transmission systems." *J. Lightwave Tech.,* vol. 15, pp. 43–52, Jan. 1997.

65. B. A. Forouzan, *Data Communication Networking*, McGraw-Hill, Burr Ridge, IL, 4th ed., 2007.

66. A. Leon-Garcia and I. Widjaja, *Communication Networks,* McGraw-Hill, Burr Ridge, IL, 2nd ed., 2004.

67. G. Keiser, *Local Area Networks,* McGraw-Hill, Burr Ridge, IL, 2nd ed., 2002.

68. S. Lin and D. J. Costello, *Error Control Coding,* Prentice-Hall, 2nd ed., 2005.

69. ITU-T Recommendation G.709, *Interfaces for the Optical Transport Network* (OTN), Mar. 2003.

70. ITU-T Recommendation G.975, *Forward Error Correction for Submarine Systems,* Oct. 2000.

71. ITU-T Recommendation G.975.1, *Forward Error Correction for High Bit Rate DWDM Submarine Systems,* Feb. 2004.

72. I. W. Stanley, "A tutorial review of techniques for coherent optical fiber transmission systems," *IEEE Commun. Mag.,* vol. 23, pp. 37–53, Aug. 1985.

73. J. Salz, "Coherent lightwave communications," *AT&T Tech. J.,* vol. 64, pp. 2153–2209, Dec. 1985.

74. P. J. Winzer and R.-J. Essiambre, "Advanced optical modulation formats," *IEEE Proc.,* vol. 94, pp. 952–985, 2006.

75. P. J. Winzer and R.-J. Essiambre, "Advanced optical modulation formats," chap. 2, in I. P. Kaminow, T. Li, and A. E. Willner, eds., *Optical Fiber Telecommunications–V,* vol. B, Academic, New York, 2008.

76. K. Kikuchi, "Coherent optical communication systems," chap. 3, in I. P. Kaminow, T. Li, and A. E. Willner, eds., *Optical Fiber Telecommunications – V,* vol. B, Academic, New York, 2008.

77. S. Betti, G. De Marchis, and E. Iannone, *Coherent Optical Communication Systems,* Wiley, Hoboken, NJ, 1995.

78. A. H. Gnauck and P. J. Winzer, "Optical phase-shift-keyed transmission," *J. Lightwave Tech.,* vol. 23, pp. 66–79, Jan. 2005.

79. D. S. Ly-Gagnon, S. Tsukamoto, K. Katoh, and K. Kikuchi, "Coherent detection of optical quadrature phase-shift keying signals with

carrier phase estimation," *J. Lightwave Tech.,* vol. 24, pp. 12–21, Jan. 2006.

80. X. Liu, S. Chandrasethar, and A. Leven, "Self-coherent optical transport systems," chap. 4, in I. P. Kaminow, T. Li, and A. E. Willner, eds., *Optical Fiber Telecommunications – V,* vol. B, Academic, New York, 2008.

81. G. Bosco and P. Poggiolini, "On the joint effect of receiver impairments on direct-detection DQPSK systems," *J. Lightwave Tech.,* vol. 24, pp. 1323–1333, Mar. 2006.

82. S. Ryu, *Coherent Lightwave Communication Systems,* Artech House, Norwood, MA, 1995.

83. R. A. Linke, "Optical heterodyne communication systems," *IEEE Commun. Mag.,* vol. 27, pp. 36–41, Oct. 1989.

84. K.-P. Ho, *Phase-Modulated Optical Communication Systems,* Springer, New York, 2005.

85. M. Seimetz, *High-Order Modulation for Optical Fiber Transmission,* Springer, New York, 2009.

86. G. Charlet, "Progress in optical modulation formats for high-bit rate WDM transmissions," *IEEE J. Sel. Topics Quantum Electron.,* vol. 12, no. 4, pp. 469–483, July/Aug. 2006.

87. F. Vacondio, A. Ghazisaeidi, A. Bononi, and L. A. Rusch, "DQPSK: When is a narrow filter receiver good enough?," *J. Lightwave Tech.,* vol. 23, no. 22, pp. 5106–5114, Nov. 2009.

88. C.R.S. Fludger, T. Duthel, D. van den Borne, C. Schulien, E.-D. Schmidt, T. Wuth, J. Geyer, E. De Man, G. D. Khoe, and H. de Waardt, "Coherent equalization and POLMUX-RZ-DQPSK for robust 100-GE transmission," *J. Lightwave Tech.,* vol. 26, pp. 64–72, Jan. 2008.

89. P. J. Winzer, G. Raybon, H. Song, A. Adamiecki, S. Corteselli, A. H. Gnauck, D. A. Fishman, C. R. Doerr, S. Chandrasekhar, L. L. Buhl, T. J. Xia, G. Wellbrock, W. Lee, B. Basch, T. Kawanishi, K. Higuma, and Y. Painchaud, "100-Gb/s DQPSK transmission: From laboratory experiments to field trials," *J. Lightwave Tech.,* vol. 26, no. 20, pp. 3388–3402, Oct. 2008.

90. M. Daikoku, I. Morita, H. Taga, H. Tanaka, T. Kawanishi, T. Sakamoto, T. Miyazaki, and T. Fujita, "100-Gb/s DQPSK transmission experiment without OTDM for 100G Ethernet transport," *J. Lightwave Tech.,* vol. 25, no. 1, pp. 139–145, Oct. 2007.

91. Z. Li, L. Cheng, Y. Yang, C. Lu, A.P.T. Lau, H. Tam, P.K.A. Wai, C. Wang, X. Xu, J. Deng, and Q. Xiong, "1500-km SSMF transmission of mixed 40-Gb/s CS-RZ duobinary and 100-Gb/s CS-RZ DQPSK signals," *Photonics Technol. Lett.,* vol. 21, no. 16, pp. 1148–1150, Aug. 2009.

92. J. X. Cai, C. R. Davidson, M. Nissov, H. Li, W. T. Anderson, Y. Cai, L. Liu, A. N. Pilipetskii, D. G. Foursa, W. W. Patterson, P. C. Corbett, A. J. Lucero, and N. S. Bergano, "Transmission of 40-Gb/s WDM signals over transoceanic distance using conventional NZ-DSF with receiver dispersion slope compensation," *J. Lightwave Tech.,* vol. 24, pp. 191–200, Jan. 2006.

CHAPTER 9

Analog Links

In telecommunication networks the trend has been to link telephone exchanges with digital circuits. A major reason for this was the introduction of digital integrated-circuit technology, which offered a reliable and economic method of transmitting both voice and data signals. Since the initial applications of fiber optics were to telecommunication networks, its first widespread usage has involved digital links. However, in many instances it is more advantageous to transmit information in its original analog form instead of first converting it to a digital format. Some examples of this are microwave-multiplexed signals,[1] subscriber services using hybrid fiber/coax (HFC), video distribution, radio-over-fiber antenna remoting,[1–10] and radar signal processing. Most analog applications use laser diode transmitters, so we concentrate on this optical source here.

When implementing an analog fiber optic system, the main parameters to consider are the carrier-to-noise ratio, bandwidth, spur free dynamic range, and signal distortion resulting from nonlinearities in the transmission system. Section 9.1 describes the general operational aspects and components of an analog fiber optic link. Traditionally, in an analog system a carrier-to-noise ratio analysis is used instead of a signal-to-noise ratio analysis, since the information signal is normally superimposed on a radio-frequency (RF) carrier. Thus in Sec. 9.2 we examine carrier-to-noise ratio requirements. This is first done for a single channel under the assumption that the information signal is directly modulated onto an optical carrier.

For transmitting multiple signals over the same channel, a subcarrier modulation SCM technique can be used. In this method, which is described in Sec. 9.3, the information signals are first superimposed on ancillary RF subcarriers. These carriers then are combined and the resulting electrical signal is used to modulate the optical carrier. A limiting factor in these systems is the signal impairment arising from harmonic and intermodulation distortions.

In SCM techniques the RF subcarriers are multiplexed in the electrical domain and then are superimposed on an optical carrier. For example, these could be the 6-MHz video signals used in CATV systems. As a result of the emerging use of broadband wireless communication devices, schemes have been investigated and implemented for using analog optical fiber links for distributing broadband microwave-frequency signals in a variety of applications. The methods for transmitting microwave analog signals in the 0.3 to 300-GHz range over an optical fiber link have become known as *RF-over-fiber* techniques. Section 9.4 examines the basics of these techniques. Section 9.5 gives an example of radio-over-fiber links used for in-building distributed antenna systems to provide wireless LAN and mobile telephony services over a single fiber.

To enable the efficient application of RF-over-fiber techniques, the field of *microwave photonics* came into existence. Research in this field encompasses the study and applications of photonic devices operating at microwave frequencies. In addition to device developments, microwave photonics also addresses optical signal processing at microwave speeds and the design and implementation of RF photonic transmission systems. Section 9.6 gives a brief overview of microwave-photonic components and their uses.

9.1 | Overview of Analog Links

Figure 9.1 shows the basic elements of an analog link. The transmitter contains either an LED or a laser diode optical source. As noted in Sec. 4.5 and shown in Fig. 4.39, in analog applications, one first sets a bias point on the source approximately at the midpoint of the linear output region. The analog signal can then be sent using one of several modulation techniques. The simplest form for optical fiber links is direct intensity modulation, wherein the optical output from the source is modulated simply by varying the current around the bias point in proportion to the message signal level. Thus the information signal is transmitted directly in the baseband.

A somewhat more complex but often more efficient method is to translate the baseband signal onto an electrical subcarrier prior to intensity modulation of the source. This is done using standard amplitude-modulation (AM), frequency-modulation (FM), or phase-modulation (PM) techniques.[11] No matter which method is implemented, one must pay careful attention to signal impairments in the optical source. These include harmonic distortions, intermodulation products, relative intensity noise (RIN) in the laser, and laser clipping.[12]

In relation to the fiber optic element shown in Fig. 9.1, one must take into account the frequency dependence of the amplitude, phase, and group delay in the fiber. Thus the fiber should have a flat amplitude and group-delay response within the passband required to send the signal free of linear distortion. In addition, since modal-distortion-limited bandwidth is difficult to equalize, it is best to choose a single-mode fiber. The fiber attenuation is also important, since the carrier-to-noise performance of the system will change as a function of the received optical power.

The use of an optical amplifier in the link leads to additional noise, known as amplified spontaneous emission (ASE), which is described in Chapter 11. In the optical receiver, the principal impairments are quantum or shot noise, APD gain noise, and thermal noise.
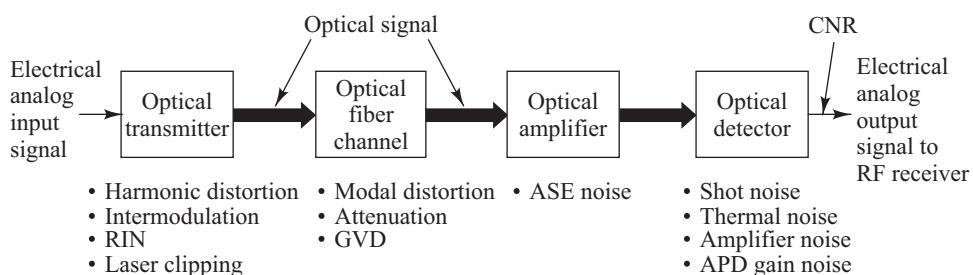
Fig. 9.1 *Basic elements of an analog link and the major noise contributors*

9.2 Carrier-to-Noise Ratio

In analyzing the performance of analog systems, one usually calculates the ratio of rms carrier power to rms noise power at the input of the RF receiver following the photodetection process. This is known as the *carrier-to-noise ratio* (CNR). Let us look at some typical CNR values for digital and analog data. For digital data, consider the use of frequency-shift keying (FSK). In this modulation scheme, the amplitude of a sinusoidal carrier remains constant, but the phase shifts from one frequency to another to represent binary signals. For FSK, BERs of 10^{-9} and 10^{-15} translate into CNR values of 36 (15.6 dB) and 64 (18.0 dB), respectively. The analysis for analog signals is more complex, since it sometimes depends on user perception of the signal quality, such as in viewing a television picture. A widely used analog signal is a 525-line studio-quality television signal. Using amplitude modulation (AM) for such a signal requires a CNR of 56 dB, since the need for bandwidth efficiency leads to a high signal-to-noise ratio. Frequency modulation (FM), on the other hand, only needs CNR values of 15–18 dB.

If CNR_i represents the carrier-to-noise ratio related to a particular signal contaminant (e.g., shot noise), then for N signal-impairment factors the total CNR is given by

$$\frac{1}{\text{CNR}} = \sum_{i=1}^{N} \frac{1}{\text{CNR}_i} \tag{9.1}$$

For links in which only a single information channel is transmitted, the important signal impairments include laser intensity noise fluctuations, laser clipping, photodetector noise, and optical-amplifier noise. When multiple message channels operating at different carrier frequencies are sent simultaneously over the same fiber, then harmonic and intermodulation distortions arise. Furthermore, the inclusion of an optical amplifier gives rise to ASE noise. In principle, the three dominant factors that cause signal impairments in a fiber link are shot noise, optical-amplifier noise, and laser clipping. Most other degradation effects can be sufficiently reduced or eliminated.

In this section, we first examine a simple single-channel amplitude-modulated signal sent at baseband frequencies. Section 9.3 addresses multichannel systems in which intermodulation noise becomes important. Problem 9.10 gives expressions for the effects of laser clipping and ASE noise.

9.2.1 Carrier Power

To find the carrier power, let us first look at the signal generated at the transmitter. As shown in Fig. 9.2, the drive current through the optical source is the sum of the fixed bias current and a time-varying sinusoid. The source acts as a square-law device, so that the envelope of the output optical power $P(t)$ has the same form as the input drive current. If the time-varying analog drive signal is $s(t)$, then

$$P(t) = P_t[1 + ms(t)] \tag{9.2}$$

where P_t is the optical output power at the bias current level and the modulation index m is defined by Eq. (4.56). In terms of optical power, the modulation index is given by

$$m = \frac{P_{\text{peak}}}{P_t} \tag{9.3}$$

where P_{peak} and P_t are defined in Fig. 9.2. Typical values of m for analog applications range from 0.25 to 0.50.
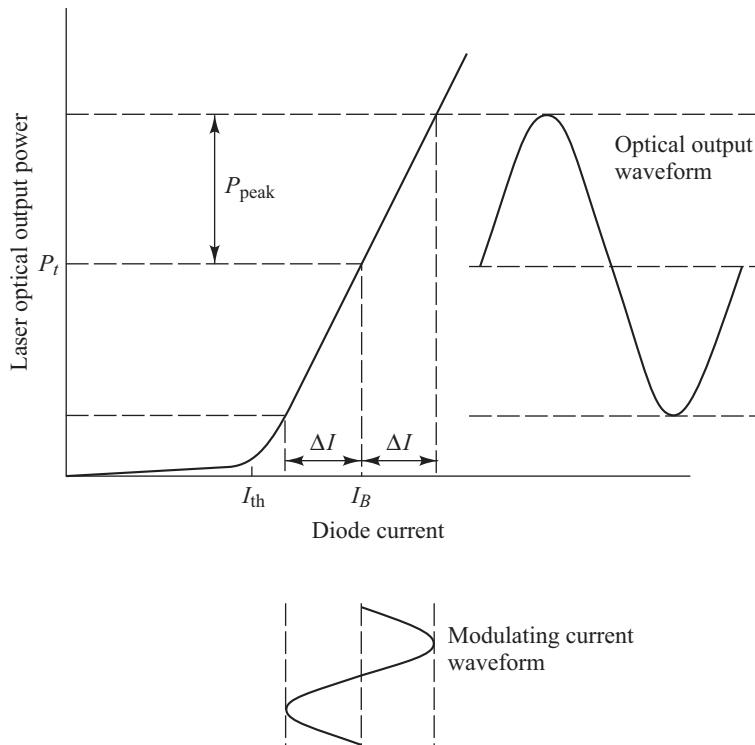
Fig. 9.2 *Biasing conditions of a laser diode and its response to analog signal modulation*

For a sinusoidal received signal, the carrier power C at the output of the receiver (in units of A^2) is

$$C = \frac{1}{2}(m\mathcal{R}\, M\, \bar{P})^2 \tag{9.4}$$

where $\mathcal{R}$ is the unity gain responsivity of the photodetector, M is the photodetector gain ($M = 1$ for *pin* photodiodes), and $\bar{P}$ is the average received optical power.

9.2.2 Photodetector and Preamplifier Noises

The expressions for the photodiode and preamplifier noises are given by Eqs. (6.16) and (6.17), respectively. That is, for the photodiode noise we have

$$\left\langle i_N^2 \right\rangle = \sigma_N^2 \approx 2q(I_p + I_D)M^2 F(M)B_e \tag{9.5}$$

Here, as defined in Chapter 6, $I_p = \mathcal{R}_0\,\bar{P}$ is the primary photocurrent, I_D is the detector bulk dark current, M is the photodiode gain with $F(M)$ being its associated noise figure, and B_e is the receiver bandwidth. Then, the CNR for the photodetector only is $\text{CNR}_{\text{det}} = C/\sigma_N^2$.

Generalizing Eq. (6.17) for the preamplifier noise, we have

$$\left\langle i_T^2 \right\rangle = \sigma_T^2 = \frac{4k_B T}{R_{eq}} B_e F_t \tag{9.6}$$

Here, R_{eq} is the equivalent resistance of the photodetector load and the preamplifier, and F_t is the noise factor of the preamplifier. Then, the CNR for the preamplifier only is $\text{CNR}_{preamp} = C/\sigma_T^2$.

9.2.3 Relative Intensity Noise (RIN)

Within a semiconductor laser, fluctuations in the amplitude or intensity of the output produce optical intensity noise. These fluctuations could arise from temperature variations or from spontaneous emission contained in the laser output. The noise resulting from the random intensity fluctuations is called *relative intensity noise* (RIN), which may be defined in terms of the mean-square intensity variations. The resultant mean-square noise current is given by

$$\left\langle i_{RIN}^2 \right\rangle = \sigma_{RIN}^2 = \text{RIN}(\mathscr{R}\,\bar{P})^2\, B_e \tag{9.7}$$

Then, the CNR due to laser amplitude fluctuations only is $\text{CNR}_{RIN} = C/\sigma_{RIN}^2$. Here, the RIN, which is measured in dB/Hz, is defined by the noise-to-signal power ratio

$$\text{RIN} = \frac{\left\langle (\Delta P_L)^2 \right\rangle}{\bar{P}_L^2} \tag{9.8}$$

where $\left\langle (\Delta P_L)^2 \right\rangle$ is the mean-square intensity fluctuation of the laser output and $\bar{P}_L$ is the average laser light intensity. This noise decreases as the injection-current level increases according to the relationship

$$\text{RIN} \propto \left(\frac{I_B}{I_{th}} - 1 \right)^{-3} \tag{9.9}$$

Vendor data sheets for 1550-nm DFB lasers typically quote RIN values of −152 to −158 dB/Hz.

Substituting the CNRs resulting from Eq. (9.4) through Eq. (9.7) into Eq. (9.1) yields the following carrier-to-noise ratio for a single-channel AM system:

$$\frac{C}{N} = \frac{\frac{1}{2}(m\mathscr{R}M\bar{P})^2}{\text{RIN}(\mathscr{R}\bar{P})^2 B_e + 2q(I_p + I_D)M^2 F(M)B_e + (4k_B T/R_{eq})B_e F_t} \tag{9.10}$$

9.2.4 Reflection Effects on RIN

In implementing a high-speed analog link, one must take special precautions to minimize optical reflections back into the laser.[2] Back-reflected signals can increase the RIN by 10–20 dB as shown in

---

Example 9.1 Figure 9.3 shows an example of Eq. (9.9) for two buried-heterostructure lasers.[13] The noise level was measured at 100 MHz. For injection currents sufficiently above threshold (i.e., for $I_B/I_{th} > 1.2$), the RIN of these index-guided lasers lies between −140 and −150 dB/Hz.
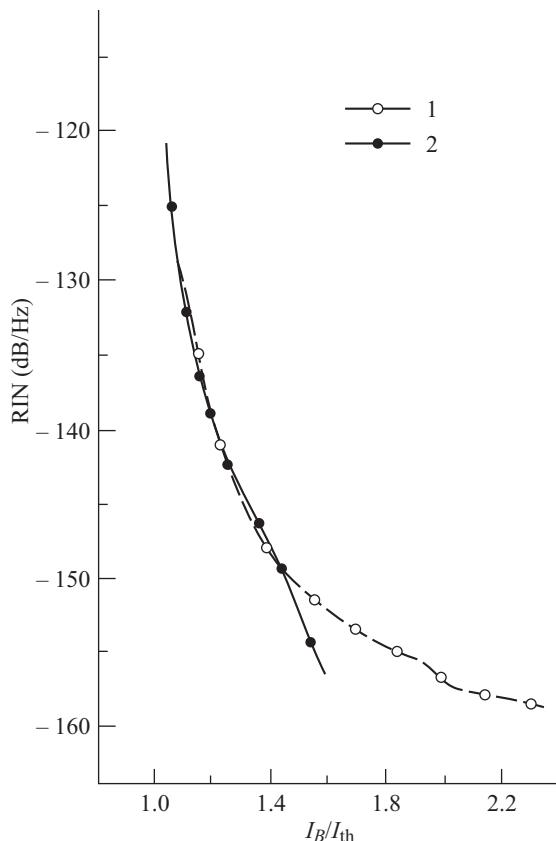
Fig. 9.3 *Example of the relative intensity noise (RIN) for two buried-heterostructure laser diodes. The noise level was measured at 100 MHz. (Reproduced with permission from Sato,[13] © 1983, IEEE.)*

Example 9.2 Figure 9.4 shows the RIN of an InGaAsP buried-heterostructure laser as a function of modulation frequency at several different bias levels.[1] The relative intensity noise is essentially independent of frequency below several hundred megahertz, and it peaks at the resonant frequency. In this case, at a bias level of 60 mA, which gives a 5-mW output, the RIN is typically less than −135 dB/Hz for modulation frequencies up to 8 GHz. For received optical signal levels of −13 dBm (50 μW) or less, the RIN of buried-heterostructure InGaAsP lasers lies sufficiently below the noise level of a 50-Ω amplifier with a 3-dB noise figure.

Fig. 9.5. These curves show the increase in relative intensity noise for bias points ranging from 1.24 to 1.62 times the threshold-current level. The feedback power ratio in Fig. 9.5 is the amount of optical power reflected back into the laser relative to the light output from the source. As an example, the dashed line shows that at $1.33I_{th}$ the feedback ratio must be less than −60 dB in order to maintain an RIN of less than −140 dB/Hz.
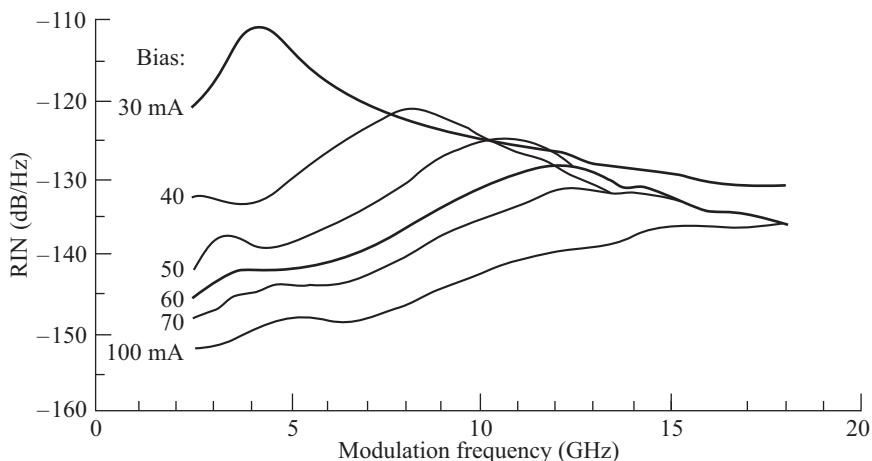
Fig. 9.4 *The RIN of an InGaAsP buried-heterostructure laser as a function of modulation frequency at several different bias levels. (Reproduced with permission from Olshansky, Lanzisera, and Hill,[1] © 1989, IEEE.)*
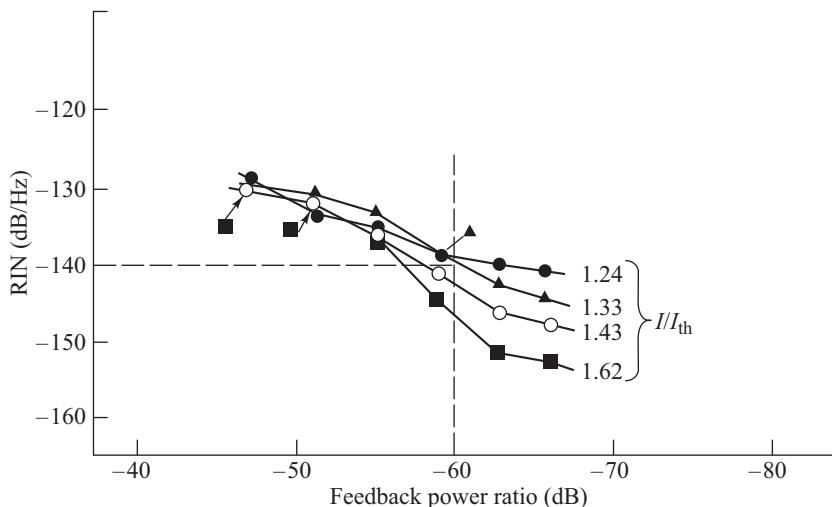
Fig. 9.5 *The increase in RIN due to back-reflected optical signals. (Reproduced with permission from Sato,[13] © 1983, IEEE.)*

9.2.5 Limiting Conditions

Let us now look at some limiting conditions. When the optical power level at the receiver is low, the preamplifier circuit noise dominates the system noise. For this, we have

$$\left(\frac{C}{N}\right)_{\text{limit 1}} = \frac{\frac{1}{2}(m\mathcal{R}M\bar{P})^2}{(4k_B T/R_{\text{eq}})B_e F_t} \tag{9.11}$$

In this case, the carrier-to-noise ratio is directly proportional to the square of the received optical power, so that for each 1-dB variation in received optical power, *C/N* will change by 2 dB.

For well-designed photodiodes, the bulk and surface dark currents are small compared with the shot (quantum) noise for intermediate optical signal levels at the receiver. Thus, at intermediate power levels the quantum noise term of the photodiode will dominate the system noise. In this case, we have

$$\left(\frac{C}{N}\right)_{\text{limit 2}} = \frac{\frac{1}{2}m^2 \mathcal{R}\overline{P}}{2qF(M)B_e} \tag{9.12}$$

so that the carrier-to-noise ratio will vary by 1-dB for every 1-dB change in the received optical power.

If the laser has a high RIN value so that the reflection noise dominates over other noise terms, then the carrier-to-noise ratio becomes

$$\left(\frac{C}{N}\right)_{\text{limit 3}} = \frac{\frac{1}{2}(mM)^2}{RIN\ B_e} \tag{9.13}$$

which is a constant. In this case, the performance cannot be improved unless the modulation index is increased.

Example 9.3 As an example of the limiting conditions, consider a link with a laser transmitter and a *pin* photodiode receiver having the following characteristics:

Transmitter	Receiver
$m = 0.25$	$\mathcal{R} = 0.6$ A/W
RIN = −143 dB/Hz	$B_e = 10$ MHz
$P_c = 0$ dBm	$I_D = 10$ nA
	$R_{\text{eq}} = 750\ \Omega$
	$F_t = 3$ dB

where P_c is the optical power coupled into the fiber. To see the effects of the different noise terms on the carrier-to-noise ratio, Fig. 9.6 shows a plot of *C/N* as a function of the optical power level at the receiver. In this case, we see that at high received powers the source noise dominates to give a constant *C/N*. At intermediate levels, the quantum noise is the main contributor, with a 1-dB drop in *C/N* for every 1-dB decrease in received optical power. For low light levels, the thermal noise of the receiver is the limiting noise term, yielding a 2-dB rolloff in *C/N* for each 1-dB drop in received optical power. It is important to note that the limiting factors can vary significantly depending on the transmitter and receiver characteristics. For example, for low-impedance amplifiers the thermal noise of the receiver can be the dominating performance limiter for all practical link lengths (see Prob. 9.1).

9.3 | Multichannel Transmission Techniques

So far, we have examined only the case of a single signal being transmitted over a channel. In broadband analog applications, such as cable television (CATV) supertrunks, one needs to send multiple analog signals over the same fiber. To do this, one can employ a multiplexing technique where a number of baseband signals are superimposed electronically on a set of *N* subcarriers that have different frequencies $f_1, f_2, ..., f_N$. These modulated subcarriers are then combined electrically through frequency-divison multiplexing (FDM) to form a composite signal that directly modulates a single optical source. Methods for achieving this include vestigial-sideband amplitude modulation (VSB-AM), frequency modulation (FM), and subcarrier multiplexing (SCM).
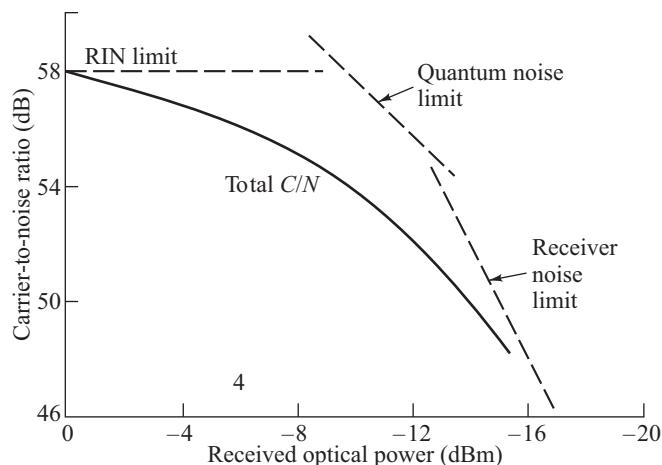
Fig. 9.6 *Carrier-to-noise ratio as a function of optical power level at the receiver. In this case, RIN dominates at high powers, quantum noise gives a 1-dB drop in C/N for each 1-dB power decrease at intermediate levels, and receiver thermal noise yields a 2-dB C/N roll-off per 1-dB drop in received power at low light levels*

Of these, AM is simple and cost-effective in that it is compatible with the equipment interfaces of a large number of CATV customers, but its signal is very sensitive to noise and nonlinear distortion. Although FM requires a larger bandwidth than AM, it provides a higher signal-to-noise ratio and is less sensitive to source nonlinearities. Microwave SCM operates at higher frequencies than AM or FM and is an interesting approach for broadband distribution of both analog and digital signals. To simplify the interface with existing coaxial cable systems, fiber links in CATV networks primarily use the AM-VSB scheme described in Sec. 9.3.1.

9.3.1 Multichannel Amplitude Modulation

The initial widespread application of analog fiber optic links, which started in the late 1980s, was to CATV networks.[14–17] These coax-based television networks operate in a frequency range from 50 to 88 MHz and from 120 to 550 MHz. The band from 88 to 120 MHz is not used, since it is reserved for FM radio broadcast. The CATV networks can deliver over 80 amplitude-modulated vestigial-sideband (AM-VSB) video channels, each having a noise bandwidth of 4 MHz within a channel bandwidth of 6 MHz, with signal-to-noise ratios exceeding 47 dB. To remain compatible with existing coax-based networks, a multichannel AM-VSB format is chosen for the fiber optic system.

Figure 9.7 depicts the technique for combining N independent messages. An information-bearing signal on channel i amplitude-modulates a carrier wave that has a frequency f_i, where $i = 1, 2..., N$. An RF power combiner then sums these N amplitude-modulated carriers to yield a composite frequency-division-multiplexed (FDM) signal that intensity-modulates a laser diode. Following the optical receiver, a bank of parallel bandpass filters separates the combined carriers back into individual channels. The individual message signals are recovered from the carriers by standard RF techniques.

For a large number of FDM carriers with random phases, the carriers add on a power basis. Thus, for N channels the optical modulation index m is related to the per-channel modulation index m_i by
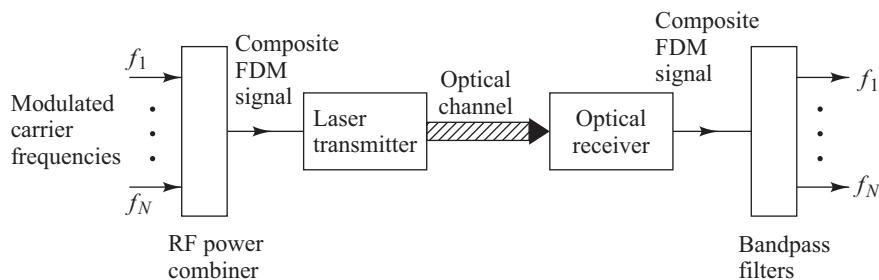
Fig. 9.7 *Standard technique for frequency-division multiplexing of N independent information-bearing signals*

$$m = \left(\sum_{i=1}^{N} m_i^2 \right)^{1/2} \tag{9.14a}$$

If each channel modulation index m_i has the same value m_c, then

$$m = m_c N^{0.5} \tag{9.14b}$$

As a result, when *N* signals are frequency-multiplexed and used to modulate a single optical source, the carrier-to-noise ratio of a single channel is degraded by 10 log *N*. If only a few channels are combined, the signals will add in voltage rather than power, so that the degradation will have a 20 log *N* characteristic.

When multiple carrier frequencies pass through a nonlinear device such as a laser diode, signal products other than the original frequencies can be produced. As noted in Sec. 4.4, these undesirable signals are called *intermodulation products*, and they can cause serious interference in both in-band and out-of-band channels. The result is a degradation of the transmitted signal. Among the intermodulation products, generally only the second-order and third-order terms are considered because higher-order products tend to be significantly smaller.

Third-order intermodulaton (IM) distortion products at frequencies $f_i + f_j - f_k$ (which are known as *triple-beat IM products*) and $2f_i - f_j$ (which are known as *two-tone third-order IM products*) are the most dominant, since many of these fall within the bandwidth of a multichannel system. For example, a 50-channel CATV network operating over a standard frequency range of $55.25-373.25$ MHz has 39 second-order IM products at 54.0 MHz and 786 third-order IM tones at 229.25 MHz. The amplitudes of the triple-beat products are 3-dB higher than the two-tone third-order IM products. In addition, since there are $N(N-1)(N-2)/2$ triple-beat terms compared with $N(N-1)$ two-tone third-order terms, the triple-beat products tend to be the major source of IM noise.

If a signal passband contains a large number of equally spaced carriers, several IM terms will exist at or near the same frequency. This so-called *beat stacking* is additive on a power basis. For example, for *N* equally spaced equal-amplitude carriers, the number of third-order IM products that fall right on the *r*th carrier is given by[18,19]

$$D_{1,2} = \frac{1}{2}\left\{ N - 2 - \frac{1}{2}[1 - (-1)^N](-1)^r \right\} \tag{9.15}$$

Table 9.1 *Distribution of the number of third-order triple-beat intermodulation products for the number of channels N ranging from 1 to 8*

	r							
N	1	2	3	4	5	6	7	8
1	0							
2	0	0						
3	0	1	0					
4	1	2	2	1				
5	2	4	4	4	2			
6	4	6	7	7	6	4		
7	6	9	10	11	10	9	6	
8	9	12	14	15	15	14	12	9

for two-tone terms of the type $2f_i - f_j$, and by

$$D_{1,1,1} = \frac{r}{2}(N - r + 1) + \frac{1}{4}\left\{(N - 3)^2 - 5 - \frac{1}{2}[1 - (-1)^N](-1)^{N+r}\right\} \qquad (9.16)$$

for triple-beat terms of the type $f_i + f_j - f_k$.

Whereas the two-tone third-order terms are fairly evenly spread through the operating passband, the triple-beat products tend to be concentrated in the middle of the channel passband, so that the center carriers receive the most intermodulation interference. Tables 9.1 and 9.2 show the distributions of the third-order triple-beat and two-tone IM products for the number of channels N ranging from 1 to 8.

Table 9.2 *Distribution of the number of third-order two-tone intermodulation products for the number of channels N ranging from 1 to 8*

	r							
N	1	2	3	4	5	6	7	8
1	0							
2	0	0						
3	1	0	1					
4	1	1	1	1				
5	2	1	2	1	2			
6	2	2	2	2	2	2		
7	3	2	3	2	3	2	3	
8	3	3	3	3	3	3	3	3

The results of beat stacking are referred to as *composite second order* (CSO) and *composite triple beat* (CTB) and describe the performance of multichannel AM links. The word *composite* means that the overall distortion is due to a collection of discrete distortions. CSO and CTO are defined as[20]

$$\text{CSO} = \frac{\text{peak carrier power}}{\text{peak power in composite 2nd-order IM tone}} \tag{9.17}$$

and

$$\text{CTB} = \frac{\text{peak carrier power}}{\text{peak power in composite 3rd-order IM tone}} \tag{9.18}$$

9.3.2 Multichannel Frequency Modulation

The use of AM-VSB signals for transmitting multiple analog channels is, in principle, straightforward and simple. However, it has a *C/N* requirement (or, equivalently, for AM, an *S/N* requirement) of at least 40 dB for each AM channel, which places very stringent requirements on laser and receiver linearity. An alternative technique is frequency modulation (FM), wherein each subcarrier is frequency-modulated by a message signal.[2,21] This requires a wider bandwidth (30 MHz versus 4 MHz for AM), but yields a signal-to-noise ratio improvement over the carrier-to-noise ratio.
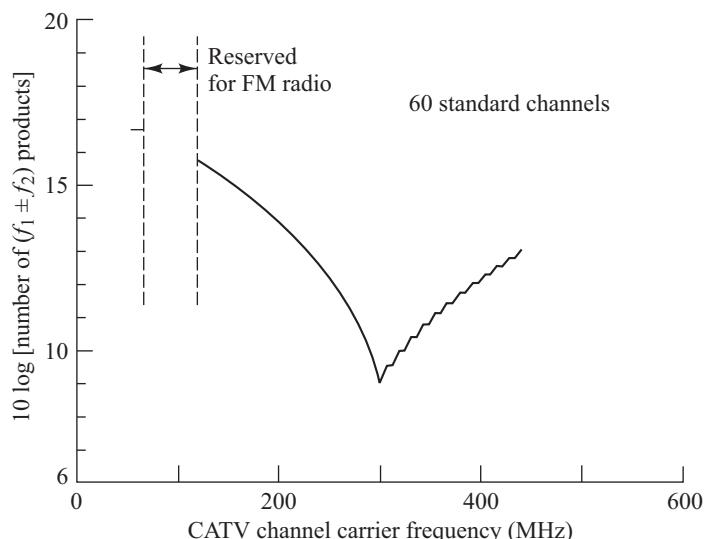
Fig. 9.8 *Predicted relative CSO performance for 60 amplitude-modulated CATV channels. The 88- to 120-MHz band is reserved for FM radio broadcast. (Reproduced with permission from Darcie, Lipson, Roxlo, and McGrath,[4] © 1990, IEEE.)*
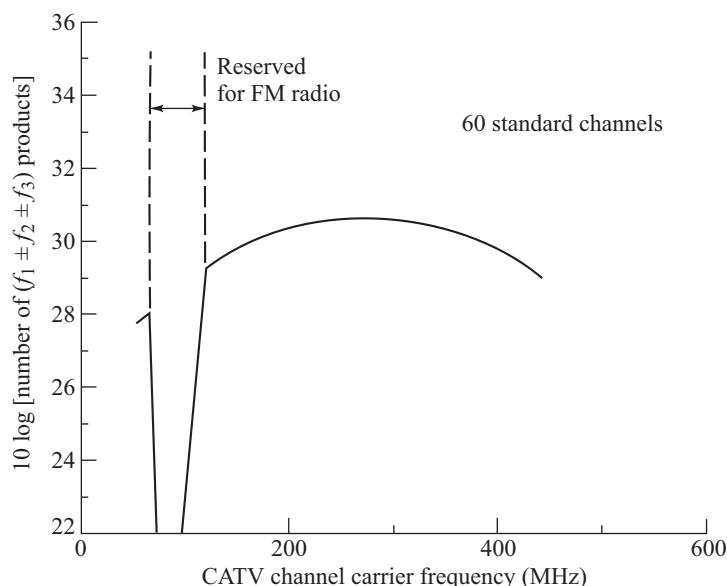
Fig. 9.9 *Predicted relative CTB performance for 60 amplitude-modulated CATV channels. The 88- to 120-MHz band is reserved for FM radio broadcast. (Reproduced with permission from Darcie, Lipson, Roxlo, and McGrath,[4] © 1990, IEEE.)*

The S/N at the output of an FM detector is much larger than the C/N at the input of the detector. The improvement is given by[2]

$$\left(\frac{S}{N}\right)_{\text{out}} = \left(\frac{C}{N}\right)_{\text{in}} + 10\log\left[\frac{3}{2}\frac{B_e}{f_v}\left(\frac{\Delta f_{\text{pp}}}{f_v}\right)^2\right] + w \tag{9.19}$$

where B_e is the required receiver bandwidth, Δf_{pp} is the peak-to-peak frequency deviation of the modulator, f_v is the highest video frequency, and w is a weighting factor used to account for the nonuniform response of the eye pattern to white noise in the video bandwidth. The total S/N improvement depends on the system design, but is generally in the range $36-44$ dB.[11,22] The reduced C/N requirements thus make an FM system much less susceptible to laser and receiver noises than an AM system.

9.3.3 Subcarrier Multiplexing

There is also great interest in using RF or *microwave subcarrier multiplexing* for high-capacity lightwave systems.[1–3,23] The term *subcarrier multiplexing* (SCM) is used to describe the capability of multiplexing both multichannel analog and digital signals within the same system.

Figure 9.12 shows the basic concept of an SCM system. The input to the transmitter consists of a mixture of N independent analog and digital baseband signals. These signals can carry either voice, data, video, digital audio, high-definition video, or any other analog or digital information. Each incoming signal $s_i(t)$ is mixed with a local oscillator (LO) having a frequency f_i. The local oscillator frequencies employed are in the 2- to 8-GHz range and are known as the *subcarriers*. Combining the modulated subcarriers gives a composite frequency-division-multiplexed signal, which is used to drive a laser diode.

Example 9.5 Figure 9.10 shows a plot of RIN versus optical modulation index per channel, comparing AM and FM broadcast TV systems.[2] The following assumptions were made in this calculation:

RIN noise dominates

$S/N = C/N + 40$ dB for the FM system

AM bandwidth per channel = 4 MHz

FM bandwidth per channel = 30 MHz

If the per-channel optical modulation index is 5 percent, then a RIN of less than -120 dB/Hz is needed for each FM TV program to have studio-quality reception, requiring $S/N \geq 56$ dB. This is easily met with a typical packaged laser diode, which has a nominal RIN value of -130 dB/Hz. On the other hand, for an AM system a laser with an RIN value of -140 dB/Hz can barely meet the CATV reception requirement of $S/N \geq 40$ dB.
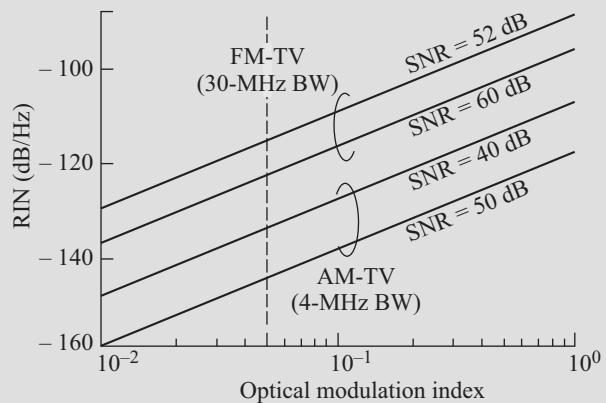
Fig. 9.10 *RIN versus the optical modulation index per channel for AM and FM video signals for several different signal-to-noise ratios (SNR). (Reproduced with permission from Way,[2] © 1989, IEEE.)*

Example 9.6 Another performance factor of AM transmission compared with FM is the limited power margin of AM. Figure 9.11 depicts the calculated power budget versus the optical modulation index (OMI) per channel for distribution of multichannel AM and FM video signals (see Sec. 4.5). The curves are given for different signal-to-noise ratios. The following assumptions were made in this calculation:

Laser power coupled into single-mode fiber = 0 dBm

RIN = -140 dB/Hz

pin Photodiode receiver with a 50-Ω front end

Preamplifer noise figure = 2 dB

AM bandwidth per channel = 4 MHz

FM bandwidth per channel = 30 MHz

Again, assuming a per-channel optical modulation index of 5 percent, the AM system has a power margin of about 10 dB for a 40-dB signal-to-noise ratio, whereas the FM system has a power margin of 20 dB for $S/N = 52$ dB.
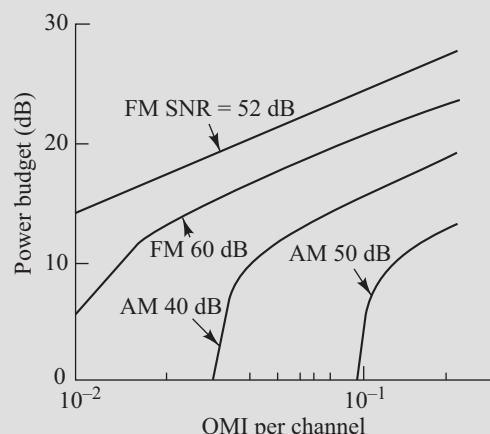
Fig. 9.11 *Power budget versus the optical modulation index (OMI) per channel for distribution of multichannel AM and FM video signals. (Reproduced with permission from Way,[2] © 1989, IEEE.)*

At the receiving end, the optical signal is directly detected with a high-speed wideband InGaAs *pin* photodiode and reconverted to a microwave signal. For long-distance links, one can also employ a wideband InGaAs avalanche photodiode with a 50- to 80-GHz gain bandwidth product or use an optical preamplifier. For amplifying the received microwave signal, one can use a commercially available wideband low-noise amplifier or a *pin*-FET receiver.
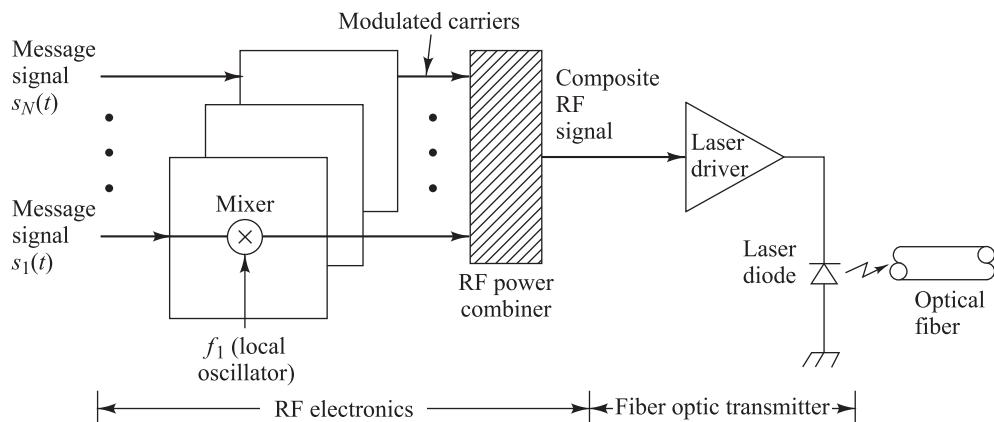
Fig. 9.12 *Basic concept of subcarrier multiplexing. One can send analog and digital signals simultaneously by frequency-division multiplexing them on different subcarrier frequencies.*

9.4 | RF over Fiber

Radio frequency (RF) signals at microwave and millimeter-wave frequencies are used in applications such as radars, satellite links, broadband terrestrial radios, and cable television networks. The signal ranges include the 0.3- to 3-GHz ultra-high frequency (UHF) band, the 3- to 30-GHz super-high frequency (SHF) region, and the 30- to 300-GHz extremely high frequency (EHF) range. Traditionally these RF systems used wireless or coaxial cable links for transporting the microwave signals from a receiving element (for example, an antenna) to a signal processing center, which could be located hundreds of meters away. As noted in Chapter 1, optical fibers offer certain advantages over coaxial cables, such as a smaller size, lower losses, wider bandwidths, and insensitivity to electromagnetic interference effects. Compared to coaxial lines, these factors allow easier deployment of fiber links over greater distances. Consequently, there has been much interest in developing and deploying high-speed optical fiber links for transporting microwave and millimeter-wave signals in their original analog format.

The methods for transmitting microwave analog signals over an optical fiber link have become known as *RF-over-fiber* techniques. This section examines the basics of RF-over-fiber techniques. Section 9.5 gives an application example of *radio-over-fiber* (ROF) systems for interconnecting antenna base stations with the central controlling office. A radio-over-fiber system also can be used as an extension of a FTTH access network for distributing RF signals to individual rooms in a home or other indoor environment. In addition to developing analog signal transmission links, much work has been done on studying the interaction of microwave and optical signals for the generation, distribution, control, and processing of microwave wave signals by means of photonics. Section 9.6 discusses this discipline, which has become known as *microwave photonics*.

9.4.1 Key Link Parameters

Figure 9.13 shows the constituents of a generic RF-over-fiber link. The three major modules are an RF-to-optical signal converting device at the transmitting end, an optical-to-RF signal converting device at the receiving end, and an optical fiber that joins these two modules. The primary parameters used to characterize the RF performance of the optical link are the gain, noise figure, and spur-free dynamic range (SFDR).
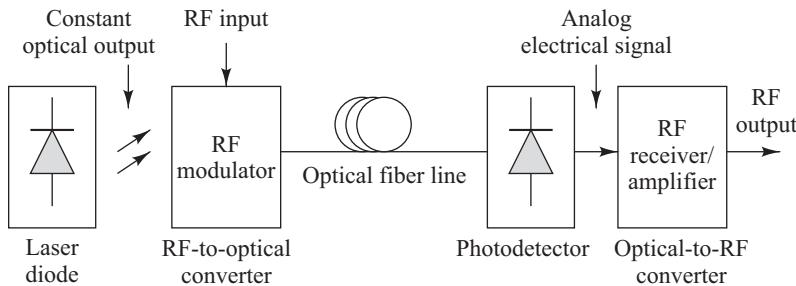
Fig. 9.13 *Basic constituents of a generic RF-over-fiber link*

The *link gain* G is defined as the ratio of the RF power P_{out} generated in the photodetector load resistor to the RF power input P_{in} to the laser transmitter. Thus, for a directly modulated link, the gain is[24,25]

$$G = \frac{P_{out}}{P_{in}} = S_M^2 \eta_{LF}^2 T_F^2 \eta_{FD}^2 \mathscr{R}^2 \frac{R_{load}}{R_M} \tag{9.20}$$

where S_M is the slope efficiency (given in watts per ampere) of the modulation device, η_{LF} is the laser-to-fiber coupling efficiency, T_F is the fiber transmission efficiency, η_{FD} is the fiber-to-detector coupling efficiency, $\mathscr{R}$ is the photodetector responsivity (given in amperes per watt), R_{load} is the detector load resistance, and R_M is the modulator resistance. Table 9.3 gives some typical values of these parameters for a 500-m link. The value of 0.7 for the fiber transmission efficiency implies a fiber loss of 3.5 dB/km at 850 nm. Gain values that are less than 1 represent a link loss. The two major contributors to the maximum value of the gain are the limits of the modulator slope efficiency and the photodetector responsivity, since the values of the other parameters do not vary significantly.

The *noise figure* (NF) represents a measure of the degradation in the signal-to-noise ratio (SNR) between the input and the output of the link. It is defined in decibels by

$$NF = 10 \log \frac{SNR_{in}}{SNR_{out}} = 10 \log \frac{\bar{N}_{out}/B_e}{k_B T G} = 10 \log \frac{N_{out}}{k_B T G} \tag{9.21}$$

where the input noise is the thermal noise power generated by a matched resistive load $k_B T$ held at $T = 290°K$. The parameter k_B is Boltzmann's constant and B_e is the noise bandwidth of the electronic

Table 9.3 *Typical parameter values for a 500-m RF-over-fiber link*

Parameter	Symbol	Value
Slope efficiency	S_M	0.3 W/A
Coupling efficiencies	η_{LF} and η_{FD}	0.80
Fiber transmission efficiency	T_F	0.7
Detector load resistance	R_{load}	50 Ω
Modulator resistance	R_M	45 Ω
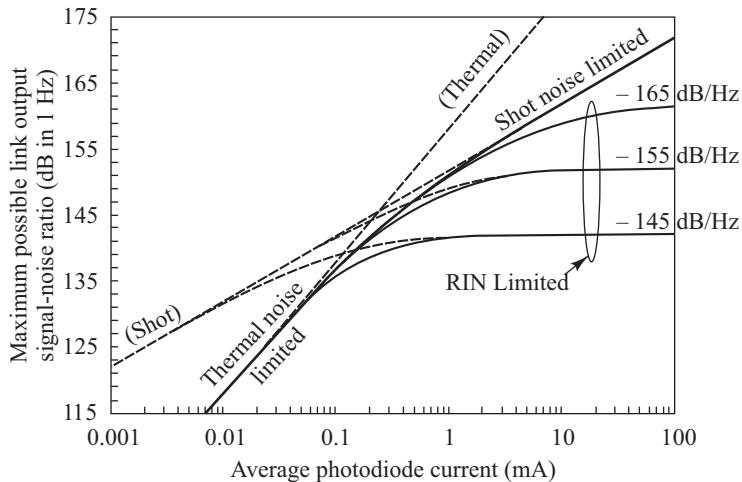
Fig. 9.14 *Limiting conditions due to thermal noise, shot noise, and RIN on the maximum possible SNR as a function of the photodetector current. (Reproduced with permission from Cox, Ackermann, Betts, and Prince,[25] © 2006, IEEE.)*

receiver. Here we use the designation $\bar{N}_{out}$ to denote the total output noise power in the bandwidth B_e. Since $\bar{N}_{out}$ is proportional to B_e, then N_{out} is the noise power per unit bandwidth, so that the NF is independent of the noise bandwidth.

At the link output the noise power is due to the laser RIN, photodetector shot noise, and thermal noise. Their contributions per unit noise bandwidth (in A^2/Hz) are given by

$$N_{out, RIN} = I_p^2 \, RIN \tag{9.22}$$

$$N_{out, shot} = 2qI_p \tag{9.23}$$

$$N_{out, thermal} = 4k_B T / R_{load} \tag{9.24}$$

where I_p is the average photodiode current in the receiver.

Figure 9.14 shows some limiting conditions on the maximum possible signal-to-noise ratio as a function of the photodetector current (that is, the received optical power level).[25] Thermal noise imposes a poorly performing limit on the SNR for low values of the photodetector current. At higher detector currents, the RIN of the laser restricts the SNR to an upper value that cannot be exceeded even if the photodiode current is increased. Figure 9.14 shows these limits with three typical values of laser RIN.

9.4.2 Spur-Free Dynamic Range

The dynamic range of an analog link is defined in relation to two-tone third-order intermodulation frequencies. First consider two large equal-power signals at fundamental frequencies f_1 and f_2, as shown in Fig. 9.15. These two signals will produce second-order modulation products at $2f_1$, $2f_2$, and $f_1 \pm f_2$, and third-order intermodulation products at frequencies $2f_1 \pm f_2$ and $2f_2 \pm f_1$. The second-order terms normally fall outside of the passband of a system, so they can be ignored. However, the third-order products are of concern because they could fall on a signal frequency within the system bandwidth and cannot be
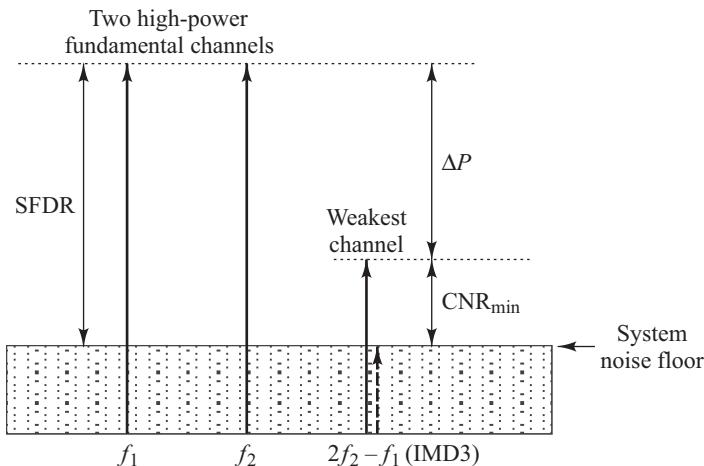
Fig. 9.15   *The relationship of third-order intermodulation products (dashed line) to system operating requirements*

removed by a simple filtering technique. To determine system operating requirements, consider the case shown in Fig. 9.15. Here a third-order intermodulation product (dashed line) resulting from the two strongest fundamental carriers falls at the frequency where the weakest channel operates. The parameter ΔP is the power difference between the strongest and weakest channels, and CNR_{min} is the minimum required carrier-to-noise ratio for the weakest signal. For the case shown in Fig. 9.15, the intermodulation products resulting from the strongest equal-power fundamental carriers are equal to the noise floor.

For standard analog links the third-order intermodulation distortion varies as the cube of the RF input power. Figure 9.16 shows this relationship and also shows the linear relationship of the output power of the fundamentals as a function of the RF input power. The *spur-free dynamic range* (SFDR) is defined as the ratio between the powers in the fundamental carrier and the third-order intermodulation (designated by IMD3) at that power level where the IMD3 is equal to the noise floor. This means that the SFDR is the usable dynamic range before spurious noise interferes with or distorts the fundamental signal. Thus, referring back to Fig. 9.15, the SFDR must be larger than $CNR_{min} + \Delta P$.

In Fig. 9.16 the point IP3 designates the input power at which the IMD3 is equal to the output carrier power. From the curves in Fig. 9.16 the SFDR is given by

$$\text{SFDR} = \frac{2}{3} 10 \log \frac{IP3}{N_{\text{out}}\, R_{\text{load}}} \tag{9.25}$$

The SFDR is measured in units of $dB \cdot Hz^{2/3}$. Many different measurements of SFDR that have been reported in the literature are summarized in Ref. 25. The general trends of these measurement values as a function of frequency are that directly modulated microwave links can have a large SFDR (up to $125\ dB \cdot Hz^{2/3}$ at 1 GHz), but the SFDR decreases significantly as the frequency increases beyond about 1 GHz. This is due to inherent distortion effects in the laser, which get worse as the operating frequency gets closer to the relaxation-oscillation peak (see Fig. 4.32). The SFDR for externally modulated links is not as high below 1 GHz as for direct modulation setups, but it remains at the upper level out to higher frequencies. For example, links using a Mach–Zehnder interferometer-based modulator are able to maintain a $112\text{-}dB \cdot Hz^{2/3}$ SFDR out to about 17 GHz.
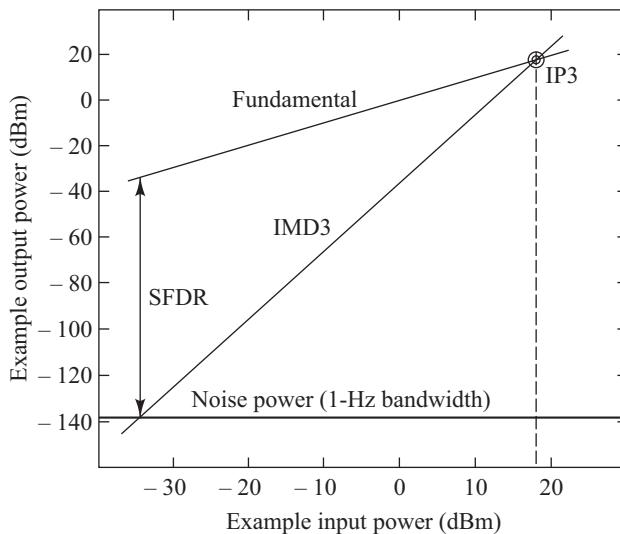
Fig. 9.16 *SFDR is the ratio of the power in the fundamental carrier to the IMD3 at the input power level where the IMD3 power equals the noise level.*

9.5 | Radio-over-Fiber Links

The transition in the use of wireless devices from pure voice communications to a wide selection of broadband services created much interest in developing radio-over-fiber (ROF) links.[26–43] The convergence of optical and wireless access networks is driven by the need to have seamless connectivity between an access network and both stationary and fast-moving mobile users at data rates of at least 2.5 Gb/s. The implementations of ROF links include the interconnection of antenna base stations with a central controlling office in a wireless access network, access to wireless services for indoor environments (e.g., large office buildings, airport departure lounges, hospitals, conference centers, hospitals, and hotel rooms), and connections to personal area networks in homes. Services of interest to mobile users include broadband Internet access, fast peer-to-peer file transfers, high-definition video, and online multiparty gaming. Depending on the network type, the optical links can use single-mode fibers, 50- or 62.5-μm core-diameter multimode glass fibers, or large-core multimode polymer optical fibers.

9.5.1 Antenna Base Station ROF Network

One application of RF-over-fiber technology is in broadband wireless access networks for interconnecting antenna base stations (BSs) with the central controlling office. Figure 9.17 shows the basic network architecture for such a scheme. Here a collection of antenna base stations provide wireless connectivity to subscribers by means of millimeter-wave frequencies. Subscribers are located up to 1 km from a local base station. The transmission range around a BS is called a *microcell* (diameter less than 1 km) or a *picocell* or *hotspot* (radii ranging from 5 to 50 m). The BSs are connected to a microcell control station (CS) in the central office, which is responsible for functions such as RF modulation and demodulation, channel control, and switching and routing of customer calls.

Due to the advantageous transmission characteristics of optical fibers (low loss, high immunity to electromagnetic interference, and a wide bandwidth), a great interest arose in using fibers to connect
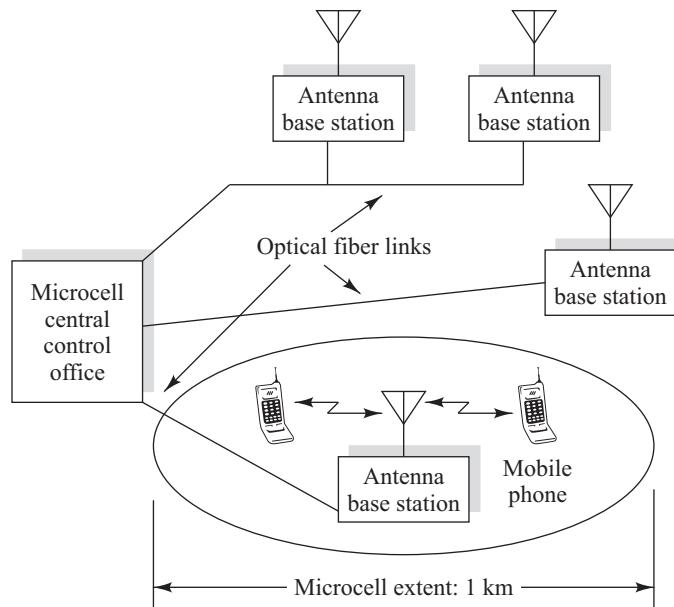
Fig. 9.17 *Radio-over-fiber concept of a broadband wireless access network for interconnecting antenna base stations with the central controlling office*

the base stations to the central office. In addition, individual BSs can be connected independently to the microcell CS through wavelength division multiplexing (WDM) techniques by using a separate unique wavelength for each BS. The network deployment enables all the equipment that performs complex signal processing and call routing to be located in the central office. This structure thus distributes the cost of the central equipment among inexpensive base stations that need to do only amplification plus electrical-to-optical and optical-to-electrical signal conversion.

9.5.2 Radio-over-Multimode Fiber

Fiber-to-the-home (FTTH) access networks based on passive optical network (PON) techniques have brought optical fiber links up to the doorstep of family dwellings (see Chapter 13). The next challenge was to extend the broadband capacity offered by optical fibers into family residences and other indoor environments. Figure 9.18 shows the concept for an in-building residential network using optical fibers to distribute radio frequency signals to individual rooms by means of a residential gateway (RG). Such a network can accommodate both fixed terminals (e.g., desktop PCs, high-definition TV sets, scanners, printers, and home servers) and wireless devices (e.g., phones, netbooks, USB modems, and gaming headsets). The wireless distribution system shown in Fig. 9.18 uses radio carriers in the millimeter-wave range. These frequencies have high propagation losses and are not able to penetrate walls. Consequently, the signals from individual antennas are easily confined to a single room. Each room thus becomes a picocell that does not interfere with an adjacent room, which is another picocell.

The indoor network cabling commonly uses 50- or 62.5-μm core-diameter multimode glass fiber or multimode polymer optical fiber (POF) with core diameters ranging from 50 μm to 1 mm. The advantages of POF include easy cable pulling because of its flexibility and durability and easy optical connector attachment to the fiber. A polymer optical fiber can be cut with a razor blade and then a metal
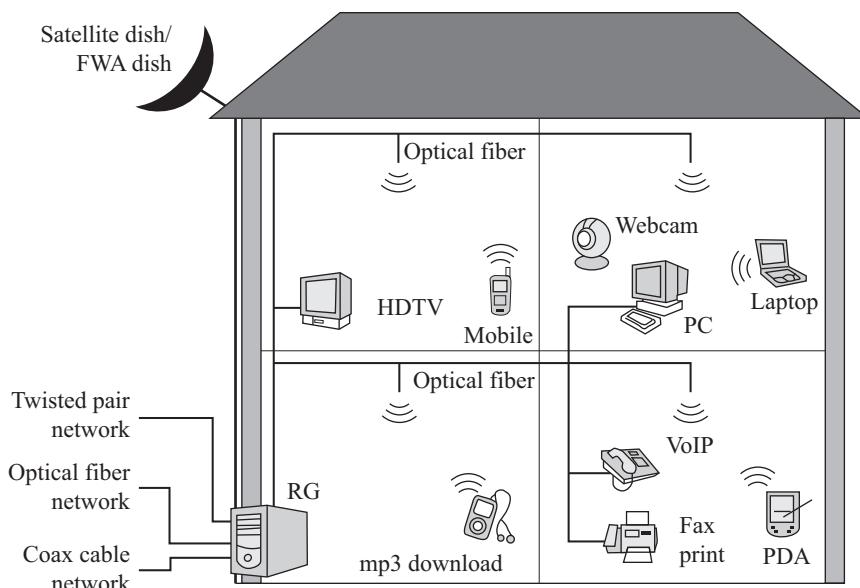
Fig. 9.18   *In-building ROF network using optical fiber. (Reproduced with permission from Koonen and Larrodé,[36] © 2008, IEEE.)*

connector ferrule can be crimped directly onto the fiber. This results in a simple low-cost installation. Standard step-index polymer optical fiber is used mainly with visible-wavelength light sources and has an attenuation of about 0.1 dB/m at 560 nm. Perfluorinated graded-index polymer optical fibers exhibit attenuations of around 10 dB/km in the 850 to 1300-nm region and support data rates up to 10 Gb/s over 100-m distances.

9.6   Microwave Photonics

The field of microwave photonics encompasses the study and applications of photonic devices operating at microwave frequencies. The key components being developed and applied include the following:

- High-frequency low-loss external optical modulators that have linear transfer functions and can withstand continuous-wave optical powers up to 60 mW
- Optical sources with high slope efficiencies and low RIN that can be modulated at tens of GHz
- High-speed photodiodes and optical receivers that can respond to signal frequencies of 20 to 60 GHz
- Microwave photonic filters that perform the same tasks as standard RF filters

    In addition to device developments, microwave photonics also addresses optical signal processing at microwave speeds and the design and implementation of RF photonic transmission systems. For example, applications of photonic signal processing at multiple-gigahertz sampling frequencies include signal filtering, analog-to-digital conversion, frequency conversion and mixings, signal correlation, generation of arbitrary waveforms, and beam-forming methodologies for phased array radars. The literature contains a number of reviews about these and other emerging techniques.[43–57]

PROBLEMS

9.1 Commercially available wideband receivers have equivalent resistances $R_{eq} = 75\ \Omega$. With this value of R_{eq} and letting the remaining transmitter and receiver parameters be the same as in Example 9.3, plot the total carrier-to-noise ratio and its limiting expressions, as given by Eqs. (9.10) through (9.13), for received power levels ranging from 0 to -16 dBm. Show that the thermal noise of the receiver dominates over the quantum noise at all power levels when $R_{eq} = 75\ \Omega$.

9.2 Consider a five-channel frequency-division-multiplexed (FDM) system having carriers at $f_1, f_2 = f_1 + \Delta, f_3 = f_1 + 2\Delta, f_4 = f_1 + 3\Delta$, and $f_5 = f_1 + 4\Delta$, where Δ is the spacing between carriers. On a frequency plot, show the number and location of the triple-beat and two-tone third-order intermodulation products.

9.3 Suppose we want to frequency-division multiplex 60 FM signals. If 30 of these signals have a per-channel modulation index $m_i = 3$ percent and the other 30 signals have $m_i = 4$ percent, find the optical modulation index of the laser.

9.4 Consider an SCM system having 120 channels, each modulated at 2.3 percent. The link consists of 12 km of single-mode fiber having a loss of 1 dB/km, plus a connector having a 0.5-dB loss on each end. The laser source couples 2 mW of optical power into the fiber and has RIN $= -135$ dB/Hz. The *pin* photodiode receiver has a responsivity of 0.6 A/W, $B_e = 5$ GHz, $I_D = 10$ nA, $R_{eq} = 50\ \Omega$, and $F_t = 3$ dB. Find the carrier-to-noise ratio for this system.

9.5 What is the carrier-to-noise ratio for the system described in Prob. 9.4 if the *pin* photodiode is replaced with an InGaAs avalanche photodiode having $M = 10$ and $F(M) = M^{0.7}$?

9.6 Consider a 32-channel FDM system with a 4.4-percent modulation index per channel. Let RIN $= -135$ dB/Hz, and assume the *pin* photodiode receiver has a responsivity of 0.6 A/W, $B_e = 5$ GHz, $I_D = 10$ nA, $R_{eq} = 50\ \Omega$, and $F_t = 3$ dB.

(*a*) Find the carrier-to-noise ratio for this link if the received optical power is -10 dBm.

(*b*) Find the carrier-to-noise ratio if the modulation index is increased to 7 percent per channel and the received optical power is decreased to -13 dBm.

9.7 For a fiber optic link using a single-longitudinal-mode laser with a 3-dB linewidth of Δv and having two fiber connectors with reflectivities R_1 and R_2, the worst-case RIN occurs when the direct and doubly reflected optical fields interfere in quadrature.[51] If τ is the light-travel time in the fiber, this is described by

$$\text{RIN}(f) = \frac{4R_1 R_2}{\pi} \frac{\Delta v}{f^2 + \Delta v^2}$$
$$\times [1 + e^{-4\pi\Delta v\tau} - 2e^{-2\pi\Delta v\tau} \cos(2\pi f\tau)]$$

where f is the carrier frequency. Show that this expression reduces to

$$\text{RIN}(f) = \frac{16R_1 R_2}{\pi} \Delta v\, \tau^2 \quad \text{for} \quad \Delta v \cdot \tau \ll 1$$

and

$$\text{RIN}(f) = \frac{4R_1 R_2}{\pi} \frac{\Delta v}{f^2 + \Delta v^2} \quad \text{for} \quad \Delta v \cdot \tau \gg 1$$

9.8 A typical DFB laser has a linewidth 1 MHz $< \Delta v < 40$ MHz, and with a 1- to 10-m optical jumper cable we have $0.005 < \Delta v \cdot \tau < 2$. With a 1-m jumper fiber and letting $\Delta v = f = 10$ MHz, use the expression in Prob. 9.7 to show that to achieve an RIN of less than -140 dB/Hz, the average reflectivity per connector should be less than -30 dB.

9.9 It is possible to achieve CATV supertrunk applications with lengths greater than 40 km by cascading standard fiber-optic CATV transmitters, which act as amplifiers. These systems are typically limited by the required CTB performance.[52] When two amplifiers are cascaded, their individual CTB products add as

$$\text{CTB}_{\text{cascade}} = x \log(10^{\text{CTB}_1/x} + 10^{\text{CTB}_2/x})$$

where CTB_i is the composite triple beat of amplifier i.

(a) For identical amplifiers, $x = 20$. What is the CTB power penalty in this case?

(b) When the amplifiers are different, the value of x can vary from 0 (cancellation of beats between the two amplifiers) to 20 (all beats of the two amplifiers are in phase). Find the values of x for the following experimental measurements on dissimilar amplifiers where dBc is the power relative to the carrier:

CTB_1 (dBc)	CTB_2 (dBc)	$CTB_{cascade}$ (dBc)
−75.2	−69.9	−70.5
−74.7	−71.4	−71.0
−72.1	−71.3	−66.7

9.10 Consider a subcarrier-multiplexed CATV distribution system that has N channels. The CNRs of the three fundamental irreducible sources of signal degradation in a fiber link containing an optical amplifier are as follows:

(1) Nonlinear distortion resulting from clipping of the laser output,

$$\text{CNR}_{\text{clip}} = \sqrt{2\pi}\,\frac{(1 + 6\mu^2)}{\mu^3}\,e^{1/2\mu^2}$$

Here, the rms modulation index $\mu = m\sqrt{N/2}$ where m is the modulation depth per channel. This distortion arises when the modulation depth m is increased to the point where the signal starts getting clipped at the lasing threshold level.

(2) Shot noise in the photodetector,

$$\text{CNR}_{\text{shot}} = \frac{\mathscr{R}\,m^2 G P_{\text{in}} L}{4q B_e}$$

where $\mathscr{R}$ is the photodetector responsivity, G is the optical amplifier gain, P_{in} is the input signal power, L is the postamplifier loss, q is the electron charge, and B_e is the receiver noise bandwidth.

(3) Signal-spontaneous emission beat noise,

$$\text{CNR}_{\text{sig-sp}} = \frac{m^2 G P_{\text{in}}}{8 h v n_{\text{sp}} (G - 1) B_e}$$

where n_{sp} is the population-inversion factor in the optical amplifier and $v = c/\lambda$ is the signal frequency. As discussed in Chapter 11, this noise arises when the amplified-spontaneous-emission (ASE) noise generated by the optical amplifier beats with the optical signal in the photodetector.

The total CNR at the receiver is then given by Eq. (9.1). For simplicity of calculation, this can be written in terms of the noise-to-carrier ratio NCR = 1/CNR, so that Eq. (9.1) becomes

$$\text{NCR}_{\text{total}} = \text{NCR}_{\text{clip}} + \text{NCR}_{\text{shot}} + \text{NCR}_{\text{sig-sp}}$$

Using the parameter values given in Table P 9.10, plot expressions (a) to (c) below on the same graph as a function of the rms modulation index over the range $0.04 \le \mu \le 0.36$:

(a) NCR_{clip}

(b) $\text{NCR}_{\text{shot}} + \text{NCR}_{\text{sig-sp}}$

(c) $\text{NCR}_{\text{total}}$

(d) Explain where the minimum NCR (maximum CNR) occurs on the plot.

Table P9.10

Parameter	Value
P_{in}	1 mW
m	$0.68\sqrt{N}$
G	100 (20 dB)
n_{sp}	2
λ	1551 nm
B_e	1×10^9 Hz
$\mathscr{R}$	0.6 A/W
L	40/3
N	60

9.11 (*a*) Calculate the gain for a directly modulated analog link using the parameter values given in Table 9.3 and letting $R = 0.6$ A/W at 850 nm.

(*b*) What is the gain if the link length is increased to 1.0 km?

9.12 Consider a medical facility housed in a four-story building with 15 rooms per floor. The facility offers general and emergency patient care, various radiological procedures, laboratory and diagnostic services, and specialized health treatments. Design a radio-over-fiber network for this facility that can be used for a wide variety of functions, such as locating physicians, requesting diagnostic results, checking records of patients, and calling for mobile test devices.

9.13 The external modulator is a key device in a photonic microwave link, since it usually determines how well the link performs. From the literature or using online resources, discuss the following characteristics of a lithium niobate (LiNbO$_3$) Mach–Zehnder interferometer-based modulator: voltage sensitivity, impedance, optical loss, optical power-handling capability, linearity, and environmental stability. Reference 25 provides a good starting point for examining the literature.

9.14 The literature describes various types of microwave photonic filters that are based on the use of spectral-slicing techniques of broadband optical sources. Choose one slicing technique and describe the tuning and reconfiguration capabilities of the resulting microwave photonic filter. Reference 43 provides a good starting point for examining the literature.

REFERENCES

1. R. Olshansky, V. A. Lanzisera, and P. M. Hill, "Subcarrier multiplexed lightwave systems for broadband distribution," *J. Lightwave Tech.,* vol. 7, pp. 1329–1342, Sept. 1989.

2. W. I. Way, "Subcarrier multiplexed lightwave system design considerations for subscriber loop applications," *J. Lightwave Tech.,* vol. 7, pp. 1806–1818, Nov. 1989.

3. W. I. Way, *Broadband Hybrid Fiber/Coax Access System Technologies*. Academic, New York, 1998.

4. T. E. Darcie, J. Lipson, C. B. Roxlo, and C. J. McGrath, "Fiber optic device technology for broadband analog video systems," *IEEE Mag. Lightwave Commun.*, vol. 1, pp. 46–52, Feb. 1990.

5. A.S. Andrawis and I. Jacobs, "A new compound modulation technique for multichannel analog video transmission on fiber," *J. Lightwave Tech.,* vol. 11, pp. 49–54, Jan. 1993.

6. J. J. Lee, R. Y. Loo, S. Livingston, V. I. Jones, J. B. Lewis, H. W. Yen, G. L. Tangonan, and M. Wechsberg, "Photonic wideband array antennas," *IEEE Trans. Antennas Propagat.*, vol. 43, pp. 966–982, Sept. 1995.

7. R. Hui, B. Zhu, R. Huang, C. T. Allen, K. R. Demarest, and D. Richards, "Subcarrier multiplexing for high-speed optical transmission," *J. Lightwave Tech.,* vol. 20, pp. 417–427, Mar. 2002.

8. E. I. Ackerman and A. S. Daryoush, "Broadband external modulation fiber-optic links for antenna-remoting applications," *IEEE Trans. Microwave Theory Tech.*, vol. 45, pp. 1436–1442, Aug. 1997.

9. S. Ovadia, *Broadband Cable TV Access Networks: From Technologies to Applications*, Prentice Hall, Upper Saddle River, NJ, 2001.

10. Special Issues on Microwave Photonics: (*a*) *IEEE Microwave Theory Tech.*, vol. 49, part II, Oct. 2001; (*b*) *J. Lightwave Tech.,* vol. 21, Dec. 2003; (*c*) *IEEE Microwave Theory Tech.*, vol. 54, part II, Feb. 2006.

11. L. W. Couch II, *Digital and Analog Communication Systems*, Prentice Hall, Upper Saddle River, NJ, 7th ed., 2007.

12. A. J. Rainal, "Laser clipping distortion in analog and digital channels," *J. Lightwave Tech.*, vol. 15, pp. 1805–1807, Oct. 1997.

13. K. Sato, "Intensity noise of semiconductor laser diodes in fiber optic analog video transmission," *IEEE J. Quantum Electron.*, vol. 19, pp. 1380–1391, Sept. 1983.

14. M. R. Phillips and T. E. Darcie, "Lightwave video transmission," in I. P. Kaminov and T. L. Koch, eds., *Optical Fiber Telecommunications–IIIA*, Academic, New York, 1997.

15. J. P. Lee, T. L. Tsao, and H. W. Tsao, "Broadband hybrid analog/digital CATV trunk network," *IEEE Trans. Broadcasting*, vol. 45, pp. 339–347, Sept. 1999.

16. R. Rabbat and K. Y. Sui, "QoS support for integrated service over CATV," *IEEE Commum. Mag.*, vol. 37, pp. 64–68, Jan. 1999.

17. S. Dravida, D. Gupta, S. Nanda, K. Rege, J. Strombosky, and M. Tandon, "Broadband access over cable for next-generation services: A distributed switch architecture," *IEEE Commum. Mag.*, vol. 40, pp. 116–124, Aug. 2002.

18. J. H. Schaffner and W. B. Bridges, "Intermodulation distortion in high dynamic range microwave fiber-optic links with linearized modulators," *J. Lightwave Tech.*, vol. 11, pp. 3–6, Jan. 1993.

19. C. Cox, *Analog Optical Links*. Cambridge University Press, Cambridge, U.K., 2004.

20. *NCTA Recommended Practices for Measurements on Cable Television Systems*, National Cable Television Association, 3rd ed., 2002.

21. R. Olshansky and V. A. Lanzisera, "60-channel FM video subcarrier-multiplexed optical communication system." *Electron. Lett.*, vol. 23. pp. 1196–1198, 1987.

22. F. V. C. Mendis and P. A. Rosher, "CNR requirements for subcarrier-multiplexed multichannel video FM transmission in optical fibre," *Electron. Lett.*, vol. 25, pp. 72–74, Jan. 1989.

23. W. H. Chen and W. I. Way, "Multichannel single-sideband SCM/DWDM transmission systems," *J. Lightwave Tech.*, vol. 22, pp. 1679–1693, July 2004.

24. C. H. Cox, E. I. Ackerman, R. Helkey, and G. E. Betts, "Techniques and performance of intensity-modulation direct-detection analog optical links," *IEEE Trans. Microwave Theory Tech.*, vol. 45, pp. 1375–1383, Aug. 1997.

25. C. H. Cox, E. I. Ackerman, G. E. Betts, and J. L. Prince, "Limits on the performance of RF-over-fiber links and their impact on device design," *IEEE Microwave Theory Tech.*, vol. 54, pp. 906–920, Feb. 2006.

26. C. Carlsson, A. Larsson, and A. Alping, "RF transmission over multimode fibers using VCSELs—Comparing standard and high-bandwidth multimode fibers," *J. Lightwave Tech.*, vol. 22, pp. 1694–1700, July 2004.

27. T. Kuri and K.-I. Kitayama, "Optical heterodyne detection technique for densely multiplexed millimeter-wave-band radio-on-fiber systems," *J. Lightwave Tech.*, vol. 21, pp. 3167–3179, Dec. 2003.

28. T. Kurniawan, A. Nirmalathas, C. Lim, D. Novak, and R. Waterhouse, "Performance analysis of optimized millimeter-wave fiber radio links," *IEEE Microwave Theory Tech.*, vol. 54, pp. 921–928, Feb. 2006.

29. T. Niiho, M. Nakaso, K. Masuda, H. Sasai, K. Utsumi, and M. Fuse, "Transmission performance of multichannel wireless LAN system based on radio-over-fiber techniques," *IEEE Microwave Theory Tech.*, vol. 54, pp. 980–989, Feb. 2006.

30. P. K. Tang, L. C. Ong, A. Alphones, B. Luo, and M. Fujise, "PER and EVM measurements of a radio-over-fiber network for cellular WLAN system applications," *J. Lightwave Tech.*, vol. 22, pp. 2370–2376, Nov. 2004.

31. H. Al-Raweshidy and S. Komaki, eds., *Radio over Fiber Technologies*, Artech House, Boston, 2002.

32. M. Sauer, A. Kobyakov, and J. George, "Radio over fiber for picocellular network architectures," *J. Lightwave Tech.*, vol. 25, no. 11, pp. 3301–3320, Nov. 2007.

33. J. J. Vegas Olmos, T. I. Kuri, and K.-I. Kitayama, "Dynamic reconfigurable WDM 60-GHz millimeter-waveband radio-over-fiber access network: Architectural considerations and experiment," *J. Lightwave Tech.*, vol. 25, no. 11, pp. 3374–3380, Nov. 2007.

34. M. J. Crisp, S. Li, A. Watts, R. V. Penty, and I. H. White, "Uplink and downlink coverage improvements of 802.11g signals using a distributed antenna network," *J. Lightwave Tech.*, vol. 25, no. 11, pp. 3388–3395, Nov. 2007.

35. N. J. Gomes, A. Nkansah, and D. Wake, "Radio-over-MMF techniques—Part I: RF to microwave frequency systems," *J. Lightwave Tech.*, vol. 26, no. 15, pp. 2388–2395, Aug. 2008.

36. A.M.J. Koonen and M. G. Larrodé, "Radio-over-MMF techniques—Part I: Microwave to millimeter-wave systems," *J. Lightwave Tech.,* vol. 26, no. 15, pp. 2396–2408, Aug. 2008.

37. C. Lethien, C. Loyez, J.-P. Vilcot, R. Kassi, N. Rolland, C. Sion, and P.-A. Rolland, "Review of glass and polymer multimode fibers used in a Wimedia ultrawideband MB-OFDM radio over fiber system," *J. Lightwave Tech.,* vol. 27, no. 10, pp. 1320–1331, May 2009.

38. M. Morant, T. F. Alves, R. Llorente, A.V.T. Cartaxo, and J. Marti, "Experimental comparison of transmission performance of multichannel OFDM-UWB signals on FTTH networks," *J. Lightwave Tech.,* vol. 27, no. 10, pp. 1408–1414, May 2009.

39. B. Huiszoon, T. Spuesens, E. Tangdiongga, H. de Waardt, G. D. Khoe, and A.M.J. Koonen, "Hybrid radio-over-fiber and OCDMA architecture for fiber to the personal area network," *J. Lightwave Tech.,* vol. 27, no. 12, pp. 1904–1911, June 2009.

40. L. Chen, J. G. Yu, S. Wen, J. Lu, Z. Dong, M. Huang, and G. K. Chang, "A novel scheme for seamless integration of ROF with centralized lightwave OFDM-WDM-PON system," *J. Lightwave Tech.,* vol. 27, no. 14, pp. 2786–2791, July 2009.

41. C. Lim, A. Nirmalathas, M. Bakaul, K.-L. Lee, D. Novak, and R. Waterhouse, "Mitigation strategy for transmission impairments in millimeter-wave radio-over-fiber networks [Invited]," *J. Opt. Network.,* vol. 8, pp. 201–214, 2009.

42. F. Vacondio, M. Mirshafiei, J. Basak, A. Liu, L. Liao, M. Paniccia, and L. A. Rusch, "A silicon modulator enabling RF over fiber for 802.11 OFDM signals," *IEEE J. Sel. Topics Quantum Electron.,* vol. 16, no.1, pp. 141–148, Jan/Feb. 2010.

43. A. J. Seeds, "Microwave photonics," *IEEE Microwave Theory Tech.*, vol. 50, pp. 877–887, Mar. 2002.

44. J. Kim, Y. S. Kang, Y. D. Chung, and K. S. Choi, "Development and RF characteristics of analog 60-GHz electro-absorption modulator module for RF/optic conversion," *IEEE Trans. Microwave Theory Tech.*, vol. 54, pp. 780–787, Feb. 2006.

45. L. Chrostowski, X. Zhao, and C. J. Chang-Hasnain, "Microwave performance of optically injection-locked VCSELs," *IEEE Trans. Microwave Theory Tech.*, vol. 54, pp. 788–796, Feb. 2006.

46. A. Stöhr, A. Malcoci, A. Sauerwald, I. C. Mayorga, R. Güsten, and D. S. Jäger, "Ultra-wide-band travelling-wave photodetectors for photonic local oscillators," *J. Lightwave Tech.,* vol. 21, pp. 3062–3070, Dec. 2003.

47. J. Capmany, B. Ortega, and D. Pastor, "A tutorial on microwave photonic filters," *J. Lightwave Tech.*, vol. 24, pp. 201–229, Jan. 2006.

48. R. A. Minasian, "Photonic signal processing of microwave signals," *IEEE Trans. Microwave Theory Tech.*, vol. 54, pp. 832–846, Feb. 2006.

49. P. W. Juodawlkis, J. J. Hargreaves, R. D. Younger, G. W. Titi, and J. C. Twichell, "Optical down-sampling of wide-band microwave signals," *J. Lightwave Tech.,* vol. 21, pp. 3116–3124, Dec. 2003.

50. S. Tonda-Goldstein, D. Dolfi, A. Monsterleet, S. Formont, J. Chazelas, and J.-P. Huignard, "Optical signal processing in radar systems," *IEEE Trans. Microwave Theory Tech.*, vol. 54, pp. 847–853, Feb. 2006.

51. R. W. Tkach and A. R. Chraplyvy, "Phase noise and linewidth in an InGaAsP DFB laser," *J. Lightwave Tech.,* vol. 4, pp. 1711–1716, Nov. 1986.

52. K. D. LaViolette, "CTB performance of cascaded externally modulated and directly modulated CATV transmitters," *IEEE Photonics Technol. Lett.*, vol. 8, pp. 281–283, Feb. 1996.

53. A. J. Seeds and K. J. Williams, "Microwave photonics," *J. Lightwave Tech.,* vol. 24, no. 12, pp. 4628-4641, Dec. 2006.

54. J. Yao, "Microwave photonics," *J. Lightwave Tech.,* vol. 27, no. 3, pp. 314–335, Feb. 2009.

55. W. S. C. Chang, *RF Photonic Technology in Optical Fiber Links*, Cambridge University Press, New York, 2007.

56. S. Iezekiel, *Microwave Photonics: Devices and Applications*, Wiley, Hoboken, NJ, 2009.

57. A. Vilcot, B. Cabon, and J.Chazelas, *Microwave Photonics: From Components to Applications and Systems*, Springer, New York, 2010.

CHAPTER 10

WDM Concepts and Components

A distinctive operational feature of an optical fiber is that there is a wide spectral region in which optical signals can be transmitted efficiently. For full-spectrum fibers this region includes the O-band through the L-band, which ranges from about 1260 to 1675 nm. The light sources used in high-capacity optical fiber communication systems emit in a narrow wavelength band of less than 1 nm, so many different independent optical channels can be used simultaneously in different segments of this wavelength range. The technology of combining a number of such independent information-carrying wavelengths onto the same fiber is known as *wavelength division multiplexing* or WDM.[1–10]

Section 10.1 addresses the operating principles of WDM, examines the functions of a generic WDM link, and discusses the internationally standardized spectral grids for two different wavelength multiplexing schemes. Sections 10.2 through 10.6 describe various categories of passive optical components that are needed to combine wavelengths at the transmitting end and separate them into individual channels at the destination. An important factor in deploying these components in a WDM system is to ensure that optical signal power from one channel does not drift into the spectral territory occupied by adjacent channels.[10]

The design of complex wavelength division multiplexed links requires choosing many optical sources with narrow spectral emission bands. The most straightforward method is to select a series of individual lasers, each of which emits at a specific wavelength. This selection process is adequate for a small number of wavelength channels but can be cumbersome for links carrying many wavelengths. To simplify this process, a variety of wavelength-tunable components and schemes have been investigated concerning the administration of channels in a WDM transmission system. Among these are tunable light sources, which Sec. 10.7 addresses, and tunable optical signal-conditioning and receiving devices, which Sec. 10.8 describes. Chapters 11 and 13 define additional active components used for signal amplification and for switching and routing of individual wavelengths.

Applications of WDM techniques are found in all levels of communication links including long-distance terrestrial and undersea transmission systems, metro networks, and fiber-to-the-premises (FTTP) networks. Chapter 13 shows how WDM methodologies apply to several categories of communication networks.

10.1 Overview of WDM

The original optical fiber links that were deployed around 1980 consisted of simple point-to-point connections. These links contained a single fiber with one light source at the transmitting end and one photodetector at the receiving end. In these early systems, signals from different light sources used separate and uniquely assigned optical fibers. Since the spectral width of a typical laser source occupies only a narrow slice of optical bandwidth, these simplex systems greatly underutilize the large bandwidth capacity of a fiber. The first use of WDM was to upgrade the capacity of installed point-to-point transmission links. This was achieved with wavelengths that were separated from several tens up to 200 nanometers in order not to impose strict wavelength-tolerance requirements on the different laser sources and the wavelength-separating components at the receiving end.

With the advent of high-quality light sources with extremely narrow spectral emission widths (less than 1 nm), many independent wavelength channels spaced less than a nanometer apart could be placed on the same fiber. For example, the modulated output of a DFB laser has a frequency spectrum of 10 to 50 MHz, which is equivalent to a spectral emission width of less than 10^{-3} nm.

With such light sources the use of WDM allows a dramatic increase in the capacity of an optical fiber compared to the original simple point-to-point link that carried only a single wavelength. For example, if each wavelength supports an independent transmission rate of 10 Gb/s, then each additional channel provides the fiber with significantly more capacity. Another advantage of WDM is that the various optical channels can support different transmission formats. Thus, by using separate wavelengths, differently formatted signals at any data rate can be sent simultaneously and independently over the same fiber without the need for a common signal structure.

10.1.1 Operational Principles of WDM

A characteristic of WDM is that the discrete wavelengths form an *orthogonal set* of carriers that can be separated, routed, and switched without interfering with each other. This isolation between channels holds as long as the total optical power intensity is kept sufficiently low to prevent nonlinear effects such as stimulated Brillouin scattering and four-wave mixing processes from degrading the link performance (see Chapter 12).

The implementation of sophisticated WDM networks requires a variety of *passive* and *active* devices to combine, distribute, isolate, and amplify optical power at different wavelengths. *Passive devices* require no external control for their operation, so they are somewhat limited in their application flexibility. These components are mainly used to split and combine or tap off optical signals. The wavelength-dependent performance of *active devices* can be controlled electronically or optically, thereby providing a large degree of network flexibility. Active WDM components include tunable optical filters, tunable sources, and optical amplifiers.

Figure 10.1 shows the implementation of passive and active components in a typical WDM link containing various types of optical amplifiers (see Chapter 11). At the transmitting end there are several independently modulated light sources, each emitting signals at a unique wavelength. Here a *multiplexer* is needed to combine these optical outputs into a continuous spectrum of signals and couple them onto a single fiber. At the receiving end a *demultiplexer* is required to separate the optical signals into appropriate detection channels for signal processing.

As Fig. 10.2 shows, there are many independent operating regions across the spectrum ranging from the O-band through the L-band in which narrow-linewidth optical sources can be used simultaneously. We can view these regions either in terms of *spectral width* (the wavelength band occupied by the light
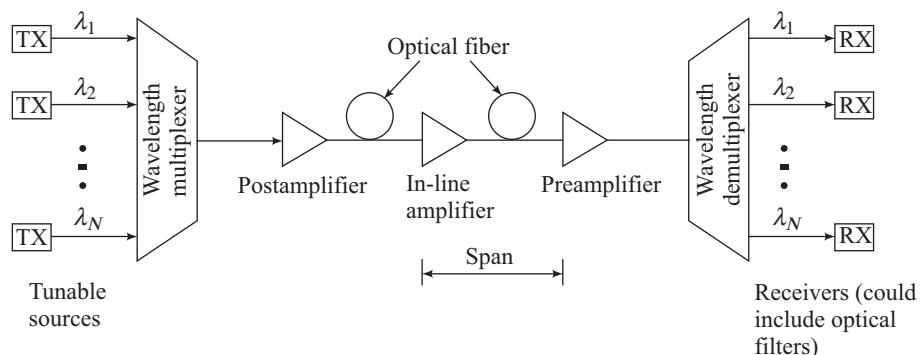
Fig. 10.1 *Implementation of a typical WDM network containing various types of optical amplifiers*

signal) or by means of *optical bandwidth* (the frequency band occupied by the light signal). To find the optical bandwidth corresponding to a particular spectral width in these regions, we use the fundamental relationship $c = \lambda \nu$, which relates the wavelength λ to the carrier frequency ν, where c is the speed of light. Differentiating this we have for $\Delta\lambda << \lambda^2$

$$\left|\Delta\nu\right| = \frac{c}{\lambda^2}\left|\Delta\lambda\right| \qquad (10.1)$$

where the frequency deviation $\Delta\nu$ corresponds to the wavelength deviation $\Delta\lambda$ around λ.
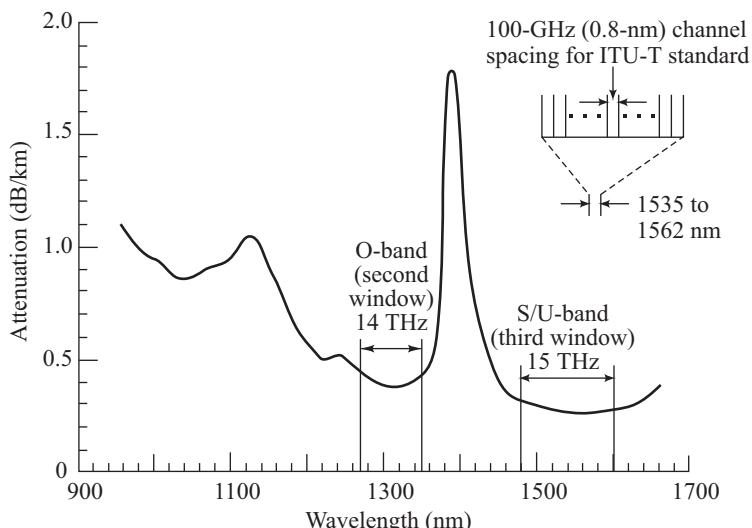
Fig. 10.2 *The transmission-band widths in the O- and C-bands (the 1310-nm and 1550-nm windows) allow the use of many simultaneous channels for sources with narrow spectral widths. The ITU-T G.692 standard for WDM specifies channels with 100-GHz spacings.*

Example 10.1 Consider a fiber that has the attenuation characteristic shown in Fig. 10.2. What are the usable spectral bands (*a*) in the O-band centered at 1420 nm; (*b*) in the combined S- and C-bands with a 1520-nm center wavelength?

Solution: (*a*) From Eq. (10.1) the optical bandwidth is $\Delta v = 14$ THz for a usable spectral band $\Delta\lambda = 100$ nm covering the O-band with a 1420-nm center wavelength.

(*b*) Similarly, $\Delta v = 14$ THz for a usable spectral band $\Delta\lambda = 105$ nm in the low-loss region covering the S- and C-bands with a 1520-nm center wavelength.

The operational frequency band allocated to a particular light source normally ranges from 25 to 100 GHz (or equivalently, a spectral band of 0.25 to 0.8 nm at a 1550-nm wavelength). The exact width of the frequency or spectral band that is selected needs to take into account possible drifts in the peak wavelength emitted by the laser and temporal variations in the wavelength response of other link components. These parameter changes can result from effects such as *component aging* or *temperature variations*.

Depending on the frequency bands chosen for the optical transmission link, many operational regions are available in the various spectral bands. The engineering challenge for using such a large number of light sources, each of which is emitting at a different wavelength, is to ensure that each source is spaced sufficiently far from its neighbors so as not to create interference. This means that highly stabilized optical transmitters are needed so that the integrities of the independent message streams from each source are maintained for subsequent conversion back to electrical signals at the receiving end.

Example 10.2 Consider a spectral band of 0.8 nm (or equivalently, a mean frequency spacing of 100 GHz at a 1550-nm wavelength) within which lasers with narrow linewidths are transmitting. How many of such signal channels fit into (*a*) the C band, and (*b*) the combined S- and C-bands?

Solution: (*a*) Since the C-band ranges from 1530 to 1565 nm, one can have $N = (35$ nm$)/(0.8$ nm per channel$) = 40$ independent signal channels.

(*b*) Since the combined S- and C-bands cover the 1460-to-1565 nm range, one can have $N = (105$ nm$)/(0.8$ nm per channel$) = 123$ independent signal channels.

10.1.2 WDM Standards

Since WDM is essentially frequency division multiplexing at optical carrier frequencies, the WDM standards developed by the International Telecommunication Union (ITU) specify channel spacing in terms of frequency.[11–13] A key reason for selecting a fixed frequency spacing, rather than a constant wavelength spacing, is that when locking a laser to a particular operating mode, it is the frequency of the laser that is fixed. Recommendation G.692 was the first ITU-T specification for WDM.[11] This document specifies selecting the channels from a grid of frequencies referenced to 193.100 THz (1552.524 nm) and spacing them 100 GHz (about 0.8 nm at 1550 nm) apart. Suggested alternative spacings in G.692 include 50 GHz and 200 GHz, which correspond to spectral widths of 0.4 and 1.6 nm, respectively, at 1550 nm.

Historically the term *dense WDM* (DWDM) generally referred to small wavelength separations, such as those denoted by ITU-T G.692. In 2002 the ITU-T released Recommendation G.694.1, which is aimed specifically at DWDM. This document specifies WDM operation in the S-, C-, and L-bands for high-quality, high-rate metro area network (MAN) and wide area network (WAN) services. It calls for narrow frequency spacings of 100 to 12.5 GHz (or, equivalently, 0.8 to 0.1 nm at 1550 nm). This implementation requires the use of stable, high-quality, temperature-controlled and wavelength-controlled (frequency-locked) laser diode light sources. For example, the wavelength-drift tolerances for 25-GHz channels are ± 0.02 nm.

Table 10.1 *Portion of the ITU-T G.694.1 dense WDM grid for 100- and 50-GHz spacings in the L- and C-bands*

L-band				C-band			
100-GHz		50-GHz offset		100-GHz		50-GHz offset	
THz	nm	THz	nm	THz	nm	THz	nm
186.00	1611.79	186.05	1611.35	191.00	1569.59	191.05	1569.18
186.10	1610.92	186.15	1610.49	191.10	1568.77	191.15	1568.36
186.20	1610.06	186.25	1609.62	191.20	1576.95	191.25	1567.54
186.30	1609.19	186.35	1608.76	191.30	1567.13	191.35	1566.72
186.40	1608.33	186.45	1607.90	191.40	1566.31	191.45	1565.90
186.50	1607.47	186.55	1607.04	191.50	1565.50	191.55	1565.09
186.60	1606.60	186.65	1606.17	191.60	1564.68	191.65	1564.27
186.70	1605.74	186.75	1605.31	191.70	1563.86	191.75	1563.45
186.80	1604.88	186.85	1604.46	191.80	1563.05	191.85	1562.64
186.90	1604.03	186.95	1603.60	191.90	1562.23	191.95	1561.83

Table 10.1 lists part of the ITU-T G.694.1 dense WDM frequency grid for 100- and 50-GHz spacings in the L- and C-bands. The column labeled "50-GHz offset" means that for the 50-GHz grid one uses the 100-GHz spacings with these 50-GHz values interleaved. For example, the 50-GHz channels in the L-band would be at 186.00 THz, 186.05 THz, 186.10 THz, and so on. Note that when the frequency spacing is uniform, the wavelengths are not spaced uniformly because of the relationship given in Eq. (10.1).

To designate which C-band channel is under consideration in 100-GHz applications, the ITU-T uses a *channel numbering convention*. For this, the frequency 19N.M THz is designated as ITU channel number NM. For example, the frequency 194.3 THz is ITU channel 43.

The concept of *coarse WDM* (CWDM) emerged from the combination of the production of full-spectrum (low-water-content) G.652C and G.652D fibers, the development of relatively inexpensive optical sources, and the desire to have low-cost optical links operating in access networks and local area networks.[14] In 2002 the ITU-T released Recommendation G.694.2, which defines the spectral grid for CWDM. As shown in Fig. 10.3, the CWDM grid is made up of 18 wavelengths defined within the range 1270 nm to 1610 nm (O- through L-bands) spaced by 20 nm with wavelength-drift tolerances of ± 2 nm. This can be achieved with inexpensive light sources that are not temperature-controlled.
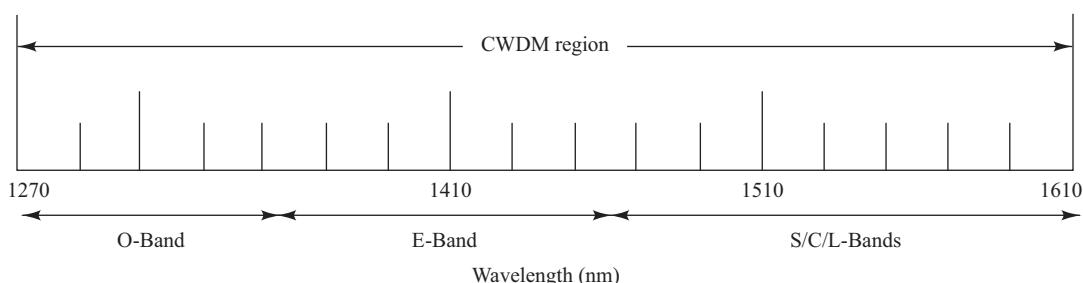
Fig. 10.3 *Spectral grid for coarse wavelength division multiplexing (CWDM)*

The ITU-T Recommendation G.695 released in 2004 outlines optical interface specifications for multiple-channel CWDM over distances of 40 and 80 km. Both unidirectional and bidirectional systems (such as used in passive optical network applications) are included in the recommendation. The applications for G.695 cover all or part of the 1270-to-1610-nm range. The main deployments are for single-mode fibers, such as those specified in ITU-T Recommendations G.652 and G.655.

10.2 | Passive Optical Couplers

Passive devices operate completely in the optical domain to split and combine light streams. They include $N \times N$ couplers (with $N \geq 2$), power splitters, power taps, and star couplers. These components can be fabricated either from optical fibers or by means of planar optical waveguides using material such as lithium niobate ($LiNbO_3$), InP, silica, silicon oxynitride, or various polymers.

Basically, most passive WDM devices are variations of a star-coupler concept. Figure 10.4 shows a generic star coupler, which can perform both power combining and splitting. In the broadest application, star couplers combine the light streams from two or more input fibers and divide them among several output fibers. In the general case, the splitting is done uniformly for all wavelengths, so that each of the N outputs receives $1/N$ of the power entering the device. A common fabrication method for an $N \times N$ splitter is to fuse together the cores of N single-mode fibers over a length of a few millimeters. The optical power inserted through one of the N fiber entrance ports gets divided uniformly into the cores of the N output fibers through evanescent power coupling in the fused region (e.g., see Fig. 10.5).

Any size star coupler can be made, in principle, provided that all fibers can be heated uniformly during the coupler-fabrication process. Couplers with 64 inputs and outputs are possible, although, more commonly, the size tends to be less than 10. One simple device is a power tap. Taps are nonuniform 2×2 couplers used to extract a small portion of optical power from a fiber line for monitoring signal quality.

The three fundamental technologies for making passive components are based on optical fibers, integrated optical waveguides, and bulk micro-optics.[15–21] The next sections describe the physical principles of several simple examples of fiber-based and integrated-optic devices to illustrate the fundamental operating principles. Couplers using micro-optic designs are not widely used because the strict tolerances required in the fabrication and alignment processes affect their cost, performance, and robustness. See Prob.10.16 for an example of a micro-optic-based multiplexer using a plane reflection grating.

10.2.1 The 2×2 Fiber Coupler

When discussing couplers and splitters, it is customary to refer to them in terms of the number of input and output ports on the device. For example, a device with two inputs and two outputs would be called a "2×2 coupler." In general, an $N \times M$ coupler has N inputs and M outputs.
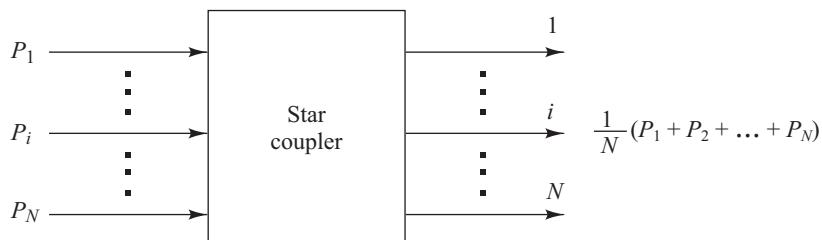
Fig. 10.4 *Basic star coupler concept for combining or splitting optical powers*
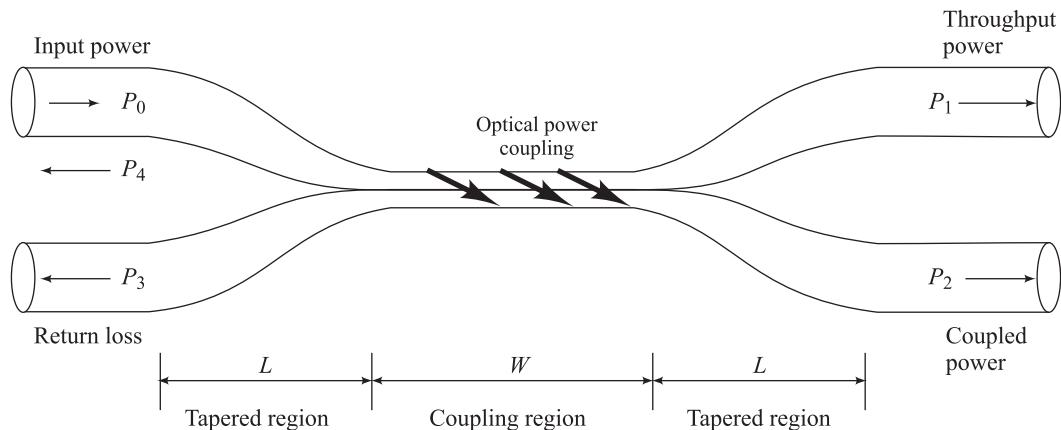
Fig. 10.5 *Cross-sectional view of a fused-fiber coupler having a coupling region W and two tapered regions of length L. The total span $\mathscr{L} = 2L + W$ is the coupler draw length*

The 2×2 coupler[18,19] is a simple fundamental device that we will use here to demonstrate the operational principles. A common construction is the fused-fiber coupler. This is fabricated by twisting together, melting, and pulling two single-mode fibers so they get fused together over a uniform section of length W, as shown in Fig. 10.5. Each input and output fiber has a long tapered section of length L, since the transverse dimensions are gradually reduced down to that of the coupling region when the fibers are pulled during the fusion process. The total draw length is $\mathscr{L} = 2L + W$. This device is known as a *fused biconical tapered coupler*. Here, P_0 is the input power, P_1 is the throughout power, and P_2 is the power coupled into the second fiber. The parameters P_3 and P_4 are extremely low signal levels (−50 to −70 dB below the input level) resulting from backward reflections and scattering due to bending in and packaging of the device.[20]

As the input light P_0 propagates along the taper in fiber 1 and into the coupling region W, there is a significant decrease in the V number owing to the reduction in the ratio r/λ [see Eq. (2.58)], where r is the reduced fiber radius. Consequently, as the signal enters the coupling region, an increasingly larger portion of the input field now propagates outside the core of the fiber. Depending on the dimensioning of the coupling region, any desired fraction of this decoupled field can be recoupled into the other fiber. By making the tapers very gradual, only a negligible fraction of the incoming optical power is reflected back into either of the input ports. Thus these devices are also known as *directional couplers*.

The optical power coupled from one fiber to another can be varied through three parameters: the axial length of the coupling region over which the fields from the two fibers interact; the size of the reduced radius r in the coupling region; and Δr, the amount of lateral overlap of the fiber cores in the coupling region. In making a fused fiber coupler, the coupling length W is normally fixed by the width of the heating flame, so that only L and r change as the coupler is elongated. Typical values for W and L are a few millimeters, the exact values depending on the coupling ratios desired for a specific wavelength, and $\Delta r/r$ is around 0.015. Assuming that the coupler is lossless, the expression for the power P_2 coupled from one fiber to another over an axial distance z is[15,16]

$$P_2 = P_0 \sin^2(\kappa z) \tag{10.2}$$

where κ is the *coupling coefficient* describing the interaction between the fields in the two fibers. By conservation of power, for identical-core fibers we have

$$P_1 = P_0 - P_2 = P_0 [1 - \sin^2(\kappa z)] = P_0 \cos^2(\kappa z) \tag{10.3}$$

This shows that the phase of the driven fiber always lags 90° behind the phase of the driving fiber, as Fig. 10.6*a* illustrates. Thus, when power is launched into fiber 1, at $z = 0$ the phase in fiber 2 lags 90° behind that in fiber 1. This lagging phase relationship continues for increasing z, until at a distance that satisfies $\kappa z = \pi/2$, all of the power has been transferred from fiber 1 to fiber 2. Now fiber 2 becomes the driving fiber, so that for $\pi/2 \leq \kappa z \leq \pi$ the phase in fiber 1 lags behind that in fiber 2, and so on. As a result of this phase relationship, the 2×2 coupler is a *directional coupler*. That is, no energy can be coupled into a wave traveling backward in the negative-z direction in the driven waveguide.

Figure 10.6*b* shows experimental measurements of how the normalized coupled power ratios P_2/P_0 and P_1/P_0 vary with wavelength for a 15-mm-long coupler. Couplers with different performances can be made by varying the parameters W, L, r, and Δr for a specific wavelength.

Example 10.3   The coupling coefficient κ is a complex parameter that depends on a variety of factors, such as the wavelength, the refractive indices of the fibers, the fiber radius a, and the axes spacing d between the two coupled fibers. A simplified and accurate empirical expression for κ in a directional coupler made from two identical step-index fibers is given by[22]

$$\kappa = \frac{\pi}{2} \frac{\sqrt{\delta}}{a} \exp[-(A + Bx + Cx^2)]$$

where $x = d/a$

$$\delta = \frac{n_1^2 - n_2^2}{n_1^2}$$

$A = 5.2789 - 3.663V + 0.3841V^2$
$B = -0.7769 + 1.2252V - 0.0152V^2$
$C = -0.0175 - 0.0064V - 0.0009V^2$

with V defined by Eq. (2.27). Consider two fibers for which $n_1 = 1.4532$, $n_2 = 1.4500$, and $a = 5.0\ \mu$m. If the spacing between the centers of the fibers is $d = 12\ \mu$m, what is the coupling coefficient κ at a 1300-nm wavelength?

Solution: Using Eq. (2.27) we find that $V = 2.329$. From the above equation we then find

$$\kappa = 20.8 \exp[-(-1.1693 + 1.9945x - 0.0373x^2)]$$
$$= 0.694\ \text{mm}^{-1}$$

for $x = 12/5 = 2.4$.

In specifying the performance of an optical coupler, one usually indicates the percentage division of optical power between the output ports by means of the *splitting ratio* or *coupling ratio*. Referring to Fig. 10.5, with P_0 being the input power and P_1 and P_2 the output powers, then

$$\text{Splitting ratio} = \left(\frac{P_2}{P_1 + P_2} \right) \times 100\% \tag{10.4}$$

By adjusting the parameters so that power is divided evenly, with half of the input power going to each output, one creates a *3-dB coupler*. A coupler could also be made in which almost all the optical power at 1500 nm goes to one port and almost all the energy around 1300 nm goes to the other port (see Prob. 10.5).

In the above analysis, we have assumed, for simplicity, that the device is lossless. However, in any practical coupler some light is always lost when a signal goes through it. The two basic losses are excess loss and insertion loss. The *excess loss* is defined as the ratio of the input power to the total output power.
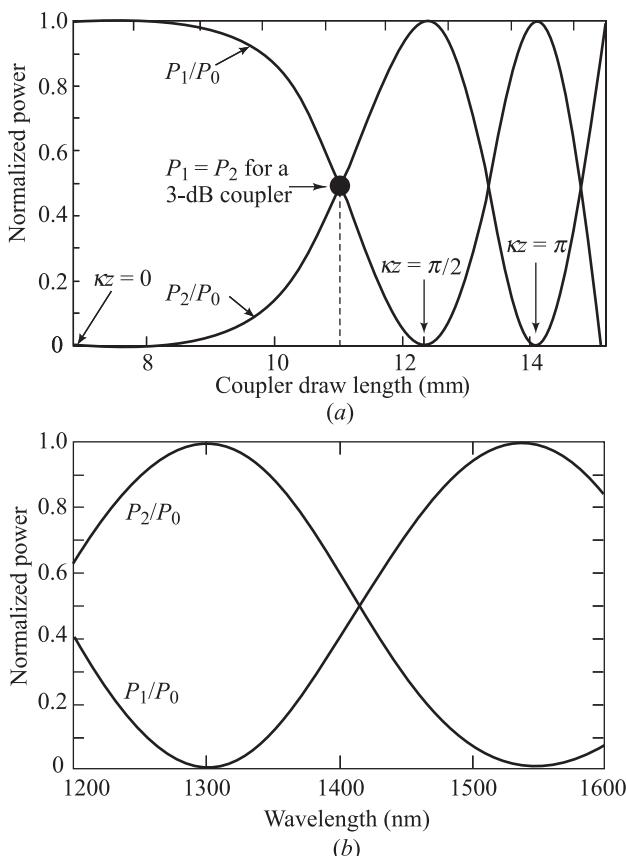
Fig. 10.6 *(a) Normalized coupled powers P_2/P_0 and P_1/P_0 as a function of the coupler draw length for a 1300-nm power level P_0 launched into fiber 1. (b) dependence on wavelength of the coupled powers in a 15-mm-long coupler. (Adapted with permission from Eisenmann and Weidel,[19] © 1988, IEEE.)*

Thus in decibels, the excess loss for a 2×2 coupler is

$$\text{Excess loss} = 10 \log \left(\frac{P_0}{P_1 + P_2} \right) \tag{10.5}$$

The *insertion loss* refers to the loss for a particular port-to-port path. For example, for the path from input port i to output port j, we have, in decibels,

$$\text{Insertion loss} = 10 \log \left(\frac{P_i}{P_j} \right) \tag{10.6}$$

Another performance parameter is *crosstalk* or *return loss,* which measures the degree of isolation between the input at one port and the optical power scattered or reflected back into the other input port.

That is, it is a measure of the optical power level P_3 shown in Fig. 10.5.

$$\text{Return loss} = 10 \log \left(\frac{P_3}{P_0} \right) \tag{10.7}$$

Example 10.4 A 2 × 2 biconical tapered fiber coupler has an input optical power level of $P_0 = 200 \ \mu W$. The output powers at the other three ports are $P_1 = 90 \ \mu W$, $P_2 = 85 \ \mu W$, and $P_3 = 6.3 \ nW$. What are the coupling ratio, excess loss, insertion losses, and return loss for this coupler?

Solution: From Eq. 10.4, the coupling ratio is

$$\text{Coupling ratio} = \left(\frac{85}{90 + 85} \right) \times 100\% = 48.6\%$$

From Eq. 10.5, the excess loss is

$$\text{Excess loss} = 10 \log \left(\frac{200}{90 + 85} \right) = 0.58 \text{ dB}$$

Using Eq. 10.6, the insertion losses are

$$\text{Insertion loss (port 0 to port 1)} = 10 \log \left(\frac{200}{90} \right) = 3.47 \text{ dB}$$

$$\text{Insertion loss (port 0 to port 2)} = 10 \log \left(\frac{200}{85} \right) = 3.72 \text{ dB}$$

The return loss is given by Eq. 10.7 as

$$\text{Return loss} = 10 \log \left(\frac{6.3 \times 10^{-3}}{200} \right) = -45 \text{ dB}$$

Example 10.5 To monitor the light signal level or quality in a link, one can use a 2 × 2 device that has a coupling fraction of around 1 to 5 percent, which is selected and fixed during fabrication. This is known as a *tap coupler*. Nominally the tap coupler is packaged as a three-port device with one arm of the 2 × 2 coupler being terminated inside the package. Figure 10.7 shows a typical package for such a tap coupler and Table 10.2 lists some representative specifications.
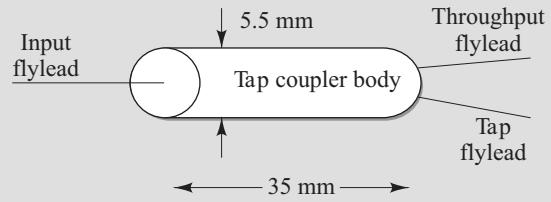
Fig. 10.7 *Typical configuration and package dimensions for a tap coupler*

Table 10.2 *Representative specifications for a 2 × 2 tap coupler*

Parameter	Unit	Specification
Tap ratio	percent	1 to 5
Insertion loss (throughput)	dB	0.5
Return loss	dB	55
Power handling	mW	1000
Flylead length	m	1
Size (diameter × length)	mm	5.5 × 35

10.2.2 Scattering Matrix Representation

One can also analyze a 2 × 2 guided-wave coupler as a four-terminal device that has two inputs and two outputs, as shown in Fig. 10.8. Either all-fiber or integrated-optics devices can be analyzed in terms of the *scattering matrix* (also called the *propagation matrix*) **S**, which defines the relationship between the two input field strengths a_1 and a_2, and the two output field strengths b_1 and b_2. By definition,[23,24]

$$\mathbf{b} = \mathbf{Sa}, \text{where } \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \qquad \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \qquad \text{and } \mathbf{S} = \begin{bmatrix} s_{11} & s_{21} \\ s_{12} & s_{22} \end{bmatrix} \tag{10.8}$$
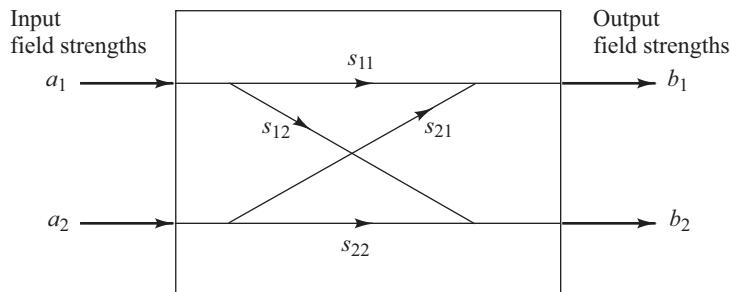
Fig. 10.8   *Generic 2 × 2 guided-wave coupler. Here a_i and b_j represent the field strengths of input port i and output port j, respectively, and the s_{ij} are the scattering matrix parameters.*

Here, $s_{ij} = |s_{ij}| \exp(j\phi_{ij})$ represents the *coupling coefficient* of optical power transfer from input port *i* to output port *j*, with $|s_{ij}|$ being the magnitude of s_{ij} and ϕ_{ij} being its phase at port *j* relative to port *i*.

For an actual physical device, two restrictions apply to the scattering matrix **S**. One is a result of the reciprocity condition arising from the fact that Maxwell's equations are invariant for time inversion; that is, they have two solutions in opposite propagating directions through the device, assuming single-mode operation. The other restriction arises from energy-conservation principles under the assumption that the device is lossless. From the first condition, it follows that

$$s_{12} = s_{21} \tag{10.9}$$

From the second restriction, if the device is lossless, the sum of the output intensities I_o must equal the sum of the input intensities I_i:

$$I_o = b_1^* b_1 + b_2^* b_2 = I_i = a_1^* a_1 + a_2^* a_2 \qquad \text{or} \qquad b^+ b = a^+ a \tag{10.10}$$

where the superscript * means the complex conjugate and the superscript + indicates the transpose conjugate. Substituting Eqs. (10.8) and (10.9) into Eq. (10.10) yields the following set of three equations:

$$s_{11}^* s_{11} + s_{12}^* s_{12} = 1 \tag{10.11}$$

$$s_{11}^* s_{12} + s_{12}^* s_{22} = 0 \tag{10.12}$$

$$s_{22}^* s_{22} + s_{12}^* s_{12} = 1 \tag{10.13}$$

If we now assume that the coupler has been constructed so that the fraction $(1 - \epsilon)$ of the optical power from input 1 appears at output port 1, with the remainder ϵ going to port 2, then we have $s_{11} = \sqrt{1 - \epsilon}$, which is a real number between 0 and 1. Here, we have assumed, without loss of generality, that the electric field at output 1 has a zero phase shift relative to the input at port 1; that is $\phi_{11} = 0$. Since we are interested in the phase change that occurs when the coupled optical power from input 1 emerges from port 2, we make the simplifying assumption that the coupler is symmetric. Then, analogous to the effect at port 1, we have $s_{22} = \sqrt{1 - \epsilon}$, with $\phi_{22} = 0$. Using these expressions, we can determine the

phases ϕ_{12} of the coupled outputs relative to the input signals and find the constraints on the composite outputs when both input ports are receiving signals.

Inserting the expressions for s_{11} and s_{22} into Eq. (10.12) and letting $s_{12} = |s_{12}| \exp(j\phi_{12})$, where $|s_{12}|$ is the magnitude of s_{12} and ϕ_{12} is its phase, we have

$$\exp(j2\phi_{12}) = -1 \tag{10.14}$$

which holds when

$$\phi_{12} = (2n + 1)\frac{\pi}{2} \qquad \text{where } n = 0, 1, 2, \dots \tag{10.15}$$

so that the scattering matrix from Eq. (10.8) becomes

$$\mathbf{S} = \begin{bmatrix} \sqrt{1-\epsilon} & j\sqrt{\epsilon} \\ j\sqrt{\epsilon} & \sqrt{1-\epsilon} \end{bmatrix} \tag{10.16}$$

Example 10.6 Assume we have a 3-dB coupler, so that half of the input power gets coupled to the second fiber. What are the output powers $P_{\text{out, 1}}$ and $P_{\text{out, 2}}$?

Solution: Since the input power is divided evenly, $\epsilon = 0.5$ and the output field intensities $E_{\text{out,1}}$ and $E_{\text{out,2}}$ can be found from the input intensities $E_{\text{in,1}}$ and $E_{\text{in,2}}$ and the scattering matrix in Eq. (10.6):

$$\begin{bmatrix} E_{\text{out,1}} \\ E_{\text{out,2}} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & j \\ j & 1 \end{bmatrix} \begin{bmatrix} E_{\text{in,1}} \\ E_{\text{in,2}} \end{bmatrix}$$

Letting $E_{\text{in, 2}} = 0$, we have $E_{\text{out,1}} = (1/\sqrt{2})E_{\text{in,1}}$ and $E_{\text{out,2}} = (j/\sqrt{2})E_{\text{in,1}}$. The output powers are then given by

$$P_{\text{out,1}} = E_{\text{out,1}}E_{\text{out,1}}^* = \frac{1}{2}E_{\text{in,1}}^2 = \frac{1}{2}P_0$$

Similarly,

$$P_{\text{out,2}} = E_{\text{out,2}}E_{\text{out,2}}^* = \frac{1}{2}E_{\text{in,1}}^2 = \frac{1}{2}P_0$$

so that half the input power appears at each output of the coupler.

It is also important to note that when we want a large portion of the input power from, say, port 1 to emerge from output 1, we need ϵ to be small. However, this, in turn, means that the amount of power at the same wavelength coupled to output 1 from input 2 is small. Consequently, if one is using the same wavelength, it is not possible, in a passive 2×2 coupler, to have all the power from both inputs coupled simultaneously to the same output. The best that can be done is to have half of the power from each input appear at the same output. However, if the wavelengths are different at each input, it is possible to couple a large portion of both power levels onto the same fiber.[18,19]

10.2.3 The 2×2 Waveguide Coupler

More versatile 2×2 couplers are possible with waveguide-type devices.[24–27] Figure 10.9 shows two types of 2×2 waveguide couplers. The uniformly symmetric device has two identical parallel guides in the coupling region, whereas the uniformly asymmetric coupler has one guide wider than the other. Analogous to fused-fiber couplers, waveguide devices have an intrinsic wavelength dependence in the coupling region, and the degree of interaction between the guides can be varied by changing the guide width w, the gap s between the guides, and the refractive index n_1 between the guides. In Fig. 10.9, the z direction lies along the coupler length and the y axis lies in the coupler plane transverse to the two waveguides.
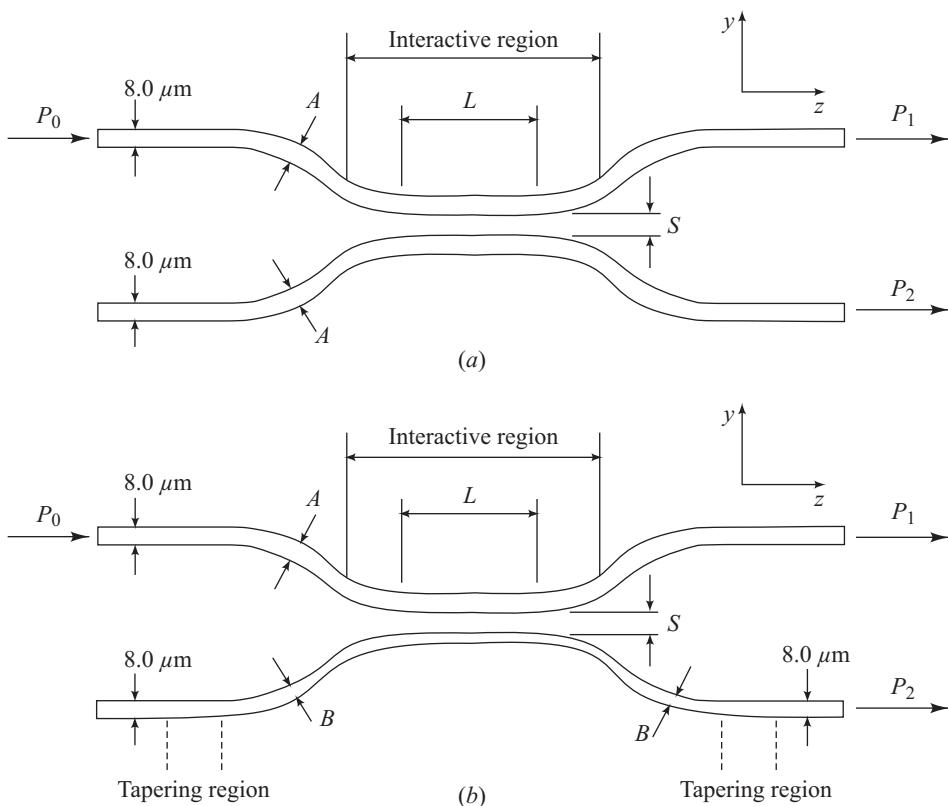
Fig. 10.9 *Cross-sectional top views of (a) a uniformly symmetric directional waveguide coupler with both guides having a width A = 8 μm, (b) a uniformly asymmetric directional coupler in which one guide has a narrower width B in the coupling region. (Adapted with permission from Takagi, Jinguji, and Kawachi,[25] © 1992, IEEE.)*

Let us first consider the symmetric coupler. In real waveguides, with absorption and scattering losses, the propagation constant β_z is a complex number given by

$$\beta_z = \beta_r + j\frac{\alpha}{2} \tag{10.17}$$

where β_r *is* the real part of the propagation constant and α is the optical loss coefficient in the guide. Hence, the total power contained in both guides decreases by a factor $\exp(-\alpha z)$ along their length. For example, losses in semiconductor and silicon oxynitride waveguide devices fall in the $0.05 < \alpha < 0.35\,\text{cm}^{-1}$ range (or, equivalently, about $0.2 < \alpha < 1.5$ dB/cm). Losses in silica waveguides are less than 0.1 dB/cm. Recall from Eq. (3.1) the relationship $\alpha(\text{dB/cm}) = 4.343\alpha(\text{cm}^{-1})$.

The transmission characteristics of the symmetric coupler can be expressed through the coupled-mode theory approach to yield[26,27]

$$P_2 = P_0 \sin^2(\kappa z)\, e^{-az} \tag{10.18}$$

where the coupling coefficient is

$$\kappa = \frac{2\beta_y^2 q e^{-qs}}{\beta_z w(q^2 + \beta_y^2)} \tag{10.19}$$

This is a function of the waveguide propagation constants β_y and β_z (in the y and z directions, respectively), the gap width d and separation s, and the extinction coefficient q in the y direction (i.e., the exponential falloff in the y direction) outside the waveguide, which is

$$q^2 = \beta_z^2 - k_1^2 \tag{10.20}$$

The theoretical power distribution as a function of the guide length is shown in Fig. 10.10 where we have used $\kappa = 0.6$ mm^{-1} and $\alpha = 0.02$ mm^{-1}. Analogous to the fused-fiber coupler, complete power transfer to the second guide occurs when the guide length L is

$$L = \frac{\pi}{2\kappa}(m+1) \qquad \text{with } m = 0, 1, 2,... \tag{10.21}$$

Since κ is found to be almost monotonically proportional to wavelength, the coupling ratio P_2/P_0 rises and falls sinusoidally from 0 to 100 percent as a function of wavelength, as Fig. 10.11 illustrates generically (assuming here, for simplicity, that the guide loss is negligible).
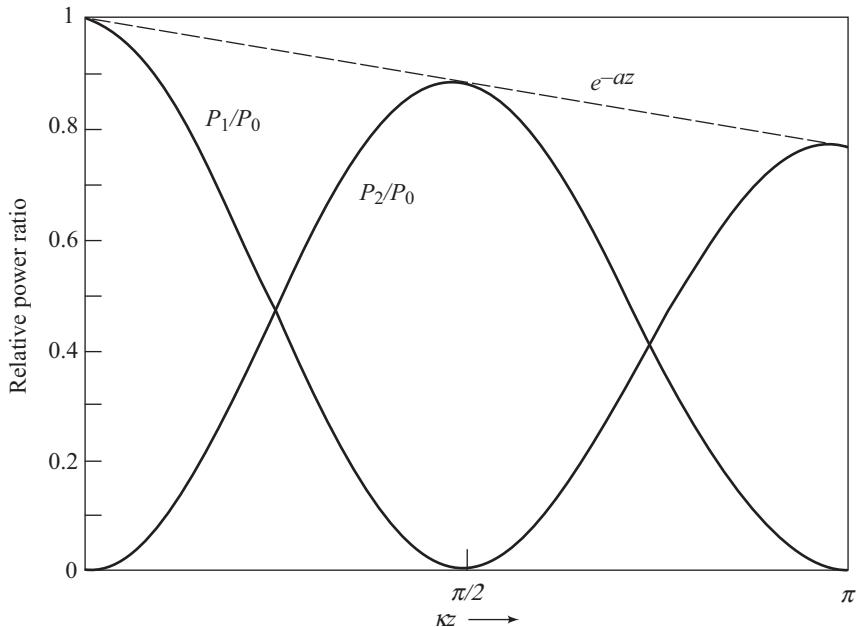
Fig. 10.10 *Theoretical through-path and coupled power distributions as a function of the guide length in a symmetric 2 × 2 guided-wave coupler with $\kappa = 0.6$ mm^{-1} and $\alpha = 0.02$ mm^{-1}*
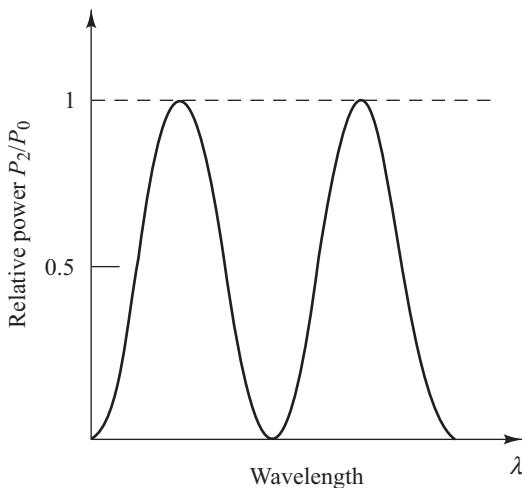
Fig. 10.11 *Wavelength response of the coupled power P_2/P_0 in the symmetric 2×2 guided-wave coupler shown in Fig. 10.9a*
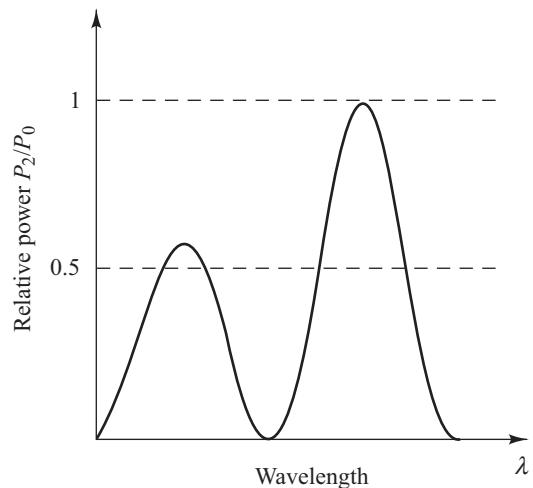
Fig. 10.12 *Wavelength response of the coupled power P_2/P_0 in the asymmetric 2×2 guided-wave coupler shown in Fig. 10.9b*

Example 10.7 A symmetric waveguide coupler has a coupling coefficient $\kappa = 0.6$ mm^{-1}. What is the coupling length?

Solution: Using Eq. (10.21), we find the coupling length for $m = 1$ to be $L = 5.24$ mm.

When the two guides do not have the same widths, as shown in Fig.10.9b, the amplitude of the coupled power is dependent on wavelength, and the coupling ratio becomes

$$P_2/P_0 = \frac{\kappa^2}{g^2} \sin^2\left(gz\right)e^{-\alpha z} \tag{10.22}$$

where

$$g^2 = \kappa^2 + \left(\frac{\Delta\beta}{2}\right)^2 \tag{10.23}$$

with $\Delta\beta$ being the phase difference between the two guides in the z direction. With this type of configuration, one can fabricate devices that have a flattened response in which the coupling ratio is less than 100 percent in a specific desired wavelength range, as shown in Fig. 10.12. The main cause of the wave-flattened response at the lower wavelength results from suppression by the amplitude term κ^2/g^2. This asymmetric characteristic can be used in a device where only a fraction of power from a specific wavelength should be tapped off. Note also that when $\Delta\beta = 0$, Eq. (10.22) reduces to the symmetric case given by Eq. (10.18).

More complex structures are readily fabricated in which the widths of the guides are tapered.[22] These nonsymmetric structures can be used to flatten the wavelength response over a particular spectral range.

It is also important to note that the above analysis based on the coupled-mode theory holds when the indices of the two waveguides are identical, but a more complex analytical treatment is needed for different indices.[24,28]

10.2.4 Star Couplers

The principal role of all star couplers is to combine the powers from N inputs and divide them equally (usually) among M output ports. Techniques for creating star couplers include fused fibers, gratings, micro-optic technologies,

Fig. 10.13 *Generic 4 × 4 fused-fiber star coupler fabricated by twisting, heating, and pulling on four fibers to fuse them together*

and integrated-optics schemes. The fiber-fusion technique has been a popular construction method for $N \times N$ star couplers. For example, 7×7 devices and 1×19 splitters or combiners with excess losses at 1300 nm of 0.4 dB and 0.85 dB, respectively, have been demonstrated.[29,30] However, large-scale fabrication of these devices for $N > 2$ is limited because of the difficulty in controlling the coupling response between the numerous fibers during the heating and pulling process. Figure 10.13 shows a generic 4×4 fused-fiber star coupler.

In an ideal star coupler, the optical power from any input is evenly divided among the output ports. The total loss of the device consists of its splitting loss plus the excess loss in each path through the star. The *splitting loss* is given in decibels by

$$\text{Splitting loss} = -10 \log\left(\frac{1}{N}\right) = 10 \log N \tag{10.24}$$

Similar to Eq. (10.5), for a single input power P_{in} and N output powers, the excess loss in decibels is given by

$$\text{Fiber star excess loss} = 10 \log\left(\frac{P_{\text{in}}}{\sum_{i=1}^{N} P_{\text{out}, i}}\right) \tag{10.25}$$

The insertion loss and return loss can be found from Eqs. (10.6) and (10.7), respectively.

An alternative is to construct star couplers by cascading 3-dB couplers. Figure 10.14 shows an example for an 8×8 device formed by using twelve 2×2 couplers. This device could be made from either fused-fiber or integrated-optic components. As can be seen from this figure, a fraction l/N of the launched power from each input port appears at all output ports. A limitation to the flexibility or modularity of this technique is that N is a multiple of 2; that is, $N = 2^n$ with the integer $n \geq 1$. The consequence is that if an extra node needs to be added to a fully connected $N \times N$ network, the $N \times N$ star needs to be replaced by a $2N \times 2N$ star, thereby leaving $2(N-1)$ new ports being unused. Alternatively, one extra 2×2 coupler can be used at one port to get $N + 1$ outputs. However, these two new ports have an additional 3-dB loss.

As can be deduced from Fig. 10.14, the number of 3-dB couplers needed to construct an $N \times N$ star is

$$N_c = \frac{N}{2} \log_2 N = \frac{N}{2} \frac{\log N}{\log 2} \tag{10.26}$$
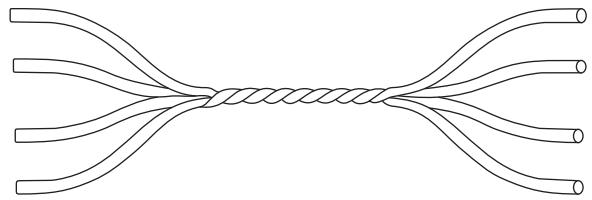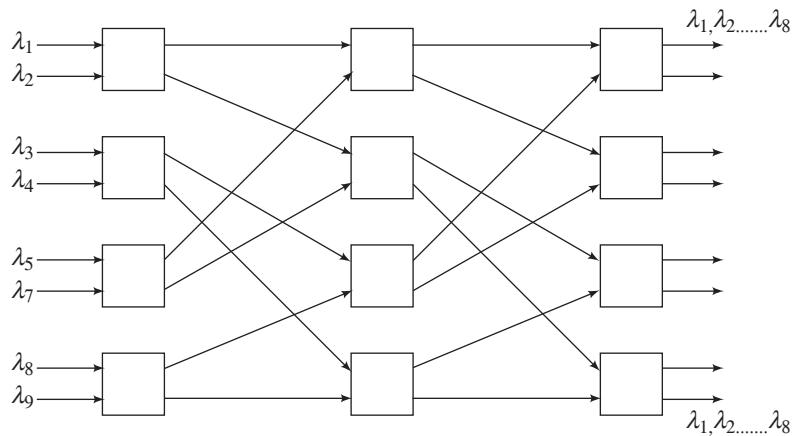
Fig. 10.14 *Example of an 8 × 8 star coupler formed by interconnecting twelve 2 × 2 couplers*

since there are $N/2$ elements in the vertical direction and $\log_2 N = \log N/\log 2$ elements horizontally. (Reminder: We use "$\log x$" to designate the base-10 logarithm of x.)

Example 10.8 A device engineer wants to construct a 32×32 coupler from a cascade of 2×2 3-dB single-mode fiber couplers. How many 2×2 elements are needed for this?

Solution: In this case there will be 16 coupler elements in the vertical direction. From Eq. (10.26), we find how many 2×2 elements are needed:

$$N_c = \frac{32}{2} \frac{\log 32}{\log 2} = 80$$

If the fraction of power traversing each 3-dB coupler element is F_T, with $0 \leq F_T \leq 1$ (i.e., a fraction $1 - F_T$ of power is lost in each 2×2 element), then the *excess loss* in decibels is

$$\text{Excess loss} = -10\log(F_T^{\log_2 N}) \tag{10.27}$$

The splitting loss for this star is, again, given by Eq. (10.24). Thus the total loss experienced by a signal as it passes through the $\log_2 N$ stages of the $N \times N$ star and gets divided into N outputs is, in decibels,

$$\text{Total loss} = \text{splitting loss} + \text{excess loss} = -10\log\left(\frac{F_T^{\log_2 N}}{N}\right)$$

$$= -10\left(\frac{\log N \log F_T}{\log 2} - \log N\right) = 10\,(1 - 3.322 \log F_T)\log N \tag{10.28}$$

This shows that the loss increases logarithmically with N.

Example 10.9 Consider a commercially available 32×32 single-mode coupler made from a cascade of 3-dB fused-fiber 2×2 couplers, where 5 percent of the power is lost in each element. What are the excess and splitting losses for this coupler?

Solution: From Eq. (10.27), the excess loss is

$$\text{Excess loss} = -10 \log (0.95^{\log 32/\log 2}) = 1.1 \text{dB}$$

and, from Eq. (10.24), the splitting loss is

$$\text{Splitting loss} = -10 \log 32 = 15 \text{dB}$$

Hence, the total loss is 16.1 db.

10.2.5 Mach-Zehnder Interferometer Multiplexers

Wavelength-dependent multiplexers can also be made using Mach-Zehnder interferometry techniques.[24, 31–34] These devices can be either active or passive. Here, we look first at passive multiplexers. Figure 10.15 illustrates the constituents of an individual Mach-Zehnder interferometer (MZI). This 2×2 MZI consists of three stages: an initial 3-dB directional coupler that splits the input signals, a central section where one of the waveguides is longer by ΔL to give a wavelength-dependent phase shift between the two arms, and another 3-dB coupler that recombines the signals at the output. As we will see in the following derivation, the function of this arrangement is that, by splitting the input beam and introducing a phase shift in one of the paths, the recombined signals will interfere constructively at one output and destructively at the other. The signals then finally emerge from only one output port. For simplicity, the following analysis does not take into account waveguide material losses or bend losses.

The propagation matrix $\mathbf{M}_{\text{coupler}}$ for a coupler of length d is

$$\mathbf{M}_{\text{coupler}} = \begin{bmatrix} \cos \kappa d & j \sin \kappa d \\ j \sin \kappa d & \cos \kappa d \end{bmatrix} \tag{10.29}$$

where κ is the coupling coefficient. Since we are considering 3-dB couplers that divide the power equally, then $2\kappa d = \pi/2$, so that

$$\mathbf{M}_{\text{coupler}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & j \\ j & 1 \end{bmatrix} \tag{10.30}$$

In the central region, when the signals in the two arms come from the same light source, the outputs from these two guides have a phase difference $\Delta\phi$ given by

$$\Delta\phi = \frac{2\pi n_1}{\lambda} L - \frac{2\pi n_2}{\lambda} (L + \Delta L) \tag{10.31}$$
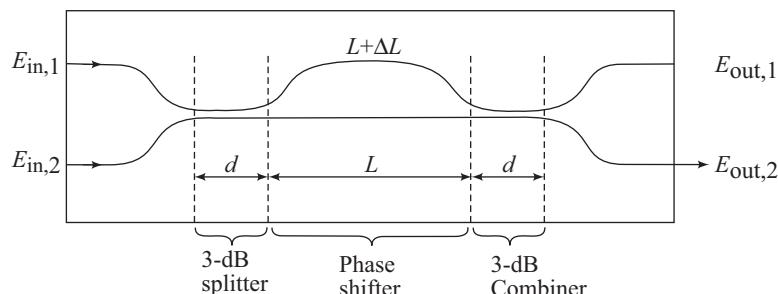
Fig. 10.15 *Layout of a basic 2×2 Mach-Zehnder interferometer*

Note that this phase difference can arise either from a different path length (given by ΔL) or through a refractive index difference if $n_1 \neq n_2$. Here, we take both arms to have the same index and let $n_1 = n_2 = n_{\text{eff}}$ (the effective refractive index in the waveguide). Then we can rewrite Eq. (10.31) as

$$\Delta \phi = k\, \Delta L \tag{10.32}$$

where $k = 2\pi n_{\text{eff}}/\lambda$.

For a given phase difference $\Delta \phi$, the propagation matrix $\mathbf{M}_{\Delta\phi}$ for the phase shifter is

$$M_{\Delta\phi} = \begin{bmatrix} \exp(jk\Delta L/2) & 0 \\ 0 & \exp(-jk\Delta L/2) \end{bmatrix} \tag{10.33}$$

The optical output fields $E_{\text{out},1}$ and $E_{\text{out},2}$ from the two central arms can be related to the input fields $E_{\text{in},1}$ and $E_{\text{in},2}$ by

$$\begin{bmatrix} E_{\text{out},1} \\ E_{\text{out},2} \end{bmatrix} = M \begin{bmatrix} E_{\text{in},1} \\ E_{\text{in},2} \end{bmatrix} \tag{10.34}$$

where

$$M = M_{\text{coupler}} \cdot M_{\Delta\phi} \cdot M_{\text{coupler}} = \begin{bmatrix} M_{11} & M_{21} \\ M_{12} & M_{22} \end{bmatrix} = j\begin{bmatrix} \sin(k\Delta L/2) & \cos(k\Delta L/2) \\ \cos(k\Delta L/2) & -\sin(k\Delta L/2) \end{bmatrix} \tag{10.35}$$

Since we want to build a multiplexer, we need to have the inputs to the MZI at different wavelengths; that is, $E_{\text{in},1}$ is at λ_1 and $E_{\text{in},2}$ is at λ_2. Then, from Eq. (10.34), the output field $E_{\text{out},1}$ and $E_{\text{out},2}$ are each the sum of the individual contributions from the two input fields:

$$E_{\text{out},1} = j[E_{\text{in},1}(\lambda_1)\sin(k_1\Delta L/2) + E_{\text{in},2}(\lambda_2)\cos(k_2\Delta L/2)] \tag{10.36}$$

$$E_{\text{out},2} = j[E_{\text{in},1}(\lambda_1)\cos(k_1\Delta L/2) - E_{\text{in},2}(\lambda_2)\sin(k_2\Delta L/2)] \tag{10.37}$$

where $k_j = 2\pi n_{\text{eff}}/\lambda_j$. The output powers are then found from the light intensity, which is the square of the field strengths. Thus,

$$P_{\text{out},1} = E_{\text{out},1}E_{\text{out},1}^* = \sin^2(k_1\Delta L/2)P_{\text{in},1} + \cos^2(k_2\Delta L/2)P_{\text{in},2} \tag{10.38}$$

$$P_{\text{out},2} = E_{\text{out},2}E_{\text{out},2}^* = \cos^2(k_1\Delta L/2)P_{\text{in},1} + \sin^2(k_2\Delta L/2)P_{\text{in},2} \tag{10.39}$$

where $P_{\text{in},j} = |E_{\text{in},j}|^2 = E_{\text{in},j} \cdot E_{\text{in},j}^*$. In deriving Eqs. (10.38) and (10.39), the cross terms are dropped because their frequency, which is twice the optical carrier frequency, is beyond the response capability of the photodetector.

From Eqs. (10.38) and (10.39), we see that if we want all the power from both inputs to leave the same output port (e.g., port 2), we need to have $k_1 \Delta L/2 = \pi$ and $k_2 \Delta L/2 = \pi/2$, or

$$(k_1 - k_2)\,\Delta L = 2\pi n_{\text{eff}}\left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right)\Delta L = \pi \tag{10.40}$$

Hence, the length difference in the interferometer arms should be

$$\Delta L = \left[2n_{\text{eff}} \left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2} \right) \right]^{-1} = \frac{c}{2n_{\text{eff}}\Delta v}$$ (10.41)

where Δv is the frequency separation of the two wavelengths.

Example 10.10   (*a*) Assume that the input wavelengths of a 2×2 silicon MZI are separated by 10 GHz (i.e., $\Delta\lambda = 0.08$ nm at 1550 nm). With $n_{\text{eff}} = 1.5$ in a silicon waveguide, we have from (Eq. 10.41) that the waveguide length difference must be

$$\Delta L = \frac{3 \times 10^8 \,\text{m/s}}{2(1.5)10^{10}\,/\text{s}} = 10 \text{ mm}$$

(*b*) If the frequency separation is 130 GHz (i.e., $\Delta\lambda = 1$ nm), then $\Delta L, = 0.77$ mm.

Using basic 2×2 MZIs, any size $N \times N$ multiplexer (with $N = 2^n$) can be constructed. Figure 10.16 gives an example for a 4×4 multiplexer.[34] Here the inputs to MZI_1 are v and $v + 2\Delta v$ (which we will call λ_1 and λ_3, respectively), and the inputs to MZI_2 are $v + \Delta v$ and $v + 3\Delta v$ (λ_2 and λ_4, respectively). Since the signals in both interferometers of the first stage are separated by $2\Delta v$, the path differences satisfy the condition

$$\Delta L_1 = \Delta L_2 = \frac{c}{2n_{\text{eff}}(2\Delta v)}$$ (10.42)

In the next stage, the inputs are separated by Δv. Consequently, we need to have

$$\Delta L_3 = \frac{c}{2n_{\text{eff}}\Delta v} = 2\Delta L_1$$ (10.43)

When these conditions are satisfied, all four input powers will emerge from port C.
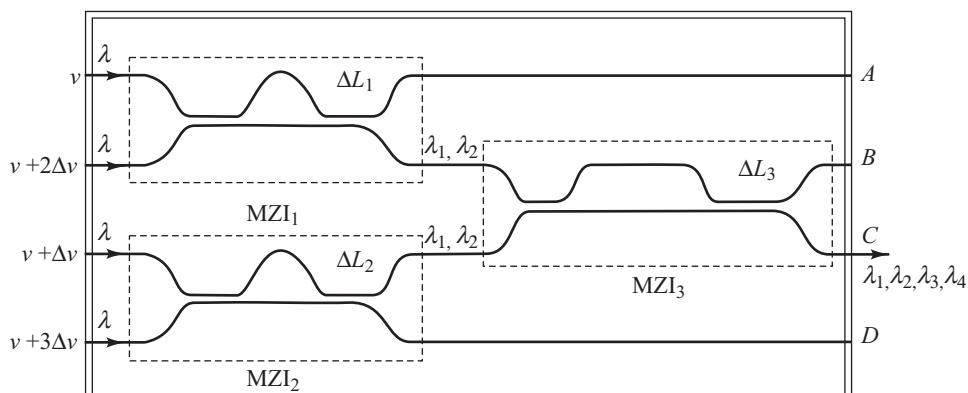
Fig. 10.16   *Example of a 4-channel wavelength multiplexer using three 2×2 MZI elements (Adapted with permission from Verbeek et al.,[34] © 1988, IEEE.)*

From this design example, we can deduce that for an N-to-1 MZI multiplexer, where $N = 2^n$ with the integer $n \geq 1$, the number of multiplexer stages is n and the number of MZIs in stage j is 2^{n-j}. The path difference in an interferometer element of stage j is thus

$$\Delta L_{\text{stage } j} = \frac{c}{2^{n-j} n_{\text{eff}} \Delta v} \tag{10.44}$$

The N-to-1 MZI multiplexer can also be used as a 1-to-N demultiplexer by reversing the light-propagation direction. For a real MZI, the ideal case given in these examples needs to be modified to have slightly different values of ΔL_1 and ΔL_2.

10.3 Isolators and Circulators

In a number of applications it is desirable to have a passive optical device that is nonreciprocal; that is, it works differently when its inputs and outputs are reversed. Two examples of such a device are isolators and circulators. To understand the operation of these devices, we need to recall some facts about polarization and polarization-sensitive components from Chapter 2:

- Light can be represented as a combination of a parallel vibration and a perpendicular vibration, which are called the two *orthogonal plane polarization states* of a lightwave.
- A *polarizer* is a material or device that transmits only one polarization component and blocks the other.
- A *Faraday rotator* is a device that rotates the state of polarization (SOP) of light passing through it by a specific angular amount.
- A device made from birefringent materials (called a *walk-off polarizer*) splits the light signal entering it into two orthogonally (perpendicularly) polarized beams, which then follow different paths through the material.
- A *half-wave plate* rotates the SOP clockwise by 45° for signals going from left to right, and counterclockwise by 45° for signals propagating in the other direction.

10.3.1 Optical Isolators

Optical isolators are devices that allow light to pass through them in only one direction. This is important in a number of instances to prevent scattered or reflected light from traveling in the reverse direction. One common application of an optical isolator is to keep such backward-traveling light from entering a laser diode and possibly causing instabilities in the optical output.

Many design configurations of varying complexity exist for optical isolators. The simple ones depend on the state of polarization of the input light. However, such a design results in a 3-dB loss (half the power) when unpolarized light is passed through the device because it blocks half of the input signal. In practice the optical isolator should be independent of the SOP because light in an optical link normally is not polarized.

Figure 10.17 shows a design for a *polarization-independent isolator* that is made of three miniature optical components. The core of the device consists of a 45° Faraday rotator that is placed between two wedge-shaped birefringent plates or walk-off polarizers. These plates consist of a material such as YVO_4 or TiO_2, as described in Chapter 3. Light traveling in the forward direction (left to right in Fig. 10.17) is separated into ordinary and extraordinary rays by the first birefringent plate. The Faraday rotator then rotates the polarization plane of each ray by 45°. After exiting the Faraday rotator, the two
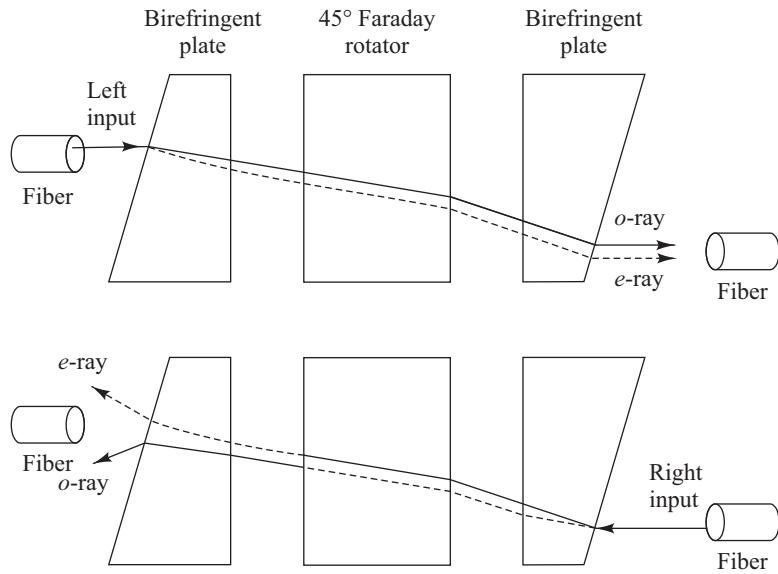
Fig. 10.17 *Design and operation of a polarization-independent isolator made of three miniature optical components*

rays pass through the second birefringent plate. The axis of this polarizer plate is oriented in such a way that the relationship between the two types of rays is maintained. Thus, when they exit the polarizer, they both are refracted in an identical parallel direction. Going in the reverse direction (right to left), the relationship of the ordinary and extraordinary rays is reversed when exiting the Faraday rotator due to the nonreciprocity of the Faraday rotation. Consequently, the rays diverge when they exit the left-hand birefringent plate and are not coupled to the fiber anymore.

Table 10.3 lists some operational characteristics of commercially available isolators. The packages have similar configurations as the tap coupler shown in Fig. 10.7.

Table 10.3 *Typical parameter values of commercially available optical isolators*

Parameter	Unit	Value
Central wavelength λc	nm	1310, 1550
Peak isolation	dB	40
Isolation at $\lambda c \pm 20$ nm	dB	30
Insertion loss	dB	< 0.5
Polarization-dependent loss	dB	< 0.1
Polarization-mode dispersion	ps	< 0.25
Size (diameter $\times$ length)	mm	6×35
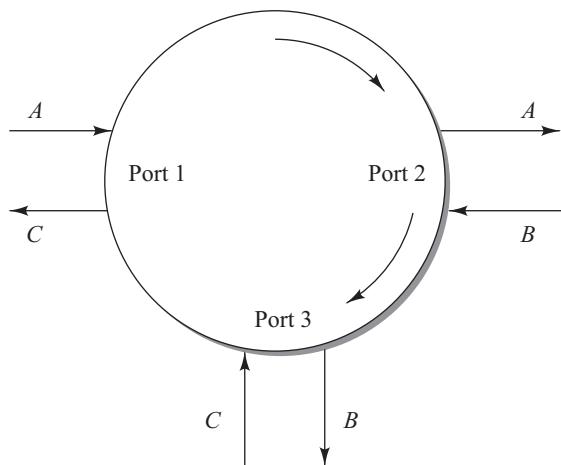
Fig. 10.18 *Operational concept of a three-port circulator*

10.3.2 Optical Circulators

An *optical circulator* is a nonreciprocal multiport passive device that directs light sequentially from port to port in only one direction. This device is used in optical amplifiers, add/drop multiplexers, and dispersion compensation modules. The operation of a circulator is similar to that of an isolator except that its construction is more complex. Typically it consists of a number of walk-off polarizers, half-wave plates, and Faraday rotators and has three or four ports, as shown in Fig. 10.18. To see how it works, consider the three-port circulator. Here an input on port 1 is sent out on port 2, an input on port 2 is sent out on port 3, and an input on port 3 is sent out on port 1.

Similarly, in a four-port device ideally one could have four inputs and four outputs if the circulator is perfectly symmetrical. However, in actual applications it usually is not necessary to have four inputs and four outputs. Furthermore, such a perfectly symmetrical circulator is rather tedious to fabricate. Therefore in a four-port circulator it is common to have three input ports and three output ports, making port 1 be an input-only port, 2 and 3 being input and output ports, and port 4 be an output-only port.

A variety of circulators are available commercially. These devices have low insertion loss, high isolation over a wide wavelength range, minimal polarization-dependent loss (PDL), and low polarization-mode dispersion (PMD). Table 10.4 lists some operational characteristics of commercially available circulators.

Table 10.4 *Typical parameter values of commercially available optical circulators*

Parameter	Unit	Value
Wavelength band	nm	C-band: 1525–1565
		L-band: 1570–1610
Insertion loss	dB	< 0.6
Channel isolation	dB	> 40
Optical return loss	dB	> 50
Operating power	mW	< 500
Polarization-dependent loss	dB	< 0.1
Polarization-mode dispersion	ps	< 0.1
Size (diameter × length)	mm	5.5 × 50

10.4 Fiber Grating Filters

A grating is an important element in WDM systems for combining and separating individual wavelengths. Basically, a grating is a periodic structure or perturbation in a material. This variation in the material has the property of reflecting or transmitting light in a certain direction depending on the wavelength. Thus, gratings can be categorized as either reflecting or transmitting gratings.

10.4.1 Grating Basics

Figure 10.19 defines various parameters for a reflection grating. Here, θ_i is the incident angle of the light, θ_d is the diffracted angle, and Λ is the *period of the grating* (the periodicity of the structural variation in the material). In a transmission grating consisting of a series of equally spaced slits, the spacing between two adjacent slits is called the *pitch* of the grating. Constructive interference at a wavelength λ occurs in the imaging plane when the rays diffracted at the angle θ_d satisfy the *grating equation* given by

$$\Lambda (\sin \theta_i - \sin \theta_d) = m\lambda \tag{10.45}$$

Here, m is called the *order* of the grating. In general, only the first-order diffraction condition $m = 1$ is considered. (Note that in some texts the incidence and refraction angles are defined as being measured from the same side of the normal to the grating. In this case, the sign in front of the term $\sin \theta_d$ changes.) A grating can separate individual wavelengths since the grating equation is satisfied at different points in the imaging plane for different wavelengths.

10.4.2 Fiber Bragg Grating

A Bragg grating constructed within an optical fiber constitutes a high-performance device for accessing individual wavelengths in the closely spaced spectrum of dense WDM systems.[35–41] Since this is an all-fiber device, its main advantages are low cost, low loss (around 0.3 dB), ease of coupling with other
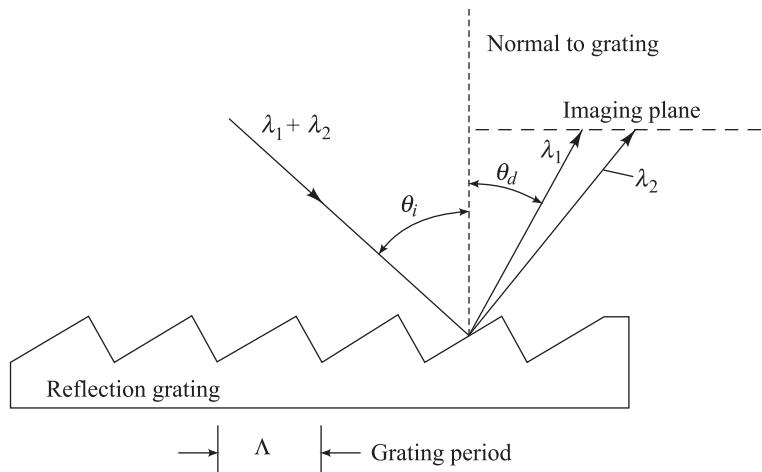
Fig. 10.19 *Basic parameters in a reflection grating*
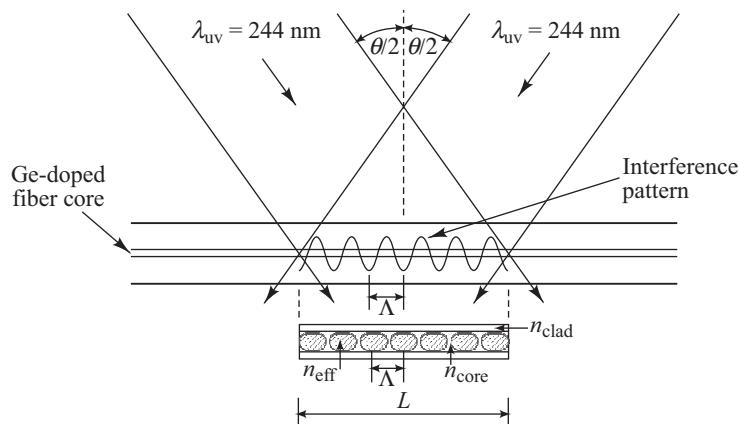
Fig. 10.20 *Formation of a Bragg grating in a fiber core by means of two intersecting ultraviolet light beams*

fibers, polarization insensitivity, low temperature coefficient (< 0.7 pm/°C for an athermal device), and simple packaging. A fiber grating is a narrowband reflection filter that is fabricated through a photoimprinting process. The technique is based on the observation that germanium-doped silica fiber exhibits high photosensitivity to ultraviolet light.[35] This means that one can induce a change in the refractive index of the core by exposing it to ultraviolet radiation such as 244 nm.

Several methods can be used to create a fiber phase-grating. Figure 10.20 demonstrates the so-called *external-writing technique.* The grating fabrication is accomplished by means of two ultraviolet beams transversely irradiating the fiber to produce an interference pattern in the core. Here, the regions of high intensity (denoted by the shaded ovals) cause an increase in the local refractive index of the photosensitive core, whereas it remains unaffected in the zero-intensity regions. A permanent reflective Bragg grating is thus written into the core. When a multi-wavelength signal encounters the grating, those wavelengths that are phase-matched to the Bragg reflection condition are reflected and all others are transmitted.

Using the standard grating equation given by Eq. (10.45), with λ being the wavelength of the ultraviolet light λ_{uv}, the period Λ of the interference pattern (and hence the period of the grating) can be calculated from the angle θ between the two interfering beams of free-space wavelength λ_{uv}. Note from Fig. 10.20 that θ is measured outside of the fiber.

The imprinted grating can be represented as a uniform sinusoidal modulation of the refractive index along the core:

$$n(z) = n_{core} + \delta n \left[1 + \cos\left(\frac{2\pi z}{\Lambda} \right) \right] \qquad (10.46)$$

where n_{core} is the unexposed core refractive index and δn is the photoinduced change in the index.

The maximum reflectivity R of the grating occurs when the Bragg condition holds; that is, at a reflection wavelength λ_{Bragg} where

$$\lambda_{Bragg} = 2\Lambda \, n_{eff} \qquad (10.47)$$

and n_{eff} is the effective index of the core. At this wavelength, the peak reflectivity R_{max} for the grating of length L and coupling coefficient κ is given by (see Prob. 10.17)

$$R_{max} = \tanh^2 (\kappa L) \tag{10.48}$$

The full bandwidth $\Delta\lambda$ over which the maximum reflectivity holds is[37]

$$\Delta\lambda = \frac{\lambda_{Bragg}^2}{\pi n_{eff} L}[(\kappa L)^2 + \pi^2]^{1/2} \tag{10.49}$$

An approximation for the full-width half-maximum (FWHM) bandwidth is

$$\Delta\lambda_{FWHM} \approx \lambda_{Bragg} s \left[\left(\frac{\delta n}{2 n_{core}} \right)^2 + \left(\frac{\Lambda}{L} \right)^2 \right]^{1/2} \tag{10.50}$$

where $s \approx 1$ for strong gratings with near 100 percent reflectivity, and $s \approx 0.5$ for weak gratings.

For a uniform sinusoidal modulation of the index throughout the core, the coupling coefficient κ is given by

$$\kappa = \frac{\pi \delta n \eta}{\lambda_{Bragg}} \tag{10.51}$$

with η being the fraction of optical power contained in the fiber core. Under the assumption that the grating is uniform in the core, η can be approximated by

$$\eta \approx 1 - V^{-2} \tag{10.52}$$

where V is the V number of the fiber. A more precise evaluation is needed for nonuniform or nonsinusoidal index variations.[42]

Example 10.11   (*a*) The table below shows the values of R_{max} as given by Eq. (10.48) for different values of κL:

kL	R_{max} (%)
1	58
2	93
3	98

(*b*) Consider a fiber grating with the following parameters: $L = 0.5$ cm, $\lambda_{Bragg} = 1530$ nm, $n_{eff} = 1.48$, $\delta n = 2.5 \times 10^{-4}$, and $\eta = 82$ percent. From Eq. (10.51) we have $\kappa = 4.2$ cm^{-1}. Substituting this into Eq. (10.49) then yields $\Delta\lambda = 0.38$ nm.

Fiber Bragg gratings are available in a wide range of reflection bandwidths varying from 25 GHz and higher. Table 10.5 lists some operational characteristics of commercially available 25-, 50-, and 100-GHz fiber Bragg gratings for use in optical communication systems.

In the fiber Bragg grating (FBG) illustrated in Fig. 10.20, the grating spacing is uniform along its length. It is also possible to have the spacing vary along the length of the fiber, which means that a range

Table 10.5 *Typical parameter values of commercially available fiber Bragg gratings*

Parameter	Typical value for three channel spacings		
	25 GHz	*50 GHz*	*100 GHz*
Reflection bandwidth	> 0.08 nm @ −0.5 dB	> 0.15 nm @ − 0.5 dB	> 0.3 nm @ −0.5dB
	< 0.2 nm @ −3 dB	< 0.4 nm @ − 3 dB	< 0.75 nm @ −3 dB
	< 0.25 nm @ −25 dB	< 0.5 nm @ − 25 dB	< 1 nm @ −25 dB
Transmission bandwidth	> 0.05 nm @ −25 dB	> 0.1 nm @ − 25 dB	> 0.2 nm @ −25dB
Adjacent channel isolation	> 30 dB		
Insertion loss	< 0.25 dB		
Central λ tolerance	< ± 0.05 nm @ 25°C		
Thermal λ drift	< 1 pm/°C (for an athermal design)		
Package size	5 mm (diameter) × 80 mm (length)		

of different wavelengths will be reflected by the FBG. This is the basis of what is known as a *chirped grating*. Chapter 13 describes the construction and application of such a device in more detail in the discussion on chromatic dispersion compensation.

10.4.3   FBG Applications

Figure 10.21 shows a simple concept of a demultiplexing function using a fiber Bragg grating. To extract the desired wavelength, a circulator is used in conjunction with the grating. Here the four wavelengths enter through port 1 of the circulator and leave from port 2. All wavelengths except λ_2 pass through the grating. Since λ_2 satisfies the Bragg condition of the grating, it gets reflected, enters port 2 of the circulator, and exits at port 3.

To create a device for combining or separating *N* wavelengths, one needs to cascade *N*−1 FBGs and *N*−1 circulators. Figure 10.22 illustrates a multiplexing function for the four wavelengths λ_1, λ_2, λ_3, and λ_4 using three FBGs and three circulators (labeled C_2, C_3, and C_4). The fiber grating filters labeled FBG_2, FBG_3, and FBG_4 are constructed to reflect wavelengths λ_2, λ_3, and λ_4, respectively, and to pass all others.
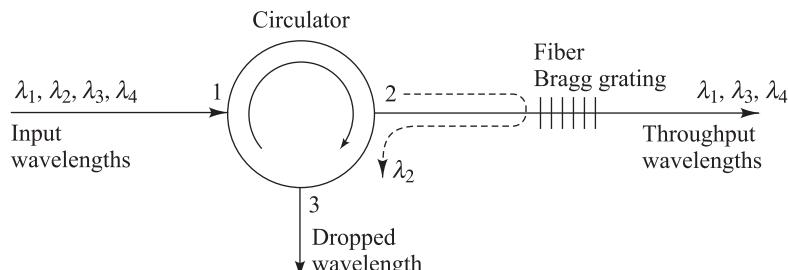
Fig. 10.21   *Simple concept of a demultiplexing function using a fiber grating and an optical circulator*
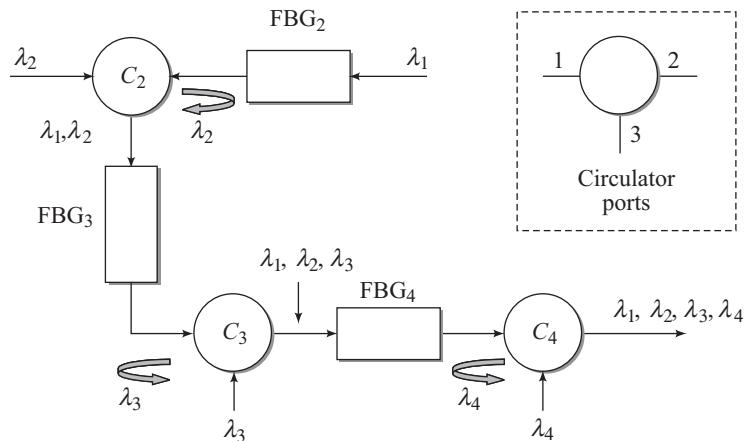
Fig. 10.22 *Multiplexing of four wavelengths using three FBG devices and three circulators*

To see how the multiplexer functions, first consider the combination of circulator C_2 and fiber filter FBG_2. Here filter FBG_2 reflects wavelength λ_2 and allows wavelength λ_1 to pass through. After wavelength λ_1 passes through FBG_2 it enters port 2 of circulator C_2 and exits from port 3. Wavelength λ_2 enters port 1 of circulator C_2 and exits from port 2. After being reflected from FBG_2 it enters port 2 of circulator C_2 and exits from port 3 together with wavelength λ_1. Next at circulator C_3 wavelength λ_3 enters port 3 of circulator C_3 and exits from port 1 and travels toward FBG_3. After being reflected from FBG_3 it enters port 1 of circulator C_3 and exits from port 2 together with wavelengths λ_1 and λ_2. After a similar process takes place at circulator C_4 and filter FBG_4 to insert wavelength λ_4, the four wavelengths all exit together from port 2 of circulator C_4 and can be coupled easily into a fiber.

The coupler size limitation when using fiber Bragg gratings is that one filter is needed for each wavelength and normally the operation is sequential with wavelengths being transmitted by one filter after another. Therefore the losses are not uniform from channel to channel, since each wavelength goes through a different number of circulators and fiber gratings, each of which adds loss to that channel. This may be acceptable for a small number of channels, but the loss differential between the first and last inserted wavelengths is a restriction for large channel counts.

10.5 Dielectric Thin-Film Filters

A dielectric *thin-film filter* (TFF) is used as an *optical bandpass filter*.[43–50] This means that it allows a particular very narrow wavelength band to pass straight through it and reflects all others. The basis of these devices is a classical Fabry-Perot filter structure, which is a cavity formed by two parallel highly reflective mirror surfaces, as shown in Fig. 10.23. This structure is called a *Fabry-Perot interferometer* or an *etalon*. It also is known as a *thin-film resonant cavity filter*.

To see how it works, consider a light signal that is incident on the left surface of the etalon. After the light passes through the cavity and hits the inside surface on the right, some of the light leaves the cavity and some is reflected. The amount of light that is reflected depends on the reflectivity R of the surface. If the roundtrip distance between the two mirrors is an integral multiple of a wavelength λ (i.e., λ, 2λ, 3λ, etc.),
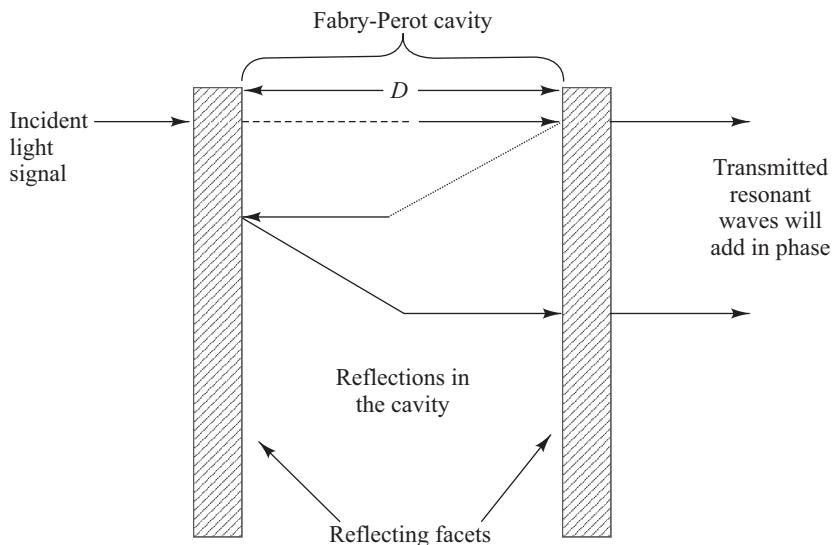
Fig. 10.23 *Two parallel light-reflecting mirrored surfaces define a Fabry-Perot resonator cavity or an etalon.*

then all light at those wavelengths that pass through the right facet *add in phase*. This means that these wavelengths *interfere constructively* in the device output beam so they add in intensity. These wavelengths are called the *resonant wavelengths* of the cavity. The etalon rejects all other wavelengths.

10.5.1 Etalon Theory

The transmission *T* of an ideal etalon in which there is no light absorption by the mirrors is an *Airy function* given by

$$T = \left[1 + \frac{4R}{(1-R)^2}\sin^2\left(\frac{\phi}{2}\right)\right]^{-1} \tag{10.53}$$

where *R* is the *reflectivity* of the mirrors (the fraction of light reflected by the mirror) and ϕ is the roundtrip phase change of the light beam. If one ignores any phase change at the mirror surface, then the phase change for a wavelength λ is

$$\phi = \frac{2\pi}{\lambda}2nD\cos\theta \tag{10.54}$$

where *n* is the refractive index of the dielectric layer that forms the mirror, *D* is the distance between the mirrors, and θ is the angle to the normal of the incoming light beam.

Figure 10.24 gives a generalized plot of Eq. (10.53) over the range $-3\pi \le \varphi \le 3\pi$. Since φ is proportional to the optical frequency $f = 2\pi/\lambda$, Fig. 10.24 shows that the power transfer function *T* is periodic in *f* (or λ). The peaks of the spacings, called the *passbands*, occur at those wavelengths that satisfy the condition $N\lambda = 2nD$, where *N* is an integer. Thus in order for a single wavelength to be selected by the filter from a particular spectral range, all the wavelengths must lie within a passband of the filter transfer function.
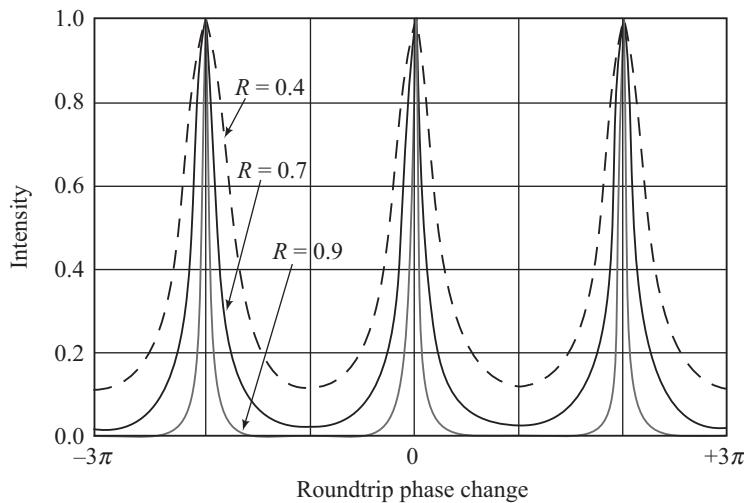
Fig. 10.24 *The behavior of the resonant wavelengths in a Fabry-Perot cavity for three values of the mirror reflectivity based on the Airy function*

If some wavelengths lie outside this range, then the filter would transmit several wavelengths. The distance between adjacent peaks is called the *free spectral range* or FSR. This is given by

$$\text{FSR} = \frac{\lambda^2}{2nD} \tag{10.55}$$

Another important parameter is the measure of the full width of the passband at its half-maximum value, which is designated by FWHM (*full-width half-maximum*). This is of interest in WDM systems for determining how many wavelengths can lie within the FSR of the filter. The ratio FSR/FWHM gives an approximation of the number of wavelengths that a filter can accommodate. This ratio is known as the *finesse F* of the filter and is given by

$$F = \frac{\pi\sqrt{R}}{1 - R} \tag{10.56}$$

A typical TFF consists of multilayer thin-film coatings of alternating low-index and high-index materials, such as SiO_2 and Ta_2O_5, as shown in Fig. 10.25. The layers usually are deposited on a glass
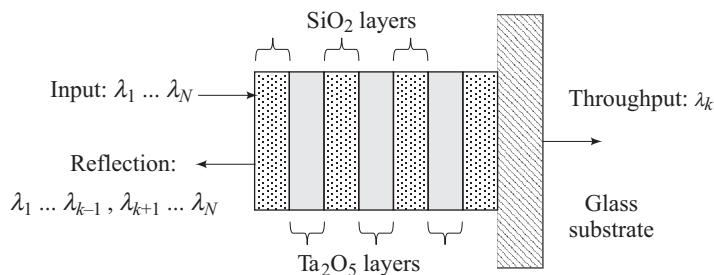
Fig. 10.25 *A multilayer optical thin-film filter consists of a stack of several dielectric thin films*

Table 10.6 *Typical parameter values of commercially available 50-GHz thin-film filters*

Parameter	Unit	Value
Channel passband	GHz	$> \pm 10$ at 0.5 dB
Insertion loss at $f_c \pm 10$ GHz	dB	< 3.5
Polarization-dependent loss	dB	< 0.20
Isolation (adjacent channels)	dB	> 25
Isolation (nonadjacent channels)	dB	> 40
Optical return loss	dB	> 45
Polarization-mode dispersion	ps	< 0.2
Chromatic dispersion	ps/nm	< 50

substrate. Each dielectric layer acts as a nonabsorbing reflecting surface, so that the structure is a series of resonance cavities each of which is surrounded by mirrors. As the number of cavities increases, the passband of the filter sharpens up to create a flat top for the filter, which is a desirable characteristic for a practical filter. In Fig. 10.25 the filter is made such that if the input spectrum contains wavelengths λ_1 through λ_N, then only λ_k passes through the device. All the other wavelengths are reflected.

Thin-film filters are available in a wide range of passbands varying from 50 GHz to 800 GHz and higher for widely spaced channels. Table 10.6 lists some operational characteristics of commercially available 50-GHz multilayer dielectric thin-film filters for use in fiber optic communication systems.

10.5.2 TFF Applications

To create a wavelength multiplexing device for combining or separating N wavelength channels, one needs to cascade $N - 1$ thin-film filters. Figure 10.26 illustrates a multiplexing function for the four wavelengths λ_1, λ_2, λ_3, and λ_4. Here the filters labeled TFF$_2$, TFF$_3$, and TFF$_4$ pass wavelengths λ_2, λ_3, and λ_4, respectively, and reflect all others. The filters are set at a slight angle in order to direct light from one
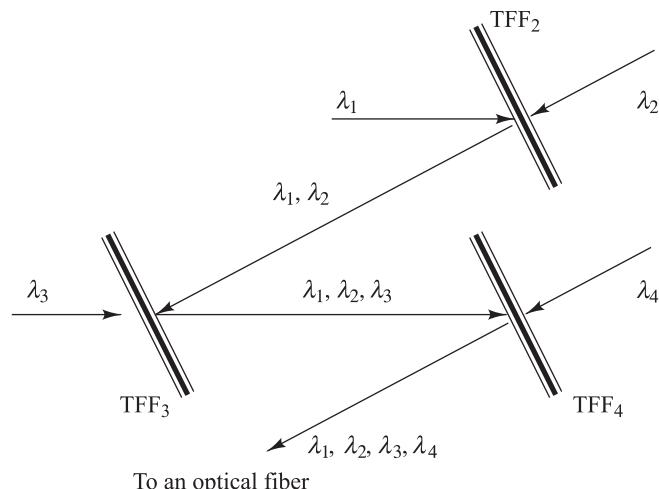
Fig. 10.26 *Multiplexing four wavelengths using thin-film filters*

Table 10.7 *Typical performance parameters for 8-channel DWDM and CWDM multiplexers based on thin-film filter technology*

Parameter	50-GHz DWDM	100-GHzDWDM	20-nm CWDM
Center wavelength accuracy	± 0.1 nm	± 0.1 nm	± 0.3 nm
Channel passband @ 0.5-dB bandwidth	± 0.20 nm	± 0.11 nm	± 6.5 nm
Insertion loss	≤ 1.0 dB	≤ 1.0 dB	≤ 2.0 dB
Ripple in passband	≤ 0.5 dB	≤ 0.5 dB	≤ 0.5 dB
Adjacent channel isolation	≥ 23 dB	≥ 20 dB	≥ 15 dB
Directivity	≥ 50 dB	≥ 55 dB	≥ 50 dB
Optical return loss	≥ 40 dB	≥ 50 dB	≥ 45 dB
Polarization dependent loss	≤ 0.1 dB	≤ 0.1 dB	≤ 0.1 dB
Thermal wavelength drift	< 0.001 nm/°C	< 0.001 nm/°C	< 0.003 nm/°C
Optical power capability	500 mW	500 mW	500 mW

TFF to another. First filter TFF_2 reflects wavelength λ_1 and allows wavelength λ_2 to pass through. These two signals then are reflected from filter TFF_3 where they are joined by wavelength λ_3. After a similar process at filter TFF_4 the four wavelengths can be coupled into a fiber by means of a lens mechanism.

To separate the four wavelengths from one fiber into four physically independent channels, the directions of the arrows in Fig. 10.26 are reversed. Since a light beam loses some of its power at each TFF because the filters are not perfect, this multiplexing architecture works for only a limited number of channels. This usually is specified as being 16 channels or less.

Table 10.7 lists typical performance parameters for commercially available wavelength multiplexers based on thin-film filter technology. The parameters address 8-channel DWDM devices with 50- and 100-GHz channel spacings and an 8-channel CWDM module.

10.6 Phased-Array-Based Devices

A versatile WDM device is based on using an arrayed waveguide grating.[51–60] This device can function as a multiplexer, a demultiplexer, a drop-and-insert element, or a wavelength router. The arrayed waveguide grating is a generalization of the 2×2 Mach-Zehnder interferometer multiplexer. As Fig. 10.27 shows,
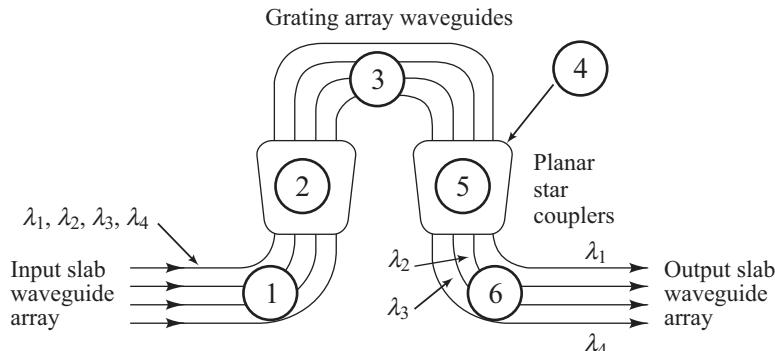
Fig. 10.27 *Top view of a typical arrayed waveguide grating and designation of its various operating regions*
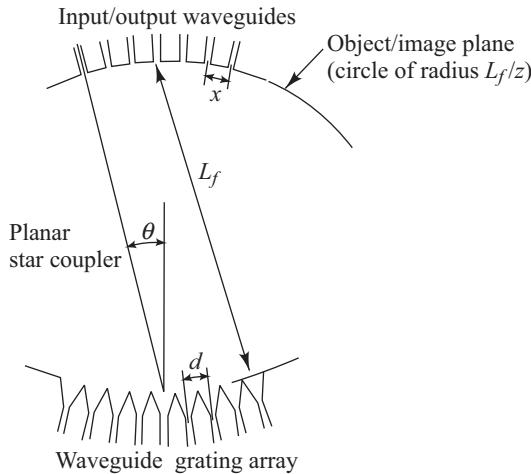
Fig. 10.28 *Geometry of the star coupler used in the arrayed waveguide grating WDM device*

the design consists of M_{in} input and M_{out} output slab waveguides designated as regions 1 and 6, respectively. The slab waveguides interface to two identical focusing planar star couplers located in regions 2 and 5. An array of N uncoupled waveguides that have a propagation constant β connect the star couplers together. In the grating array region, the path length of each waveguide differs by a very precise amount ΔL from the lengths in adjacent arms, so that the array forms a Mach-Zehnder-type grating. For a pure multiplexer, we can take $M_{in} = N$ and $M_{out} = 1$. The reverse holds for a demultiplexer, that is $M_{in} = 1$ and $M_{out} = N$. In the case of a network routing application, we can have $M_{in} = M_{out} = N$.

Figure 10.28 depicts the geometry of the star coupler. The coupler acts as a lens of focal length L_f so that the object and image planes are located at a distance L_f from the transmitter and receiver slab waveguides, respectively. Both the input and output waveguides are positioned on the focal lines, which are circles of radius $L_f/2$. In Fig. 10.28, the parameter x is the center-to-center spacing between the input waveguides and the output waveguides at the star coupler interfaces, d is the spacing between the grating array waveguides, and θ is the diffraction angle in the input or output slab waveguide. The refractive indices of the star coupler and the grating array waveguides are n_s and n_c, respectively.

From a top-level viewpoint of Fig. 10.27, the AWG functions are as follows:

- Starting from the left, the input slab waveguides in region 1 are connected to a planar star coupler (region 2) that acts as a lens.
- The lens distributes the entering optical power among the different waveguides in the grating array in region 3.
- Adjacent waveguides of the grating array in region 3 differ in path length by a precise length ΔL. The path length differences ΔL can be chosen such that all input wavelengths emerge at point 4 with different phase delays $\Delta\Phi = 2\pi n_c \Delta L/\lambda$.
- The second lens in region 5 refocuses the light from all the grating array waveguides onto the output slab waveguide array in region 6.
- Thus each wavelength is focused into a different output waveguide in region 6.

From the phase-matching condition, the light emitted from the output channel waveguides must satisfy the grating equation

$$n_s d \sin\theta + n_c \Delta L = m\lambda \tag{10.57}$$

where the integer m is the diffraction order of the grating.

Focusing is achieved by making the path-length difference ΔL between adjacent array waveguides, measured inside the array, to be an integer multiple of the central design wavelength of the demultiplexer

$$\Delta L = m\frac{\lambda_c}{n_c} \tag{10.58}$$

where λ_c is the central wavelength in vacuum; that is, it is defined as the pass wavelength for the path from the center input waveguide to the center output waveguide.

To determine the channel spacing, we need to find the angular dispersion. This is defined as the incremental lateral displacement of the focal spot along the image plane per unit frequency change, and is found by differentiating Eq. (10.57) with respect to frequency. Doing so, and considering the result in the vicinity of $\theta = 0$, yields

$$\frac{d\theta}{d\nu} = -\frac{m\lambda^2}{n_s cd}\frac{n_g}{n_c} \tag{10.59}$$

where the group index of the grating array waveguide is defined as

$$n_g = n_c - \lambda\frac{dn_c}{d\lambda} \tag{10.60}$$

In terms of frequency, the channel spacing $\Delta\nu$ is

$$\Delta\nu = \frac{x}{L_f}\left(\frac{d\theta}{d\nu}\right)^{-1} = \frac{x}{L_f}\frac{n_s cd}{m\lambda^2}\frac{n_c}{n_g} \tag{10.61}$$

or, in terms of wavelength,

$$\Delta\lambda = \frac{x}{L_f}\frac{n_s d}{m}\frac{n_c}{n_g} = \frac{x}{L_f}\frac{\lambda_0 d}{\Delta L}\frac{n_s}{n_g} \tag{10.62}$$

Equations (10.61) and (10.62) thus define the pass frequencies or wavelengths for which the multiplexer operates, given that it is designed for a central wavelength λ_c. We also note that by making ΔL large, the device can multiplex and demultiplex optical signals with very small wavelength spacings.

Example 10.12   Consider an $N \times N$ waveguide grating multiplexer having $L_f = 10$ mm, $x = d = 5$ μm, $n_c = 1.45$, and a central design wavelength $\lambda_c = 1550$ nm. What is the channel spacing $\Delta\lambda$ for $m = 1$?

Solution: For $m = 1$, the waveguide length difference from Eq. (10.58) is

$$\Delta L = m\frac{\lambda_c}{n_c} = \frac{1550}{1.45} = 1.069 \ \mu m$$

If $n_s = 1.45$ and $n_g = 1.47$, then from Eq. (10.62) we have

$$\Delta\lambda = \frac{x}{L_f}\frac{n_s d}{m}\frac{n_c}{n_g} = \frac{5}{10^4}\frac{(1.45)(5)}{1}\frac{1.45}{1.47} \ \mu m = 3.58 \ nm$$

Equation (10.57) shows that the phased array is periodic for each path through the device, so that after every change of 2π in θ between adjacent waveguides the field will again be imaged at the same spot. The period between two successive field maxima in the frequency domain is called the *free spectral*

range (FSR) and can be represented by the relationship[52]

$$\Delta \nu_{\text{FSR}} = \frac{c}{n_g(\Delta L + d\sin\theta_i + d\sin\theta_o)} \tag{10.63}$$

where θ_i and θ_o are the diffraction angles in the input and output waveguides, respectively. These angles are generally measured from the center of the array, so that we have $\theta_i = jx/L_f$ and $\theta_o = kx/L_f$ for the jth input port and the kth output port, respectively, on either side of the central port. This shows that the FSR depends on which input and output ports the optical signal utilizes. When the ports are across from each other, so that $\theta_i = \theta_o = 0$, then

$$\Delta \nu_{\text{FSR}} = \frac{c}{n_g \Delta L} \tag{10.64}$$

Alternatively, the FSR can be expressed in terms of a wavelength separation $\Delta\lambda_{\text{FSR}}$ as

$$\text{FSR} = \Delta\lambda_{\text{FSR}} = \lambda_c^2 \big/ (\Delta L n_c) \tag{10.65}$$

Example 10.13   As shown in Fig 10.29, suppose an AWG is designed to separate light in the 4-THz-wide frequency range in the C-band running from 195.00 THz (1537.40 nm) to 191.00 THz (1569.59 nm) into forty 100-GHz channels. Then it also will separate the next higher-frequency 4-THz spectral segment in the S-band and lower-frequency 4-THz spectral segment in the L-band into the same forty output fibers. The free spectral range

$\Delta\lambda_{\text{FSR}}$ can be determined from Eq. (10.65). For the 4-THz frequency range denoted here, the center wavelength λ_c is 1550.5 nm, the free spectral range $\Delta\lambda_{\text{FSR}}$ should be at least 32.2 nm in order to separate all the wavelengths into distinct fibers, and the effective refractive index n_c is nominally 1.45 in silica. Then the length difference between adjacent array waveguides is $\Delta L = 51.49 \ \mu$m.

The passband shape of the AWG filter versus wavelength can be altered by the design of the input and output slab waveguides. Two common passband shapes are shown in Fig. 10.30. On the left is the *normal* or *gaussian* passband. This passband shape exhibits the lowest loss at the peak, but the fact that it rolls

| 187 THz | 191 THz | 195 THz | 199 THz |
| (1603.16 nm) | (1569.59 nm) | (1537.40 nm) | (1506.49 nm) |

4-THz FSR

Forty 100-GHz
L-band channels
go into
fibers 1 to 40

Forty 100-GHz
C-band channels
go into
fibers 1 to 40

Forty 100-GHz
S-band channels
go into
fibers 1 to 40
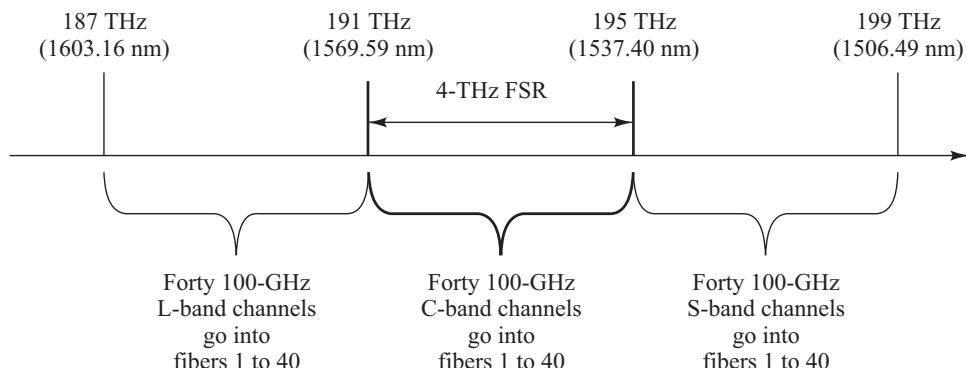
Fig. 10.29   *The FSR specifies the spectral width that will be separated across the output waveguides of an AWG*
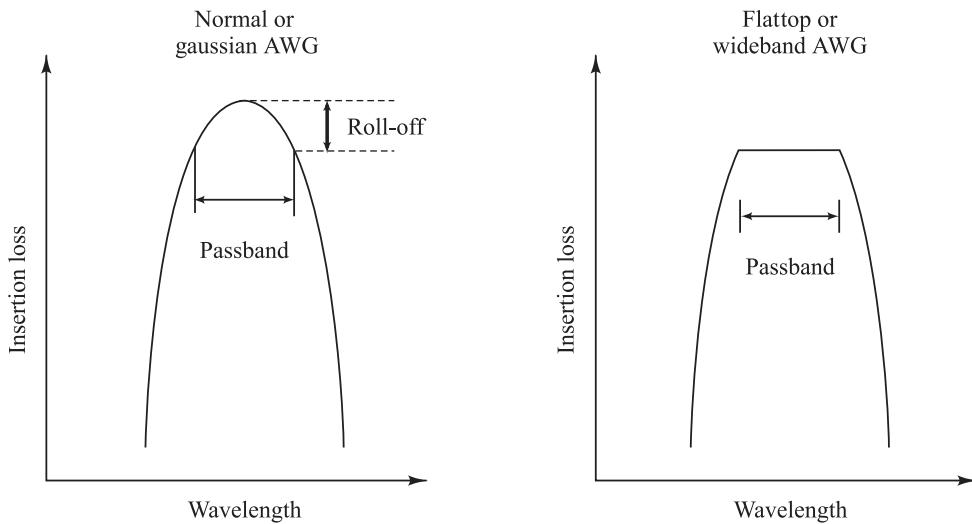
Fig. 10.30 *Two common optical-filter passband shapes: normal or gaussian and flat-top or wideband*

off quickly on either side of the peak means that it requires a high stabilization of the laser wavelength. Furthermore, for applications where the light passes through several AWGs, the accumulative effect of the filtering function reduces the passband to an extremely small value. An alternative to the gaussian passband shape is the *flattop* or *wideband* shape, as shown on the right in Fig. 10.30. This wideband device has a uniform insertion loss across the passband, and is therefore not as sensitive to laser drift or the sensitivity of cascaded filters as is the gaussian passband. However, the loss in a flattop device is usually 2 to 3 dB higher than that in a gaussian AWG. Table 10.8 compares the main operating characteristics of these two designs for a typical 40-channel AWG.

Table 10.8 *Performance characteristics of typical 40-channel arrayed waveguide gratings (AWGs)*

Parameter	Gaussian	Wideband
Channel spacing	100 GHz	100 GHz
1-dB bandwidth	> 0.2 nm	> 0.4 nm
3-dB bandwidth	> 0.4 nm	> 0.6 nm
Insertion loss	< 5dB	< 7 dB
Polarization dependent loss	< 0.25 dB	< 0.15 dB
Adjacent channel crosstalk	30 dB	30 dB
Passband ripple	1.5 dB	0.5 dB
Optical return loss	45 dB	45 dB
Size ($L \times W \times H$)	$130 \times 65 \times 15$ (mm)	$130 \times 65 \times 15$ (mm)

Incident light

$\lambda_1, \lambda_2, \lambda_3, \lambda_4 ...$

λ_1

λ_2

λ_N

Reflection grating

Optical fibers
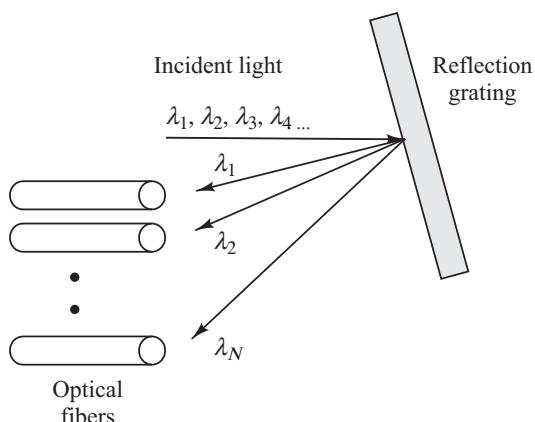
Fig. 10.31 *The angle at which reflected light leaves a reflection grating depends on its wavelength.*

10.7 Diffraction Gratings

Another DWDM technology is based on diffraction gratings.[61–63] A *diffraction grating* is a conventional optical device that spatially separates the different wavelengths contained in a beam of light. The device consists of a set of diffracting elements, such as narrow parallel slits or grooves, separated by a distance comparable to the wavelength of light. These diffracting elements can be either reflective or transmitting, thereby forming a reflection grating or a transmission grating, respectively. Separating and combining wavelengths with diffraction gratings is a parallel process, as opposed to the serial process that is used with the fiber-based Bragg gratings.

Reflection gratings are fine ruled or etched parallel lines on some type of reflective surface. With these gratings, light will bounce off the grating at an angle. The angle at which the light leaves the grating depends on its wavelength, so the reflected light fans out in a spectrum. For DWDM applications, the lines are spaced equally and each individual wavelength will be reflected at a slightly different angle, as shown in Fig. 10.31. There can be a reception fiber at each of the positions where the reflected light gets focused. Thus individual wavelengths will be directed to separate fibers. The reflective diffraction grating works reciprocally; that is, if different wavelengths come into the device on the individual input fibers, all of the wavelengths will be focused back into one fiber after traveling through the device. One also could have a photodiode array in place of the receiving fibers for functions such as power-per-wavelength monitoring.

$d \longrightarrow$

Structural support layers

Λ

α

$\lambda_1, \lambda_2, \lambda_3, \lambda_4 ...$
Incident light

Transmission grating

λ_1

λ_2

λ_N

Optical fibers

Fig. 10.32 *Each wavelength emerges at a slightly different angle after passing through a transmission grating*

One type of *transmission grating*, which is known as a *phase grating,* consists of a periodic variation of the refractive index of the grating. These may be characterized by a *Q*-parameter, which is defined as

$$Q = \frac{2\pi\lambda d}{n_g \Lambda^2 \cos\alpha} \tag{10.66}$$

where λ is wavelength, d is the thickness of the grating, n_g is the refractive index of the material, Λ is the grating period, and α is the incident angle, as shown in Fig. 10.32. The phase grating is called *thin* for $Q < 1$ and *thick* for $Q > 10$. After a spectrum of wavelength channels passes through the grating, each wavelength emerges at a slightly different angle and can be focused into a receiving fiber.

10.8 Active Optical Components

Active components require some type of external energy either to perform their functions or to be used over a wider operating range than a passive device, thereby offering greater application flexibility. These devices include variable optical attenuators, tunable optical filters, dynamic gain equalizers, optical add/drop multiplexers, polarization controllers, and dispersion compensators. Many types of active optical components are based on using micro-electrical-mechanical systems or MEMS technology. This is the topic of Sec. 10.8.1. The remainder of the section describes various active devices.

10.8.1   MEMS Technology

MEMS is the acronym for *micro electro-mechanical systems*.[64,65] These are miniature devices that can combine mechanical, electrical, and optical components to provide sensing and actuation functions. MEMS devices are fabricated using integrated-circuit compatible batch-processing techniques and range in size from micrometers to millimeters. The control or actuation of a MEMS device is done through electrical, thermal, or magnetic means such as micro gears or movable levers, shutters, or mirrors. The devices are used widely in automobile air-bag deployment systems, in ink-jet printer heads, for monitoring mechanical shock and vibration during transportation of sensitive goods, for monitoring the condition of moving machinery for preventative maintenance, and in biomedical applications for patient activity monitoring and pacemakers. MEMS technologies also are finding applications in lightwave systems for variable optical attenuators, tunable optical filters, tunable lasers, optical add-drop multiplexers, optical performance monitors, dynamic gain equalizers, optical switches, and other optical components and modules.

Figure 10.33 shows a simple example of a MEMS *actuation method*. At the top of the device there is a thin suspended polysilicon beam that has typical length, width, and thickness dimensions of 80 μm, 10 μm, and 0.5 μm, respectively. At the bottom there is a silicon ground plane that is covered by an insulator material. There is a gap of nominally 0.6 μm between the beam and the insulator. When
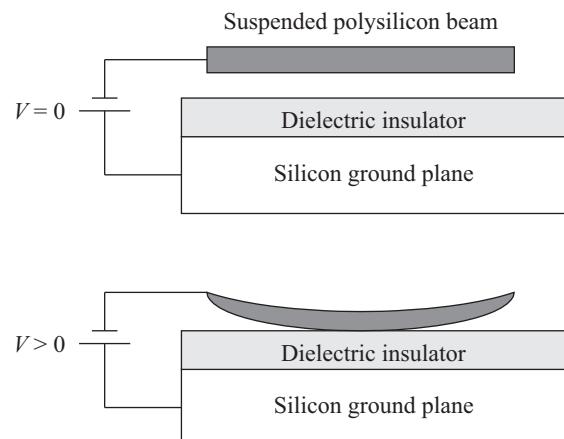
Fig. 10.33   *A simple example of a MEMS actuation method. The top shows an "off" position and the bottom shows an "on" position.*

a voltage is applied between the silicon ground plane and the polysilicon beam, the electric force pulls the beam down so that it makes contact with the lower structure.

Initially MEMS devices were based on standard silicon technology, which is a very stiff material. Since some type of electric force typically is used to bend or deflect one of the MEMS layers to produce the desired mechanical motion, stiffer materials require higher voltages to achieve a given mechanical deflection. To reduce these required forces, current MEMS devices are being made with highly compliant polymeric materials that are as much as six orders of magnitude less stiff than silicon. This class of components is referred to as *compliant MEMS* or CMEMS. This technology employs a soft, rubberlike material called an *elastomer* (from the words *elastic* and *polymer*). Elastomeric material can be stretched as much as 300 percent, as opposed to less than 1 percent for silicon. As a result, CMEMS devices require much lower voltages to achieve a given mechanical deflection, and for equivalent voltages their mechanical range of motion is much larger than with silicon MEMS.

10.8.2 Variable Optical Attenuators

Precise active signal-level control is essential for proper operation of DWDM networks. For example, all wavelength channels exiting an optical amplifier need to have the same gain level, certain channels may need to be blanked out to perform network monitoring, span balancing may be needed to ensure that all signal strengths at a user location are the same, and signal attenuation may be needed at the receiver to prevent photodetector saturation. A *variable optical attenuator* (VOA) offers such dynamic signal-level control. This device attenuates optical power by various means to control signal levels precisely without disturbing other properties of a light signal. That means they should be polarization independent, attenuate light independent of the wavelength, and have a low insertion loss. In addition, they should have a dynamic range of 15 to 30 dB (a control factor ranging from 30 to 1000).

The control methods include *mechanical*, *thermo-optic*, *MEMS*, or *electro-optic techniques*. The mechanical control methods are reliable but have a low dynamic range and a slow response time. Thermo-optic methods have a high dynamic range, but are slow and require the use of a thermoelectric cooler (TEC), which may not be desirable. The two most popular control methods are MEMS-based and electro-optic-based techniques. For MEMS techniques an electrostatic actuation method is the most common and well-developed, since integrated-circuit processes offer a wide selection of conductive and insulating materials. In this method a voltage change across a pair of electrodes provides an electrostatic actuation force. This requires lower power levels than other methods and is the fastest.

When wavelengths are added, dropped, or routed in a WDM system, a VOA can manage the optical power fluctuations of these wavelengths and other simultaneously propagating wavelength signals. Table 10.9 shows some representative operational parameter values for a VOA.

Table 10.9 *Representative operational parameter values for a typical VOA*

Parameter	Specification
Insertion loss	< 1.8 dB
Attenuation range	> 15 dB (up to 60 dB possible)
PDL @ 25 dB attenuation	< 0.3 dB
Maximum optical power per channel	> 150 mW (up to 500 mW possible)
Optical return loss	> 42 dB

10.8.3 Tunable Optical Filters

Tunable optical filters are key components for dense WDM optical networks. Two main technologies to make a tunable filter are MEMS-based and Bragg-grating-based devices. MEMS actuated filters have the advantageous characteristics of a wide tuning range and design flexibility. One such filter is a tunable variation on the classical structure that has been used widely for interferometer applications. The MEMS-based device consists of two sets of epitaxially grown semiconductor layers that form a single Fabry-Perot cavity. The device operation is based on allowing one of the two mirrors to be moved precisely by an actuator. This enables a change in the distance between the two cavity mirrors, thereby resulting in the selection of different wavelengths to be filtered (see Sec. 10.5).

Fiber Bragg gratings are wavelength-selective reflective filters with steep spectral profiles, as shown in Fig. 10.34. Tunable optical filters based on fiber Bragg gratings involve a stretching and relaxation process of the spacing in the fiber grating, that is, in the periodic variation in the refractive index along the core. Since glass is a slightly stretchable medium, as an optical fiber is stretched with the grating inside of it, the spacing of the index perturbations and the refractive index will change. This process induces a change in the Bragg wavelength thereby changing the center wavelength of the filter. Before it is stretched, the center wavelength λ_c of a fiber Bragg grating filter is given by $\lambda_c = 2n_{\text{eff}}\Lambda$, where n_{eff} is the effective index of the fiber containing the grating and Λ (*lambda*) is the period of the index variation of the grating. When elongating the fiber grating by a distance $\Delta\Lambda$, the corresponding change in the center wavelength is $\Delta\lambda_c = 2n_{\text{eff}}\Delta\Lambda$. Such optical filters can be made for the S-, C-, and L-bands and for operation in the 1310-nm region.

The stretching can be done by thermo-mechanical, piezoelectric, or stepper-motor means, as shown in Fig. 10.35. The thermo-mechanical methods might use a bimetal differential-expansion element that changes its shape as its temperature varies. In the figure the high-expansion bar changes its length more with temperature than the low-expansion frame, thereby leading to temperature-induced length variations in the fiber grating. This method is inexpensive but it is slow, takes time to stabilize, and has a limited tuning range. The *piezoelectric technique* uses a material that changes its length when a voltage is applied. Although this method provides precise wavelength resolution, it is more expensive,
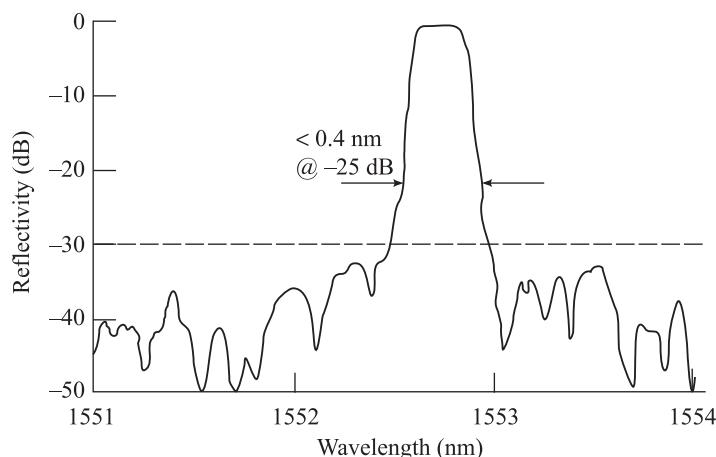
Fig. 10.34 *Example of the reflection band and steep spectral profiles for a 50-GHz fiber Bragg grating filter*
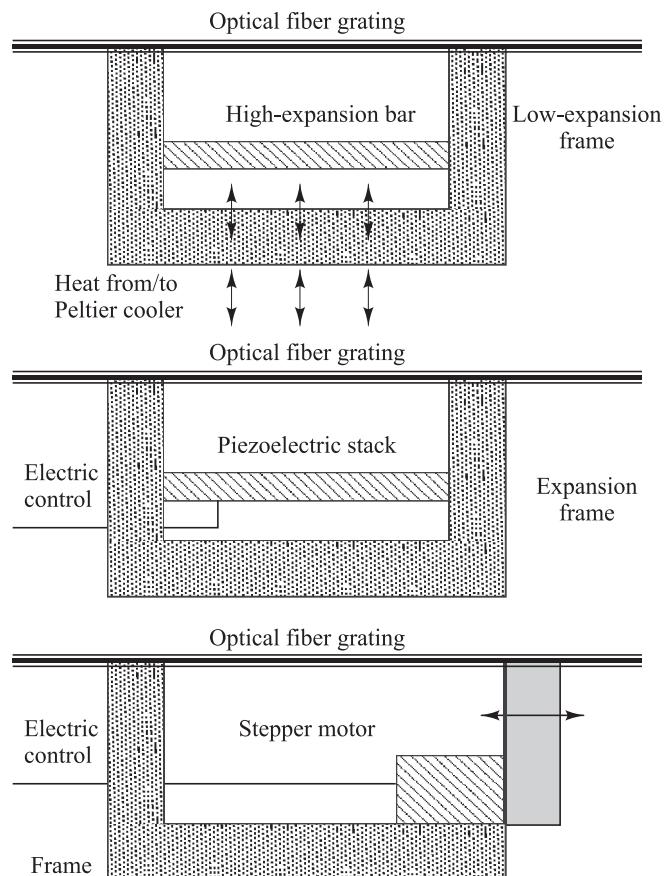
Fig. 10.35   *Three methods for adjusting the wavelength of a tunable Bragg grating*

complex to implement, and has a limited tuning range. The stepper-motor method changes the length of the fiber grating by pulling or relaxing one end of the structure. It has a moderate cost, is reliable, and has a reasonable tuning speed.

Table 10.10 lists representative performance parameters of a tunable optical filter. Applications of these devices include gain-tilt monitoring in optical fiber amplifiers, optical performance monitoring in central offices, channel selection at the receive side of a WDM link, and suppression of amplified spontaneous emission (ASE) noise in optical amplifiers (see Chapter 11).

10.8.4   Dynamic Gain Equalizers

A *dynamic gain equalizer* (DGE) is used to reduce the attenuation of the individual wavelengths within a spectral band. These devices also are called *dynamic channel equalizers* (DCE) or *dynamic spectral equalizers*. The function of a DGE is equivalent to filtering out individual wavelengths and equalizing them on a channel-by-channel basis. Their applications include flattening the nonlinear gain profile of an optical amplifier (such as an EDFA or the Raman amplifier described in Chapter 11), compensation for variation in transmission losses on individual channels across a given spectral band within a link, and

Table 10.10 *Typical performance parameters of a tunable optical filter*

Parameter	Specification
Tuning range	40 nm typical
Channel selectivity	100, 50, and 25 GHz
Bandwidth	< 0.2 nm
Insertion loss	< 3 dB across tuning range
Polarization dependent loss (PDL)	< 0.2 dB across tuning range
Tuning speed	Technology dependent
Tuning voltage	12 to 40 V

attenuating, adding, or dropping selective wavelengths. For example, the gain profile across a spectral band containing many wavelengths usually changes and needs to be equalized when one of the wavelengths is suddenly added or dropped on a WDM link. Note that component vendors sometimes distinguish between a DGE for flattening the output of an optical amplifier and a DCE, which is used for channel equalization or add/drop functions. Depending on the application, certain operational parameters such as the channel attenuation range may be different.

These devices operate by having individually tunable attenuators, such as a series of VOAs, control the gain of a small spectral segment across a wide spectral band, such as the C- or L-band. For example, within a 4-THz spectral range (around 32 nm in the C-band) a DGE can individually attenuate the optical power of 40 channels spaced at 100 GHz or 80 channels spaced at 50 GHz. For example, Fig. 10.36 shows how a DGE equalizes the gain profile of an erbium-doped fiber amplifier. The operation of these devices can be controlled electronically and configured by software residing in a microprocessor. This control is based on feedback information received from a performance-monitoring card that provides the parameter values needed to adjust and adapt to required link specifications. This allows a high degree of agility in responding to optical power fluctuations that may result from changing network conditions.

10.8.5   Optical Add/Drop Multiplexers

The function of an *optical add/drop multiplexer* (OADM) is to insert (add) or extract (drop) one or more selected wavelengths at a designated point in an optical network. Figure 10.37 shows a simple OADM configuration that has four input and four output ports. In this case, the add-and-drop functions are controlled by MEMS-based miniature mirrors that are activated selectively to connect the desired fiber
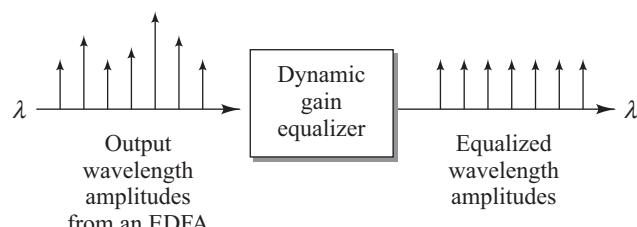
Fig. 10.36 *Example of how a DGE equalizes the gain profile of an erbium-doped fiber amplifier (EDFA)*
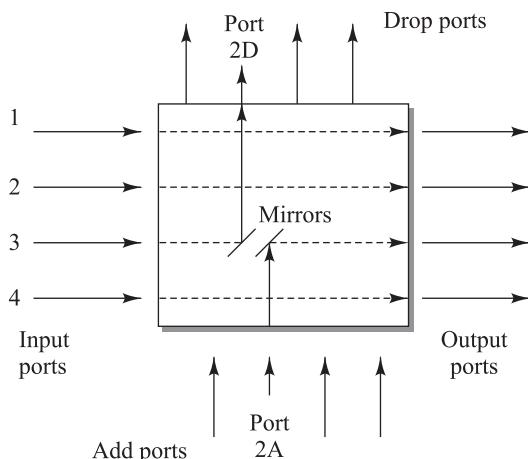
Fig. 10.37 *Example of adding and dropping wavelengths with a 4 × 4 OADM device that uses miniature switching mirrors*

paths. When no mirrors are activated, each incoming channel passes through the switch to the output port. Incoming signals can be dropped from the traffic flow by activating the appropriate mirror pair. For example, to have the signal carried on wavelength λ_3 entering port 3 be dropped to port 2D, the mirrors are activated as shown in Fig. 10.37. When an optical signal is dropped, another path is established simultaneously allowing a new signal to be added from port 2A to the traffic flow.

There are many variations on optical add/drop device configurations depending on the switching technology that is used. However, in each case the operation is independent of wavelength, data rate, and signal format. Chapter 13 describes applications of these devices to routing and switching of optical signals.

10.8.6   Polarization Controllers

Polarization controllers offer high-speed real-time polarization control in a closed-loop system that includes a polarization sensor and control logic. These devices dynamically adjust any incoming state of polarization to an arbitrary output state of polarization. For example, the output could be a fixed, linearly polarized state. Nominally this is done through electronic control voltages that are applied independently to adjustable polarization-retardation plates.

Applications of polarization controllers include *polarization mode dispersion* (PMD) compensation, polarization scrambling, and polarization multiplexing.

10.8.7   Chromatic Dispersion Compensators

A critical factor in optical links operating above 2.5 Gb/s is compensating for chromatic dispersion effects.[66–69] This phenomenon causes pulse broadening, which leads to increased bit-error rates. An effective means of meeting the strict narrow dispersion tolerances for such high-speed networks is to start with a first-order dispersion management method, such as a dispersion compensating fiber, which operates across a wide spectral range. Then fine tuning can be carried out by means of a tunable dispersion compensator that works over a narrow spectral band to correct for any residual and variable dispersion.

The device for achieving this fine tuning is referred to as a *dispersion compensating module* (DCM). Similar to many other devices, this module can be tuned manually, remotely, or dynamically. *Manual tuning* is done by a network technician prior to or after installation of the module in telecommunications racks. By using network management software it can be adjusted *remotely* from a central management console by a network operator if this feature is included in its design. *Dynamic tuning* is done by the module itself without any human intervention.

One method of achieving dynamic chromatic dispersion is through the use of a chirped fiber Bragg grating (FBG), as shown in Fig. 10.38. Here the grating spacing varies linearly over the length of the grating, which creates what is known as a *chirped grating*. This results in a range of wavelengths that satisfy the Bragg condition for reflection. In the configuration shown, the spacing decreases along the fiber, which means that the Bragg wavelength decreases with distance along the grating length. Consequently, the

Fig. 10.38 *Dynamic chromatic dispersion may be accomplished with a chirped fiber Bragg grating*

shorter-wavelength components of a pulse travel farther into the fiber before being reflected. Thereby they experience more delay in going through the grating than the longer-wavelength components. The relative delays induced by the grating on the different frequency components of the pulse are the opposite of the delays caused by the fiber. This results in dispersion compensation because it compresses the pulse.

10.9 Tunable Light Sources

Many different laser designs have been proposed to generate the spectrum of wavelengths needed for DWDM networks. The use of discrete single-wavelength DFB or DBR lasers is the simplest method. Here one hand-selects individual sources, each of which operates at a different wavelength. Although it is straightforward, this method can be expensive because of the high cost of individual lasers. In addition, the sources must be carefully controlled and monitored to ensure that their wavelengths do not drift with time and temperature into the spectral region of adjacent sources.

A more flexible implementation is to have a tunable laser.[70–75] The fundamental concept to making such a laser is to change the cavity length in which the lasing occurs in order to have the device emit at different wavelengths. The basic tuning options include the following:
- Wavelength tuning of a laser by means of temperature or current variations
- Use of a specially designed wavelength tunable (or frequency tunable) laser, such as a multiple-section laser or an external cavity laser
- Frequency locking to a particular lasing mode in a Fabry-Perot laser
- Spectral slicing by means of a fixed or tunable narrow-band optical filter and a broadband LED

With a frequency-tunable laser, one needs only this one source. These devices are based on DFB or DBR structures, which have a waveguide-type grating filter in the lasing cavity (as is described in Chapter 4). Frequency tuning is achieved either by changing the temperature of the device (since the wavelength changes approximately 0.1 nm/°C), or by altering the injection current into the active (gain) section or the passive section (yielding a wavelength change of 0.8×10^{-2} to 4.0×10^{-2} nm/mA, or, equivalently, 1 to 5 GHz/mA). The latter method is generally used. This results in a change in the effective refractive index, which causes a shift in the peak output wavelength. The maximum tuning range depends on the optical output power, with a larger output level resulting in a narrower tuning range. Figure 10.39 illustrates the tuning range of an injection-tunable three-section DBR laser.

The tuning range $\Delta\lambda_{tune}$ can be estimated by

$$\frac{\Delta\lambda_{tune}}{\lambda} = \frac{\Delta n_{eff}}{n_{eff}} \qquad (10.67)$$
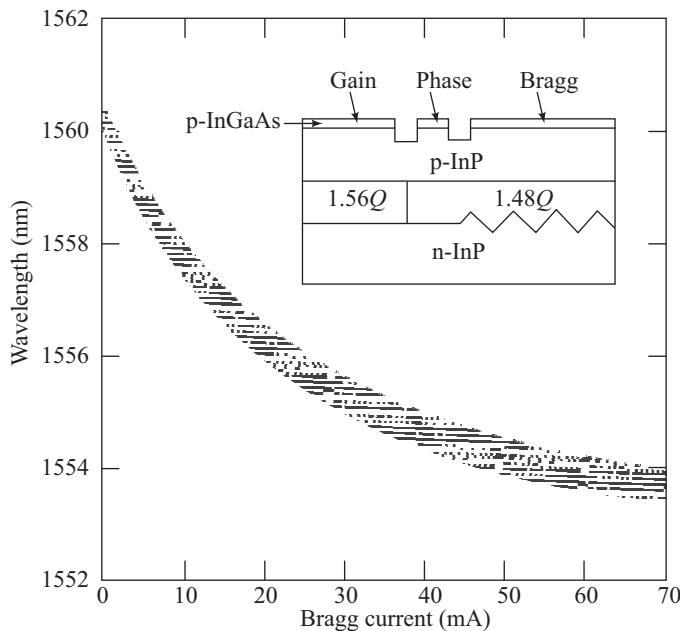
Fig. 10.39 *Tuning range of an injection-tunable three-section DBR laser. (Reproduced with permission from Staring et al.,[75] © 1994, IEEE.)*

where Δn_{eff} is the change in the effective refractive index. Practically, the maximum index change is around 1 percent, resulting in a tuning range of 10–15 nm. Figure 10.40 depicts the relationships between tuning range, channel spacing, and source spectral width. To avoid crosstalk between adjacent channels, a channel spacing of 10 times the source spectral width $\Delta\lambda_{\text{signal}}$ is often specified. That is,

$$\Delta\lambda_{\text{channel}} \approx 10\Delta\lambda_{\text{signal}} \tag{10.68}$$

Thus, the maximum number of channels N that can be placed in the tuning range $\Delta\lambda_{\text{tune}}$ is

$$N \approx \frac{\Delta\lambda_{\text{tune}}}{\Delta\lambda_{\text{channel}}} \tag{10.69}$$
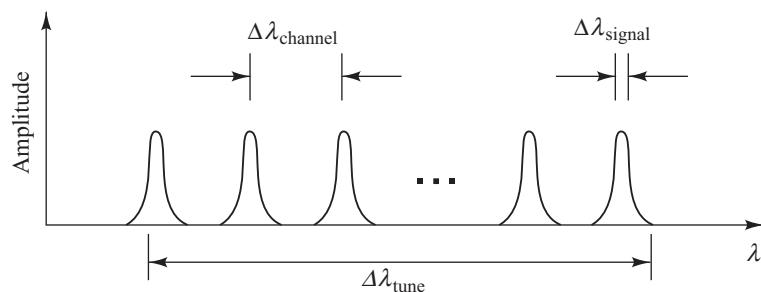
Fig. 10.40 *Relationship between tuning range, channel spacing, and source spectral width*

Example 10.14   Suppose that the maximum index change of a particular DBR laser operating at 1550 nm is 0.65 percent. Then, the tuning range is

$$\Delta\lambda_{\text{tune}} = \lambda \frac{\Delta n_{\text{eff}}}{n_{\text{eff}}} = (1550 \text{ nm})(0.0065) = 10 \text{ nm}$$

If the source spectral width $\Delta\lambda_{\text{signal}}$ is 0.02 nm for a 2.5-Gb/s signal, then using Eqs. 10.68 and 10.69 the number of channels that can operate in this tuning range is

$$N = \frac{\Delta\lambda_{\text{tune}}}{\Delta\lambda_{\text{channel}}} = \frac{10 \text{ nm}}{10 \,(0.02 \text{ nm})} = 50$$

External-cavity laser designs include the use of Littman and Littrow cavities. The *Littman cavity* scheme uses a grating and a MEMS-based tuning mirror to deliver a high level of side-mode suppression (typically 60 dB) with a narrow linewidth (0.3–5 MHz). The *Littrow cavity* method uses a grating to offer an increase in optical output power but with a slight reduction in side-mode suppression (40 dB). In both devices coarse tuning is achieved by manual adjustment of a high-precision adjuster and further fine tuning is achieved by means of a piezoelectric actuator. Various multiple-section tunable lasers have been examined. These designs can include a distributed Bragg reflector, a gain portion, a passive phase-correction section, and a coarse-tuning section. Modulating the Bragg-grating reflector provides a series, or comb, of wavelength peaks. By using an external control current, the coarse tuner then selects one of these peaks. Such a device can be tuned over a 32-nm range, which covers the entire C-band.

Other designs utilize an integrated combination of an optical source (either a broadband laser diode or LED), a waveguide grating multiplexer, and an optical amplifier.[76–80] In this method, which is known as *spectral slicing*, a broad spectral output (for example, from an amplified LED) is spectrally sliced by the waveguide grating to produce a comb of precisely spaced optical frequencies, which become an array of constant-output sources. These spectral slices are then fed into a sequence of individually addressable wavelength channels that can be externally modulated.

PROBLEMS

10.1 A DWDM optical transmission system is designed to have 100-GHz channel spacings. How many wavelength channels can be utilized in the 1536-to-1556-nm spectral band?

10.2 Assume that a 32-channel DWDM system has uniform channel spacings of $\Delta\nu = 100$ GHz and let the frequency ν_n correspond to the wavelength λ_n. Using this correspondence, let the wavelength $\lambda_1 = 1550$ nm. Calculate the wavelength spacing between the first two channels (between channels 1 and 2) and between the last two channels (between channels 31 and 32). From the result, what can be concluded about using an equal-wavelength spacing definition in this wavelength band instead of the standard equal-frequency channel spacing specification?

10.3 Assume that for a given tap coupler the throughput and coupled powers are 230 and 5 μW, respectively, for an input power of 250 μW.
(a) What is the coupling ratio?
(b) What are the insertion losses?
(c) Find the excess loss of the coupler.

10.4 A product sheet for a 2×2 single-mode biconical tapered coupler with a 40/60 splitting ratio states that the insertion losses are 2.7 dB for the 60-percent channel and 4.7 dB for the 40-percent channel.
(a) If the input power $P_0 = 200$ μW, find the output levels P_1 and P_2.
(b) Find the excess loss of the coupler.
(c) From the calculated values of P_1 and P_2, verify that the splitting ratio is 40/60.

10.5 Consider the coupling ratios as a function of pull lengths shown in Fig. 10.41 for a fused biconical tapered coupler. The performances are given for 1310-nm and 1540-nm operation.
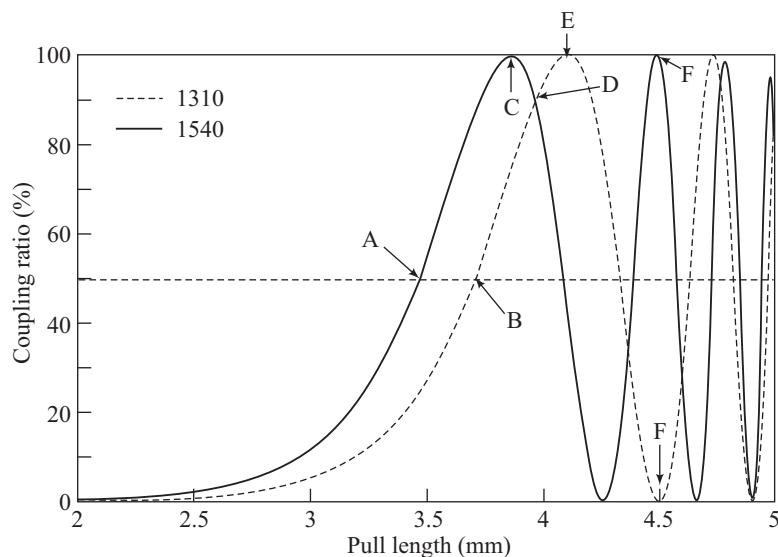
Fig. 10.41 *Pull-length-dependent coupling ratios*

Discuss the behavior of the coupler for each wavelength if its pull length is stopped at the following points: A, B, C, D, E, and each F.

10.6 Consider the 2×2 coupler shown in Fig. 10.8, where **a** and **b** are the matrices representing the field strengths of the input and output propagating waves, respectively. For a given input a_1, we impose the condition that there is no power emerging from the second input port; that is, $a_2 = 0$. Find expressions for the transmissivity T and the reflectivity R in terms of the elements s_{ij} in the scattering matrix **S** given in Eq. (10.8).

10.7 A 2×2 waveguide coupler has $\kappa = 0.4$ mm^{-1}, $\alpha = 0.06$ mm^{-1}, and $\Delta\beta = 0$. How long should the coupler be to make a 3-dB power divider? If that length is doubled, what fraction of the input power emerges from the second channel?

10.8 Suppose we have two 2×2 waveguide couplers (couplers A and B) that have identical channel geometries and spacings, and are formed on the same substrate material. If the index of refraction of coupler A is larger than that of coupler B, which device has a larger

coupling coefficient κ? What does this imply about the device lengths needed in each case to form a 3-dB coupler?

10.9 Measurements on a 7×7 star coupler yield the insertion losses from input port 1 to each output port shown in Table 10.11. Find the total excess loss through the coupler for inputs to port 1.

Table 10.11

Exit port no.	1	2	3	4	5	6	7
Insertion loss (dB)	9.33	7.93	7.53	9.03	9.63	8.64	9.04

10.10 Consider an optical fiber transmission star coupler that has seven inputs and seven outputs. Suppose the coupler is constructed by arranging the seven fibers in a circular pattern (a ring of six with one in the center) and putting them against the end of a glass rod that serves as the mixing element.

(*a*) If the fibers have 50-μm core diameters and 125-μm outer cladding diameters, what is the coupling loss resulting from light escaping between the output fiber cores?

Let the rod diameter be 300 μm. Assume the fiber cladding is not removed.

(b) What is the coupling loss if the fiber ends are arranged in a row and a 50-μm × 800-μm glass plate is used as the star coupler?

10.11 Repeat Prob. 10.10 for seven fibers that have 200-μm core diameters and 400-μm outer cladding diameters. What should the sizes of the glass rod and the glass plate be in this case?

10.12 Suppose an $N \times N$ star coupler is constructed of n 3-dB 2×2 couplers, each of which has a 0.1-dB excess loss. Find the maximum value of n and the maximum size N if the power budget for the star coupler is 30 dB.

10.13 Using Eq. (10.29) for the 2×2 coupler propagation matrix, derive the expressions for M_{11}, M_{12}, M_{21}, and M_{22} in Eq (10.35). From this, find the more general expressions for the output powers given by Eqs. (10.38) and (10.39).

10.14 Consider the 4×4 multiplexer shown in Fig. 10.16.
(a) If $\lambda_1 = 1548$ nm and $\Delta v = 125$ GHz, what are the four input wavelengths?
(b) If $n_{eff} = 1.5$, what are the values of ΔL_1 and ΔL_3?

10.15 Following the same line of analysis as in Example 10.10, use 2×2 Mach-Zehnder interferometers to design an 8-to-1 multiplexer that can handle a channel separation of 25 GHz. Let the shortest wavelength be 1550 nm. Specify the value of ΔL for the 2×2 MZIs in each stage.

10.16 A plane reflection grating can be used as a wavelength-division multiplexer when mounted as shown in Fig. 10.42. The angular properties of this grating are given by the grating equation

$$\sin \phi - \sin \theta = \frac{k\lambda}{n\Lambda}$$

where Λ is the grating period, k is the interference order, n is the refractive index of the medium between the lens and the grating, and ϕ and θ are the angles of the incident and reflected beams, respectively, measured normal to the grating.

(a) Using the grating equation, show that the angular dispersion is given by

$$\frac{d\theta}{d\lambda} = \frac{k}{n\Lambda\cos\theta} = \frac{2\tan\theta}{\lambda}$$

(b) If the fractional beam spread S is given by

$$S = 2(1 + m)\frac{\Delta\lambda}{\lambda}\tan^2\theta$$

where m is the number of wavelength channels, find the upper limit on θ for beam spreading of less than 1 percent given that $\Delta\lambda = 26$ nm, $\lambda = 1350$ nm, and $m = 3$.
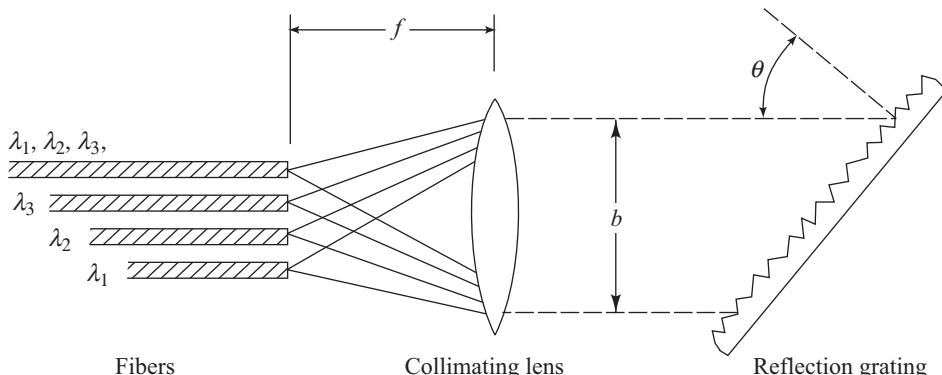
Fig. 10.42 *Wavelength multiplexing with a reflection grating*

10.17 On the same graph, make plots of the reflectivity R_{max} given by Eq. (10.48) and the transmissivity $T = 1 - R_{max}$ for a fiber Bragg grating as a function of κL for $0 \le \kappa L \le 4$. If $\kappa = 0.75$ mm^{-1}, at what grating length does one get 93 percent reflectivity?

10.18 Consider a wavelength multiplexer made from circulators and fiber Bragg gratings as shown in Fig. 10.22. Assume the input optical power of each wavelength λ_1 through λ_4 is 1 mW. Let both the insertion loss and the throughput loss of each FBG be 0.25 dB, and let the insertion loss of a circulator be 0.6 dB. What is the power level of each wavelength λ_1 through λ_4 emerging from the final circulator C_4?

10.19 Based on coupled-mode theory, the reflectivity of a fiber grating is given by[36]

$$R = \frac{(\kappa L)^2 \sinh^2(SL)}{(\delta\beta L)^2 \sinh^2(SL) + (SL)^2 \cosh^2(SL)}$$
$$\text{for } (\kappa L)^2 > (\delta\beta L)^2$$

and

$$R = \frac{(\kappa L)^2 \sin^2(QL)}{(\delta\beta L)^2 - (\kappa L)^2 \cos^2(QL)}$$
$$\text{for } (\kappa L)^2 < (\delta\beta L)^2$$

where

$$SL = (\delta\beta L)\left[\left(\frac{\kappa L}{\delta\beta L}\right)^2 - 1\right]^{1/2}$$

and

$$QL = (\delta\beta L)\left[1 - \left(\frac{\kappa L}{\delta\beta L}\right)^2\right]^{1/2}$$

Here, $\delta\beta = \beta - p\pi/\Lambda = 2\pi n_{eff}/\lambda - p\pi/\Lambda$, with Λ being the grating period and p an integer. For values of $\kappa L = 1, 2, 3,$ and 4, plot $R(\kappa L)$ as a function of $\delta\beta L$ for the range $-10 \le \delta\beta L \le 10$. Note that R_{max} is found by setting $\delta\beta = 0$.

10.20 Using the expression for $R(\kappa L)$ given in Prob. 10.19, derive Eq. (10.49), which gives the full bandwidth $\Delta\lambda$ measured between the zeros on either side of R_{max}.

10.21 A 0.5-cm-long fiber Bragg grating is constructed by irradiating a single-mode fiber with a pair of 244-nm ultraviolet light beams. The fiber has $V = 2.405$ and $n_{eff} = 1.48$. The half-angle between the two beams is $\theta/2 = 13.5°$. If the photo-induced index change is 2.5×10^{-4}, find the following:
(a) the grating period,
(b) the Bragg wavelength,
(c) the coupling coefficient,
(d) the full bandwidth $\Delta\lambda$ measured between the zeros on either side of R_{max}
(e) the maximum reflectivity.

10.22 Consider a wavelength multiplexer made from a series of thin-film filters as shown in Fig. 10.26. Assume the input optical power of each wavelength λ_1 through λ_4 is 1 mW. For each TFF let the throughput loss equal 1.0 dB and let the reflection loss be 0.4 dB. What is the power level of each wavelength λ_1 through λ_4 emerging from the final thin-film filter TFF$_4$?

10.23 Show that Eq. (10.59) follows from the differentiation of Eq. (10.57) with respect to frequency.

10.24 Consider a waveguide grating multiplexer that has the values for the operational variables listed in Table 10.12.
(a) Find the waveguide length difference.
(b) Calculate the channel spacing $\Delta\nu$ and the corresponding pass wavelength differential $\Delta\lambda$.
(c) What is the free spectral range for diagonally opposite ports in this device?
(d) Letting $\theta_i = jx/L_f$ and $\theta_o = kx/L_f$, what is the FSR for $j = 2$ and $k = 8$?

Table 10.12

Symbol	Parameter	Value
L_f	Focal length	9.38 mm
λ_0	Center wavelength	1554 nm
n_c	Array channel index	1.451
n_g	Group index for n_c	1.475
n_s	Slab waveguide index	1.453
x	Input/output waveguide spacing	25 μm
d	Grating waveguide spacing	25 μm
m	Diffraction order	118

10.25 Consider a tunable DBR laser operating at 1550 nm that has a linewidth (frequency spread) of 1.25 GHz. If the maximum index change is 0.55 percent, how many wavelength channels can this laser provide if the channel spacing is 10 times the source spectral width?

10.26 Consider a tunable 2×2 MZI that is constructed on an electro-optical crystal having an effective index of 1.5.

(*a*) If the MZI is used to combine two wavelength channels separated by 0.2 nm, find the required ΔL if the wavelength is centered at 1550 nm.

(*b*) Assume the parameter ΔL is varied by electrically modulating the refractive index of the electro-optical crystal. What is the index change needed if the waveguide length is 100 mm? Note that the effective optical path length L_{eff} in a waveguide is given by $L_{\text{eff}} = n_{\text{eff}} L$.

10.27 The transmitted optical intensity I_t in a tunable fiber Fabry-Perot filter is given by

$$I_t = \frac{I_i}{1 + \dfrac{4R\sin^2 kd}{(1-R)^2}}$$

where I_i is the incident intensity, R is the reflectivity of the dielectric mirrors, k is the propagation constant of the lightwave, and d is the spacing of the mirrors. Then, $2kd$ is the phase of the wave for one roundtrip in the Fabry-Perot cavity. Plot the intensity I_t as a function of kd for values of $R = 0.04$, 0.3, and 0.9 over the range $0 \leq kd \leq 2\pi$.

REFERENCES

1. C. A. Brackett, "Dense wavelength division multiplexing networks: Principles and applications," *IEEE J. Select Areas Commun*. vol. 8, pp. 948–964, Aug. 1990.

2. G. E. Keiser, "A review of WDM technology and applications," *Optical Fiber Tech.*, vol. 5, pp. 3–39, Jan. 1999.

3. O. Krauss, *DWDM and Optical Networks*, Wiley, Hoboken, NJ, 2002.

4. M. Maier, *Metropolitan Area WDM Networks*, Springer, New York, 2003.

5. J. Zyskind and A. Srivastava, *Optically Amplified WDM Networks*, Academic, New York, 2010.

6. J. Zheng and M. T. Mouftah, *Optical WDM Networks: Concepts and Design Principles*, Wiley, Hoboken, NJ, 2004.

7. S. Dixit, ed., *IP over WDM: Building the Next-Generation Optical Network*, Wiley, Hoboken, NJ, 2004.

8. B. Mukherjee, *Optical WDM Networks*, Springer, New York, 2006.

9. G. Keiser, *FTTX Concepts and Applications*, Wiley, Hoboken, NJ, 2006.

10. M. Pfennigbauer and P. J. Winzer, "Choice of MUX/DEMUX filter characteristics for NRZ, RZ, and CSRZ DWDM systems," *J. Lightwave Tech.*, vol. 24, pp. 1689–1696, Apr. 2006.

11. ITU-T Recommendation G.692, *Optical Interfaces for Multichannel Systems with Optical Amplifiers*, Oct. 1998; Amendment 1, Jan. 2005.

12. (a) ITU-T Recommendation G.694.1, *Dense Wavelength Division Multiplexing (DWDM)*, June 2002.

    (b) ITU-T Recommendation G.694.2, *Coarse Wavelength Division Multiplexing (CWDM)*, Dec. 2003.

13. ITU-T Recommendation G.695, *Optical Interfaces for Coarse Wavelength Division Multiplexing Applications*, Feb. 2004.

14. Y. Liu, A. R. Davies, J. D. Ingham, R. V. Penty, and I. H. White, "Uncooled DBR laser directly modulated at 3.125 Gb/s as an athermal transmitter for low-cost WDM systems," *IEEE Photonics Technol. Lett.*, vol. 17, pp. 2026–2028, Oct. 2005.

15. E. Pennings, G.-D. Khoe, M. K. Smit, and T. Staring, "Integrated-optic versus micro optic devices for fiber-optic telecommunication systems: A comparison," *IEEE J. Selected*

Topics Quantum Electronics, vol. 2, pp. 151–164, June 1996.

16. J. M. Liu, *Photonic Devices*, Cambridge University Press, 2009.

17. A. K. Dutta, N. K. Dutta, and M. Fujiwara, eds., *WDM Technologies: Passive Optical Components*, Elsevier, 2003.

18. V. J. Tekippe, "Passive fiber optic components made by the fused biconical taper process," *Fiber & Integrated Optics*, vol. 9, no. 2, pp. 97–123, 1990.

19. M. Eisenmann and E. Weidel, "Single-mode fused biconical couplers for WDM with channel spacing between 100 and 300 nm," *J. Lightwave Tech.*, vol. 6, pp. 113–119, Jan. 1988.

20. R.W.C. Vance and J. D. Love, "Back reflection from fused biconic couplers," *J. Lightwave Tech.*. vol. 13, pp. 2282–2289, Nov. 1995.

21. R. Hui, Yu. Wan, J. Li, S. Jin, J. Lin, and H. Jiang, "III-nitride-based planar lightwave circuits for long wavelength optical communications," *IEEE J. Quantum Electron.*, vol. 41, pp. 100–110, Jan 2005.

22. R. Tewari and K. Thyagarajan, "Analysis of tunable single-mode fiber directional couplers using simple and accurate relations," *J. Lightwave Tech.*, vol. 4, pp. 386–390, Apr. 1986.

23. J. Pietzsch, "Scattering matrix analysis of 3×3 fiber couplers," *J. Lightwave Tech.,* vol. 7, pp. 303–307, Feb. 1989.

24. C.-L. Chen, *Foundations for Guided Wave Optics*, Wiley, Hoboken, NJ, 2007.

25. A. Takagi, K. Jinguji, and M. Kawachi, "Wavelength characteristics of 2×2 optical channel-type directional couplers with symmetric or nonsymmetric coupling structures," *J. Lightwave Tech.*, vol. 10, pp. 735–746, June 1992.

26. R. G. Hunsperger, *Integrated Optics: Theory and Technology*, 6th ed., Springer, New York, 2009.

27. B.E.A. Saleh and M. Teich, *Fundamentals of Photonics,* Wiley, Hoboken, NJ, 2nd ed., 2007.

28. S. Srivastava, N. Gupta, M. Saini, and E. K. Sharma, "Power exchange in coupled optical waveguides," *J. Optical Common.*, vol. 18, pp. 5–9, no. 1, 1997.

29. J. W. Arkwright and D. B. Mortimore, "7×7 monolithic single-mode star coupler," *Electron. Lett.*, vol. 26, pp. 1534–1535, Aug. 1990.

30. J. W. Arkwright, D. B. Mortimore, and R. M. Adams, "Monolithic 1×19 single-mode fused fiber couplers," *Electron. Lett.*, vol. 27, pp. 737–738, Apr. 1991.

31. D. Mechin, P. Yvernault, L. Brilland, and D. Pureur, "Influence of Bragg gratings phase mismatch in a Mach–Zehnder-based add–drop multiplexer," *J. Lightwave Tech.*, vol. 21, pp. 1411–1416, May 2003.

32. K. W. Fussgaenger and R. H. Rossberg, "Uni- and bidirectional $4 \lambda \times 560$ Mb/s transmission systems using WDM devices based on wavelength-selective fused single-mode fiber couplers," *IEEE J. Selected Areas Commun.*, vol. 8, pp. 1032–1042, Aug. 1990.

33. P. Hariharan, *Basics of Interferometry,* Academic, New York, 2007.

34. B. H. Verbeek, C. H. Henry, N. A. Olsson, K. J. Orlowsky, R. F. Kazarinov, and B. H. Johnson, "Integrated four-channel Mach-Zehnder multi/demultiplexer fabricated with phosphorous doped SiO_2 waveguides on Si," *J. Lightwave Tech.*, vol. 6, pp. 1011–1015, June 1988.

35. K. O. Hill, B. Malo, F. Bilodeau, and D. C. Johnson, "Photosensitivity in optical fibers," *Annu. Rev. Mater. Sci.*, vol. 23, pp. 125–157, 1993.

36. R. Kashyap, *Fiber Bragg Gratings*, 2nd ed., Academic, New York, 2010.

37. I. Bennion, J. A. R. Williams, L. Zhang, K. Sugden, and N. J. Doran, "UV-written in-fibre Bragg gratings: A tutorial review," *Optical Quantum Electronics*, vol. 28, pp. 93–135, Feb. 1996.

38. C. S. Goh, M. R. Mokhtar, S. A. Butler, S. Y. Set, K. Kikuchi, and M. Ibsen, "Wavelength tuning of fiber Bragg gratings over 90 nm using a simple tuning package," *IEEE Photonics Technol. Lett.*, vol. 15, pp. 557–559, Apr. 2003.

39. X. Dong, X. Yang, P. Shum, and C. C. Chan, "Tunable WDM filter with 0.8-nm channel spacing using a pair of long-period fiber gratings," *IEEE Photonics Technol. Lett.*, vol. 17, pp. 795–797, Apr. 2005.

40. D. B. Hunter and L.V.T. Nguyen, "Widely tunable RF photonic filter using WDM and

a multichannel chirped fiber grating," *IEEE Trans. Microwave Theory Tech.*, vol. 54, pp. 900–905, Feb. 2006.

41. A. Ozcan, M.J.F. Digonnet, L. Lablonde, D. Pureur, and G. S. Kino, "A new iterative technique to characterize and design transmission fiber Bragg gratings," *J. Lightwave Tech.*, vol. 24, pp. 1913–1921, Apr. 2006.

42. G. P. Agrawal, *Nonlinear Fiber Optics*, Academic, New York, 4th ed., 2006.

43. C. K. Madsen and J. H. Zhao, *Optical Filter Design and Analysis: A Signal Processing Approach*, Wiley, Hoboken, NJ, 1999.

44. H. A. Macleod, *Thin-Film Optical Filters*, Taylor & Francis, London, U.K., 4th ed., 2010.

45. M. Ohring, *Materials Science of Thin Films Deposition and Structure*, Academic, San Diego, CA, 2nd ed., 2002.

46. V. Kochergin, *Omnidirectional Optical Filters,* Springer, New York, 2003.

47. L. Domash, M. Wu, N. Nemchuk, and E. Ma, "Tunable and switchable multiple-cavity thin film filters," *J. Lightwave Tech.*, vol. 22, pp. 126–135, Jan. 2004.

48. J. Jiang, J. J. Pan, Y. H. Guo, and G. Keiser, "Model for analyzing manufacturing-induced internal stresses in 50-GHz DWDM multilayer thin film filters and evaluation of their effects on optical performances," *J. Lightwave Tech.*, vol. 23, pp. 495–503, Feb. 2005.

49. J. H. Song, K.-Y. Kim, J. Cho, D. Han, J. Lee, Y. S. Lee, S. Jung, Y. Oh, D.-H. Jang, and K. S. Lee, "Thin film filter-embedded triplexing-filters based on directional couplers for FTTH networks," *IEEE Photonics Technol. Lett.*, vol. 17, pp. 1668–1670, Aug. 2005.

50. S. Sumriddetchkajorn and K. Chaitavon, "A thin-film filter-based 1×2 reconfigurable fiber-optic add-drop module with a quadruple fiber-optic collimator," *IEEE Photonics Technol. Lett.*, vol. 18, pp. 676–678, Mar. 2006.

51. M. K. Smit and C. van Dam, "PHASAR-based WDM devices: Principles, design and applications," *IEEE J. Selected Topics Quantum Electron.*, vol. 2, pp. 236–250, June 1996.

52. H. Takahashi, K. Oda, H. Toba, and Y. Inoue, "Transmission characteristics of arrayed waveguide $N \times N$ wavelength multiplexers," *J. Lightwave Tech.*, vol. 13, pp. 447–455, Mar. 1995.

53. L. H. Spiekman, M. R. Amersfoort, A. H. de Vreede, F.P.G.M. van Ham, A. Kuntze, J. W. Pedersen, P. Demeester, and M. K. Smit, "Design and realization of polarization independent phased array wavelength demultiplexers," *J. Lightwave Tech.*, vol. 14, pp. 991–995, June 1996.

54. Special Issue on "Arrayed grating routers/WDM mux demuxs and related applications/uses" and "Integrated optics and optoelectronics," *IEEE J. Selected Topics Quantum Electron.*, vol. 8, Nov./Dec. 2002.

55. K. Takada, M. Abe, T. Shibata, and K. Okamoto, "1-GHz-spaced 16-channel arrayed-waveguide grating for a wavelength reference standard in DWDM network systems," *J. Lightwave Tech.*, vol. 20, pp. 850–853, May 2002.

56. Z. Zhang, G. Z. Xiao, J. Liu, C. P. Grover, S. Nikumb, and H. W. Reshef, "A cost-effective solution for packaging the arrayed waveguide grating (AWG) photonic components," *IEEE Trans. Components Packaging Technol.*, vol. 28, pp. 564–570, Sept. 2005.

57. Y.-H. Lin and S.-L. Tsao, "'Improved design of a 64×64 arrayed waveguide grating based on silicon-on-insulator substrate," *IEE Proc.–Optoelectron.*, vol. 153, pp. 57–62, Apr. 2006.

58. S. Bregni, A. Pattavina, and G. Vegetti, "Architectures and performance of AWG-based optical switching nodes for IP networks," *IEEE J. Sel. Areas Commun.*, vol. 21, pp. 1113–1121, Sept. 2003.

59. H. Uetsuka, "AWG technologies for dense WDM applications," *IEEE J. Sel. Topics Quantum Electron.*, vol. 10, pp. 393–402, Mar./Apr. 2004.

60. P. Muñoz, D. Pastor, J. Capmany, D. Ortega, A. E. Pujol, and J. R. Bonar, "AWG model validation through measurement of fabricated devices," *J. Lightwave Tech.*, vol. 22, pp. 2763–2777, Dec. 2004.

61. D. Lazikov, C. Greiner, and T. W. Mossberg, "Apodizable integrated filters for coarse WDM and FTTH-type applications," *J. Lightwave Tech.*, vol. 22, pp. 1402–1407, May 2004.

62. E. G. Loewen and E. Popov, *Diffraction Gratings and Applications*, Taylor & Francis, London, 1997.

63. C.-F. Lin, *Optical Components for Communications: Principles and Applications*, Springer, New York, 2004.

64. A.-Q. Liu, *Photonic MEMS Devices: Design, Fabrication and Control*, CRC Press, Boca Raton, FL, 2009.

65. P. DeDobbelaere, K. Falta, L. Fan, S. Gloeckner, and S. Patra, "Digital MEMS for optical switching," *IEEE Commun. Mag.*, vol. 40, pp. 88–95, March 2002.

66. B. Jopson and A. H. Gnauck, "Dispersion compensation for optical fiber systems," *IEEE Commun. Mag.*, vol. 33, pp. 96–102, June 1995.

67. D. C. Kilper, R. Bach, D. J. Blumenthal, D. Einstein, T. Landolsi, L. Ostar, M. Preiss, and A. E. Willner, "Optical performance monitoring," *J. Lightwave Tech.,* vol. 22, pp. 294–304, Jan. 2004.

68. T. Inui, T. Komukai, K. Mori, and T. Morioka, "160-Gb/s adaptive dispersion of equalization using an asynchronous dispersion-induced chirp monitor," *J. Lightwave Tech.*, vol. 23, pp. 2039–2045, June 2005.

69. D. M. Marom, C. R. Doerr, M. A. Cappuzzo, E. Y. Chen, A. Wong-Foy, L. T. Gomez, and S. Chandrasekhar, "Compact colorless tunable dispersion compensator with 1000-ps/nm tuning range for 40-Gb/s data rates," *J. Lightwave Tech.*, vol. 24, pp. 237–241, Jan. 2006.

70. J. Buus, M.-C. Amann, and D. J. Blumenthal, *Tunable Laser Diodes and Related Optical Sources*, Wiley-IEEE Press, Hoboken, NJ, 2nd ed., 2005.

71. L. A. Coldren, G. A. Fish, Y. Akulova, J. S. Barton, L. Johansson, and C. W. Coldren, "Tunable semiconductor lasers," *J. Lightwave Tech.*, vol. 22, pp. 193–202, Jan. 2004.

72. J. Kani and K. Iwatsuki, "A wavelength-tunable optical transmitter using semiconductor optical amplifiers and an optical tunable filter for metro/access DWDM applications" *J. Lightwave Tech.*, vol. 23, pp. 1164–1169, Mar. 2005.

73. J. Buus and E. J. Murphy, "Tunable lasers in optical networks," *J. Lightwave Tech.*, vol. 24, pp. 5–11, Jan. 2006.

74. Y.-H. Lee, C.-L. Yang, M.-H. Chuang, H.-W. Tseng, Y.-S. Chou, H.-W. Tsao, and S.-L. Lee, "Analysis and selection of optimum driving current combinations for tunable wavelength laser," *Microwave Optical Technol. Lett.*, vol. 48, pp. 1417–1423, July 2006.

75. A.A.M. Staring, J.J.M. Binsma, P. I. Kuindersma, E. J. Jansen, P.J.A. Thijs, T. van Dongen, and G.F.G. Depovere, "Wavelength-independent output power from an injection-tunable DBR laser," *IEEE Photonics Tech. Lett.*, vol. 6, pp. 147–149, Feb. 1994.

76. V. Arya and I. Jacobs, "Optical preamplifier receiver for spectrum-sliced WDM," *J. Lightwave Tech.*, vol. 15, pp. 576–583, Apr. 1997.

77. H. F. Taylor, "Tunable spectral slicing filters for dense wavelength-division multiplexing," *J. Lightwave Tech.*, vol. 21, pp. 837–847, Mar. 2003.

78. S. Kaneko, J.-I. Kani, K. Iwatsuki, A. Ohki, M. Sugo, and S. Kamei, "Scalability of spectrum-sliced DWDM transmission and its expansion using forward error correction," *J. Lightwave Tech.*, vol. 24, pp. 1295–1301, Mar. 2006.

79. H. Kirn, S. Kim, S. Hwang, and Y. Oh, "Impact of dispersion, PMD, and PDL on the performance of spectrum-sliced incoherent light sources using gain-saturated semiconductor optical amplifiers," *J. Lightwave Tech.*, vol. 24, pp. 775–785, Feb. 2006.

80. Y. J. Chai, C. G. Leburn, A. A. Lagatsky, C.T.A. Brown, R. V. Penty, I. H. White, and W. Sibbett, "1.36-Tb/s spectral slicing source based on a Cr-YAG femtosecond laser," *J. Lightwave Tech.,* vol. 23, pp. 1319–1324, Mar. 2005.

CHAPTER 11

Optical Amplifiers

Traditionally, when setting up an optical link, one formulates a power budget and adds repeaters when the path loss exceeds the available power margin. To amplify an optical signal with a conventional repeater, one performs photon-to-electron conversion, electrical amplification, retiming, pulse shaping, and then electron-to-photon conversion. Although this process works well for moderate-speed single-wavelength operation, it can present a data transmission bottleneck for high-speed multiple-wavelength systems. Thus to eliminate the transmission delay problem a great deal of effort has been expended to develop all-optical amplifiers. These devices operate completely in the optical domain to boost the power levels of multiple lightwave signals over spectral bands of 30 nm and more.[1–5]

This chapter first looks at the basic usage of optical amplifiers and classifies the three fundamental amplifier types. These are *semiconductor optical amplifiers* (SOAs), *doped-fiber amplifiers* (DFAs), and *Raman amplifiers*. Section 11.2 discusses SOAs, which are based on the same operating principles as laser diodes. This discussion includes external pumping principles and gain mechanisms. Next, Sec. 11.3 gives details on *erbium-doped fiber amplifiers* (EDFAs), which are widely used in the C-band (1530 to 1565 nm) for optical communication networks. Noise effects generated in the amplification process are discussed in Sec. 11.4. The concept of optical signal-to-noise ratio (OSNR) and its relation to bit-error rate is discussed in Sec. 11.5. The topic of Sec. 11.6 concerns system applications of EDFAs when they are used in three basic locations. Section 11.7 addresses the operation and use of optical amplifiers based on the Raman scattering mechanism. Finally, Sec. 11.8 describes wideband optical amplifiers that operate over several wavelength bands simultaneously.

11.1 Basic Applications and Types of Optical Amplifiers

Optical amplifiers have found widespread use in diverse applications ranging from ultra-long undersea links to short links in access networks. In long-distance undersea and terrestrial point-to-point links, the traffic patterns are relatively stable, so that the input power levels to an optical amplifier do not vary significantly. However, since many closely spaced wavelength channels are being transported over these links, the amplifier must have a wide spectral response range and be highly reliable. Usually fewer wavelengths are carried on metro and access network links, but the traffic patterns can be bursty and wavelengths often can be added or dropped depending on customers' demand for service. Optical amplifiers for these applications thus need to be able to recover quickly from rapid input power variations.

Although these diverse applications offer different optical amplifier design challenges, all devices share some basic operational requirements and performance characteristics, which are given in this section.

11.1.1 General Applications

Figure 11.1 shows general applications of the following three classes of optical amplifiers:

In-line Optical Amplifiers

In a single-mode link, the effects of fiber dispersion may be small so that the main limitation to repeater spacing is fiber attenuation. Since such a link does not necessarily require a complete regeneration of the signal, simple amplification of the optical signal is sufficient. Thus an in-line optical amplifier can be used to compensate for transmission loss and increase the distance between regenerative repeaters, as Fig. 11.1*a* illustrates.

Preamplifier

Figure 11.1*b* shows an optical amplifier being used as a front-end preamplifier for an optical receiver. Thereby, a weak optical signal is amplified before photodetection so that the signal-to-noise ratio
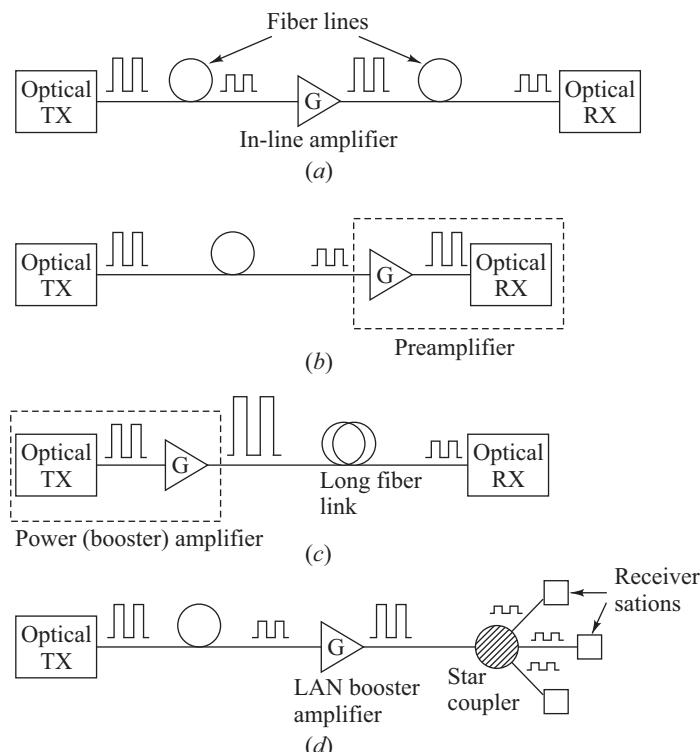
Fig. 11.1 *Four possible applications of optical amplifiers: (a) in-line amplifier to increase transmission distance, (b) preamplifier to improve receiver sensitivity, (c) booster of transmitted power, (d) booster of signal level in a local area network*

degradation caused by thermal noise in the receiver electronics can be suppressed. Compared with other front-end devices such as avalanche photodiodes or optical heterodyne detectors, an optical preamplifier provides a larger gain factor and a broader bandwidth.

Power Amplifier

Power or booster amplifier applications include placing the device immediately after an optical transmitter to boost the transmitted power, as Fig. 11.1*c* shows. This serves to increase the transmission distance by 10–100 km depending on the amplifier gain and fiber loss. As an example, using this technique together with an optical preamplifier at the receiving end can enable repeaterless undersea transmission distances of 200–250 km. One can also employ an optical amplifier in a local area network to compensate for coupler-insertion loss and power-splitting loss. Figure 11.1*d* shows an example for boosting the optical signal in front of a passive star coupler so sufficient power arrives at each receiver.

11.1.2 Amplifier Types

The three main optical amplifier types can be classified as semiconductor optical amplifiers (SOAs), active-fiber or doped-fiber amplifiers (DFAs) and Raman amplifiers. Here we first give a concise overview of these amplifier types; details are in the following sections. All optical amplifiers increase the power level of incident light through a stimulated emission or an optical power transfer process. In SOAs and DFAs the mechanism for creating the population inversion that is needed for stimulated emission to occur is the same as is used in laser diodes. Although the structure of such an optical amplifier is similar to that of a laser, it does not have the optical feedback mechanism that is necessary for lasing to take place. Thus an optical amplifier can boost incoming signal levels, but it cannot generate a coherent optical output by itself. The basic operation is shown in Fig. 11.2. Here, the device absorbs energy supplied from an external source called the *pump*. The pump supplies energy to electrons in an active medium, which raises them to higher energy levels to produce a population inversion. An incoming signal photon will trigger these excited electrons to drop to lower levels through a stimulated-emission process. Since one incoming trigger photon stimulates a cascade effect in which many excited electrons emit photons of equal energy as they drop to the ground state, the result is an amplified optical signal. As Sec. 11.7 describes, in contrast to the amplification mechanisms used in an SOA or DFA, in Raman amplification there is a transfer of optical power from a high-power pump wavelength (e.g., 500 mW at 1480 nm) to lightwave signals at longer wavelengths (e.g., a −25-dBm signal around 1550 nm). This Raman amplification mechanism is done without the need for a population-inversion process.

Alloys of semiconductor materials from groups III and V (e.g., phosphorous, gallium, indium, and arsenic) make up the active medium in SOAs. The attractiveness of SOAs is that devices can be made to work in the O-band (around 1310 nm) as well as in the C-band. They can be integrated easily on
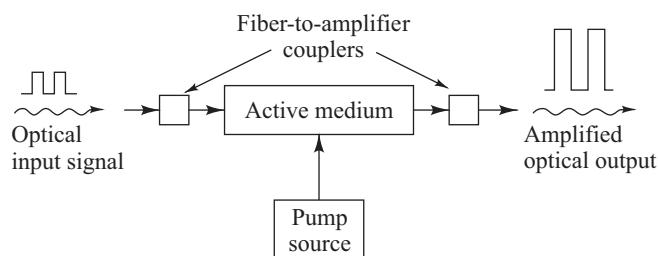
Fig. 11.2 *Basic operation of a generic optical amplifier*

the same substrate as other optical devices and circuits (e.g., couplers, optical isolators, and receiver circuits), and compared with DFAs they consume less electrical power, have fewer components, and are more compact. SOAs have a more rapid gain response, which is on the order of 1 to 100 ps. This rapid response results in both advantages and limitations. The advantage is that SOAs can be implemented when both switching and signal-processing functions are called for in optical networks. The limitation is that the rapid carrier response causes the gain at a particular wavelength to fluctuate with the signal rate for speeds up to several Gb/s. Since this fluctuation affects the overall gain, the signal gain at other wavelengths also fluctuates. Thus the rapid gain response gives rise to crosstalk effects when a broad spectrum of wavelengths must be amplified.

In DFAs, the active medium for operation in the S-, C-, and L-bands is created by lightly doping a silica (silicon dioxide) or tellurite (tellurium oxide) fiber core with rare-earth elements such as thulium (Tm), erbium (Er), or ytterbium (Yb). The DFAs for the O-band are achieved through doping fluoride-based fibers (rather than silica fibers) with elements such as neodymium (Nd) and praseodymium (Pr). The important features of DFAs include the ability to pump the devices at several different wavelengths, low coupling loss to the compatible-sized fiber transmission medium, and very low dependence of gain on light polarization. In addition, DFAs are highly transparent to signal format and bit rate, since they exhibit slow gain dynamics, with carrier lifetimes on the order of 0.1–10 ms. The result is that, in contrast to SOAs, the gain responses of DFAs are basically constant for signal modulations greater than a few kilohertz. Consequently, they are immune from interference effects (such as crosstalk and intermodulation distortion) between different optical channels when wavelength channels in a broad spectrum (e.g., in a 30-nm spectral band ranging from 1530 to 1560 nm) are injected simultaneously into the amplifier.

A Raman optical amplifier is based on a nonlinear effect called *stimulated Raman scattering* (SRS), which occurs in fibers at high optical powers. Whereas a DFA requires a specially constructed optical fiber for its operation, Raman amplification takes place within a standard transmission fiber. The Raman gain mechanism can be achieved through either a lumped (or discrete) amplifier or a distributed amplifier. In the *lumped Raman amplifier* configuration, a spool of about 80 m of small-core fiber along with appropriate pump lasers is inserted into the transmission path as a distinct packaged unit. For the *distributed Raman amplifier* application, one or more Raman pump lasers convert the final 20 to 40 km of the transmission fiber into a preamplifier. Since the Raman gain in a particular spectral range is derived from the SRS-induced transfer of optical power from shorter pump wavelengths to longer signal wavelengths, these amplifiers can be designed for use in any wavelength band.

Table 11.1 lists some possible optical amplifier structures and their operating ranges. The following sections give details on their characteristics.

Table 11.1   *Various optical amplifier structures and their operating regions*

Acronym	*Structure*	*Operating band*
GC-SOA	Gain-clamped semiconductor optical amplifier	O- or C-band
PDFFA	Praseodymium-doped fluoride fiber amplifier	O-band
TDFA	Thulium-doped fiber amplifier	S-band
EDFA	Erbium-doped fiber amplifier	C-band
GS-EDFA	Gain-shifted EDFA	L-band
ETDFA	Er/Tm-doped tellurite (tellurium oxide) glass fiber	C- and L-bands
RFA	Raman fiber amplifier	1260 to 1650 nm

11.2 Semiconductor Optical Amplifiers

A *semiconductor optical amplifier* is essentially an InGaAsP laser that is operating below its threshold point.[3,6–9] Analogous to the construction of a laser diode, the gain peak of an SOA can be selected in any narrow wavelength band extending from 1280 nm in the O-band to 1650 nm in the U-band by varying the composition of the active InGaAsP material. Most SOAs belong to the *traveling-wave* (TW) *amplifier* category. This means that in contrast to the laser feedback mechanism where the optical signal makes many passes through the lasing cavity in the SOA, the optical signal travels through the device only once. During this single passage the signal gains energy and emerges intensified at the other end of the amplifier.

The SOA construction is similar to a resonator cavity structure of a laser diode. The SOA has an active region of length L, width w, and height d. The end facets have reflectivities R_1 and R_2. However, in contrast to a semiconductor laser diode in which the reflectivities are around 0.3, R_1 and R_2 for an SOA are dramatically lower in order for the optical signal to pass through the amplification cavity only once. Low reflectivities of about 10^{-4} are achieved by depositing thin layers of silicon oxide, silicon nitride, or titanium oxide on the SOA end facets.

11.2.1 External Pumping

External current injection is the pumping method used to create the population inversion needed for having a gain mechanism in SOAs. This is similar to the operation of laser diodes (see Sec. 4.3). Thus, from Eq. (4.31), the sum of the injection, stimulated-emission, and spontaneous-recombination rates gives the rate equation that governs the carrier density $n(t)$ in the excited state[3,8]

$$\frac{\partial n(t)}{\partial t} = R_p(t) - R_{st}(t) - \frac{n(t)}{\tau_r} \tag{11.1}$$

where

$$R_p(t) = \frac{J(t)}{qd} \tag{11.2}$$

is the external pumping rate from the injection current density $J(t)$ into an active layer of thickness d, τ_r is the combined time constant coming from spontaneous emission and carrier-recombination mechanisms, and

$$R_{st}(t) = \Gamma a V_g(n - n_{th}) N_{ph} \equiv g V_g N_{ph} \tag{11.3}$$

is the net stimulated emission rate. Here, V_g is the group velocity of the incident light, Γ is the optical confinement factor, a is a gain constant (which depends on the optical frequency v), n_{th} is the threshold carrier density, N_{ph} is the photon density, and g is the overall gain per unit length. Given that the active area of the optical amplifier has a width w and a thickness d, then for an optical signal of power P_s with photons of energy hv and group velocity V_g, the photon density is

$$N_{ph} = \frac{P_s}{V_g(hv)(wd)} \tag{11.4}$$

In the steady state, $\partial n(t)/\partial t = 0$, so that Eq. (11.1) becomes

$$R_p = R_{st} + \frac{n}{\tau_r} \tag{11.5}$$

We now substitute Eq. (11.2) for R_p, the second equality in Eq. (11.3) for R_{st}, and the first equality in Eq. (11.3) solved for n into Eq. (11.5). Solving for g then yields the *steady-state gain per unit length*

$$g = \frac{\dfrac{J}{qd} - \dfrac{n_{th}}{\tau_r}}{V_g N_{ph} + 1/\Gamma a \tau_r} = \frac{g_0}{1 + N_{ph} / N_{ph;sat}} \tag{11.6}$$

where

$$N_{ph;sat} = \frac{1}{\Gamma a V_g \tau_r} \tag{11.7a}$$

is defined as the *saturation photon density*, and

$$g_0 = \Gamma a \tau_r \left(\frac{J}{qd} - \frac{n_{th}}{\tau_r} \right) \tag{11.7b}$$

is the medium gain per unit length in the absence of signal input (when the photon density is zero), which is called the *zero-signal* or *small-signal gain per unit length.*

Example 11.1 Consider an InGaAsP SOA with $w = 5\ \mu m$ and $d = 0.5\ \mu m$. Given that $V_g = 2 \times 10^8$ m/s, if a 1.0-μW optical signal at 1550 nm enters the device, what is the photon density?

Solution: From Eq. (11.4) the photon density is

$$N_{ph} = \frac{1 \times 10^{-6}\,\text{W}}{(2 \times 10^8\,\text{m/s})\dfrac{(6.626 \times 10^{-34}\,\text{J·s})(3 \times 10^8\ \text{m/s})}{1.55 \times 10^{-6}\ \text{m}}(5\,\mu m)(0.5\,\mu m)}$$

$$= 1.56 \times 10^6\ \text{photons/m}^3$$

Example 11.2 Consider the following parameters for a 1300-nm InGaAsP SOA:

Symbol	Parameter	Value
w	Active area width	$3\ \mu m$
d	Active area thickness	$0.3\ \mu m$
L	Amplifier length	$500\ \mu m$
Γ	Confinement factor	0.3
τ_r	Time constant	1 ns
a	Gain coefficient	$2 \times 10^{-20}\,\text{m}^2$
n_{th}	Threshold density	$1.0 \times 10^{24}\,\text{m}^{-3}$

(a) What is the pumping rate for the SOA?
(b) What is the zero-signal gain?

Solution:

(a) If a 100-mA bias current is applied to the device, then, from Eq. (11.2), the pumping rate is

$$R_p = \frac{J}{qd} = \frac{1}{qdwL}$$

$$= \frac{0.1\text{A}}{(1.6 \times 10^{-19}\,\text{C})(0.3\ \mu m)(3\ \mu m)(500\ \mu m)}$$

$$= 1.39 \times 10^{33}\ (\text{electrons/m}^3)/\text{s}$$

(b) From Eq. (11.7b), the zero-signal gain is

$$g_0 = 0.3(2.0 \times 10^{-20}\ \text{m}^2)\,(1\ \text{ns})$$

$$\times \left(1.39 \times 10^{33}\,\text{m}^{-3}\text{s}^{-1} - \frac{1.0 \times 10^{24}\,\text{m}^{-3}}{1.0\ \text{ns}} \right)$$

$$= 2340\ \text{m}^{-1} = 23.4\ \text{cm}^{-1}$$

11.2.2 Amplifier Gain

One of the most important parameters of an optical amplifier is the *signal gain* or *amplifier gain G*, which is defined as

$$G = \frac{P_{s,\text{out}}}{P_{s,\text{in}}} \tag{11.8}$$

where $P_{s,\text{in}}$ and $P_{s,\text{out}}$ are the input and output powers, respectively, of the optical signal being amplified. As noted in Chapter 4, the radiation intensity at a photon energy $h\nu$ varies exponentially with the distance traversed in a lasing cavity. Hence, using Eq. (4.23), the single-pass gain in the active medium of the SOA is

$$G = \exp[\Gamma(g_m - \bar{\alpha})L] \equiv \exp[g(z)L] \tag{11.9}$$

where Γ is the optical confinement factor in the cavity, g_m is the material gain coefficient, $\bar{\alpha}$ is the effective absorption coefficient of the material in the optical path, L is the amplifier length, and $g(z)$ is the overall gain per unit length.

Equation (11.9) shows that the gain increases with device length. However, the internal gain is limited by gain saturation.[3,10,11] This occurs because the carrier density in the gain region of the amplifier depends on the optical input intensity. As the input signal level is increased, excited carriers (electron–hole pairs) are depleted from the active region. When there is a sufficiently large optical input power, further increases in the input signal level no longer yield an appreciable change in the output level because there are not enough excited carriers to provide an appropriate level of stimulated emission. We note here that the carrier density at any point z in the amplifying cavity depends on the signal level $P_s(z)$ at that point. In particular, near the input where z is small, incremental portions of the device may not have reached saturation at the same time as the sections farther down the device, where incremental portions may be saturated because of higher values of $P_s(z)$.

An expression for the gain G as a function of the input power can be derived by examining the gain parameter $g(z)$ in Eq. (11.9). This parameter depends on the carrier density and the signal wavelength. Using Eqs. (11.4) and (11.6), we have that at a distance z from the input end, $g(z)$ is given by

$$g(z) = \frac{g_0}{1 + \dfrac{P_s(z)}{P_{\text{amp,sat}}}} \tag{11.10}$$

where g_0 is the unsaturated medium gain per unit length in the absence of signal input, $P_s(z)$ is the internal signal power at point z, and $P_{\text{amp,sat}}$ is the *amplifier saturation power*, which is defined as the internal power level at which the gain per unit length has been halved. Thus the gain given by Eq. (11.9) decreases with increasing signal power. In particular, the gain coefficient in Eq. (11.10) is reduced by a factor of 2 when the internal signal power is equal to the amplifier saturation power.

Given that $g(z)$ is the gain per unit length, in an incremental length dz the light power increases by

$$dP = g(z)\, P_s(z)dz \tag{11.11}$$

Substituting Eq. (11.10) into Eq. (11.11) and rearranging terms gives

$$g_0(z)dz = \left(\frac{1}{P_s(z)} + \frac{1}{P_{\text{amp,sat}}} \right)dP \tag{11.12a}$$
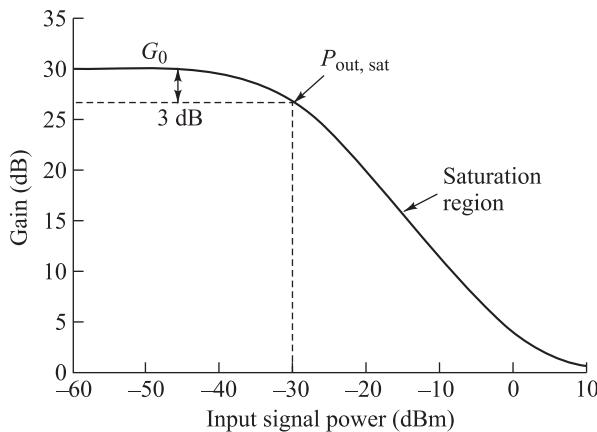
Fig. 11.3   *Typical dependence of the single-pass gain on optical input power for a small-signal gain of $G_0 = 30$ dB (a gain of 1000)*

Integrating this equation from $z = 0$ to $z = L$ yields

$$\int_0^L g_0 \, dz = \int_{P_{s,\text{in}}}^{P_{s,\text{out}}} \left(\frac{1}{P_s(z)} + \frac{1}{P_{\text{amp,sat}}} \right) dp \tag{11.12b}$$

Defining the single-pass gain in the absence of light to be $G_0 = \exp(g_0 L)$, and using Eq. (11.8), we then have

$$G = 1 + \frac{P_{\text{amp,sat}}}{P_{s,\text{in}}} \ln\left(\frac{G_0}{G} \right) = G_0 \exp\left(-\frac{G-1}{G} \frac{P_{s,\text{out}}}{P_{\text{amp,sat}}} \right) \tag{11.13}$$

Figure 11.3 illustrates the dependence of the gain on the input power. Here, the zero-signal gain (or small-signal gain) is $G_0 = 30$ dB, which is a gain factor of 1000. The curve shows that as the input signal power is increased, the gain first stays near the small-signal level and then starts to decrease. After decreasing linearly in the gain saturation region, it finally approaches an asymptotic value of 0 dB (a unity gain) for high input powers. Also shown is the *output saturation power*, which is the point at which the gain is reduced by 3 dB (see Prob. 11.4).

The wavelength at which the SOA has a maximum gain can be tailored to occur anywhere between about 1200 and 1700 nm by changing the composition of the active InGaAsP material. As an example, Fig. 11.4 shows a typical gain versus wavelength characteristic for a device with a peak gain of 25 dB at 1530 nm. The wavelength span over which the gain decreases by less than 3 dB with respect to the maximum gain is known as the *gain bandwidth* or the *3-dB optical bandwidth*. In the example shown in Fig. 11.4 the 3-dB optical bandwidth is 85 nm. Values of up to 100 nm can be achieved.

11.2.3   SOA Bandwidth

A general expression for the cavity gain G_c as a function of signal frequency f is given by[8]

$$G_c(f) = \frac{(1 - R_1)(1 - R_2)G}{(1 - \sqrt{R_1 R_2}\, G)^2 + 4\sqrt{R_1 R_2}\, G \sin^2 \varphi} \tag{11.14}$$
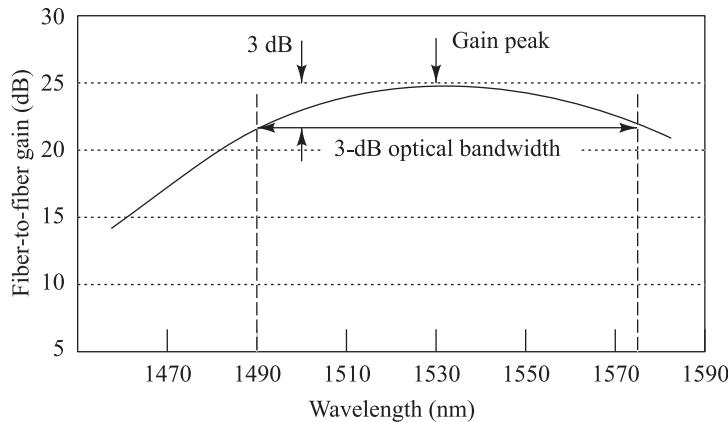
Fig. 11.4 *Typical gain-versus-wavelength characteristic for an SOA with a peak gain of 25 dB at 1530 nm. The definition of the 3-dB optical bandwidth is illustrated*

where G is the single-pass gain, R_1 and R_2 are the input and output facet reflectivities, respectively, and φ is the single-pass phase shift through the amplifier. The phase can be expressed as $\varphi = \pi(f - f_0)/\Delta f_{FSR}$, where f_0 is the cavity resonance frequency and Δf_{FSR} is the free spectral range of the SOA (see Sec. 10.5.1).

From Eq. (11.14) the 3-dB spectral bandwidth B_{SOA} of an SOA can be expressed by[8]

$$B_{SOA} = 2(f - f_0) = \frac{2\Delta f_{FSR}}{\pi} \sin^{-1}\left[\frac{1 - \sqrt{R_1 R_2}\, G}{2(\sqrt{R_1 R_2}\, G)^{1/2}}\right] = \frac{c}{\pi n L} \sin^{-1}\left[\frac{1 - \sqrt{R_1 R_2}\, G}{2(\sqrt{R_1 R_2}\, G)^{1/2}}\right] \quad (11.15)$$

Here L is the length of the amplifier and n is its refractive index.

11.3 Erbium-Doped Fiber Amplifiers

The active medium in an optical fiber amplifier consists of a nominally 10- to 30-m length of optical fiber that has been lightly doped (e.g., 1000 parts per million weight) with a rare-earth element, such as erbium (Er), ytterbium(Yb), thulium (Tm), or praseodymium (Pr). The host fiber material can be standard silica, a fluoride-based glass, or a tellurite glass.

The operating regions of these devices depend on the host material and the doping elements. A popular material for long-haul telecommunication applications is a silica fiber doped with erbium, which is known as an *erbium-doped fiber amplifier* or EDFA.[4,12–17] In some cases Yb is added to increase the pumping efficiency and the amplifier gain.[18,19] The operation of a standard EDFA normally is limited to the 1530 to l565 nm region. Actually the fact that an EDFA operates in this spectral band is the origin of the term *C-band* or *conventional band* (see Chapter 1). However, various techniques have been proposed and used to extend the operation to the S-band and the L-band. Section 11.8 describes some of these techniques for creating wideband optical amplifiers.

11.3.1 Amplification Mechanism

Whereas semiconductor optical amplifiers use external current injection to excite electrons to higher energy levels, optical fiber amplifiers use *optical pumping*. In this process, one uses photons to directly
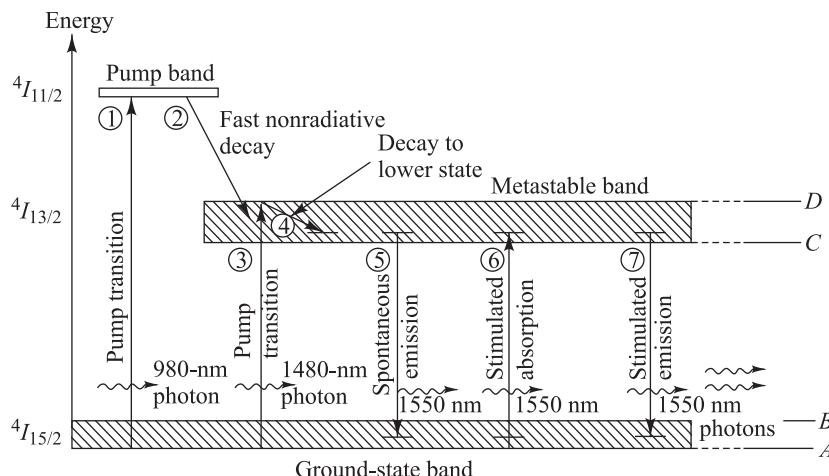
Fig. 11.5   *Simplified energy-level diagrams and various transition processes of Er³⁺ ions in silica*

raise electrons into excited states. The optical pumping process requires three or more energy levels. The top energy level to which the electron is elevated initially must lie energetically above the desired final emission level. After reaching its initial excited state, the electron must quickly release some of its energy and drop to a slightly lower energy level. A signal photon can then trigger the excited electron sitting in this new lower level into stimulated emission, whereby the electron releases its remaining energy in the form of a new photon with a wavelength identical to that of the signal photon. Since the pump photon must have a higher energy than the signal photon, the pump wavelength is shorter than the signal wavelength.

To get a phenomenological understanding of how an EDFA works, we need to look at the energy-level structure of erbium.[4,12–17] The erbium atoms in silica are Er^{3+} ions, which are erbium atoms that have lost three of their outer electrons. In describing the transitions of the outer electrons in these ions to higher energy states, it is common to refer to the process as "raising the ions to higher energy levels." Figure 11.5 shows a simplified energy-level diagram and various energy-level transition processes of these Er^{3+} ions in silica glass. The two principal levels for telecommunication applications are a *metastable level* (the so-called $^4I_{13/2}$ level) and the $^4I_{11/2}$ *pump level*. The term "metastable" means that the lifetimes for transitions from this state to the ground state are very long compared with the lifetimes of the states that led to this level. (Note that, by convention, the possible states of a multielectron atom are referred to by the symbol $^{2S+1}L_J$, where $2S+1$ is the spin multiplicity, L is the orbital angular momentum, and J is the total angular momentum.) The metastable, the pump, and the ground-state levels are actually bands of closely spaced energy levels that form a manifold due to the effect known as *Stark splitting*. Furthermore, each Stark level is broadened by thermal effects into an almost continuous band.

To understand the various energy transitions and photon emission ranges, consider the following conditions:

- The pump band shown in the top left of Fig. 11.5 exists at a 1.27-eV separation from the bottom of the $^4I_{15/2}$ ground state. This energy corresponds to a 980-nm wavelength.
- The top of the $^4I_{13/2}$ metastable band (level D in Fig. 11.5) is separated from the bottom of the $^4I_{15/2}$ ground state band (level A in Fig. 11.5) by 0.841 eV. This energy corresponds to a 1480-nm wavelength.

- The bottom of the $^4I_{13/2}$ metastable band (level C in Fig. 11.5) is separated from the bottom of the $^4I_{15/2}$ ground state band (level A in Fig. 11.5) by 0.814 eV. This energy corresponds to a 1530-nm wavelength.
- The bottom of the $^4I_{13/2}$ metastable band (level C in Fig. 11.5) is separated from the top of the $^4I_{15/2}$ ground state band (level B in Fig. 11.5) by about 0.775 eV. This energy corresponds to a 1600-nm wavelength.

This means that possible pump wavelengths are 980 and 1480 nm. The photons emitted during transitions of electrons between possible energy levels in the metastable and ground-state bands can range from 1530 to 1600 nm.

In normal operation, a pump laser emitting 980-nm photons is used to excite ions from the ground state to the pump level, as shown by transition process 1 in Fig. 11.5. These excited ions decay (relax) very quickly (in about 1 μs) from the pump band to the metastable band, shown as transition process 2. During this decay, the excess energy is released as phonons or, equivalently, mechanical vibrations in the fiber. Within the metastable band, the electrons of the excited ions tend to populate the lower end of the band. Here, they are characterized by a very long fluorescence time of about 10 ms.

Another possible pump wavelength is 1480 nm. The energy of these pump photons is very similar to the signal-photon energy, but slightly higher. The absorption of a 1480-nm pump photon excites an electron from the ground state directly to the lightly populated top of the metastable level, as indicated by transition process 3 in Fig. 11.5. These electrons then tend to move down to the more populated lower end of the metastable level (transition 4). Some of the ions sitting at the metastable level can decay back to the ground state in the absence of an externally stimulating photon flux, as shown by transition process 5. This decay phenomenon is known as *spontaneous emission* and adds to the amplifier noise.

Two more types of transitions occur when a flux of signal photons that have energies corresponding to the bandgap energy between the ground state and the metastable level passes through the device. First, a small portion of the external photons will be absorbed by ions in the ground state, which raises these ions to the metastable level, as shown by transition process 6. Second, in the stimulated emission process (transition process 7) a signal photon triggers an excited ion to drop to the ground state, thereby emitting a new photon of the same energy, wavevector, and polarization as the incoming signal photon. The widths of the metastable and ground-state levels allow high levels of stimulated emissions to occur in the 1530-to-1560-nm range.

The absorption and emission responses of an EDFA depend on the composition of the host glass and on the types of dopants, such as Ge and Al, in the glass. Figure 11.6 gives an example for a Ge-doped silica glass that has Al codopants added. The Al atoms help to absorb the Er ions in the glass, and they broaden the amplifier gain spectrum. Note that there is a peak gain response at 1532 nm and a 20- to 30-nm region at longer wavelengths in which the gain is relatively flat. Beyond 1560 nm the gain decreases steadily until it reaches 0 dB (unity gain) at around 1616 nm.

11.3.2   EDFA Architecture

An optical fiber amplifier consists of a doped fiber, one or more pump lasers, a passive wavelength coupler, optical isolators, and tap couplers, as shown in Fig. 11.7. The dichroic (two-wavelength) coupler handles either 980/1550-nm or 1480/1550-nm wavelength combinations to couple both the pump and signal optical powers efficiently into the fiber amplifier. The tap couplers are wavelength-insensitive with typical splitting ratios ranging from 99:1 to 95:5. They are generally used on both sides of the amplifier to compare the incoming signal with the amplified output. The optical isolators prevent the amplified signal from reflecting back into the device, where it could increase the amplifier noise and decrease the amplifier efficiency.
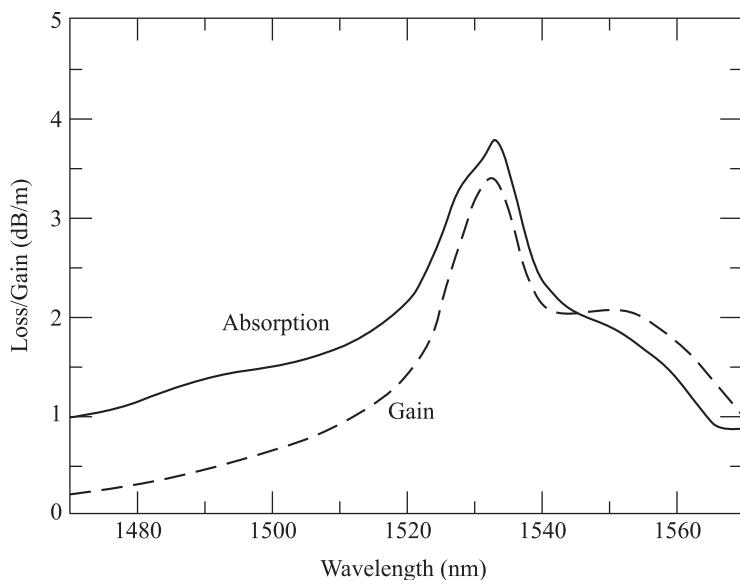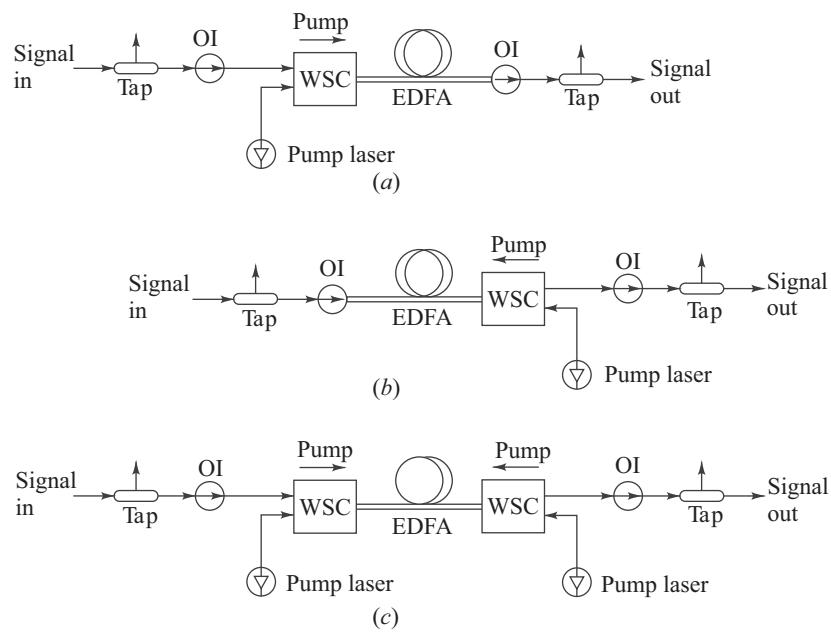
Fig. 11.6 *Absorption and gain spectra for erbium ions in a silica fiber doped with Ge and Al. (Modified with permission from Giles and Desurvire,[20]© 1991, IEEE.)*

OI: Optical isolator
WSC: Wavelength-selective coupler

Fig. 11.7 *Three possible configurations of an EDFA: (a) codirectional pumping, (b) counterdirectional pumping, (c) dual pumping*

The pump light is usually injected from the same direction as the signal flow. This is known as *codirectional pumping*. It is also possible to inject the pump power in the opposite direction to the signal flow, which is known as *counterdirectional pumping*. As shown in Fig. 11.7, one can employ either a single pump source or use *dual-pump schemes*, with the resultant gains typically being +17 dB and +35 dB, respectively. Counterdirectional pumping allows higher gains, but codirectional pumping gives better noise performance. In addition, pumping at 980 nm is preferred, since it produces less noise and achieves larger population inversions than pumping at 1480 nm.

11.3.3 EDFA Power-Conversion Efficiency and Gain

As is the case with any amplifier, as the magnitude of the output signal from an EDFA increases, the amplifier gain eventually starts to saturate. The reduction of gain in an EDFA occurs when the population inversion is reduced significantly by a large signal, thereby yielding the typical gain-versus-power performance curve shown in Fig. 11.3.

The input and output signal powers of an EDFA can be expressed in terms of the principle of energy conservation:[4,12]

$$P_{s,\text{out}} \leq P_{s,\text{in}} + \frac{\lambda_p}{\lambda_s} P_{p,\text{in}} \tag{11.16}$$

where $P_{p,\text{in}}$ is the input pump power, and λ_p and λ_s are the pump and signal wavelengths, respectively. The fundamental physical principle here is that the amount of signal energy that can be extracted from an EDFA cannot exceed the pump energy that is stored in the device. The inequality in Eq. (11.16) reflects the possibility of effects such as pump photons being lost due to various causes (such as interactions with impurities) or pump energy lost due to spontaneous emission.

From Eq. (11.16), we see that the maximum output signal power depends on the ratio λ_p/λ_s. For the pumping scheme to work, we need to have $\lambda_p < \lambda_s$, and, to have an appropriate gain, it is necessary that $P_{s,\text{in}} \leq P_{p,\text{in}}$. Thus, the *power conversion efficiency* (PCE), defined as

$$\text{PCE} = \frac{P_{s,\text{out}} - P_{s,\text{in}}}{P_{p,\text{in}}} \approx \frac{P_{s,\text{out}}}{P_{p,\text{in}}} \leq \frac{\lambda_p}{\lambda_s} \leq 1 \tag{11.17}$$

is less than unity. The maximum theoretical value of the PCE is λ_p/λ_s. For absolute reference purposes, it is helpful to use the *quantum conversion efficiency* (QCE), which is wavelength-independent and is defined by[4,12]

$$\text{QCE} = \frac{\lambda_s}{\lambda_p} \text{PCE} \tag{11.18}$$

The maximum value of QCE is unity, in which case all the pump photons are converted to signal photons.

We can also rewrite Eq. (11.16) in terms of the amplifier gain G. Assuming there is no spontaneous emission, then

$$G = \frac{P_{s,\text{out}}}{P_{s,\text{in}}} \leq 1 + \frac{\lambda_p P_{p,\text{in}}}{\lambda_s P_{s,\text{in}}} \tag{11.19}$$

This shows an important relationship between signal input power and gain. When the input signal power is very large so that $P_{s,\text{in}} \gg (\lambda_s/\lambda_s)P_{p,\text{in}}$, then the maximum amplifier gain is unity. This means that the device

is transparent to the signal. From Eq. (11.19), we also see that in order to achieve a specific maximum gain G, the input signal power cannot exceed a value given by

$$P_{s,\text{in}} \leq \frac{(\lambda_p/\lambda_s)P_{p,\text{in}}}{G-1} \tag{11.20}$$

Example 11.3 Consider an EDFA being pumped at 980 nm with a 30-mW pump power. If the gain at 1550 nm is 20 dB, what are the maximum input and output powers?

Solution: From Eq. (11.20), the maximum input power is

$$P_{s,\text{in}} \leq \frac{(980/1550)(30\,\text{mW})}{100-1} = 190\,\mu\text{W}$$

From Eq. (11.16), the maximum output power is

$$P_{s,\text{out}}(\text{max}) = P_{s,\text{in}}(\text{max}) + \frac{\lambda_p}{\lambda_s}P_{p,\text{in}}$$

$$= 190\,\mu\text{W} + 0.63(30\,\text{mW})$$

$$= 19.1\,\text{mW} = 12.8\,\text{dBm}$$

In addition to pump power, the gain also depends on the fiber length. The maximum gain in a three-level laser medium of length L, such as an EDFA, is given by

$$G_{\text{max}} = \exp(\rho\sigma_e L) \tag{11.21}$$

where σ_e is the signal-emission cross section and ρ is the rare-earth element concentration. When determining the maximum gain, Eqs. (11.19) and (11.21) must be considered together. Consequently, the maximum possible EDFA gain is given by the lowest of the two gain expressions:

$$G \leq \min\left\{\exp(\rho\sigma_e L), 1 + \frac{\lambda_p}{\lambda_s}\frac{P_{p,\text{in}}}{P_{s,\text{in}}}\right\} \tag{11.22}$$

Since $G = P_{s,\text{out}}/P_{s,\text{in}} = \exp(\rho\sigma_e L)$, it follows similarly that the maximum possible EDFA output power is given by the minimum of the two expressions:

$$P_{s,\text{out}} \leq \min\left\{P_{s,\text{in}}\exp(\rho\sigma_e L), P_{s,\text{in}} + \frac{\lambda_p}{\lambda_s}P_{p,\text{in}}\right\} \tag{11.23}$$

Figure 11.8 illustrates the onset of gain saturation for various doped-fiber lengths as the pumping power increases.[20] The gain was measured for a 100-nW input signal at 1550 nm. As the fiber length increases for low pumping powers, the gain starts to decrease after a certain length because the pump does not have enough energy to create a complete population inversion in the downstream portion of the amplifier. In this case, the unpumped region of the fiber absorbs the signal, thus resulting in signal loss rather than gain in that section.

Since an electron state in the metastable level in an EDFA has a relatively long lifetime, it is possible to obtain very high saturated output powers. The *saturated output power* (the power at which gain saturation occurs) is defined as the 3-dB compression point of the small-signal gain.[21] For large signal operation, the saturated gain increases linearly with pump power, as can be inferred from Fig. 11.9. This figure shows that as the input power increases for a given pump level, the amplifier gain remains constant until saturation occurs.
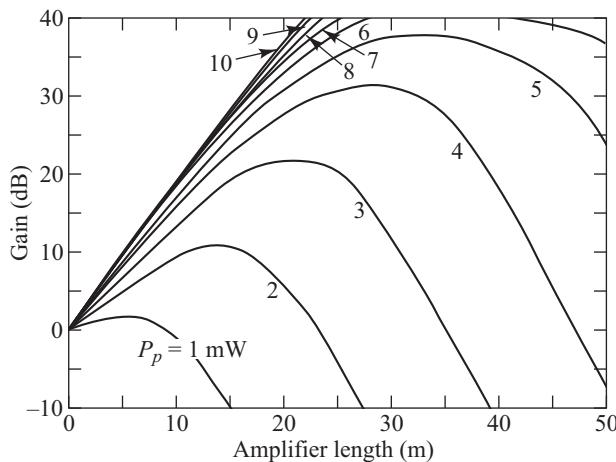
Fig. 11.8 *Calculation of the dependence of EDFA gain on fiber length and pump power for a 1480-nm pump and a 1550-nm signal. (Reproduced with permission from Giles and Desurvire,[20] © 1991, IEEE.)*

Example 11.4 From Fig. 11.8 we see that for 1480-nm pumping, a 20-dB gain can be achieved for a 100-nW signal at 1550 nm with a pump power of 3 mW for an amplifier length of about 22 m. Since the 3-mW pump does not have enough energy to create population inversion for longer amplifier fibers, the gain decreases to about 12 dB for an amplifier length of 30 m.
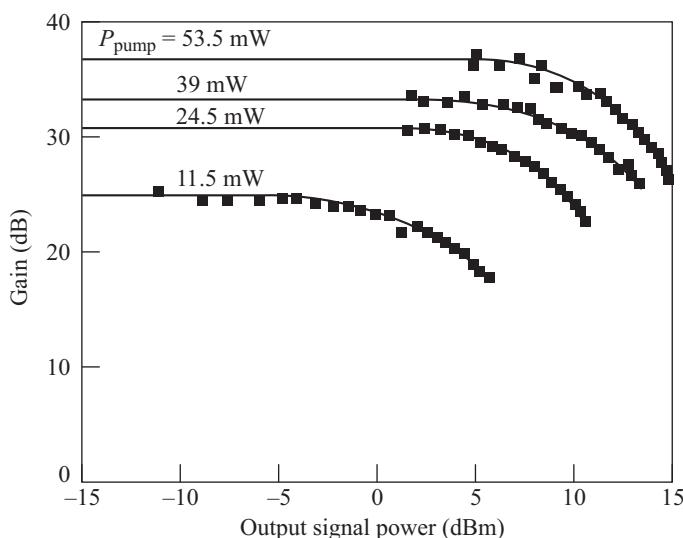
Fig. 11.9 *Gain behavior of an EDFA as a function of output signal power for various pump levels. (Reproduced with permission from Li,[21] © 1993, IEEE.)*
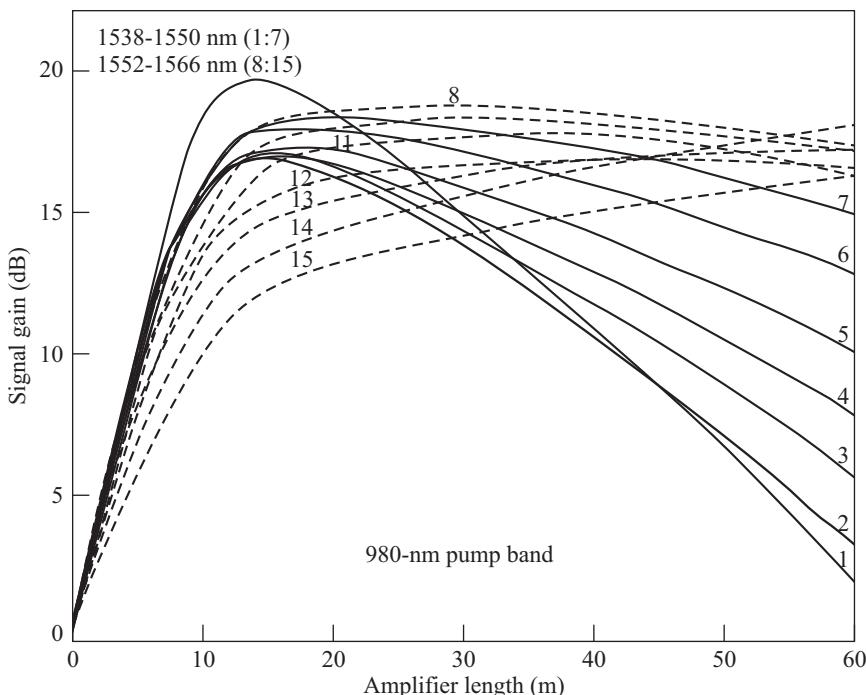
Fig. 11.10 *EDFA gain versus amplifier length for 15 WDM signals that are codirectionally pumped with 50 mW of 980-nm pump power. (Reproduced with permission from Ali, Elrefaie, Wagner, and Ahmed,[22] © 1996, IEEE.)*

Note that for simplicity of calculation to demonstrate the amplifier gain versus length relationships, the curves shown in Fig. 11.8 were done for the small-signal case and did not take into account internal amplifier noise that can induce amplifier saturation. Figures 11.10 and 11.11 show more detailed power-gain versus amplifier-length calculations for 980- and 1480-nm pumping, respectively.[22] In this case the EDFA amplified fifteen signal wavelengths spaced 2 nm apart in the 1538- to 1566-nm range. The calculations assume a forward pumping scheme with 50 mW of pump power for each EDFA. The total optical signal input power for the EDFA was −3 dBm (0.5 mW) or −15 dBm/channel (0.03 mW/wavelength). Several points can be noted from the curves:

- The amplifier length that yields a maximum gain becomes longer with increasing signal wavelength, because photons at longer wavelengths have less energy and thus need less power to have the same gain as photons at shorter wavelengths.
- If a specific amplifier length is chosen, for example, 30 m, then the EDFA will amplify each wavelength differently, again because photon energy is wavelength dependent. This leads to gain skew among different wavelengths, as described in Sec. 11.6.4.
- The 980-nm pumping yields a complete population inversion (maximum gain) at shorter amplifier lengths than 1480-nm pumping. This leads to a lower amplifier noise figure when using 980-nm pumping.
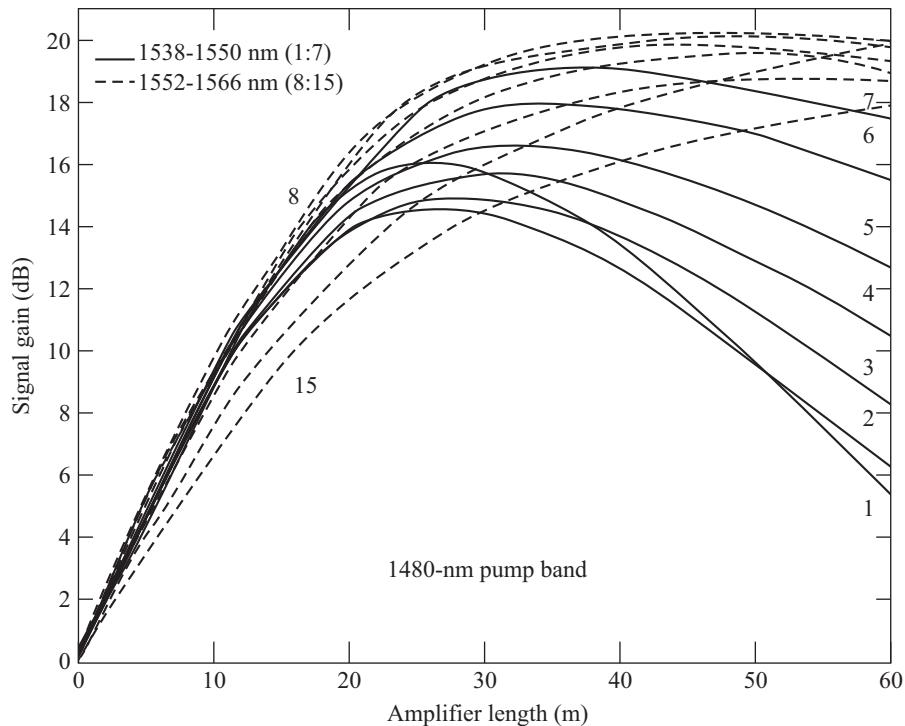
Fig. 11.11 *EDFA gain versus amplifier length for 15 WDM signals that are codirectionally pumped with 50 mW of 1480-nm pump power. (Reproduced with permission from Ali, Elrefaie, Wagner, and Ahmed,[22] © 1996, IEEE.)*

11.4 Amplifier Noise

The dominant noise generated in an optical amplifier is called *amplified spontaneous emission* (ASE) noise. The origin of this noise is the spontaneous recombination of electrons and holes in the amplifier medium (transition 5 in Fig. 11.5). This recombination occurs over a wide range of electron-hole energy differences and thus gives rise to a broad spectral background of noise photons that get amplified along with the optical signal as they travel through the EDFA. This is shown in Fig. 11.12 for an EDFA amplifying a signal at 1540 nm. The spontaneous noise can be modeled as a stream of random infinitely short pulses that are distributed all along the amplifying medium. Such a random process is characterized by a noise power spectrum that is flat with frequency. The power spectral density of the ASE noise is[17]

$$S_{ASE}(f) = h v n_{sp}[G(f) - 1] = P_{ASE}/B_o \tag{11.24}$$

where P_{ASE} is the ASE noise power in one polarization state in an optical bandwidth B_o and n_{sp} is the *spontaneous-emission* or *population-inversion factor* defined as

$$n_{sp} = \frac{n_2}{n_2 - n_1} \tag{11.25}$$
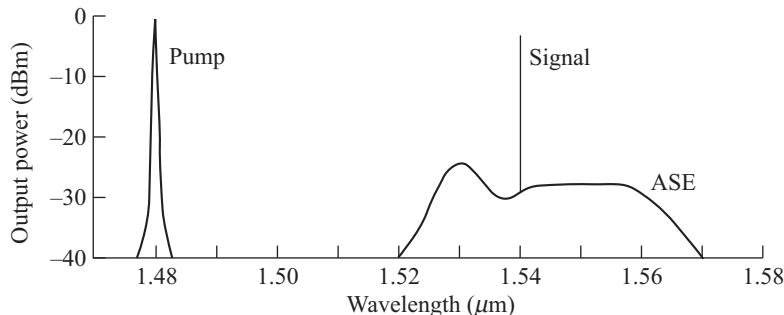
Fig. 11.12 *Representative 1480-nm pump spectrum and a typical output signal at 1540 nm with the associated amplified-spontaneous-emission (ASE) noise*

where n_1 and n_2 are the fractional densities or populations of electrons in a lower state 1 and an upper state 2, respectively. Thus, n_{sp} denotes how complete the population inversion is between two energy levels. From Eq. (11.25) $n_{sp} \geq 1$, with equality holding for an ideal amplifier when the population inversion is complete. Typical values range from 1.4 to 4, depending on the wavelength and the pumping rate.

An important point is that the expression for P_{ASE} is for an individual spatial mode and an individual polarization state. For single-mode fibers the center term in Eq. (11.24) must be multiplied by a factor 2 to get the total ASE power, since a single-mode fiber has one spatial mode and two polarization modes. If an EDFA is made from a multimode fiber, P_{ASE} becomes much larger since such a fiber has many spatial modes.

The ASE noise level depends on whether codirectional or counterdirectional pumping is used. Figure 11.13 shows experimental and calculated data on ASE noise versus pump power for different EDFA lengths.[23]

Since ASE originates ahead of the photodiode, it gives rise to three different noise components in an optical receiver in addition to the normal thermal noise of the photodetector. This occurs because the photocurrent consists of a number of beat signals between the signal and the optical noise fields, in addition to the squares of the signal field and the spontaneous-emission field. If the total optical field is the sum of the signal field E_s and the ASE noise field E_n, then the total photodetector current i_{tot} is proportional to the square of the electric field of the composite optical signal: $i_{tot} \propto (E_s + E_n)^2 = E_s^2 + E_n^2 + 2E_s \cdot E_n$. Here the first two terms arise purely from the signal and noise, respectively. The third term is a mixing component (a *beat signal*) between the signal and noise, which can fall within the bandwidth of the receiver and degrade the signal-to-noise ratio. First, taking the ASE photons into account, the optical power incident on the photodetector becomes

$$P_{in} = GP_{s,in} + P_{ASE} = GP_{s,in} + S_{ASE} B_o \qquad (11.26)$$

Note that B_o can be reduced significantly if an optical filter precedes the photodetector. Substituting P_{in} into Eq. (6.6) and inserting the resulting expression for the photocurrent I_p into Eq. (6.13) to find the shot noise then yields the total mean-square signal-plus-ASE shot-noise current

$$\left\langle i_{shot}^2 \right\rangle = \sigma_{shot}^2 = \sigma_{shot\text{-}s}^2 + \sigma_{shot\text{-}ASE}^2 = 2q\mathcal{R}GP_{s,in}B_e + 2q\mathcal{R}S_{ASE}B_oB_e \qquad (11.27)$$

where B_e is the front-end receiver electrical bandwidth.
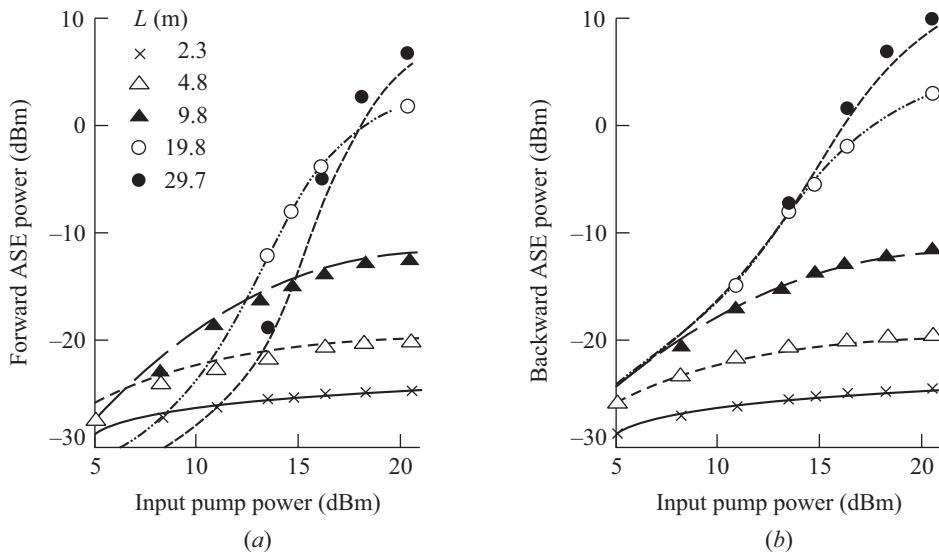
Fig. 11.13 *Experimental and theoretical ASE noise powers versus input pump power for various EDFA lengths arising from (a) codirectional (forward) pumping and (b) counterdirectional (backward) pumping. (Reproduced with permission from Pedersen et al.,[23] © 1990, IEEE.)*

The other two noises arise from the mixing of the different optical frequencies contained in the light signal and the ASE, which generates two sets of beat frequencies. Since the signal and the ASE have different optical frequencies, the beat noise of the signal with the ASE noise that is in the same polarization state as the signal is

$$\sigma^2_{s\text{-ASE}} = 4(\mathcal{R}GP_{s,\text{in}})(\mathcal{R}S_{\text{ASE}}B_e) \tag{11.28}$$

In addition, since the ASE spans a wide optical frequency range, it can beat against itself giving rise to the noise current

$$\sigma^2_{\text{ASE-ASE}} = \mathcal{R}^2 S^2_{\text{ASE}}(2B_o - B_e)B_e \tag{11.29}$$

The total mean-square receiver noise current then becomes

$$\left\langle i^2_{\text{total}} \right\rangle = \sigma^2_{\text{total}} = \sigma^2_T + \sigma^2_{\text{shot-}s} + \sigma^2_{\text{shot-ASE}} + \sigma^2_{s\text{-ASE}} + \sigma^2_{\text{ASE-ASE}} \tag{11.30}$$

where the thermal noise variance σ^2_T is given by Eq. (6.17).

The last four terms in Eq. (11.30) tend to be of similar magnitudes when the optical bandwidth B_o is taken to be the optical bandwidth of the spontaneous emission noise, which covers a 30-nm spectrum (see Prob. 11.7). However, one generally uses a narrow optical filter at the receiver, so that B_o is on the order of 125 GHz (a 1-nm spectral width at 1550 nm) or less. In that case, we can simplify Eq. (11.30) by examining the magnitudes of the various noise components. First, the thermal noise can generally

be neglected when the amplifier gain is large enough. Furthermore, since the amplified signal power $GP_{s,\text{in}}$ is much larger than the ASE noise power $S_{\text{ASE}} B_o$, the ASE–ASE beat noise given by Eq. (11.29) is significantly smaller than the signal–ASE beat noise. This also means that the second term in Eq. (11.27) is small compared to the first term, so that

$$\sigma_{\text{shot}}^2 \approx 2q\mathcal{R}GP_{s,\text{in}}B_e \tag{11.31}$$

Using these results together with the expression for S_{ASE} from Eq. (11.24) yields the following approximate signal-to-noise ratio (S/N) at the photodetector output:

$$\left(\frac{S}{N}\right)_{\text{out}} = \frac{\sigma_{\text{ph}}^2}{\sigma_{\text{total}}^2} = \frac{\mathcal{R}^2 G^2 P_{s,\text{in}}^2}{\sigma_{\text{total}}^2} \approx \frac{\mathcal{R}P_{s,\text{in}}}{2qB_e} \frac{G}{1 + 2\eta n_{\text{sp}}(G-1)} \tag{11.32}$$

where η is the quantum efficiency of the photodetector and, from Eq. (6.11), the mean-square input photocurrent is

$$\left\langle i_{\text{ph}}^2 \right\rangle = \sigma_{\text{ph}}^2 \approx \mathcal{R}^2 G^2 P_{s,\text{in}}^2 \tag{11.33}$$

Example 11.5 Figure 11.14 shows measured values of the noise figure for an EDFA under gain saturation for both codirectional and counterdirectional pumping.[24] The pump wavelength was 1480 nm and the signal wavelength was 1558 nm with an input power to the amplifier of −60 dBm. Under small-signal conditions, the codirectional pumping noise was about 5.5 dB, which included a 1.5-dB input coupling loss. The noise figure of the optical amplifier itself was thus 4 dB, compared with the theoretical minimum of 3 dB with complete population inversion. The noise figure in the counterdirectional pumping case was about 1 dB higher.
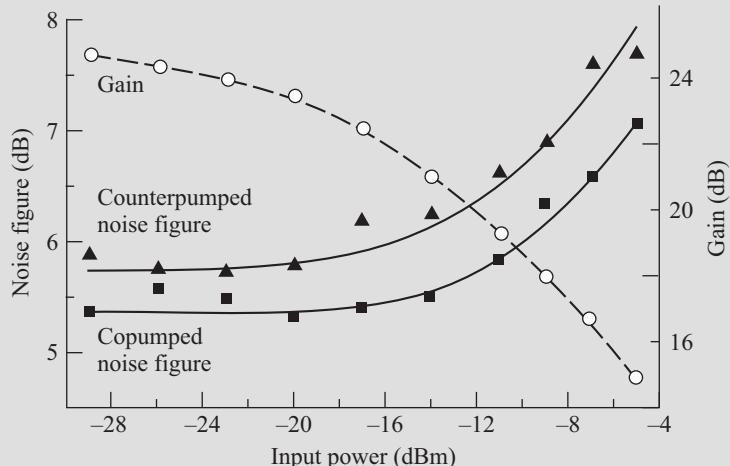
Fig. 11.14 *Measured EDFA noise figure under gain saturation for codirectional pumping and counterdirectional pumping at 1480 nm. The gain was similar for both pump directions. (Reproduced with permission from Walker et al.,[24] © 1991, IEEE.)*

Note that the term

$$\left(\frac{S}{N}\right)_{\text{in}} = \frac{\mathscr{R}P_{s,\text{in}}}{2qB_e} \tag{11.34}$$

in Eq. (11.32) is the signal-to-noise ratio of an ideal photodetector at the input to the optical amplifier. From Eq. (11.32) we can then find the noise figure of the optical amplifier, which is a measure of the S/N degradation experienced by a signal after passing through the amplifier. Using the standard definition of *noise figure* as the ratio between the S/N at the input and the S/N at the amplifier output, we have for an ideal photodetector with $\eta = 1$

$$\text{Noise figure} = F_{\text{EDFA}} = \frac{(S/N)_{\text{in}}}{(S/N)_{\text{out}}} = \frac{1 + 2n_{\text{sp}}(G-1)}{G} \tag{11.35}$$

When G is large, this becomes $2n_{\text{sp}}$. A perfect amplifier would have $n_{\text{sp}} = 1$, yielding a noise figure of 2 (or 3 dB), assuming $\eta = 1$. That is, using an ideal receiver with a perfect amplifier would degrade the S/N by a factor of 2. In a real EDFA, for example, n_{sp} is around 2, so the input S/N gets reduced by a factor of about 4. This results in a noise figure of 4 to 5 dB for a practical EDFA.[25]

11.5 Optical SNR

When analyzing a transmission link that has a series of optical amplifiers in it, an important point is that the light signal entering the optical receiver may contain a significant level of ASE noise that has been added by the cascade of optical amplifiers. In this case one has to evaluate the *optical signal-to-noise ratio* (OSNR).[4] This parameter is defined as the ratio of the average EDFA optical signal output power P_{ave} to the unpolarized ASE optical noise power P_{ASE}. The OSNR is given by

$$\text{OSNR} = \frac{P_{\text{ave}}}{P_{\text{ASE}}} \tag{11.36a}$$

or, in decibels,

$$\text{OSNR(dB)} = 10 \log \frac{P_{\text{ave}}}{P_{\text{ASE}}} \tag{11.36b}$$

In practice, the OSNR can be measured with an *optical spectrum analyzer* (OSA), which is described in Chapter 14. The OSNR does not depend on factors such as the data format, pulse shape, or optical filter bandwidth, but only on the average optical signal power P_{ave} and the average optical noise power. OSNR is a metric that can be used in the design and installation of networks, as well as to check the health and status of individual optical channels. Sometimes an optical filter is used to significantly reduce the total ASE noise seen by the receiver. Typically such a filter has an optical bandwidth that is large compared to the signal, so that it does not affect the signal, yet that is narrow compared to the bandwidth associated with the ASE background. The ASE noise filter does not change the OSNR. However, it reduces the total power in the ASE noise to avoid overloading the receiver front end.

In order to have a meaningful assessment of what the OSNR tells about system performance, a connection is needed between OSNR and bit-error rate (BER). A number of different relationships have been proposed for this in the literature. Furthermore, there are different approaches for interpreting the results of an OSA measurement, which can lead to a difference in results of several decibels. Using the expression for Q given by Eq. (7.13) and the definition of OSNR from Eq. (11.36b), the following relation between Q and OSNR can be derived[12]

$$Q = \frac{2\sqrt{2}\,\text{OSNR}}{1 + \sqrt{1 + 4\,\text{OSNR}}} \tag{11.37}$$

Solving Eq. (11.37) for OSNR yields

$$\text{OSNR} = \frac{1}{2}Q\left(Q + \sqrt{2}\right) \tag{11.38}$$

Example 11.6 In Chapter 7 it is shown that to achieve a BER = 10^{-9} the factor Q must be 6. What is the OSNR for this BER?

Therefore if an OSA measures an OSNR ≤ 13.5 dB, then the corresponding error rates are equal to or higher than BER = 10^{-9}.

Solution: Using Eq. (11.38) then yields

OSNR (BER = 10^{-9}) = 0.5(6)(6 + $\sqrt{2}$) = 22.24 ≈ 13.5 dB

Let us look at the case for a single EDFA. Converting the ASE noise power given by Eq. (11.24) into a decibel format, for unpolarized ASE noise we have for large G

$$10 \log P_{\text{ASE}} = 10 \log [(h\nu)(B_o)] + 10 \log 2n_{\text{sp}} + 10 \log G \tag{11.39}$$

Here, $h\nu$ is the photon energy, and B_o is the optical frequency range in which the OSNR is measured, which typically is 12.5 GHz (a 0.1-nm spectral width at 1550 nm). At 1550 nm, we have that $(h\nu)(B_o) = 1.58 \times 10^{-6}$ mW, so that $10 \log (h\nu)(B_o) = -58$ dBm. Assuming that $G \gg 1$ and taking $F_{\text{EDFA}}(\text{dB}) = 10 \log 2n_{\text{sp}}$ from Eq. (11.35) as the amplifier noise figure, then in decibels

$$P_{\text{ASE}}(\text{dBm}) = -58 \text{ dBm} + F_{\text{EDFA}}(\text{dB}) + G(\text{dB}) \tag{11.40}$$

Using Eq. (11.36) and taking the EDFA output power to be G times the optical input power, $P_{\text{out}} = GP_{\text{in}}$, we then have the requirement that in order to have an acceptable BER, the OSNR must be at least

$$\text{OSNR}(\text{dB}) = P_{\text{in}}(\text{dBm}) + 58 \text{ dBm} - F_{\text{EDFA}}(\text{dB}) \tag{11.41}$$

11.6 System Applications

In designing an optical fiber link that requires optical amplifiers, there are three possible locations where the amplifiers can be placed, as shown in Fig. 11.1. Although the physical amplification process is the same in all three configurations, the various uses require operation of the device over different input power ranges. This, in turn, implies use of different amplifier gains. The complete analysis of the signal-to-noise

ratios, taking into account factors such as detailed photon statistics and discrete amplifier configurations, are fairly involved.[4,12] Here, we will look at a simple conceptual analysis and present generic operational values for the three possible locations of EDFAs in an optical link.

11.6.1   Power Amplifiers

For the power amplifier, the input power is high because the device immediately follows an optical transmitter. High pump powers are normally required for this application.[26] The amplifier inputs are generally −8 dBm or greater, and the power amplifier gain must be greater than 5 dB in order for it to be more advantageous than using a preamplifier at the receiver.

Example 11.7   Consider an EDFA used as a power amplifier with a 10-dB gain. Assume the amplifier input is a 0-dBm level from a 1540-nm laser diode transmitter. If the pump wavelength is 980 nm, what is the pump power?

Solution: From Eq. (11.16), for a 10-dBm output at 1540 nm, the pump power must be at least

$$P_{p,\text{in}} \geq \frac{\lambda_s}{\lambda_p}(P_{s,\text{out}} - P_{s,\text{in}}) = \frac{1540}{980}(10 \text{ mW} - 1 \text{ mW})$$

$$= 14 \text{ mW}$$

11.6.2   In-Line Amplifiers

In a long transmission system, optical amplifiers are needed to periodically restore the power level after it has decreased due to attenuation in the fiber. Normally, the gain of each EDFA in this amplifier chain is chosen to compensate exactly for the signal loss incurred in the preceding fiber section of length L, that is, $G = \exp(+ \alpha L)$. The accumulated ASE noise is the dominant degradation factor in such a cascaded chain of amplifiers.

Example 11.8   Consider Fig. 11.15, which shows the values of the per-channel signal power, the per-channel ASE noise, and the SNR along a chain of seven optical amplifiers in a WDM link. The input signal level starts out at 6 dBm and decays due to fiber attenuation as it travels along the link. When its power level has dropped to −24 dBm, it gets boosted back to 6 dBm by an optical amplifier. For a given channel transmitted over the link, the SNR starts out at a high level and then decreases at each amplifier as the ASE noise accumulates through the length of the link. For example, following amplifier number 1, the SNR is 28 dB for a 6-dBm amplified signal level and a −22-dBm ASE noise level. After amplifier number 4, the SNR is 22 dB for a 6-dBm amplified signal level and a −16-dBm ASE noise level. The higher the gain in the amplifier, the faster the ASE noise builds up. However, although the SNR decreases quickly in the first few amplifications, the incremental effect of adding another EDFA diminishes rapidly with an increasing number of amplifiers. As a consequence, although the SNR drops by 3 dB when the number of EDFAs increases from one to two, it also drops by 3 dB when the number of amplifiers is increased from two to four, and by another 3 dB when the number of amplifiers is further increased to eight.

To compensate for the accumulated ASE noise, the signal power must increase at least linearly with the length of the link in order to keep a constant signal-to-noise ratio. If the total system length is $L_{\text{tot}} = NL$ and the system contains N optical amplifiers each having a gain $G = \exp(+ \alpha L)$, then, using Eq. (11.24), the path-averaged ASE power along a chain of optical amplifiers is[4,27]

$$\left\langle P_{\text{ASE}} \right\rangle_{\text{path}} = \frac{NP_{\text{ASE}}}{L} \int_0^L \exp(-\alpha z)dz = \alpha L_{\text{tot}} h\nu \, n_{\text{sp}} F_{\text{path}}(G)B_o \tag{11.42}$$
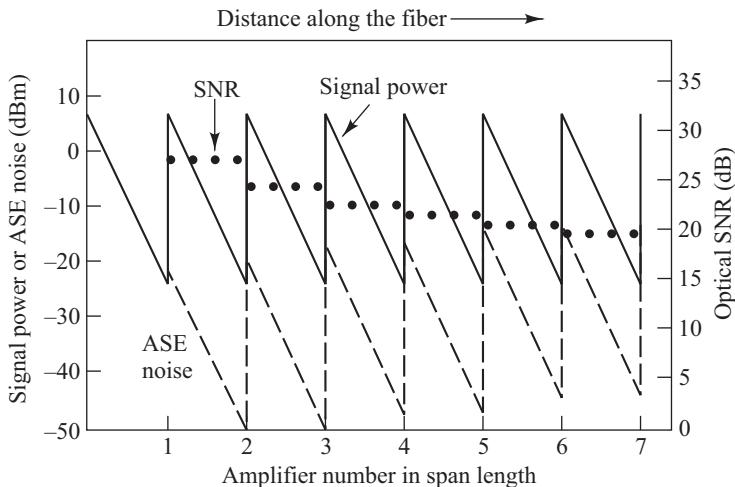
Fig. 11.15 *SNR degradation as a function of link distance over which the ASE noise increases with the number of amplifiers. The curves show the signal level (solid lines), the ASE noise level (dashed lines), and the SNR (dotted lines) for a single channel in a WDM link.*

where α is the fiber attenuation and the noise figure $F_{path}(G)$ is a power penalty defined as

$$F_{path}(G) = \frac{1}{G}\left(\frac{G-1}{\ln G}\right)^2 \tag{11.43}$$

Basically, $F_{path}(G)$ gives the factor by which the path-average signal energy must be increased (as G increases) in a chain of N cascaded optical amplifiers to maintain a fixed SNR. For long-haul networks these optical amplifiers can be placed uniformly along the transmission path to yield the best combination of overall gain and final SNR. The input power levels for these in-line amplifiers nominally ranges from −26 dBm (2.5 μW) to −9 dBm (125 μW) with gains ranging from 8 to 20 dB. For metro networks typically only a single optical amplifier is needed to compensate for the path loss between two successive nodes.[28,29]

Example 11.9   Consider an optical transmission path containing N cascaded optical amplifiers each having a 30-dB gain. If the fiber has a loss of 0.2 dB/km, then the span between optical amplifiers is 150 km if there are no other system impairments. As an example, for a 900-km link we would need five amplifiers. From Eq. (11.43), the noise penalty factor over the total path is (in decibels)

$$10 \log F_{path}(G) = 10 \log\left[\frac{1}{1000}\left(\frac{1000-1}{\ln 1000}\right)^2\right]$$

$$= 10 \log 20.9 = 13.2 \text{ dB}$$

If we reduce the gain to 20 dB, then the impairment-free transmission distance is 100 km for which we need eight amplifiers. In this case, the noise-penalty factor is

$$10 \log F_{path}(G) = 10 \log\left[\frac{1}{100}\left(\frac{100-1}{\ln 100}\right)^2\right]$$

$$= 10 \log 4.62 = 6.6 \text{ dB}$$

11.6.3 Preamplifiers

An optical amplifier can be used as a preamplifier to improve the sensitivity of direct-detection receivers that are limited by thermal noise.[30–32] First, assume the receiver noise is represented by the electrical power level N. Let S_{min} be the minimum value of the electrical signal power S that is required for the receiver to perform with a specific acceptable bit-error rate. The acceptable signal-to-noise ratio then is S_{min}/N. If we now use an optical preamplifier with gain G, the electrical received signal power is $G^2 S'$ and the signal-to-noise ratio is

$$\left(\frac{S}{N}\right)_{preamp} = \frac{G^2 S'}{N + N'}$$

(11.44)

where the noise term N' is the spontaneous emission from the optical preamplifier that gets converted by the photodiode in the receiver to an additional background noise. If S'_{min} is the new minimum detectable electrical signal level needed to maintain the same signal-to-noise ratio, then we need to have

$$\frac{G^2 S'_{min}}{N + N'} = \frac{S_{min}}{N}$$

(11.45)

For an optical preamplifier to enhance the received signal level, we must have $S'_{min} < S_{min}$, so that

$$\frac{S_{min}}{S'_{min}} = G^2 \frac{N}{N + N'} > 1$$

(11.46)

The ratio S_{min} to S'_{min} is the *improvement of minimum detectable signal* or *detector sensitivity.*

Example 11.10 Consider an EDFA used as an optical preamplifier. Assume that N is due to thermal noise and that the noise N' introduced by the preamplifier is dominated by signal–ASE beat noise. Under what conditions does Eq. (11.46) hold?

Solution: For sufficiently high gain G, Eq. (11.46) becomes

$$G^2 - 1 \approx G^2 > \frac{N'}{N} \approx \frac{\sigma^2_{s\text{-}ASE}}{\sigma^2_T}$$

Substituting Eqs. (6.17) and (11.28) into this expression, using Eq. (11.24) for S_{ASE}, and solving for $P_{s,in}$ yields

$$P_{s,in} < \frac{k_B T \, h\nu}{R n_{sp} \eta^2 q^2}$$

If $T = 300$ K, $R = 50\ \Omega$, $\lambda = 1550$ nm, $n_{sp} = 2$, and $\eta = 0.65$, then $P_{s,in} < 490\ \mu$W. This level is much higher than any expected received signal, so the condition in Eq. (11.46) is always satisfied. However, note that this only specifies the upper bound on $P_{s,in}$. It does not mean that by making G sufficiently high, the improvement in sensitivity can be made arbitrarily large, since there is a minimum received optical power level that is needed to achieve a specific BER (see Prob. 11.15).

11.6.4 Multichannel Operation

An advantage of both semiconductor optical amplifiers and EDFAs is their ability to amplify multiple optical channels, provided the bandwidth of the multichannel signal is smaller than the amplifier bandwidth.[33] For both SOAs and EDFAs, this bandwidth ranges from 1 to 5 THz. A disadvantage of SOAs is their sensitivity to interchannel crosstalk arising from carrier-density modulation due to beating of signals from adjacent optical channels.[34] For SOAs, this beating occurs whenever the channel spacing is less than 10 GHz.

This crosstalk does not occur in EDFAs, as long as the channel spacing is greater than 10 kHz, which holds in practice. Thus EDFAs are ideally suited for multichannel amplification. For multichannel operation in an EDFA, the signal power for N channels is given by

$$P_s = \sum_{i=1}^{N} P_{s,i} \tag{11.47}$$

where $P_{s,i}$ is the signal power in channel i; that is, at the optical carrier frequency v_i.

Another characteristic of an EDFA is that its gain is wavelength-dependent in its normal 1530-to-1560-nm operating window.[35] Figures 11.16 and 11.17 show this behavior for 980-nm and 1480-nm pumping at a 50 mW level, respectively.[22] In the 980-nm pumping illustration, the gains were measured for six wavelengths that were spaced 2 nm apart in a 1542-to-1552-nm spectral band after they passed through a cascade of 15 EDFAs. The signal strengths were −10 dBm (0.1 mW) for each wavelength channel at the input to the first amplifier. The wavelength values are list in the figure inset. For the 1480-nm pumping case, the gains of the same six wavelengths were measured after a cascade of 27 optical amplifiers. If this gain variation is not equalized over the spectral range of operation in a multiple-wavelength system, it will create a large signal-to-noise ratio differential among the channels after they pass through a series of several EDFAs.

Numerous static and dynamic techniques have been considered for equalizing the optical amplifier outputs at different wavelengths.[36-39] These techniques include the use of fiber Bragg gratings, Fabry-Perot etalon filters, liquid crystal devices, and multilayered thin-film filters.

In the preceding discussions and analysis, we have looked only at transmission in one direction along a fiber. However, bidirectional propagation is also possible in a fiber link carrying multiple wavelengths
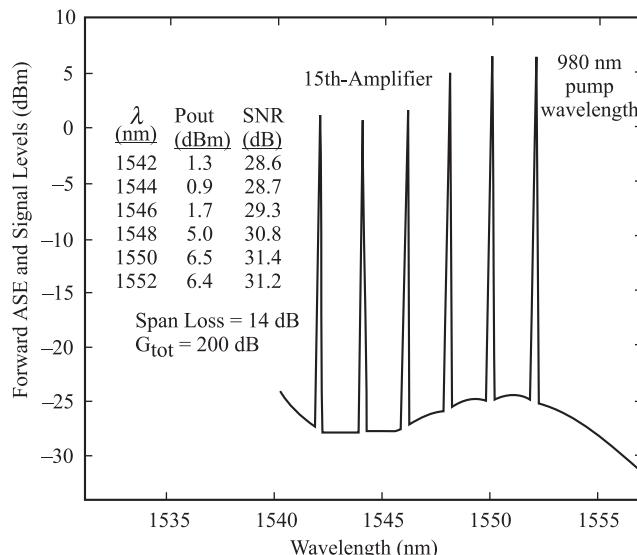
Fig. 11.16 *Signal level and ASE spectrum in a 0.2-nm optical bandwidth for 6 WDM signals after a cascade of 15 amplifiers. The wavelengths were spaced 2 nm apart and each was pumped with 50 mW of 980-nm pump power. (Reproduced with permission from Ali, Elrefaie, Wagner, and Ahmed,[22] © 1996, IEEE.)*
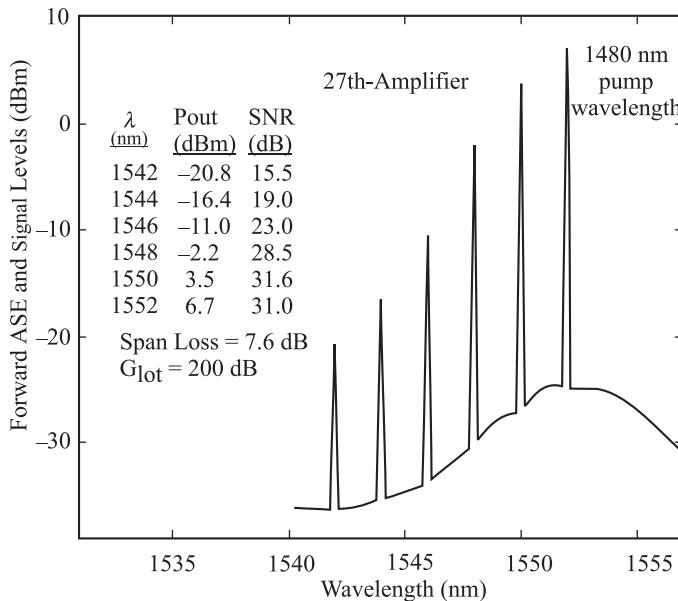
Fig. 11.17 *Signal level and ASE spectrum in a 0.2-nm optical bandwidth for 6 WDM signals after a cascade of 27 amplifiers. The wavelengths were spaced 2 nm apart and each was pumped with 50 mW of 1480-nm pump power. (Reproduced with permission from Ali, Elrefaie, Wagner, and Ahmed,[22] © 1996, IEEE.)*

in each direction through a cascaded chain of EDFAs.[40] This bidirectional method is used without any optical amplifier for fiber-to-the-premises (FTTP) networks, as Sec. 13.8 describes.[41]

11.6.5 In-Line Amplifier Gain Control

In any fiber transmission system using optical amplifiers, the dynamic characteristics of network traffic, particularly in a metro network, can create large, rapid fluctuations in the input optical power level to an EDFA. These optical power fluctuations can arise from factors such as the addition or deletion of wavelength channels, the rerouting of traffic when there is congestion, or the initiation of protection switching when links fail. The result is that when the overall input optical power to an amplifier increases, for example, from the addition of more channels, the gain for each channel will decrease. Likewise, when the total input optical power level decreases, the individual channels experience more gain so that the signals arrive at the receivers with higher power. This effect is particularly troublesome in a cascade of amplifiers, where sudden power peaks or drops can overload a receiver or can cause the received power level to fall below the required minimum, thereby leading to burst errors in the receiver.

Thus to avoid a degradation in system performance, it is necessary to keep the output power per channel of the in-line amplifiers constant regardless of the input power level. This is not an easy task, especially in cascaded chains of EDFAs in which the signal power overshoots owing to variations of the input optical power could occur at a much faster rate than for an individual EDFA.[42,43] Numerous methods have been tried for implementing schemes to mitigate these transient effects.[44–51]

One method uses an *automatic gain control* (AGC) circuit, which monitors the level of the optical signal power entering the EDFA and adjusts the pump power so that the output always remains at the same level. This is effective only if an automatic channel-by-channel gain and power regulation function is used with the AGC. Another alternative related to this is to insert another independent wavelength into the EDFA amplification band, which is referred to as the *control channel*. In this procedure the input and output powers of the control channel are monitored and the gain of the channel is calculated. A control circuit then adjusts the pump level to maintain a constant gain for the control channel.

Several different gain-clamping methods based on electrical or optical feedback, or a combination of the two, also have been examined. The AGC technologies using optical feedback tend to be expensive and require a complex control loop. The electrical AGC mechanisms generally have a lower cost and a high long-term reliability. In addition, it has been noted that amplitude transients depend on the length of the erbium-doped fiber and the erbium-ion concentration.[51] Thus by careful selection of the EDFA design parameters, it is possible to greatly reduce the effects of gain transients. Another possibility for mitigating the effects of power transients is through the use of an *erbium-doped wave guide amplifier* (EDWA) instead of an EDFA.[52–53] These type of devices are significantly smaller than an EDFA, can be made at a lower cost, and allow for integration of other electronic functions onto the same basic device substrate.

11.7 Raman Amplifiers

A *Raman optical amplifier* is based on a nonlinear effect called *stimulated Raman scattering* (SRS), which occurs in fibers at high optical powers.[54–62] Chapter 12 describes this characteristic in more detail. The SRS effect is due to an interaction between an optical energy field and the vibrational modes of the lattice structure in a material. Basically what happens is that an atom first absorbs a photon at a particular energy and then releases another photon at a lower energy, that is, at a longer wavelength than that of the absorbed photon. The energy difference between the absorbed and the released photons is transformed into a *phonon*, which is a vibrational mode of the material. The power transfer to higher wavelengths occurs over a broad spectral range of 80 to 100 nm. The shift to a particular longer wavelength is referred to as the *Stokes shift* for that wavelength. Figure 11.18 shows the Raman gain spectrum for a pump laser
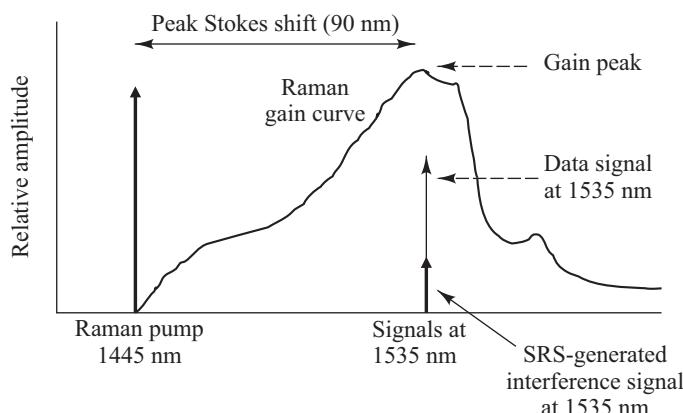
Fig. 11.18 *Stokes shift and the resulting Raman gain spectrum from a pump laser operating at 1445 nm*

operating at 1445 nm and illustrates the SRS-induced power transfer to a signal at 1535 nm, which is 90 nm away from the pump wavelength. Depending on the link architecture, the SRS-generated signal can act as either an intentional amplification of a particular data wavelength or it could be an unwanted interference signal at that wavelength (see Chapter 12). The gain curve is given in terms of the *Raman gain coefficient g_R* units of 10^{-14} m/W.

Whereas an EDFA requires a specially constructed optical fiber for its operation, a Raman amplifier makes use of the standard transmission fiber itself as the amplification medium. The Raman gain mechanism can be achieved through either a lumped (or discrete) amplifier or a distributed amplifier. In the *lumped Raman amplifier* configuration, a spool of about 80 m of small-core fiber along with appropriate pump lasers is inserted into the transmission path as a distinct packaged unit.

For the *distributed Raman amplifier* application, optical power from one or more Raman pump lasers is inserted into the end of the transmission fiber toward the transmitting end. This process converts the final 20 to 40 km of the transmission fiber into a preamplifier. Hence the word *distributed* is used because the gain is spread out over a wide distance. Figure 11.19 shows this effect on a single wavelength for several different pump levels. In this figure the term *on-off Raman gain* is defined as the increase in signal power at the receiver amplifier output when the Raman pump lasers are turned on. As the optical power from the pumps travels upstream (from the receiver toward the sender), the SRS effect progressively transfers power from shorter pump wavelengths to longer signal wavelengths. This occurs over the characteristic Raman-gain length $L_G = g_R P/A_{\text{eff}}$, where P is the pump laser power and A_{eff} is the effective area of the transmission fiber, which is approximately equal to the actual fiber cross-sectional area (see Chapter 12 for a detailed definition). In general, noise factors limit the practical gain of a distributed Raman amplifier to less than 20 dB.

In practice, link design engineers use several pump lasers to generate a flat wideband gain spectrum. Figure 11.20 shows a typical Raman gain spectrum for six pump lasers at different wavelengths. As indicated in the figure, when using several pump lasers, it is important to remember that there is a
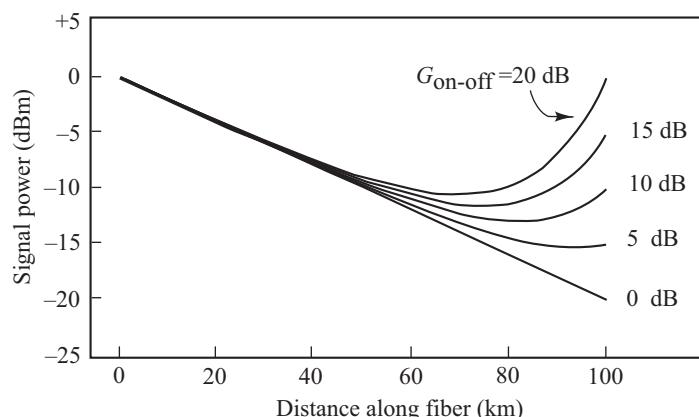
Fig. 11.19 *Simulated signal power evolution along a 100-km fiber link for different values of on-off Raman gain, which is the increase in signal power at the amplifier output when the pump lasers are turned on. (Reproduced with permission from Bromage,[5] © 2004, IEEE.)*
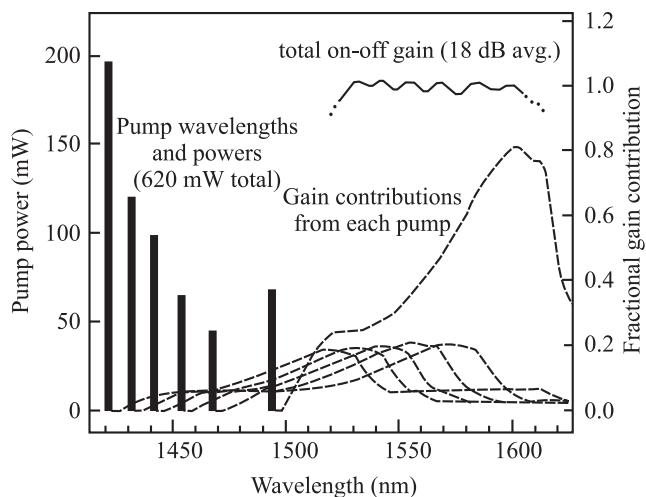
Fig. 11.20 *Numerical example of broadband Raman gain in a nonzero-dispersion shifted fiber when using six pump lasers. The bars show the pump wavelengths and their input powers. The solid line shows the total small-signal on-off gain and the dashed lines show the fractional gain contribution from each pump wavelength. (Reproduced with permission from Bromage,[5] © 2004, IEEE.)*

strong Raman interaction between the pumps themselves. Since pumps at short wavelengths amplify the power of longer-wavelength pumps, the shorter-wavelength pumps typically need to have higher power levels. In addition, note that as a result of the asymmetry of the Raman gain efficiency spectrum, the power level and wavelength of the longest-wavelength pump needs to be calculated carefully. For example, in the case shown in Fig. 11.20, although only 10 percent of the total pump power comes from the laser operating at 1495 nm, since it is amplified by the other pumps it contributes 80 percent of the
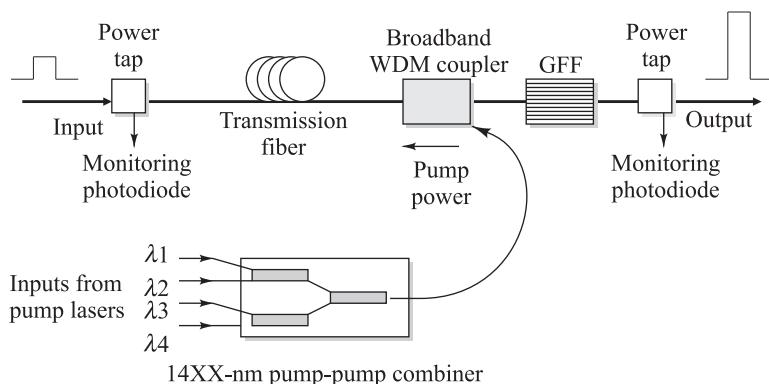
Fig. 11.21 *Setup for a typical Raman amplification system*

Table 11.2 *Performance parameters of a 14XX-nm pump-pump combiner based on fused-fiber coupler technology*

Parameter	Performance value
Device technology	Fused-fiber coupler
Wavelength range	1420 to 1500 nm
Channel spacing	Customized: 10 to 40 nm
	Standard: 10, 15, 20 nm
Insertion loss	< 0.8 dB
Polarization dependent loss	< 0.2 dB
Directivity	>55 dB
Optical power capability	3000 mW

gain for the longest signal wavelengths. The top right curve of Fig. 11.20 illustrates that by using an appropriate combination of pump wavelengths and pump powers, it is possible to achieve a fairly flat gain over a wide spectral range.

Pump lasers with high output powers in the 1400-to-1500-nm region are required for Raman amplification of C- and L-band signals. Lasers that provide fiber-launch powers of up to 300 mW are available in standard 14-pin butterfly packages. Figure 11.21 shows the setup for a typical Raman amplification system. Here a pump combiner multiplexes the outputs from four pump lasers operating at different wavelengths (examples might be 1425, 1445, 1465, and 1485 nm) onto a single fiber. These pump-power couplers are referred to popularly as *14XX-nm pump-pump combiners*. Table 11.2 lists the performance parameters of a pump combiner based on fused-fiber coupler technology. This combined pump power then is coupled into the transmission fiber in a counterpropagating direction through a broadband WDM coupler, such as those listed in Table 11.3. The differences in the power levels measured between the two monitoring photodiodes shown in Fig. 11.21 gives the amplification gain. The *gain-flattening filter* (GFF) is used to equalize the gains at different wavelengths.

Table 11.3 *Performance parameters of broadband WDM couplers for combining 14XX-nm pumps and C-band or L-band signals*

Parameter	Performance value	Performance value
Device technology	Micro-optics	Thin-film filter
Reflection channel λ range	1420 to 1490 nm	1440 to 1490 nm
Pass channel λ range	1505 to 1630 nm	1528 to 1610 nm
Reflection channel insertion loss	0.30 dB	0.6 dB
Pass channel insertion loss	0.45 dB	0.8 dB
Polarization dependent loss	0.05 dB	0.10 dB
Polarization mode dispersion	0.05 ps	0.05 ps
Optical power capability	2000 mW	500 mW
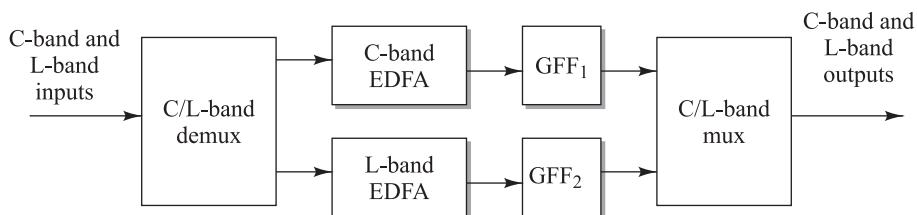
Fig. 11.22 *Representation of two different band optical amplifiers in parallel*

11.8 Wideband Optical Amplifiers

The ever growing demand for more bandwidth created an interest in developing wideband optical amplifiers that operate over several wavelength bands to handle a large number of WDM channels simultaneously. For example, a combination of two amplifier types can provide effective amplification in both the C- and L-bands or in the S- and C-bands. Extending this concept further, use of three amplifier types can provide signal gains in the S-, C-, and L-bands, in the C-, L-, and U-bands, or some other combination. The individual amplifiers could be based on thulium-doped silica fibers for the S-band, standard EDFAs for the C-band, gain-shifted EDFAs for the L-band, and different versions of Raman amplifiers.[63–69]

The amplifier combinations can be in parallel or in series, as shown in Figs. 11.22 and 11.23, respectively. In the parallel design a wideband demultiplexer splits the incoming signal spectrum into two wavelength bands. The two bands then pass through corresponding optical amplifiers after which a wideband multiplexer recombines the two spectral bands. This setup requires the use of a guard band spanning several nanometers between the two spectral regions. This guard band prevents amplification overlap between the different paths and prevents noise power originating in one amplifier from interfering with signal amplification in an adjacent amplifier. In addition to having this unusable wavelength band, another disadvantage of the parallel configuration is that the two WDM devices needed before and after each amplifier add to the system insertion loss.

The series configuration is known as a *seamless wideband optical amplifier* because it does not require splitting the signals into separate paths. It also avoids the noise figure degradations of wavelength couplers and the additional costs of the couplers themselves. These amplifiers can be constructed either from a concatenation of two or more doped-fiber amplifiers or from a combination of a fiber amplifier and a Raman amplifier. However, the impact on the amplifier design due to nonlinear effects and amplification of Rayleigh scattering need to be considered for hybrid fiber amplifiers consisting of a combination of an EDFA and a Raman amplifier. These effects are not as strong for a hybrid optical amplifier consisting of a series of concatenated doped-fiber amplifiers, but in this case the gain characteristics of the different amplifier segments need to be matched carefully.[66]
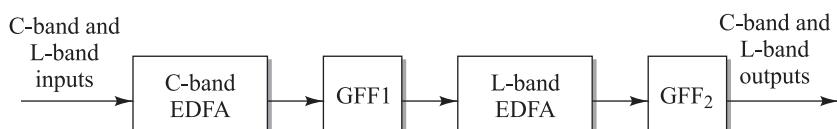
Fig. 11.23 *Representation of two different band optical amplifiers in series*

PROBLEMS

11.1 Consider an InGaAsP semiconductor optical amplifier that has the following parameter values:

Symbol	Parameter	Value
w	Active area width	$5\ \mu m$
d	Active area thickness	$0.5\ \mu m$
L	Amplifier length	$200\ \mu m$
Γ	Confinement factor	0.3
τ_r	Time constant	$1\ ns$
a	Gain coefficient	$1 \times 10^{-20}\ m^2$
V_g	Group velocity	$2.0 \times 10^8\ m/s$
n_{th}	Threshold density	$1.0 \times 10^{24}\ m^{-3}$

If a 100-mA bias current is applied, find (*a*) the pumping rate R_p, (*b*) the maximum (zero-signal) gain, (*c*) the saturation photon density, (*d*) the photon density if a l-μW signal at 1310 nm enters the amplifier. Compare the results of (*c*) and (*d*).

11.2 Verify that the gain expression in Eq. (11.13) follows from the integral relationship Eq. (11.12*b*).

11.3 Solving Eq. (11.13) numerically, make plots of the normalized amplifier gain (G/G_0) versus the normalized output power ($P_{s,out}/P_{amp,sat}$) for unsaturated amplifier gain values of $G_0 =$ 30 dB, 15 dB, and 10 dB.

11.4 The *output saturation power* $P_{out,sat}$ is defined as the amplifier output power for which the amplifier gain G is reduced by 3 dB (a factor of 2) from its unsaturated value G_0. Assuming $G_0 \gg 1$, show the amplifier output saturation power is

$$P_{out,sat} = \frac{G_0 \ln 2}{(G_0 - 2)} P_{amp,sat}$$

11.5 Since the gain constant a depends on the frequency, the amplifier gain is also frequency-dependent. The 3-dB bandwidth (full-width at half-maximum, FWHM) is defined as the frequency for which the power gain $G(v)$ is reduced by a factor of 2. Assume the gain parameter g has a gaussian profile

$$g(v) = \frac{g_0}{1 + 4(v - v_0)^2 \big/ (\Delta v)^2}$$

where Δv is the optical bandwidth (the spectral width of the gain profile) and v_0 is the maximum-gain frequency. Starting with Eq. (11.9), show that the ratio of the 3-dB bandwidth $2(v - v_0)$ to the optical bandwidth Δv is

$$\frac{2(v - v_0)}{\Delta v} = [\log_2(G_0/2)]^{-1/2}$$

where $\log_2 X$ is the base-2 logarithm of X. What does this equation show concerning the relationship between the amplifier gain and the optical bandwidth?

11.6 Assume the gain profile of an optical amplifier is

$$g(\lambda) = g_0 e^{-(\lambda - \lambda_0)^2 \big/ 2(\Delta \lambda)^2}$$

where λ_0 is the peak-gain wavelength and $\Delta \lambda$ is the spectral width of the amplifier gain. If $\Delta \lambda =$ 25 nm, find the FWHM (the 3-dB gain) of the amplifier gain if the peak gain at λ_0 is 30 dB.

11.7 (*a*) Compare the maximum theoretical power conversion efficiency (PCE) for 980-nm and 1475-nm pumping in an EDFA for a 1545-nm signal. Contrast this with actual measured results of PCE = 50.0 percent and 75.6 percent for 980-nm and 1475-nm pumping, respectively.

(*b*) Using the actual results for PCE given in (*a*), plot the maximum signal output power as a function of pump power for $0 \leq P_{p,in} \leq$ 200 mW for pump wavelengths of 980 nm and 1475 nm.

11.8 Assume we have an EDFA power amplifier that produces $P_{s,out} =$ 27 dBm for an input level of 2 dBm at 1542 nm.

(*a*) Find the amplifier gain.

(*b*) What is the minimum pump power required?

11.9 (*a*) To see the relative contributions of the various noise mechanisms in an optical amplifier, calculate the values of the five noise terms in Eq. (11.30) for operational gains of $G = 20$ dB and 30 dB. Assume the optical bandwidth is equal to the spontaneous emission bandwidth (30-nm spectral width) and use the following parameter values:

Symbol	Parameter	Value
η	Photodiode quantum efficiency	0.6
$\mathcal{R}$	Responsivity	0.73 A/W
P_{in}	Input optical power	$1\,\mu$W
λ	Wavelength	1550 nm
B_o	Optical bandwidth	3.77×10^{12} Hz
B_e	Receiver bandwidth	1×10^9 Hz
n_{sp}	Spontaneous emission factor	2
R_L	Receiver load resistor	1000 Ω

(*b*) To see the effect of using a narrowband optical filter at the receiver, let $B_o = 1.25 \times 10^{11}$ Hz (125 GHz at 1550 nm) and find the same five noise terms for $G = 20$ dB and 30 dB.

11.10 Plot the penalty factor $F(G)$ given by Eq. (11.43) as a function of amplifier gain for gains ranging from 0 to 30 dB. Assuming the fiber attenuation is 0.2 dB/km, draw a distance axis parallel to the gain axis to show the transmission distances corresponding to the gain values.

11.11 Consider a cascaded chain of k fiber-plus-EDFA combinations, as shown in Fig. 11.24.
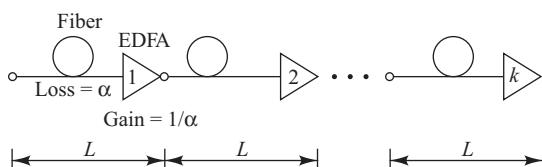
Fig. 11.24   *Cascaded EDFA chain*

(*a*) Show that the path-averaged signal power is

$$\left\langle P \right\rangle_{\text{path}} = P_{s,\text{in}} \frac{G-1}{G \ln G}$$

(*b*) Derive the path-averaged ASE power given by Eq. (11.42).

11.12 Consider a long-distance transmission system containing a cascaded chain of EDFAs. Assume each EDFA is operated in the saturation region and that the slope of the gain-versus-input power curve in this region is -0.5; that is, the gain changes by ± 3 dB for a ∓ 6-dB variation in input power. Let the link have the following operational parameters:

Symbol	Parameter	Value
G	Nominal gain	7.1 dB
$P_{s,\text{out}}$	Nominal output optical power	3.0 dBm
$P_{s,\text{in}}$	Nominal input optical power	-4.1 dBm

(*a*) Suppose there is a sudden 6-dB drop in signal level at some point in the link. Find the power output levels after the degraded signal has passed through 1, 2, 3, and 4 succeeding amplifier stages.

(*b*) Repeat part (*a*) for a signal-level drop of 12 dB.

11.13 Let the electric field of an optical signal at a carrier frequency v_i be

$$E_i(t) = \sqrt{2P_i}\,\cos(2\pi v_i t + \phi_i)$$

where P_i is the signal power at the carrier frequency v_i and ϕ_i is the carrier phase. If N optical signals each at a different frequency v_i are traveling along a fiber, show that the signal power is

$$P_s = \sum_{i=1}^{N} P_{s,i} + \sum_{j}^{N}\sum_{k \neq j}^{N} 2\sqrt{P_j P_k}\,\cos(\Omega_{jk} t + \phi_j - \phi_k)$$

where $\Omega_{jk} = 2\pi(v_j - v_k)$ represents the beat frequency at which the carrier population oscillates.

11.14 Consider an EDFA with a gain of 26 dB and a maximum output power of 0 dBm.

(*a*) Compare the output signal levels per channel for 1, 2, 4, and 8 wavelength channels, where the input power is 1 μW for each signal.

(*b*) What are the output signal levels per channel in each case if the pump power is doubled?

11.15 Recall from Eq. (7.13) that the bit-error rate (BER) can be given in terms of a Q factor, where, from Eq. (7.14),

$$Q = \frac{b_{\text{on}} - b_{\text{off}}}{\sqrt{\sigma_{\text{on}}^2} + \sqrt{\sigma_{\text{off}}^2}}$$

When thermal noise is dominant, we have $\sigma_{\text{on}}^2 = \sigma_{\text{off}}^2$. However, for an EDFA the existence of signal-ASE beat noise produces the condition $\sigma_{\text{on}}^2 > \sigma_{\text{off}}^2$. In this case, $\sigma_{\text{on}}^2 = \sigma_{\text{total}}^2$ from Eq. (11.30),

$$\sigma_{\text{off}}^2 = \sigma_T^2 + \sigma_{\text{shot-ASE}}^2 + \sigma_{\text{ASE-ASE}}^2$$

and

$$b_{\text{on}} - b_{\text{off}} = (GI_s + I_N) - I_N = G\mathscr{R}P_{s,\text{in}}$$

Noting that the receiver sensitivity P_R is half the signal power of a transmitted 1 bit for a uniform distribution of ones and zeros (i.e., $P_s = 2P_R$), show that in terms of the Q factor the receiver sensitivity is

$$P_R = hv\, B_e \left\{ FQ^2 + Q\left[\frac{2n_{\text{sp}}}{\eta}\left(\frac{G-1}{G^2} \right)\frac{B_o}{B_e} \right. \right.$$
$$\left. \left. + n_{\text{sp}}^2\left(\frac{G-1}{G} \right)^2 \left(2\frac{B_o}{B_e} - 1 \right) + \frac{4k_BT}{R_L(\eta q)^2 G^2 B_e} \right]^{1/2} \right\}$$

where F is the noise figure given by Eq. (11.35).

11.16 Using the expression for the receiver sensitivity P_R given in Prob. 11.15, plot P_R as a function of G for gain values ranging from 10 to 40 dB. Use the following values of B_o: 2.5, 12.5, 125, and 675 GHz (corresponding to

spectral bandpasses of 0.02, 0.1, 1, and 5 nm, respectively, at 1550 nm). Let $Q = 6$ and assume the following values for the parameters in the expression for P_R:

Symbol	Parameter	Value
η	Photodiode quantum efficiency	0.6
T	Temperature	300 K
λ	Wavelength	1550 nm
B_e	Receiver bandwidth	1.25 GHz
n_{sp}	Spontaneous emission factor	2
R_L	Receiver load resistor	50 Ω

11.17 Consider the expression for the receiver sensitivity P_R given in Prob. 11.15. For sufficiently large values of G, the thermal noise term is negligible and the first term in the square-root expression is small compared with the second term. In this case, and for $2B_o \gg 1$, the expression for P_R becomes

$$P_R \approx hv\, B_e F \left\{ Q^2 + \frac{Q}{2\eta}\left(\frac{B_o}{B_e} \right)^{1/2} \right\}$$

Letting $Q = 6$ (for a 10^{-9} BER), plot P_R as a function of the relative optical filter bandwidth B_o/B_e for the range $10 \le B_o/B_e \le 1000$ for the following three values of F: 3 dB, 5 dB, and 7 dB. Assume a wavelength of 1550 nm, a photodiode quantum efficiency of 0.5, and an electrical bandwidth of 1.25 GHz.

11.18. If the input power and the optical bandwidth are relatively large, the relative intensity noise inherent in the laser transmitter may affect the noise figure of an optical amplifier. In this case, the noise figure F relative to $2n_{\text{sp}}$ (the noise figure when signal–ASE beating is the dominant noise) is given by[70]

$$\frac{F}{2n_{\text{sp}}} = 1 + \frac{P_{\text{ASE}}}{P_s} + \frac{(\text{RIN})B_o}{8}\frac{2B}{B_o}$$
$$+ \frac{(\text{RIN})B_o}{4}\frac{P_s}{P_{\text{ASE}}}$$

where $P_{ASE} = n_{sp} h\nu B_o$ is the ASE noise power. Here, the third term comes from the interaction between the RIN and the ASE, and the fourth term arises from the interaction between the RIN and the signal. Plot $F/2n_{sp}$ as a function of P_s/P_{ASE} for the range -10 dB $\leq P_s/P_{ASE} \leq$ 20 dB. Use the following values of (RIN) B_o: 0.01, 0.1, 1, and 10.

REFERENCES

1. R. Giles and T. T. Li, "Optical amplifiers transform long-distance lightwave telecommunications," *Proc. IEEE*, vol. 84, pp. 870–883, June 1996.

2. D. R. Zimmerman and L. H. Spiekman, "Amplifiers for the masses: EDFA, EDWA, and SOA amplets for metro and access applications," *J. Lightwave Tech.*, vol. 22, pp. 63–70, Jan. 2004.

3. M. J. Connelly, *Semiconductor Optical Amplifiers*, Springer, New York, 2002.

4. E. Desurvire, *Erbium-Doped Fiber Amplifiers: Principles and Applications*, Wiley, New York, 2002.

5. J. Bromage, "Raman amplification for fiber communications systems," *J. Lightwave Tech.*, vol. 22, pp. 79–93, Jan. 2004.

6. A. Yariv and P. Yeh, *Photonics: Optical Electronics in Modern Communications*, Oxford University Press USA, New York, 6th ed., 2006.

7. B.E.A. Saleh and M. C. Teich, *Fundamentals of Photonics*, Wiley, Hoboken, NJ, 2nd ed., 2007.

8. M. J. O'Mahoney, "Semiconductor laser optical amplifiers for use in future fiber systems," *J. Lightwave Tech.*, vol. 6, pp. 531–544, Apr. 1988.

9. B. L. Anderson and R. L. Anderson, *Fundamentals of Semiconductor Devices*, McGraw-Hill, New York, 2005.

10. R. Ramaswami and P. Humblet, "Amplifier induced crosstalk in multichannel optical networks," *J. Lightwave Tech.*, vol. 8, pp. 1882–1896, Apr. 1990.

11. K. Morito, "Output-level control of semiconductor optical amplifier by external light injection," *J. Lightwave Tech.*, vol. 23, pp. 4332–4341, Dec. 2005.

12. E. Desurvire, D. Bayart, B. Desthieux, and S. Bigo, *Erbium-Doped Fiber Amplifiers: Devices and System Developments*, Wiley, New York, 2002.

13. W. Miniscalco, "Erbium-doped glasses for fiber amplifiers at 1500 nm," *J. Lightwave Tech.*, vol. 9, pp. 234–250, Feb. 1991.

14. W. Miniscalco, "Optical and electronic properties of rare earth ions in glasses," in M.J.F. Digonnet, ed., *Rare Earth Doped Fiber Lasers and Amplifiers*, Dekker, New York, 1993.

15. C.-C. Fan, J.-D. Peng, J.-H. Li, X. Jiang, G.-S. Wu, and B.-K. Zhou, "Theoretical and experimental investigations on erbium-doped fiber amplifiers," *Fiber & Integrated Optics*, vol. 13, no. 3, pp. 247–260, 1994.

16. F. Fontana and G. Grasso, "The erbium-doped fiber amplifier: Technology and applications," *Fiber & Integrated Optics*, vol. 13, no. 3, pp. 135–145, 1994.

17. P. C. Becker, N. A. Olsson, and J. R. Simpson, *Erbium-Doped Fiber Amplifiers*, Academic, New York, 1999.

18. C. Lester, A. Bjarklev, T. Rasmussen, and P. G. Dinesen, "Modeling of Yb-sensitized Er-doped silica waveguide amplifiers," *J. Lightwave Tech.*, vol. 13, pp. 740–743, May 1995.

19. E. Yahel, O. Hess, and A. Hardy, "Transient analysis of short, high-concentration, gain-clamped Er^{3+}–Yb^{3+} codoped fiber amplifiers," *J. Lightwave Tech.*, vol. 24, pp. 2190–2195, May 2006.

20. C. R. Giles and E. Desurvire, "Modeling erbium-doped fiber amplifiers," *J. Lightwave Tech.*, vol. 9, pp. 271–283, Feb. 1991.

21. T. Li, "The impact of optical amplifiers on long-distance telecommunications," *Proc. IEEE*, vol. 81, pp. 1568–1579, June 1993.

22. M. A. Ali, A. F. Elrefaie, R. E. Wagner, and S. A. Ahmed, "A detailed comparison of the overall performance of 980 and 1480 nm pumped EDFA cascades in WDM multiple-access lightwave networks," *J. Lightwave Tech.*, vol. 14, pp. 1436–1448, June 1996.

23. B. Pedersen, K. Dybdal, C. D. Hansen, A. Bjarklev, J. H. Povlsen, H. Vendeltorp-Pommer, and C. C. Larsen, "Detailed theoretical and experimental investigation of high-gain erbium-doped fiber amplifier," *IEEE Photonics Tech. Lett.*, vol. 2, pp. 863–865, Dec. 1990.

24. G. R. Walker, N. G. Walker, R. C. Steele, M. J. Creaner, and M. C. Brain, "Erbium-doped fiber amplifier cascade for multichannel coherent optical transmission," *J. Lightwave Tech.*, vol. 9, pp. 182–193, Feb. 1991.

25. D. M. Baney, P. Gallion, and R. S. Tucker, "Theory and measurement techniques for the noise figure of optical amplifiers," *Opt. Fiber Tech.*, vol. 6, no. 2, pp. 122–154, Apr. 2000.

26. A. Hardy and R. Oron, "Signal amplification in strongly pumped fiber amplifiers," *IEEE J. Quantum Electron.*, vol. 33, pp. 307–313, Mar. 1997.

27. J. P. Gordon and L. F. Mollenauer, "Effects of fiber nonlinearities and amplifier spacing on ultra-long distance transmission," *J. Lightwave Tech.*, vol. 9, pp. 170–173, Feb. 1991.

28. A. V. Tran, R. S. Tucker, and N. L. Boland, "Amplifier placement methods for metropolitan WDM ring networks," *J. Lightwave Tech.*, vol. 22, pp. 2509–2522, Nov. 2004.

29. J. J. Pan, K. Guan, X. Qiu, W. Wang, M. Zhang, J. Jiang, E. Zhang, and F. Q. Zhou, "Advantages of low-cost, miniature, intelligent EDFAs for next-generation dynamic metro/access networks," *Optical Fiber Tech.*, vol. 9, no. 2, pp. 80–94, Apr. 2003.

30. T. T. Ha, G. E. Keiser, and R. L. Borchart, "Bit error probabilities of OOK lightwave systems with optical amplifiers," *J. Opt. Commun.*, vol. 18, pp. 151–155, Aug. 1997.

31. Y. K. Park and S. W. Granlund, "Optical preamplifier receivers: Applications to digital long-haul transmission," *Optical Fiber Tech.*, vol. 1, pp. 59–71, Oct. 1994.

32. B. Razavi, *Design of Integrated Circuits for Optical Communications*, McGraw-Hill, New York, 2003.

33. F. Matera and M. Settembre, "Performance of optical links with optical amplifiers," *Fiber & Integ. Optics*, vol. 15, no. 2, pp. 89–107, 1996.

34. K. Inoue, "Crosstalk and its power penalty in multichannel transmission due to gain saturation in a semiconductor laser amplifier," *J. Lightwave Tech.*, vol. 7, pp. 1118–1124, July 1989.

35. A. E. Willner and S.-M. Hwang, "Transmission of many WDM channels through a cascade of EDFAs in long distance links and ring networks," *J. Lightwave Tech.*, vol. 13, pp. 802–816, May 1995.

36. J.-C. Dung, S. Chi, and S. Wen, "Gain flattening of erbium-doped fiber amplifier using fiber Bragg gratings," *Electron. Lett.*, vol. 34, pp. 555–556, Mar. 1998.

37. K. Mizuno, Y. Nishi, Y. Mimura, Y. Iida, H. Matsuura, D. Yoon, O. Aso, T. Yamamoto, T. Toratani, Y. Ono, and A. Yo, "Development of etalon-type gain-flattening filter," *Furukawa Review*, vol. 19, pp. 53–58, 2000.

38. K. Maru, T. Chiba, T. Hasegawa, K. Tanaka, H. Nonen, and H. Uetsuka, "A dynamic gain equalizer for next-generation WDM systems," *Hitachi Cable Review*, vol. 21, pp. 7–10, Aug. 2002.

39. Y.-Q. Lu, C. Wong, and S.-T. Wu, "A liquid crystal-based Fourier optical spectrum analyzer," *IEEE Photonics Tech. Lett.*, vol. 16, pp. 861–862, Mar. 2004.

40. C. Delisle and J. Conradi, "Model for bidirectional transmission in an open cascade of optical amplifiers," *J. Lightwave Tech.*, vol. 15, pp. 749–757, May 1997.

41. G. Keiser, *FTTX Concepts and Applications*, Wiley, Hoboken, NJ, 2006.

42. A. K. Srivastava, Y. Sun, J. L. Zyskind, J. W. Sulhoff, C. Wolf, and R. W. Tkach, "EDFA transient response to channel loss in WDM transmission system," *IEEE Photon. Tech. Lett.*, vol. 9, pp. 386–388, 1997.

43. S. R. Chinn, "Simplified modeling of transients in gain-clamped erbium-doped fiber amplifiers,"

J. Lightwave Tech., vol. 16, pp. 1095–1100, June 1998.

44. M. Karásek and F. W. Willems, "Channel addition/removal response in cascades of strongly inverted erbium-doped fiber amplifiers," *J. Lightwave Tech.*, vol. 16, pp. 2311–2317, Dec. 1998.

45. M. Karásek, M. Menif, and L. A. Rusch, "Output power excursions in a cascade of EDFAs fed by multichannel burst-mode packet traffic: experimentation and modeling," *J. Lightwave Tech.*, vol. 19, pp. 933–940, July 2001.

46. H. Ono, K. Shimano, M. Fukutoku, and S. Kuwano, "An EDFA gain control and power monitoring scheme for fault detection in WDM networks by employing a power-stabilized control channel," *J. Lightwave Tech.*, vol. 20, pp. 1335–1341, Aug. 2002.

47. C. Tian and S. Kinoshita, "Analysis and control of transient dynamics of EDFA pumped by 1480- and 980-nm lasers," *J. Lightwave Tech.*, vol. 21, pp. 1728–1734, Aug. 2003.

48. A. V. Tran, C. J. Chae, R. S. Tucker, and Y. J. Wen, "EDFA transient control based on envelope detection for optical burst switched networks," *IEEE Photon. Tech. Lett.*, vol. 17, pp. 226–228, Jan. 2005.

49. J. H. Ji, L. Zhan, L. L. Yi, C. C. Tang, Q. H. Ye, and Y. X. Xia, "Low noise-figure gain-clamped erbium-doped fiber ring lasing amplifier with an interleaver," *J. Lightwave Tech.*, vol. 23, pp. 1375–1379, Mar. 2005.

50. T. Yoshikawa, K. Okamura, E. Otani, T. Okaniwa, T. Uchino, M. Fukushima, and N. Kagi, "WDM burst mode signal amplification by cascaded EDFAs with transient control," *Opt. Express*, vol. 14, pp. 4650–4655, May 2006.

51. P. Chan and H. Tsang, "Minimizing gain transient dynamics by optimizing the erbium concentration and cavity length of a gain clamped EDFA," *Opt. Express*, vol. 13, pp. 7520–7526, Sept. 2005.

52. K. Ennser, G. DellaValle, M. Ibsen, J. Shmulovich, and S. Taccheo, "Erbium-doped waveguide amplifier for reconfigurable WDM metro networks," *IEEE Photonics Tech. Lett.*, vol. 17, pp. 1468–1470, July 2005.

53. A. Laliotis, E. M. Yeatman, and S. J. Al-Bader, "Modeling signal and ASE evolution in erbium-doped amplifiers with the method of lines," *J. Lightwave Tech.*, vol. 24, pp. 1589–1600, Mar. 2006.

54. K. Rottwitt and A. J. Stentz, "Raman amplification in lightwave communication systems," in *Optical Fiber Communications IV-A*, I. P. Kaminow and T. Li, eds., Academic, San Diego, 2002.

55. M. N. Islam, "Raman amplifiers for telecommunications," *IEEE J. Sel. Topics Quantum Electron.*, vol. 8, no. 3, pp. 548–559, May/June 2002.

56. M. N. Islam, ed., *Raman Amplifiers for Telecommunications 1—Physical Principles*, Springer, New York, 2004.

57. M. N. Islam, ed., *Raman Amplifiers for Telecommunications 2—Sub-Systems and Systems*, Springer, New York, 2004.

58. J. Bromage, J.-C. Bouteiller, H. J. Thiele, K. Brar, L. E. Nelson, S. Stulz, C. Headley, R. Boncek, J. Kim, A. Klein, G. Baynham, L. V. Jørgensen, L. Grüner-Nielsen, L. R. Lingle Jr., and D. J. DiGiovanni, "WDM transmission over multiple long spans with bidirectional Raman pumping," *J. Lightwave Tech.*, vol. 22, pp. 225–232, Jan. 2004.

59. H. Suzuki, N. Takacio, H. Masuda, and K. Iwatsuki, "Super-dense WDM transmission technology in the zero-dispersion region employing distributed Raman amplification," *J. Lightwave Tech.*, vol. 21, pp. 973–981, Apr. 2003.

60. T. Miyamoto, M. Tanaka, J. Kobayashi, T. Tsuzaki, M. Hirano, T. Okuno, M. Kakui, and M. Shigematsu, "Highly nonlinear fiber-based lumped fiber Raman amplifier for CWDM transmission systems," *J. Lightwave Tech.*, vol. 23, pp. 3475–3483, Nov. 2005.

61. J. Chen, C. Lu, Y. Wang, and Z. Li, "Design of multistage gain-flattened fiber Raman amplifiers," *J. Lightwave Tech.*, vol. 24, pp. 935–944, Feb. 2006.

62. J. Park, J. Park, D. Lee, N. Y. Kim, H. Lee, and N. Park, "Nonlinear phase shift scanning method for the optimal design of Raman

transmission systems," *J. Lightwave Tech.*, vol. 24, pp. 1257–1268, Mar. 2006.

63. H. Masuda, S. Kawai, K.-I. Suzuki, and K. Aida, "Ultrawide 75-nm gain-band optical amplification with erbium-doped fluoride fiber amplifiers and distributed Raman amplifiers," *IEEE Photonics Tech. Lett.*, vol. 10, pp. 516–518, Apr. 1998.

64. H. Ono, M. Yamada, and M. Shimizu, "S-band erbium-doped fiber amplifiers with a multistage configuration: Design, characterization, and gain tilt compensation," *J. Lightwave Tech.*, vol. 21, pp. 2240–2246, Oct. 2003.

65. C.-H. Yeh, C.-C. Lee, C.-Y. Chen, and S. Chi, "S-band gain-clamped erbium-doped fiber amplifier by using optical feedback method," *IEEE Photonics Tech. Lett.*, vol. 16, pp. 90–92, Jan. 2004.

66. T. Sakamoto, S.-I. Aozasa, M. Yamada, and M. Shimizu, "Hybrid fiber amplifiers consisting of cascaded TDFA and EDFA for WDM signals," *J. Lightwave Tech.*, vol. 24, pp. 2287–2295, June 2006.

67. S.W. Harun, P. Poopalan, and H. Ahmad, "Gain enhancement in L-band EDFA through a double-pass technique," *IEEE Photonics Tech. Lett.*, vol. 14, pp. 296–297, Mar. 2002.

68. L. Huang, A. Jha, S. Shen, and X. Liu, "Broadband emission in Er^{3+}-Tm^{3+} codoped tellurite fibre," *Optics Express*, vol. 12, pp. 2429–2434, 31 May 2004.

69. J. H. Lee, Y. M. Chang, Y. G. Han, H. Chung, S. H. Kim, and S. B. Lee, "A detailed experimental study on single-pumped Raman/EDFA hybrid amplifiers: Static, dynamic, and system performance comparison," *J. Lightwave Tech.*, vol. 23, pp. 3484–3493, Nov. 2005.

70. I. Jacobs, "Dependence of optical amplifier noise figure on relative-intensity noise," *J. Lightwave Tech.*, vol. 13, pp. 1461–1465, July 1995.

CHAPTER 12

Nonlinear Effects

The design of a lightwave transmission system requires careful planning of factors such as fiber selection, choice and tuning of optoelectronic components, optical amplifier placement, and path routing. The goal of this planning is to create a network that meets the design criteria, is reliable, and is easy to operate and maintain. As the previous chapters describe, the design process must take into account all power penalties associated with optical signal-degradation processes.

Intuitively, it seems natural to let the input optical power be as large as possible to overcome the power penalty effects to achieve the link design goals. However, this works only if the fiber is a linear medium; that is, if the loss and refractive index are independent of the optical signal power level. In an actual fiber several different nonlinear effects start to appear as the optical power level increases.[1–7] These nonlinearities arise when several high-strength optical fields from different signal wavelengths are present in a fiber at the same time and when these fields interact with acoustic waves and molecular vibrations. For example, if the nonlinear threshold for the total launched power into a fiber is 17 dBm (50 mW), then for a 64-channel DWDM link the power limit per wavelength is −1.0 dBm (0.78 mW). Consequences of nonlinear effects for signal levels of this magnitude include power gain or loss at different wavelengths, wavelength conversions, and crosstalk between wavelength channels. In some cases the nonlinear effects can degrade WDM system performance, while in other situations they might provide a useful application.

This chapter first gives a general overview of nonlinear processes in optical fibers in Sec. 12.1. Since the nonlinearities arise above a certain optical power threshold, the effect becomes negligible once the signal has become sufficiently attenuated after traveling a certain distance along the fiber. This gives rise to the concept of an effective length and an associated parameter called effective area, as Sec. 12.2 describes. The next five sections discuss how the major nonlinear processes physically affect system performance. These nonlinearities are *stimulated Raman scattering* (Sec. 12.3), *stimulated Brillouin scattering* (Sec. 12.4), *self-phase modulation* (Sec. 12.5), *cross-phase modulation* (Sec. 12.6), and *four-wave mixing* (Sec. 12.7). Mitigation of four-wave mixing can be achieved by means of a special fiber design or by a *chromatic dispersion-compensation method*, which is the topic of Sec. 12.8. On the other hand, nonlinear effects also can have beneficial uses. Section 12.9 describes applications of cross-phase-modulation and four-wave-mixing techniques for performing wavelength conversion in WDM networks. Another application of nonlinear effects in a silica fiber is the use of solitons for optical communications, which depends on self-phase modulation effects. Section 12.10 addresses this topic.

Table 12.1 *Summary of nonlinear effects in optical fibers*

Nonlinearity category	Single-channel	Multiple-channel
Index related	Self-phase modulation	Cross-phase modulation Four-wave mixing
Scattering related	Stimulated Brillouin scattering	Stimulated Raman scattering

12.1 General Overview of Nonlinearities

Optical nonlinearities can be classified into two general categories, which are summarized in Table 12.1. The first category encompasses the nonlinear inelastic scattering processes. These are *stimulated Raman scattering* (SRS) and *stimulated Brillouin scattering* (SBS). The second category of nonlinear effects arises from intensity-dependent variations in the refractive index in a silica fiber, which is known as the *Kerr effect*. These effects include *self-phase modulation* (SPM), *cross-phase modulation* (XPM), and *four-wave mixing* (FWM). In the literature, FWM sometimes is referred to as four-photon mixing (FPM), and XPM also is designated by CPM. Note that certain nonlinear effects are independent of the number of WDM channels.

SBS, SRS, and FWM result in gains or losses in a wavelength channel. The power variations depend on the optical signal intensity. These three nonlinear processes provide gains to some channels while depleting power from others, thereby producing crosstalk between the wavelength channels. In analog video systems, SBS significantly degrades the carrier-to-noise ratio when the scattered power is equivalent to the signal power in the fiber. SPM and XPM affect only the phase of signals, which causes chirping in digital pulses. This can worsen pulse broadening due to dispersion, particularly in very high-rate systems, such as 40 Gb/s. FWM can be suppressed through special arrangements of fibers having different dispersion characteristics. In this chromatic dispersion-compensation technique, the design engineer tailors the various fiber segments in a WDM link to have high local dispersion, but an overall low dispersion. The low average dispersion minimizes pulse spreading, whereas the high local dispersion destroys the carrier-frequency phase relationships that give rise to FWM intermodulation products.

When any of these nonlinear effects contribute to signal impairment, an additional amount of power will be needed at the receiver to maintain the same BER as in their absence. This additional power (in decibels) is known as the *power penalty* for that effect. As the following sections show, factors that influence to what degree a particular nonlinearity affects optical fiber link performance include chromatic and polarization-mode dispersions, the effective core area of the fiber, the number and spacing of wavelength channels in a WDM system, the length of the transmission link, and the light source linewidth and emitted optical power level.

12.2 Effective Length and Area

Modeling the nonlinear processes can be quite complicated because their effects depend on the transmission length, the cross-sectional area of the fiber, and the optical power level in the fiber.[7] The difficulty arises from the fact that the impact of the nonlinearity on signal fidelity increases with distance. However, this is offset by the continuous decrease in signal power along the fiber due to attenuation, as Fig. 12.1 shows. In practice, one can use a simple but sufficiently accurate model that assumes the power is constant over
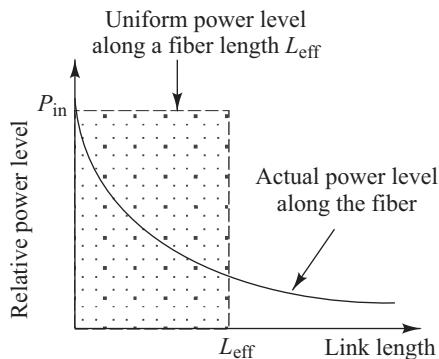
Fig. 12.1 *The effective length is modeled as the transmission length at which the shaded area equals the area under the actual power-distribution curve*
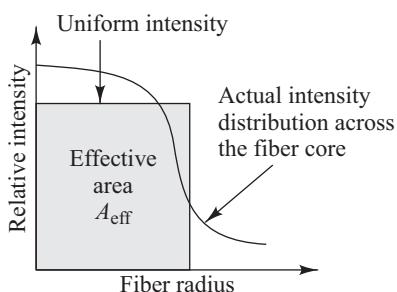
Fig. 12.2 *The effective area is modeled as a central area of the fiber core within which the intensity is assumed to be uniform*

a certain fiber length, which is less than or equal to the actual fiber length. This *effective length L_{eff}*, which takes into account power absorption along the length of the fiber (i.e., the fact that the optical power decays exponentially with length), is given by

$$L_{eff} = \frac{1 - e^{-\alpha L}}{\alpha} \qquad (12.1)$$

As Fig. 12.1 illustrates, this means that the area designated by the shaded pattern is equal to the area under the power-distribution curve. For an attentuation of 0.22 dB/km (or, equivalently, 5.07×10^{-2} km^{-1}) at 1550 nm, the effective length is about 20 km when $L_{eff} \gg 1/\alpha$. When there are optical amplifiers in a link, the signal-impairments owing to the nonlinearities do not change as the signal passes through the amplifier.

The effects of nonlinearities increase with the light intensity in a fiber. For a given optical power, this intensity is inversely proportional to the cross-sectional area of the fiber core. As Fig. 12.2 shows, although the intensity is not distributed uniformly across the fiber-core area, for practical purposes one can use an *effective cross-sectional area A_{eff}*, which assumes a uniform intensity distribution across most of the core. This area can be calculated from mode-overlap integrals and, in general, is close to the actual core area. The result is that the impact of most nonlinear effects in a fiber can be calculated based on the effective area of the fundamental mode in that fiber. For example, the effective intensity of a light pulse becomes $I_e = P/A_{eff}$, where P is the optical power contained in the pulse. Table 12.2 lists the effective areas of some single-mode fibers.

Table 12.2 *Effective area and attenuation of some single-mode fibers*

Fiber type	Attenuation (dB/km)	Effective area (μm^2)
G.652 standard single-mode	0.35 at 1310 nm	72
G.652C/D low-water-peak	0.20 at 1550 nm	72
Dispersion-compensating	0.40 at 1550 nm	21
G.655 single-mode	0.21 at 1550 nm	55

12.3 Stimulated Raman Scattering

Stimulated Raman scattering is an interaction between lightwaves and the vibrational modes of silica molecules.[3–8] If a photon with energy $h\nu_1$ is incident on a molecule having a vibrational frequency ν_m, the molecule can absorb some energy from the photon. In this interaction, the photon is scattered, thereby attaining a lower frequency ν_2 and a corresponding lower energy $h\nu_2$. The modified photon is called a *Stokes photon*. Because the optical signal wave that is injected into a fiber is the source of the interacting photons, it is often called the *pump wave* because it supplies power for the generated wave.

This process generates scattered light at a wavelength longer than that of the incident light. If another signal is present at this longer wavelength, the SRS light will amplify it and the pump-wavelength signal will decrease in power; Fig. 12.3 illustrates this effect. Consequently, SRS can severely limit the performance of a multichannel optical communication system by transferring energy from short-wavelength channels to neighboring higher-wavelength channels. This is a broadband effect that can occur in both directions. Powers in WDM channels separated by up to 16 THz (125 nm) can be coupled through the SRS effect, as Fig. 12.4 illustrates in terms of the Raman gain coefficient g_R as a function of the channel separation $\Delta\nu_s$. This shows that, owing to SRS, the power transferred from a lower-wavelength channel to a higher-wavelength channel increases approximately linearly with channel spacing up to a maximum of about $\Delta\nu_c = 16$ THz (or $\Delta\lambda_c = 125$ nm in the 1550-nm window), and then drops off sharply for larger spacings.

To see the effects of SRS, consider a WDM system that has N channels equally spaced in a 30-nm band centered at 1545 nm. Channel 0, which is at the lowest wavelength, is affected the worst because power gets transferred from this channel to all longer-wavelength channels. For simplicity, assume that the transmittted power P is the same on all channels, that the Raman gain increases linearly as shown by the dashed line in Fig.12.4, and that there is no interaction between the other channels. If $F_{out}(j)$ is the fraction of power coupled from channel 0 to channel j, then the total fraction of power coupled out of channel 0 to all the other channels is[9]

$$F_{out} = \sum_{j=1}^{N-1} F_{out}(j) = \sum_{j=1}^{N-1} g_{R,peak} \frac{j\Delta\nu_s}{\Delta\nu_c} \frac{PL_{eff}}{2A_{eff}} = \frac{g_{R,peak}\Delta\nu_s PL_{eff}}{2\Delta\nu_c A_{eff}} \frac{N(N-1)}{2} \tag{12.2}$$

The power penalty for this channel then is $-10 \log(1 - F_{out})$. To keep the penalty below 0.5 dB, we need to have $F_{out} < 0.1$. Using Eq.12.2, and with $A_{eff} = 55$ μm^2 and $g_{R,peak} = 7 \times 10^{-14}$ m/W from Fig. 12.4, gives the criterion

$$[NP][N - 1][\Delta\nu_s] L_{eff} < 5 \times 10^3 \text{ mW} \cdot \text{THz} \cdot \text{km} \tag{12.3}$$

Here, NP is the total power coupled into the fiber, $(N - 1) \Delta\nu_s$ is the total occupied optical bandwidth, and L_{eff} is the effective length, which takes into account absorption along the length of the fiber.
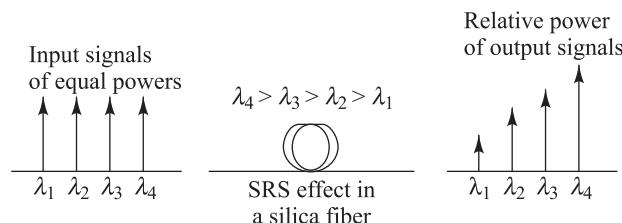
Fig. 12.3 *SRS transfers optical power from shorter wavelengths to longer wavelengths*
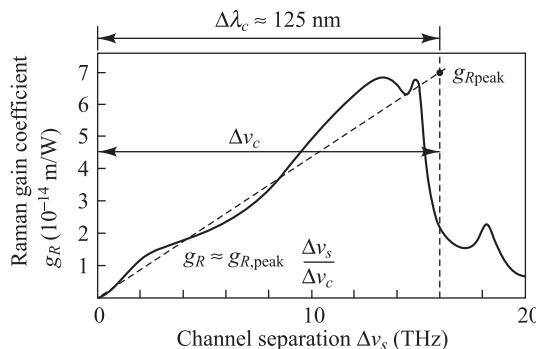
Fig. 12.4 *Simplified linear approximation for the Raman gain coefficient as a function of channel spacing. The actual function has some fluctuations about this approximation and a low-value decaying tail of less than 0.5×10^{-8} µm/W at channel separations greater than 16 THz*

Example 12.1 The limits indicated by Eq. (12.3) are illustrated in Fig. 12.5 for systems with four and eight wavelength channels. The curves show the maximum power per channel as a function of the number of wavelengths for three different channel spacings (recall that a 125-GHz frequency spacing is equivalent to a 1-nm wavelength spacing at 1550 nm), a fiber attenuation of 0.2 dB/km (or, equivalently, 4.61×10^{-2} km^{-1}), and an amplifier spacing of 75 km (which yields an effective length of $L_{\text{eff}} = 22$ km).

The results in Fig. 12.5 were calculated for the worst-case scenario caused by SRS. In general, if the optical power per channel is not excessively high (e.g., less than 1 mW each), then the effects of SRS do not contribute significantly to the eye-closure penalty as a function of transmission distance.[1,6]
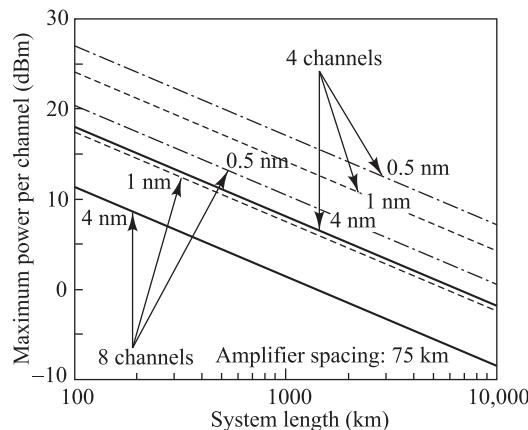
Fig. 12.5 *Maximum allowable power per wavelength channel versus transmission length for three different channel spacings. The curves are for the power levels that ensure an SRS degradation of less than 1 dB for all channels. (Reproduced with permission from O'Mahoney, Simeinidou, Yu, and Zhou,* J. Lightwave Tech., *vol. 13, pp. 817–828, © IEEE, May 1995.)*

12.4 Stimulated Brillouin Scattering

Stimulated Brillouin scattering arises when a strong optical signal generates an acoustic wave that produces variations in the refractive index.[1,3,4–10] These index variations cause lightwaves to scatter in the backward direction toward the transmitter. This backscattered light experiences gain from the forward-propagating signals, which leads to depletion of the signal power. The frequency of the scattered light experiences a Doppler shift given by

$$v_B = 2nV_s/\lambda \tag{12.4}$$

where n is the index of refraction and V_s is the velocity of sound in the material. In silica, this interaction occurs over a very narrow *Brillouin linewidth* of $\Delta v_B = 20$ MHz at 1550 nm. For $V_s = 5760$ m/s in fused silica, the frequency of the backward-propagating light at 1550 nm is downshifted by 11 GHz (0.09 nm) from the original signal. This shows that the SBS effect is confined within a single wavelength channel in a WDM system. Thus, the effects of SBS accumulate individually for each channel, and, consequently, occur at the same power level in each WDM channel, analogous to a single-channel system.

System impairment starts when the amplitude of the scattered wave is comparable to the signal power. For typical fibers, the threshold power for this process is around 10 mW for single-fiber spans. In a long fiber chain containing optical amplifiers, there are normally optical isolators to prevent backscattered signals from entering the amplifier. Consequently, the impairment due to SBS is limited to the degradation occurring in a single amplifier-to-amplifier span.

One criterion for determining at what point SBS becomes a problem is to consider the SBS threshold power P_{th}. This is defined to be the signal power at which the backscattered light equals the fiber-input power. The calculation of this expression is rather complicated, but an approximation is given by[4]

$$P_{th} \approx 21 \frac{A_{eff} b}{g_B L_{eff}} \left(1 + \frac{\Delta v_{source}}{\Delta v_B}\right) \tag{12.5}$$

Here, A_{eff} is the effective cross-sectional area of the propagating wave, Δv_{source} is the source linewidth, and the polarization factor b lies between 1 and 2 depending on the relative polarizations of the pump and Stokes waves. The effective length L_{eff} is given in Eq. (12.1) and g_B is the *Brillouin gain coefficient*, which is approximately 4×10^{-11} m/W, independent of the wavelength. Equation (12.5) shows that the SBS threshold power increases as the source linewidth becomes larger.

Example 12.2 Consider an optical source with a 40-MHz linewidth. Using the values $\Delta v_B = 20$ MHz at 1550 nm, $A_{eff} = 55 \times 10^{-12}$ m^2 (for a typical dispersion-shifted single-mode fiber), $L_{eff} = 20$ km, and assuming a value of $b = 2$, then from Eq. (12.5) we have $P_{th} = 8.6$ mW = 9.3 dBm.

Figure 12.6 illustrates the effect of SBS on unmodulated signal power once the threshold is reached. The plots give the relative Brillouin scattered power and the signal power transmitted through a fiber as a function of the input power. Below a certain signal level called the *SBS threshold*, the transmitted power increases linearly with the input level. The effect of SBS is negligible for these low power levels but becomes greater as the optical power level increases.

At the SBS threshold, the SBS process becomes nonlinear and the launched signal loses an increasingly greater percentage of its power as the signal strength becomes larger. Beyond the SBS threshold, the
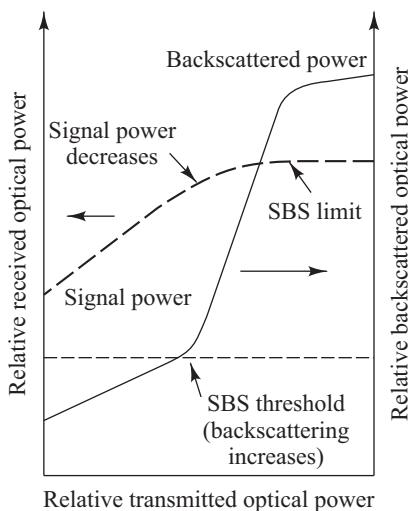
Fig. 12.6 *The effect of SBS on signal power in an optical fiber*

percentage increase in signal depletion continues to grow with signal strength until the *SBS limit* is reached. Any additional optical power launched into the fiber after this point merely is scattered backward along the fiber due to the SBS effect. Thus above the SBS limit the transmitted power remains constant for higher inputs, since all the added power is extracted from the signal to feed the scattered wave.

In standard G.652 single-mode fiber, the SBS limit restricts the maximum launched optical power to be 17 dBm. As an example of efforts to mitigate this limit, Corning Incorporated designed a G.652-compatible fiber with a 3-dB higher SBS threshold. This new design enables twice as much power (an additional 3 dB) to be launched into the fiber.

Figure 12.7 illustrates the SBS-induced impairment on the carrier-to-noise ratio (CNR) of an amplitude-modulated vestigial-sideband (AM-VSB) video signal for the same fiber as in Fig. 12.6. Here, the CNR grows with increasing fiber-injected power up to the SBS threshold. Beyond this point the CNR starts to decrease.[11]

Several schemes are available for reducing the power-penalty effects of SBS. These include[1,3,4,8]

- Keeping the optical power per WDM channel below the SBS thresholds. For long-haul systems, this may require a reduction in the amplifier spacing.
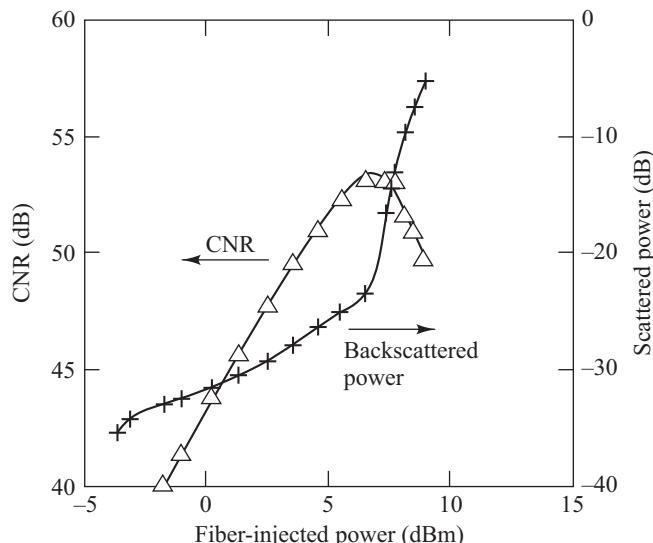
Fig. 12.7 *The SBS impairment on the CNR of an AM-VSB signal. The triangles are the CNR and the crosses represent the backscattered power. (Adapted with permission from Mao, Bodeep, Tkach, Chraplyvy, Darcie, and Dorosier,[11] © IEEE, 1992.)*

- Increasing the linewidth of the source as the gain bandwidth of SBS is very small. This can be achieved through direct modulation of the source (as opposed to external modulation), since this causes the linewidth to broaden because of chirping effects. However, a large dispersion penalty may result from this.
- Slightly dithering the laser output in frequency, at roughly 100 to 200 MHz to raise the Brillouin threshold. This is effective since SBS is a narrowband process. The dither frequency should scale as the ratio of the injected power to the SBS threshold.

12.5 Self-Phase Modulation

The refractive index n of many optical materials has a weak dependence on *optical intensity I* (equal to the optical power per effective area in the fiber) given by

$$n = n_0 + n_2 I = n_0 + n_2 \frac{P}{A_{\text{eff}}}$$ (12.6)

where n_0 is the *ordinary refractive index* of the material and n_2 is the *nonlinear index coefficient*. The factor n_2 is about 2.6×10^{-8} μm^2/W in silica, between 1.2 and 5.1×10^{-6} μm^2/W in tellurite glasses, and 2.4×10^{-5} μm^2/W in $As_{40}Se_{60}$ chalcogenide glass. The nonlinearity in the refractive index is known as the *Kerr nonlinearity*.[12] This nonlinearity produces a carrier-induced phase modulation of the propagating signal, which is called the *Kerr effect*. In single-wavelength links, this gives rise to *self-phase modulation* (SPM), which converts optical power fluctuations in a propagating lightwave to spurious phase fluctuations in the same wave.[3–5,13]

The main parameter γ, which indicates the magnitude of the nonlinear effect for SPM, is given by

$$\gamma = \frac{2\pi}{\lambda} \frac{n_2}{A_{\text{eff}}}$$ (12.7)

where λ is the free-space wavelength and A_{eff} is the effective core area. The value of γ ranges from 1 to 5 $\text{W}^{-1}\,\text{km}^{-1}$ in silica depending on the fiber type and the wavelength. For example, $\gamma = 1.3$ $\text{W}^{-1}\,\text{km}^{-1}$ at 1550 nm for a standard single-mode fiber that has an effective area equal to 72 μm^2. The frequency shift $\Delta\varphi$ arising from SPM is given by

$$\Delta\varphi = \frac{d\varphi}{dt} = \gamma L_{\text{eff}} \frac{dP}{dt}$$ (12.8)

Here L_{eff} is the effective length given by Eq. (12.1) and dP/dt is the derivative of the optical pulse power; that is, it shows that the frequency shift occurs when the optical pulse power is changing in time.

To see the effect of SPM, consider what happens to the optical pulse shown in Fig. 12.8 as it propagates in a fiber. Here the time axis is normalized to the parameter t_0, which is the pulse half-width at the $1/e$-intensity point. The edges of the pulse represent a time-varying intensity, which rises rapidly from zero to a maximum value, and then returns to zero. In a medium having an intensity-dependent refractive index, a time-varying signal intensity will produce a time-varying refractive index. Thus the index at the peak of the pulse will be slightly different than the value in the wings of the pulse. The leading edge will see a positive dn/dt, whereas the trailing edge will see a negative dn/dt.
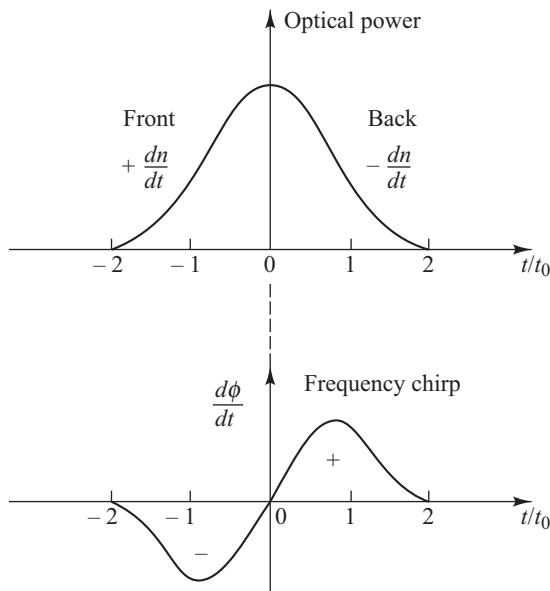
Fig. 12.8 *Phenomenological description of spectral broadening of a pulse due to self-phase modulation*

This temporally varying index change results in a temporally varying phase change, shown by $d\varphi/dt$ in Fig. 12.8. The consequence is that the instantaneous optical frequency differs from its initial value ν_0 across the pulse. That is, since the phase fluctuations are intensity-dependent, different parts of the pulse undergo different phase shifts. This leads to what is known as *frequency chirping*, in that the rising edge of the pulse experiences a red shift in frequency (toward lower frequencies or longer wavelengths), whereas the trailing edge of the pulse experiences a blue shift in frequency (toward higher frequencies or shorter wavelengths). Since the degree of chirping depends on the transmitted power, SPM effects are more pronounced for higher-intensity pulses.

For some types of fibers, the time-varying phase may result in a power penalty owing to a GVD-induced spectral broadening of the pulse as it travels along the fiber. In the *normal dispersion region* the chromatic dispersion is negative [that is, from Eq. (3.25) we have $\beta_2 > 0$] and the group delay decreases with wavelength. This means that since red light has a longer wavelength than blue, the red light travels faster in silica because $n_{\text{red}} < n_{\text{blue}}$ (see Fig. 3.12). Therefore, in the normal dispersion region the red-shifted leading edge of the pulse travels faster and thus moves away from the center of the pulse. At the same time the blue-shifted trailing edge travels slower, and thus also moves away from the center of the pulse. In this case chirping worsens the effects of GVD-induced pulse broadening. On the other hand, in the *anomalous dispersion region* where chromatic dispersion is positive so that the group delay increases with wavelength, the red-shifted leading edge of the pulse travels slower and thus moves toward the center of the pulse. Similarly, the blue-shifted trailing edge travels faster, and also moves toward the center of the pulse. In this case, SPM causes the pulse to narrow, thereby partly compensating for chromatic dispersion.

12.6 Cross-Phase Modulation

Cross-phase modulation (XPM) appears in WDM systems and has a similar origin as SPM. In this case the name derives from the fact that the refractive index nonlinearity converts optical intensity fluctuations in a particular wavelength channel to phase fluctuations in another copropagating channel.[1–4] In addition, since the refractive index seen by a particular wavelength is influenced by both the optical intensity of that wave itself and also by the optical power fluctuations of neighboring wavelengths, SPM is always present when XPM occurs. Analogous to SPM, for two interacting wavelengths the XPM-induced frequency shift $\Delta\varphi$ is given by

$$\Delta\varphi = \frac{d\varphi}{dt} = 2\gamma L_{\text{eff}} \frac{dP}{dt} \tag{12.9}$$

where the parameters are the same as for Eq. (12.8). When multiple wavelengths propagate in a fiber, the total phase shift for an optical signal with frequency ω_i is

$$\Delta\varphi_i = \gamma L_{\text{eff}} \left[\frac{dP_i}{dt} + 2\sum_{j \neq i} \frac{dP_j}{dt} \right] \qquad (12.10)$$

The first term in square brackets represents the SPM contribution and the second term arises from XPM. The factor 2 in the bracketed expression shows that the weight of XPM is twice that of SPM. However, XPM only appears when the two interacting light beams or pulses overlap in space and time. In general, pulses from two different wavelength channels will not remain superimposed because each has a different GVD. This greatly reduces the impact of XPM for direct-detection optical fiber transmission systems. However, XPM could be a problem for high-rate ultra-dense WDM systems (for example, 2.5- or 10-Gb/s systems with a wavelength spacing of 25 GHz or less) and when coherent detection schemes are used.[2,3]

12.7 Four-Wave Mixing

Dense WDM transmission in which individual wavelength channels are modulated at rates of 10 Gb/s offers capacities of $N \times 10$ Gb/s, where N is the number of wavelengths. To transmit such high capacities over long distances requires operation in the 1550-nm window of dispersion-shifted fiber. In addition, to preserve an adequate signal-to-noise ratio, a 10-Gb/s system operating over long distances and having nominal optical repeater spacings of 100 km needs optical launch powers of around 1 mW per channel. For such WDM systems, the simultaneous requirements of high launch power and low dispersion give rise to the generation of new frequencies due to four-wave mixing.[1–4,14–16]

Four-wave mixing (FWM) is a third-order nonlinearity in optical fibers that is analogous to intermodulation distortion in electrical systems. When wavelength channels are located near the zero-dispersion point, three optical frequencies (v_i, v_j, v_k) will mix to produce a fourth intermodulation product v_{ijk} given by

$$v_{ijk} = v_i + v_j - v_k \qquad \text{with} \qquad i,\ j \neq k \qquad (12.11)$$

When this new frequency falls in the transmission window of the original frequencies, it can cause severe crosstalk.

Figure 12.9 shows a simple example for two waves at frequencies v_1 and v_2. As these waves copropagate along a fiber, they mix and generate sidebands at $2v_1 - v_2$ and $2v_2 - v_1$. Similarly, three copropagating waves will create nine new optical sideband waves at frequencies given by Eq. (12.11). These sidebands will travel along with the original waves and will grow at the expense of signal-strength depletion. In general, for N wavelengths launched into a fiber, the number of generated mixing products M is

$$M = \frac{N^2}{2}(N-1) \qquad (12.12)$$
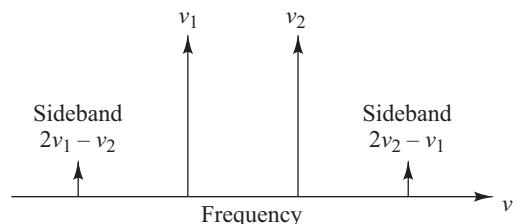
Fig. 12.9 *Two optical waves at frequencies v_1 and v_2 mix to generate two third-order sidebands*

If the channels are equally spaced, a number of the new waves will have the same frequencies as the injected signals. Thus the resultant crosstalk interference plus the depletion of the original signal waves can severely degrade multichannel system performance unless steps are taken to diminish it.

The efficiency of four-wave mixing depends on fiber dispersion and the channel spacings. Since the dispersion varies with wavelength, the signal waves and the generated waves have different group velocities. This destroys the phase matching of the interacting waves and lowers the efficiency at which power is transferred to newly generated frequencies. The higher the group velocity mismatches and the wider the channel spacings, the lower the four-wave mixing.

At the exit of a fiber of length L and attenuation α, the power P_{ijk} that is generated at frequency ν_{ijk} due to the interaction of signals at frequencies ν_i, ν_j, and ν_k that have fiber-input powers P_i, P_j, and P_k, respectively, is

$$P_{ijk}(L) = \eta(\mathscr{D}\kappa)^2 P_i(0) P_j(0) P_k(0) \exp(-\alpha L) \tag{12.13}$$

where the *nonlinear interaction constant κ* is

$$\kappa = \frac{32\pi^3 \chi_{1111}}{n_2 \lambda c}\left(\frac{L_{\text{eff}}}{A_{\text{eff}}}\right) \tag{12.14}$$

Here, χ_{1111} is the third-order nonlinear susceptibility; η is the efficiency of the four-wave mixing; n is the fiber refractive index; and $\mathscr{D}$ is the degeneracy factor, which has the value of 3 or 6 for two waves mixing or three waves mixing, respectively. The effective length L_{eff} is given by Eq. (12.1) and A_{eff} is the effective cross-sectional area of the fiber. Figure 12.10 gives examples of η as a function of channel spacing for three equally spaced frequencies for dispersion values of a conventional G.652 single-mode fiber [16 ps/(nm · km) average in the 1550-nm window] and a dispersion-shifted G.653 fiber [1 ps/(nm · km) average in the 1550-nm window]. These curves show the frequency-spacing range over which the FWM process is efficient for these two dispersion values (see Prob.12.5 for a detailed expression for η, which leads to an oscillatory behavior of P_{ijk} as a function of channel spacing). For example, in the conventional single-mode fiber, only frequencies with separations less than 20 GHz will mix efficiently. In contrast, the FWM mixing efficiencies are greater than 20 percent for channel separations up to 50 GHz for G.653 dispersion-shifted fibers.

Example 12.3 Consider a 75-km link of dispersion-shifted single-mode fiber carrying two wavelengths at 1540.0 and 1540.5 nm. The new frequencies generated due to FWM are at

$$\nu_{112} = 2\nu_1 - \nu_2 = 2(1540.0 \text{ nm}) - 1540.5 \text{ nm}$$
$$= 1539.5 \text{ nm}$$

and

$$\nu_{221} = 2\nu_2 - \nu_1 = 2(1540.5 \text{ nm}) - 1540.0 \text{ mm}$$
$$= 1541.0 \text{ nm}$$

Assume the fiber has an attenuation of $\alpha = 0.20$ dB/km = 0.0461 km^{-1}, a refractive index of 1.48, and a 9.0-μm core diameter, so that $L_{\text{eff}} = 22$ km and $A_{\text{eff}} = 6.4 \times 10^{-11}$ m^2. From Fig. 12.10 we find $\eta \approx 5$ percent for a 62-GHz (0.5-nm) channel spacing. If each channel has an input power of 1 mW, then, using the values $\chi_{1111} = 6 \times 10^{-16}$ cm^3/erg = 6×10^{-15} m^3/(W · s) and $\mathscr{D} = 3$, we find

$$P_{112} = 0.05(3)^2 \left[\frac{32\pi^3 6 \times 10^{-15}\ \dfrac{\text{m}^3}{\text{W}\cdot\text{s}}}{(1.48)^2 (1.54 \times 10^{-6}\text{m}) 3 \times 10^8 \text{ m/s}}\right]^2$$

$$\times \left(\frac{22 \times 10^3 \text{m}}{6.4 \times 10^{-11}\text{m}^2}\right)^2 (1.0 \times 10^{-3}\text{W})^3$$

$$\times \exp\left[-(0.0461/\text{km})75\text{ km}\right]$$

$$= 5.80 \times 10^{-8} \text{mW}$$

12.8 FWM Mitigation

To reduce the effects of four-wave mixing in a DWDM link with close channel separations, for example, 100-GHz or less, it is important to have a high value of chromatic dispersion throughout the link. The reason is due to the fact that the efficiency of FWM depends on matching the phase relationship between the interacting DWDM signals. When there is chromatic dispersion in the fiber, the signals at different wavelengths travel with different group velocities. This means that the propagating waves move in and out of phase with each other. This condition greatly diminishes the FWM efficiency.

If the chromatic dispersion is low, or if there are regions of both positive and negative dispersion in the DWDM operating band, then a large number of FWM terms can be generated by the DWDM signals. This effect is particularly troublesome if G.653 dispersion-shifted fibers are used for

Fig. 12.10 *Efficiencies of four-wave mixing as a function of wavelength channel spacing. The solid curve is for standard single-mode fiber with 16 ps/(nm·km) dispersion. The dashed curve is for dispersion-shifted fiber with 1-ps/(nm-km) dispersion. (Reproduced with permission from Chraplyvy,[5] © IEEE, 1990.)*

DWDM applications in the C-band. The main problem is that these fibers have both positive and negative dispersion regions around the zero-dispersion point at 1550 nm, as shown in Fig. 12.11. The consequence is
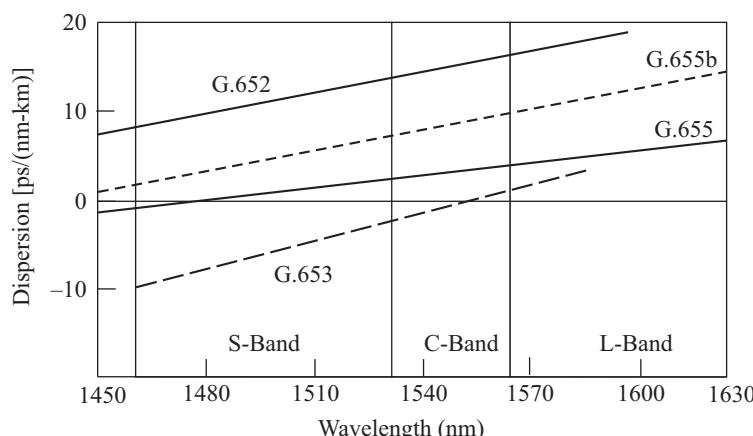
Fig. 12.11 *Chromatic dispersion as a function of wavelength in various spectral bands for several different standard fiber types*

that the DWDM channels on either side of this zero-dispersion point generate a large number of interfering in-band signals.

For standard G.652 single-mode fibers the high chromatic dispersion value of about 17 ps/(nm · km) in the C-band effectively suppresses four-wave mixing. However, for high data rates, such as 2.5 or 10 Gb/s, this high chromatic dispersion value quickly leads to large pulse-spreading effects along a fiber.

The limitations of the G.652 and G.653 fibers with respect to suppressing FWM led to the development of the G.655 fiber. As Fig. 12.11 shows, the G.655 fiber has a chromatic dispersion value ranging from about 3 to 9 ps/(nm · km) in the C-band. The ITU-T Recommendation G.655 specifies several different versions of this fiber. These include G.655B for both S- and C-band use, G.655.D low-dispersion fibers with a chromatic dispersion ranging from 2.80 to 6.2 ps/(nm · km) at 1550 nm, and G.655.E medium-dispersion fibers with a dispersion value ranging from 6.06 to 9.31 ps/(nm · km) at 1550 nm. In either case, these chromatic dispersion values are sufficient to suppress FWM effects.

A consequence of using fibers such as the G.655 type is that a large accumulation of chromatic dispersion may build up along a link. Chapter 13 describes some dispersion compensation techniques that may be used for reducing this pulse spreading effect while still maintaining sufficient FWM reduction.[17–21]

12.9 | Wavelength Converters

One beneficial application of cross-phase modulation and four-wave mixing techniques is for performing wavelength conversion in WDM networks. An optical wavelength converter is a device that can translate information on an incoming wavelength directly to a new wavelength without entering the electrical domain. Such a device is an important component in all-optical networks, since the wavelength of the incoming signal may already be in use by another information channel residing on the destined outgoing path. Converting the incoming signal to a new wavelength will allow both information channels to traverse the same outbound fiber simultaneously. This section describes two classes of wavelength converters with one example from each class.

12.9.1 Optical-Gating Wavelength Converters

A wide variety of optical-gating techniques using devices such as semiconductor optical amplifiers, semiconductor lasers, or nonlinear optical-loop mirrors have been investigated to achieve wavelength conversion. The use of a semiconductor optical amplifier (SOA) in a cross-phase modulation (XPM) mode has been one of the most successful techniques for implementing single-wavelength conversion.[22–27] The configurations for implementing this scheme include the Mach-Zehnder or the Michelson interferometer setups shown in Fig. 12.12.

The XPM scheme relies on the dependency of the refractive index on the carrier density in the active region of the SOA. As depicted in Fig. 12.12, the basic concept is that an incoming information-carrying signal at wavelength λ_s and a continuous-wave (CW) signal at the desired new wavelength λ_c (called the *probe beam*) are simultaneously coupled into the device. The two waves can be either copropagating or counterpropagating. However, the noise in the latter case is higher.[28,29] The signal beam modulates the gain of the SOA by depleting the carriers, which produces a modulation of the refractive index. When the CW beam encounters the modulated gain and refractive index, its amplitude and phase are changed, so that it now carries the same information as the input signal. As shown in Fig. 12.12, the SOAs are placed in an asymmetric configuration so that the phase change in the two amplifiers is different. Consequently, the CW light is modulated according to the phase difference. A typical splitting ratio is 69/31 percent. These types of converters readily handle data rates of at least 10 Gb/s.

(a) Mach-Zehnder interferometer

(b) Michelson interferometer
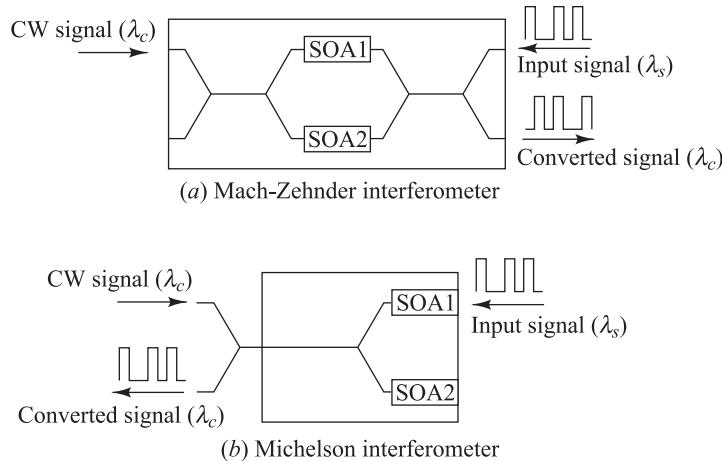
Fig. 12.12   *(a) Mach-Zehnder interferometer and (b) Michelson interferometer setups using a pair of SOAs for implementing the cross-phase modulation wavelength-conversion scheme*

A limitation of the XPM architecture is that it only converts one wavelength at a time, In addition, it has limited transparency in terms of the data format. Any information that is in the form of phase, frequency, or analog amplitude is lost during the wavelength conversion process. Consequently, this scheme is restricted to converting digital signal streams.

12.9.2   Wave-Mixing Wavelength Converters

Wavelength conversion based on nonlinear optical wave mixing offers important advantages compared to other conversion methods.[22,30–34] The advantages include a multi-wavelength conversion capability and transparency to the modulation format. The mixing arises from nonlinear interactions among optical waves traversing a nonlinear material. The outcome is the generation of another wave whose intensity is proportional to the product of the intensities of the interacting waves. The phase and frequency of the generated wave are a linear combination of these parameters of the interacting waves. Therefore, the wave mixing preserves both amplitude and phase information, and consequently is the only wavelength-conversion category that offers strict transparency to the modulation format.

Two successful wavelength-conversion schemes are four-wave mixing in either a passive waveguide or SOA, and difference-frequency generation in waveguides. For wavelength conversion, the FWM scheme employs the mixing of three distinct input waves to generate a fourth distinct output wave. In this method, an intensity pattern resulting from two input waves interacting in a nonlinear material forms a grating. For example, in SOAs there are three physical mechanisms that can form a grating. These are carrier-density modulation, dynamic carrier heating, and spectral hole burning. The third input wave in the material gets scattered by this grating, thereby generating an output wave. The frequency of the generated output wave is offset from that of the third wave by the frequency difference between the first two waves. If one of the three incident waves contains amplitude, phase, or frequency information and the other two waves are constant, then the generated wave will contain the same information.

Difference-frequency generation in waveguides is based on the mixing of two input waves. In this case, the nonlinear interaction of the material is with a pump and a signal wave. Figure 12.13 gives an
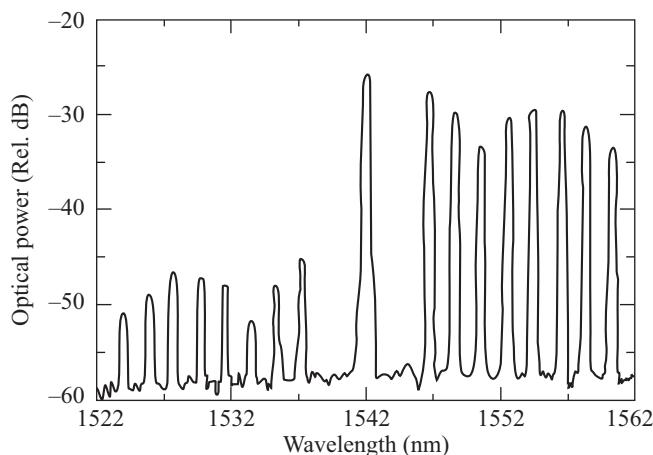
Fig. 12.13 *Simultaneous conversion of eight input wavelengths (1546, 1548, 1550, 1552, 1554, 1556, 1558, 1560 nm) to a set of eight output wavelengths (1538, 1536, 1534, 1532, 1530, 1528, 1526, 1524 nm) using difference-frequency generation. The 1542-nm spike is a second-order spectrometer response to the pump wave at 771 nm. (Reproduced with permission from Yoo, Rajhel, Caneau, Bhat, and Koza,* OFC Technical Digest, © *1997, OSA.)*

example of the simultaneous conversion of eight input wavelengths in the 1546-to-1560-nm region to a set of eight output wavelengths in the 1524-to-1538-nm region.[25]

12.10 Solitons

As Chapter 3 describes, group velocity dispersion (GVD) causes most pulses to broaden in time as they propagate through an optical fiber. However, a particular pulse shape known as a *soliton* takes advantage of nonlinear effects in silica, particularly self-phase modulation (SPM) resulting from the Kerr nonlinearity, to overcome the pulse-broadening effects of GVD.[35–43]

The term "soliton" refers to special kinds of waves that can propagate undistorted over long distances and remain unaffected after collisions with each other. John Scott Russell made the first recorded observation of a soliton in 1838, when he saw a peculiar type of wave generated by boats in narrow Scottish canals.[35] The resulting water wave was of great height and traveled rapidly and unattenuated over a long distance. After passing through slower waves of lesser height, the waves emerged from the interaction undistorted, with their identities unchanged.

In an optical communication system, solitons are very narrow, high-intensity optical pulses that retain their shape through the interaction of balancing pulse dispersion with the nonlinear properties of an optical fiber. If the relative effects of SPM and GVD are controlled just right, and the appropriate pulse shape is chosen, the pulse compression resulting from SPM can exactly offset the pulse broadening effect of GVD. Depending on the particular shape chosen, the pulse either does not change its shape as it propagates, or it undergoes periodically repeating changes in shape. The family of pulses that do not change in shape are called *fundamental solitons*, and those that undergo periodic shape changes are called *higher-order solitons.* In either case, attenuation in

the fiber will eventually decrease the soliton energy. Since this weakens the nonlinear interaction needed to counteract GVD, periodically spaced optical amplifiers are required in a soliton link to restore the pulse energy.

12.10.1 Soliton Pulses

Let us look at the soliton pulse features in more detail. No optical pulse is monochromatic, since it excites a spectrum of frequencies. For example, as Eq. (10.1) shows, if an optical source emits power in a wavelength band $\Delta\lambda$, its spectral spread is Δv. This is important, because in an actual fiber a pulse is affected by both the GVD and the Kerr nonlinearity. This is particularly significant for high-intensity optical excitations. Since the medium is dispersive, the pulse width will spread in time with increasing distance along the fiber owing to GVD. In addition, when a high-intensity optical pulse is coupled to a fiber, the optical power modulates the refractive index seen by the optical excitation. This induces phase fluctuations in the propagating wave, thereby producing a chirping effect in the pulse, as shown in Fig. 12.8. The result is that the front of the pulse (at smaller times) has lower frequencies and the back of the pulse (at later times) has higher frequencies than the carrier frequency.

When such a pulse traverses a medium with a positive GVD parameter β_2 for the constituent frequencies, the leading part of the pulse is shifted toward a longer wavelength (lower frequencies), so that the speed in that portion increases. Conversely, in the trailing half, the frequency rises so the speed decreases. This causes the trailing edge to be further delayed. Consequently, in addition to a spectral change with distance, the energy in the center of the pulse is dispersed to either side, and the pulse eventually takes on a rectangular-wave shape. Figure 12.14 illustrates these intensity changes as the pulse travels along such a fiber. The plot is in terms of the normalized time. These effects will severely limit high-speed long-distance transmission if the system is operated in this condition.

On the other hand, when a narrow high-intensity pulse traverses a medium with a negative GVD parameter for the constituent frequencies, GVD counteracts the chirp produced by SPM. Now, GVD retards the low frequencies in the front end of the pulse and advances the high frequencies at the back.
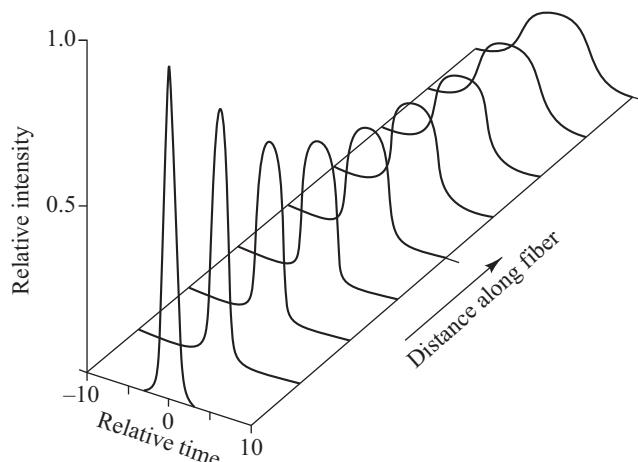
Fig. 12.14 *Temporal changes in a narrow high-intensity pulse that is subjected to the Kerr effect as it travels through a nonlinear dispersive fiber that has a positive GVD parameter*
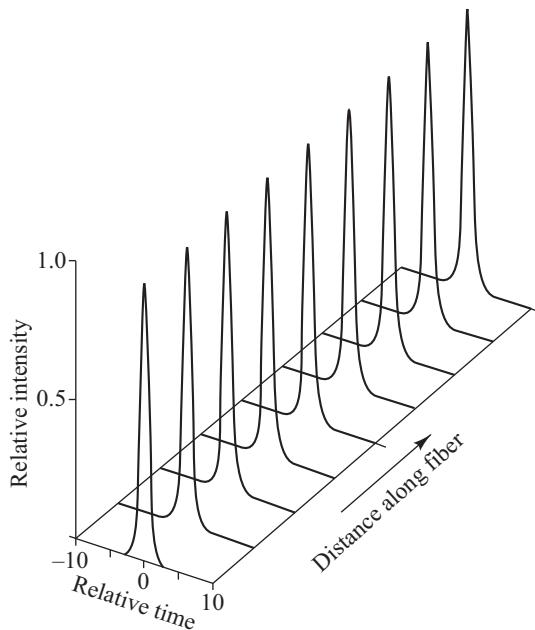
Fig. 12.15 *Characteristics of a high-intensity sharply peaked soliton pulse that is subjected to the Kerr effect as it travels through a nonlinear dispersive fiber that has a negative GVD parameter*

The result is that the high-intensity sharply peaked soliton pulse changes neither its shape nor its spectrum as it travels along the fiber. Figure 12.15 illustrates this for a fundamental soliton. Provided the pulse energy is sufficiently strong, this pulse shape is maintained as it travels along the fiber. In a standard G.652 optical fiber, there is a zero-dispersion point around 1320 nm (see Fig. 3.20). For wavelengths shorter than 1320 nm β_2 is positive, and for longer wavelengths it is negative. Thus soliton operation is limited to the region greater than 1320 nm for these types of fibers.

To derive the evolution of the pulse shape required for soliton transmission, one needs to consider the *nonlinear Schrödinger (NLS) equation.*

$$-j \frac{\partial u}{\partial z} = \frac{1}{2} \frac{\partial^2 u}{\partial t^2} + N^2 |u|^2 u - j(\alpha/2) u \qquad (12.15)$$

Here, $u(z, t)$ is the pulse envelope function, z is the propagation distance along the fiber, N is an integer designating the *order* of the soliton, and α is the coefficient of energy gain per unit length, with negative values of α representing energy loss. Following conventional notation, the parameters in Eq. (12.15) have been expressed in special soliton units to eliminate scaling constants in the equation.

These parameters (defined in Sec.12.10.2) are the *normalized time T_0*, the *dispersion length L_{disp}*, and the *soliton peak power P_{peak}*.

For the three right-hand terms in Eq. (12.15):

    1. The first term represents GVD effects of the fiber. Acting by itself, dispersion tends to broaden pulses in time.

2. The second nonlinear term denotes the fact that the refractive index of the fiber depends on the light intensity. Through the self-modulation process, this physical phenomenon broadens the frequency spectrum of a pulse.

3. The third term represents the effects of energy loss or gain; for example, due to fiber attenuation or optical amplification, respectively.

Solving the NLS equation analytically yields a pulse envelop that is either independent of z (for the fundamental soliton with $N = 1$) or that is periodic in z (for higher-order solitons with $N \geq 2$). The general theory of solitons is mathematically complex and can be found in the literature.[40–43] Here, we present the basic concepts for fundamental solitons. The solution to Eq. (12.15) for the fundamental soliton is given by

$$u(z, t) = \text{sech}(t) \exp(jz/2) \qquad (12.16)$$

Fig. 12.16 *The hyperbolic secant function used for soliton pulses. The time scale is given in units normalized to the 1/e width of the pulse*

where $\text{sech}(t)$ is the hyperbolic secant function. This is a bell-shaped pulse, as Fig. l2.16 illustrates. The time scale is given in units normalized to the l/e width of the pulse. Since the phase term $\exp(jz/2)$ in Eq. (12.16) has no influence on the shape of the pulse, the soliton is independent of z and hence is nondispersive in the time domain.

When examining the NLS equation, one finds that the first-order effects of the dispersive and nonlinear terms are just complementary phase shifts. For a pulse given by Eq. (12.16), these phase shifts are

$$d\phi_{\text{nonlin}} = |u(t)|^2 \, dz = \text{sech}^2(t) \, dz \qquad (12.17)$$

for the nonlinear process, and

$$d\phi_{\text{disp}} = \left(\frac{1}{2u} \frac{\partial^2 u}{\partial t^2} \right) dz = \left[\frac{1}{2} - \text{sech}^2(t) \right] dz \quad (12.18)$$

for the dispersion effect. Figure 12.17 shows plots of these terms and their sum, which is a constant. Upon integration, the sum simply yields a phase shift of $z/2$, which is common to the entire pulse. Since such a phase shift changes neither the temporal nor the spectral shape of a pulse, the soliton remains completely nondispersive in both the temporal and frequency domains.

Fig. 12.17 *Dispersive and nonlinear phase shifts of a soliton pulse. Their sum is a constant, which yields a common phase shift for the entire pulse.*

12.10.2 Soliton Parameters

Recall that the full-width half-maximum (FWHM) of a pulse is defined as the full width of the pulse at its *half-maximum power level* (see Fig. 12.18).
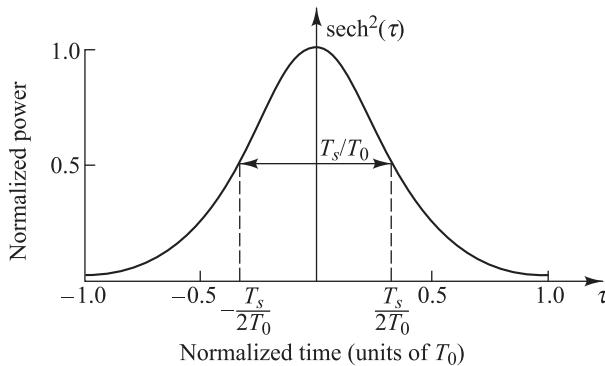
Fig. 12.18 *Definition of the half-maximum soliton width in terms of normalized time units*

For the solution to Eq. (12.15), the power is given by the square of the envelope function in Eq. (12.16). Thus the FWHM T_s of the fundamental soliton pulse in normalized time is found from the relationship $\text{sech}^2(\tau) = \frac{1}{2}$ with $\tau = T_s/(2T_0)$, where T_0 is the basic normalized time unit. This yields

$$T_0 = \frac{T_s}{2\cosh^{-1}\sqrt{2}} = \frac{T_s}{1.7627} \approx 0.567T_s$$

$$(12.19)$$

Example 12.4 Typical soliton FWHM pulse widths T_s range from 15 to 50 ps, so that the normalized time T_0 is on the order of 9–30 ps.

The *normalized distance* parameter (also called *dispersion length*) L_{disp} is a characteristic length for the effects of the dispersion term. As described later, L_{disp} is a measure of the period of a soliton. This parameter is given by

$$L_{\text{disp}} = \frac{2\pi c}{\lambda^2}\frac{T_0^2}{D} = \frac{1}{\left[2\cosh^{-1}\sqrt{2}\right]^2}\frac{2\pi c}{\lambda^2}\frac{T_s^2}{D} \approx 0.322\frac{2\pi c}{\lambda^2}\frac{T_s^2}{D}$$

$$(12.20)$$

where c is the speed of light, λ is the wavelength in vacuum, and D is the dispersion of the fiber.

Example 12.5 Consider a dispersion-shifted fiber having $D = 0.5$ ps/(nm · km) at 1550 nm. If $T_s = 20$ ps, we have

$$L_{\text{disp}} = \frac{1}{(1.7627)^2}\frac{2\pi(3\times10^8\text{ m/s})}{(1550\text{ nm})^2}\frac{(20\text{ ps})^2}{0.5\text{ ps}/(\text{nm}\cdot\text{km})}$$

$$= 202\text{ km}$$

which shows that L_{disp} is on the order of hundreds of kilometers.

The parameter P_{peak} is the *soliton peak power* and is given by

$$P_{\text{peak}} = \frac{A_{\text{eff}}}{2\pi n_2}\frac{\lambda}{L_{\text{disp}}} = \left(\frac{1.7627}{2\pi}\right)^2\frac{A_{\text{eff}}\lambda^3}{n_2 c}\frac{D}{T_s^2}$$

$$(12.21)$$

where A_{eff} is the effective area of the fiber core, n_2 is the nonlinear intensity-dependent refractive-index coefficient [see Eq. (12.6)], and L_{disp} is measured in km.
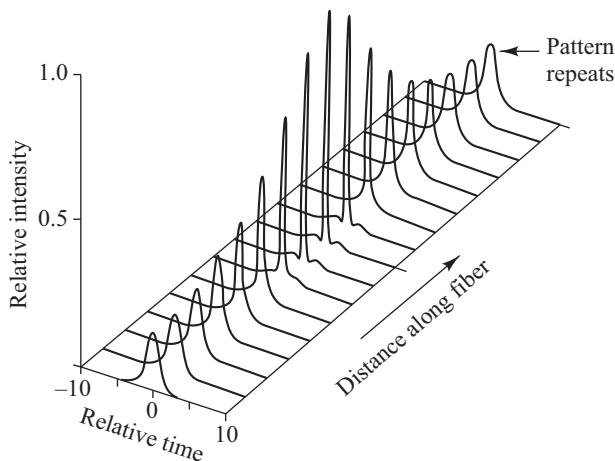
Fig. 12.19 *Propagation characteristics of a second-order soliton (N = 2)*

Example 12.6 For $\lambda = 1550$ nm, $A_{\text{eff}} = 50$ μm^2, $n_2 = 2.6 \times 10^{-16}$ cm^2/W, and with the value of $L_{\text{disp}} = 202$ km from Example 12.5, we have that the soliton peak power P_{peak} is

$$P_{\text{peak}} = \frac{A_{\text{eff}}}{2\pi n_2}\frac{\lambda}{L_{\text{disp}}} = \frac{(50\,\mu\text{m}^2)}{2\pi(2.6 \times 10^{-16}\,\text{cm}^2/\text{W})}\frac{1550\,\text{nm}}{202\,\text{km}}$$

$$= 2.35\ \text{mW}$$

This shows that when L_{disp} is on the order of hundreds of kilometers, P_{peak} is on the order of a few milliwatts.

For $N > 1$, the soliton pulse experiences periodic changes in its shape and spectrum as it propagates through the fiber. It resumes its initial shape at multiple distances of the *soliton period*, which is given by

$$L_{\text{period}} = \frac{\pi}{2}L_{\text{disp}} \tag{12.22}$$

As an example, Fig. l2.19 shows the evolution of a second-order soliton ($N = 2$).

12.10.3 Soliton Width and Spacing

The soliton solution to the NLS equation holds to a reasonable approximation only when individual pulses are well separated. To ensure this, the soliton width must be a small fraction of the bit slot. This eliminates use of the non-return-to-zero (NRZ) format that is commonly implemented in standard digital systems. Consequently, the return-to-zero (RZ) format is used. This condition thus constrains the achievable bit rate, since there is a limit on how narrow a soliton pulse can be generated.

If T_B is the width of the bit slot, then we can relate the bit rate B to the soliton half-maximum width T_s by

$$B = \frac{1}{T_B} = \frac{1}{2s_0 T_0} = \frac{1.7627}{2s_0 T_s} \tag{12.23}$$

where the factor $2s_0 = T_B/T_0$ is the *normalized separation* between neighboring solitons.

The physical explanation of the separation requirement is that the overlapping tails of closely spaced solitons create nonlinear interactive forces between them. These forces can be either attractive or repulsive, depending on the initial relative phase of the solitons. For solitons that are initially in phase and separated by $2s_0 \gg 1$, the soliton separation is periodic with an *oscillation period*:

$$\Omega = \frac{\pi}{2} \exp(s_0) \tag{12.24}$$

The mutual interactive force between in-phase solitons thus results in periodic attraction, collapse, and repulsion. The *interaction distance* is

$$L_I = \Omega L_{\text{disp}} = L_{\text{period}} \exp(s_0) \tag{12.25}$$

This interaction distance, and particularly the ratio L_I/L_{disp}, determine the maximum bit rate allowable in soliton systems.

These types of interactions are not desirable in a soliton system, since they lead to jitter in the soliton arrival times. One method for avoiding this situation is to increase s_0, since the interaction between solitons depends on their spacing. Since Eq. (12.23) is accurate for $s_0 > 3$, this equation together with the criterion that $\Omega L_{\text{disp}} \gg L_T$, where L_T is the total transmission distance, is suitable for system designs in which soliton interaction can be ignored.

Using Eq. (12.20) for L_{disp}, Eq. (12.23) for T_0, and Eq. (3.25) for D, the design condition $\Omega L_{\text{disp}} \gg L_T$ becomes

$$B^2 L_T \ll \left(\frac{2\pi}{s_0 \lambda}\right)^2 \frac{c}{16D} \exp(s_0) = \frac{\pi}{8 s_0^2 \, |\beta_2|} \exp(s_0) \tag{12.26}$$

When written in this form, Eq. (12.26) shows the effects on the bandwidth B or the total transmission distance L_T for selected values of s_0.

Example 12.7 Suppose we want to transmit information at a rate of 10 Gb/s over an 8600-km trans-Pacific soliton link.

(*a*) Since this is a high data rate over a long distance, we start by selecting a value of $s_0 = 8$. Then, from Eq. (12.24) we have $\Omega = 4682$. Given that the dispersion length is at least 100 km, then $\Omega L_{\text{disp}} > 4.7 \times 10^5$ km, which for all practical purposes satisfies the condition $\Omega L_{\text{disp}} \gg L_T = 8600$ km.

(*b*) If $D = 0.5$ ps/(nm · km) at 1550 nm then Eq. (12.26) yields

$$B^2 L_T \ll 2.87 \times 10^7 \text{ km (Gb/s)}^2$$

For a 10-Gb/s data rate, the transmission distance then must satisfy the condition

$$L_T \ll 2.87 \times 10^5 \text{ km}$$

This is satisfied, since the right-hand side is 33 times greater than the desired length.

(*c*) Using Eq. (12.23), we find the FWHM soliton pulse width to be

$$T_s = \frac{0.881}{s_0 B} = \frac{0.881}{8(10 \times 10^9 \text{ b/s})} = 11 \text{ ps}$$

(*d*) The fraction of the bit slot occupied by a soliton when $s_0 = 8$ is

$$\frac{T_s}{T_B} = \frac{0.881}{s_0} = \frac{0.881}{8} = 11\%$$

Note that for a given value of s_0, this is independent of the bit rate. For example, if the data rate is 20 Gb/s, then the FWHM pulse width is 5.5 ps, which also occupies 11 percent of the bit slot.

PROBLEMS

12.1 Consider three copropagating optical signals at frequencies v_1, v_2, and v_3.

(*a*) If these frequencies are evenly spaced so that $v_1 = v_2 - \Delta v$ and $v_3 = v_2 + \Delta v$, where Δv is an incremental frequency change, list the third-order waves that are generated due to FWM and plot them in relation to the original three waves. Note that several of these FWM-generated waves coincide with the original frequencies.

(*b*) Now examine the case when $v_1 = v_2 - \Delta v$ and $v_3 = v_2 + \dfrac{3}{2}\Delta v$. Find the FWM-generated frequencies and plot them in relation to the original three waves.

12.2 A single-mode fiber has an attenuation of 0.55 dB/km at 1310 nm and 0.28 dB/km at 1550 nm. Compare the effective lengths of this fiber at 1310 and 1550 nm.

12.3 Consider a 1550-nm optical source that has a 40-MHz linewidth. Suppose we have a single-mode fiber that has a 72-μm^2 effective area and a 0.2-dB/km attenuation at 1550 nm. Assuming that the polarization factor $b = 2$ and that the Brillouin gain coefficient is $g_B = 4 \times 10^{-11}$ m/W, what is the threshold power for stimulated Brillouin scattering at 1550 nm for a 40-km link? If the attenuation for this fiber is 0.4 dB/km at 1310 nm and all other parameters are the same, what is the threshold power for stimulated Brillouin scattering at 1310 nm?

12.4 An engineer needs to design a four-wavelength WDM system using a dispersion-shifted fiber that has a zero-dispersion point at 1550 nm. The four wavelengths need to be in the 193.1-to-194.1-THz band on a 100-GHz grid. Where should these wavelengths be placed so that none of the four-wave mixing components interfere with any transmission wavelength?

12.5 A detailed expression for the FWM efficiency η is given by[4]

$$\eta = \frac{\alpha^2}{\alpha^2 + \Delta\beta^2}\left[1 + \frac{4\exp\left(-\alpha L\right)\sin^2\left(\Delta\beta\, L/2\right)}{\left[1 - \exp\left(-\alpha L\right)\right]^2}\right]$$

where the factor $\Delta\beta$ is the difference of the propagation constants of the various waves due to dispersion, and is given by

$$\Delta\beta = \frac{2\pi\lambda^2}{c}|v_i - v_k| \times |v_j - v_k|$$
$$\times \left[D\left(v_0\right) + \frac{\lambda^2}{2c}\frac{dD}{d\lambda}\left(|v_i - v_0| + |v_j - v_0|\right)\right]$$

Here the value of the dispersion $D(v_0)$ and its slope $dD/d\lambda$, are taken at the optical frequency v_0. Using these expressions in Eq. (12.13), plot the ratio of the generated power P_{112} to the transmitted channel power P_1 as a function of the channel spacing for two +7-dBm channels. Find this ratio for the following dispersion and wavelength values:

(*a*) $D = 0$ ps/(nm·km) and $\lambda = 1556.6$ nm.

(*b*) $D = 0.13$ ps/(nm·km) and $\lambda = 1556.1$ nm.

(*c*) $D = 1.64$ ps/(nm·km) and $\lambda = 1537.2$ nm.

Let the frequency spacing of the two channels range from 0 to 250 GHz. In each case, take $dD/d\lambda = 0.08$ ps/(nm²·km), $\alpha = 0.0461$ km^{-1}, $L = 11$ km, and $A_{\text{eff}} = 55\ \mu\text{m}^2$. For χ_{1111} and $\mathscr{D}$ use the values given in Example 12.3.

12.6 A soliton transmission system operates at 1550 nm with fibers that have a dispersion of 1.5 ps/(nm · km) and an effective core area of 50 μm^2. Find the peak power required for fundamental solitons that have a 16-ps FWHM width. Use the value $n_2 = 2.6 \times 10^{-16}$ cm²/W. What are the dispersion length and the soliton period? What is the required peak power for 30-ps pulses?

12.7 A telecommunications service provider wants a single-wavelength soliton transmission system that is to operate at 40 Gb/s over a 2000-km distance. How would you design such a system? You are free to choose whatever components and design parameters are needed.

12.8 Create a cost model for the soliton system designed in Prob. 12.7, so that the service provider can determine the funding required for the project.

12.9 Consider a WDM system that utilizes two soliton channels at wavelengths λ_1 and λ_2. Since different wavelengths travel at slightly

different velocities in a fiber, the solitons of the faster channel will gradually overtake and pass through the slower-channel solitons. If the collision length L_{coll} is defined as the distance between the beginning and end of the pulse overlap at the half-power points, then

$$L_{coll} = \frac{2T_s}{D\Delta\lambda}$$

where $\Delta\lambda = \lambda_1 - \lambda_2$, T_s is the pulse FWHM, and D is the dispersion parameter.

(*a*) What is the collision length for $T_s = 16$ ps, $D = 0.5$ ps/(nm · km), and $\Delta\lambda = 0.8$ nm?

(*b*) Four-wave mixing effects arise between the soliton pulses during their collision, but then collapse to zero afterward. To avoid amplifying these effects, the condition $L_{coll} \geq 2L_{amp}$ should be satisfied, where L_{amp} is the amplifier spacing. What is the upper bound for L_{amp} for the above case?

12.10 Based on the conditions described in Prob. 12.9, what is the maximum number of allowed wavelength channels spaced 0.4 nm apart in a WDM soliton system when $L_{amp} = 25$ km, $T_s = 20$ ps, and $D = 0.4$ ps/(nm · km)?

REFERENCES

1. J. Toulouse, "Optical nonlinearities in fibers: Review, recent examples, and systems applications," *J. Lightwave Tech.*, vol. 23, pp. 3625–3641, Nov. 2005.

2. M. Wu and W. I. Way, "Fiber nonlinearity limitations in ultra-dense WDM systems," *J. Lightwave Tech.*, vol. 22, pp. 1483–1498, June 2004.

3. G. P. Agrawal, *Nonlinear Fiber Optics*, Academic, San Diego, CA, 4th ed., 2006.

4. F. Forghieri, R. W. Tkach, and A. R. Chraplyvy, "Fiber nonlinearities and their impact on transmission systems," in I. P. Kaminow and T. L. Koch, eds., *Optical Fiber Telecommunications–III*, Vol. A, Academic, New York, 1997.

5. A. R. Chraplyvy, "Limitations on lightwave communications imposed by optical-fiber nonlinearities," *J. Lightwave Tech.*, vol. 8, pp. 1548–1557, Oct. 1990.

6. X. Y. Zou, M. I. Hayee, S.-M. Hwang, and A. E. Willner, "Limitations in 10-Gb/s WDM optical fiber transmission when using a variety of fiber types to manage dispersion and nonlinearities," *J. Lightwave Tech.*, vol. 14, pp. 1144–1152, June 1996.

7. R. H. Stolen, "Nonlinear properties of optical fibers," in S. E. Miller and A. G. Chynoweth, eds., *Optical Fiber Telecommunications*, Academic, New York, 1979.

8. R. Ramaswami and K. N. Sivarajan, *Optical Networks*, Morgan Kaufmann, San Francisco, 3rd ed., 2009.

9. J. A. Buck, *Fundamentals of Optical Fibers*, Wiley, Hoboken, NJ, 2nd ed., 2004.

10. X. P. Mao, R. W. Tkach, A. R. Chraplyvy, R. M. Jopson, and R. M. Dorosier, "Stimulated Brillouin threshold dependence on fiber type and uniformity" *IEEE Photonics Tech. Lett.*, vol. 4, pp. 66–69, Jan. 1992.

11. X. P. Mao, G. E. Bodeep, R. W. Tkach, A. R. Chraplyvy, T. E. Darcie, and R. M. Dorosier, "Brillouin scattering in externally modulated lightwave AM-VSB CATV transmission systems," *IEEE Photonics Tech. Lett.*, vol. 4, pp. 287–289, Mar. 1992.

12. S. O. Kasap, *Principles of Electronic Materials and Devices*, McGraw-Hill, New York, 3rd ed., 2006.

13. N. Kikuchi and S. Sasaki, "Analytical evaluation technique of self-phase modulation effect on the performance of cascaded optical amplifier systems," *J. Lightwave Tech.*, vol. 13, pp. 868–878, May 1995.

14. N. Shibata, R. P. Braun, and R. G. Waarts, "Phase-match dependence of efficiency of wave generation through four-wave mixing in a single-mode optical fiber," *IEEE J. Quantum Electron.*, vol. 23, pp. 1205–1210, July 1987.

15. R. W. Tkach, A. R. Chraplyvy, F. Forghieri, A. H. Gnauck, and R. M. Dorosier, "Four-photon mixing and high-speed WDM systems," *J. Lightwave Tech.*, vol. 13, pp. 841–849, May 1995.

16. W. Zeiler, F. Di Pasquale, P. Bayvel, and J. E. Midwinter, "Modeling of four-wave mixing and gain peaking in amplified WDM optical communication systems and networks," *J. Lightwave Tech.*, vol. 14, pp. 1933–1941, Sept. 1996.

17. B. H. Choi, M. Attygalle, Y. J. Wen, and S. D. Dods, "Dispersion map optimization and dispersion slope mismatch issue on 40 channel × 10 Gbit/s transmission over 3000 km using standard SMF and EDFA amplification," *Optics Communications*, vol. 242. pp. 525–532, Dec. 2004.

18. B. Jopson and A. H. Gnauck, "Dispersion compensation for optical fiber systems," *IEEE Commun. Mag.*, vol. 33, pp. 96–102, June 1995.

19. L. Grüner-Nielsen, S. N. Knudsen, B. Edvold, T. Veng, D. Magnussen, C. C. Larsen, and H. Damsgaard, "Dispersion-compensating fibers," *Opt. Fiber Technology*, vol. 6, no. 2, pp. 164–180, Apr. 2000.

20. M. Suzuki and N. Edagawa, "Dispersion-managed high-capacity ultra-long-haul transmission," *J. Lightwave Tech.*, vol. 21, pp. 916–929, Apr. 2003.

21. L. Grüner-Nielsen, M. Wandel, P. Kristensen, C. Jørgensen, L. V. Jørgensen, B. Edvold, B. Pálsdóttir, and D. Jakobsen, "Dispersion-compensating fibers," *J. Lightwave Tech.*, vol. 23, pp. 3566–3579, Nov. 2005.

22. M. J. Connelly, *Semiconductor Optical Amplifiers*, Springer, New York, 2002.

23. J. Wang, *Pattern Effect Mitigation Techniques for All-Optical Wavelength Converters Based on Semiconductor Optical Amplifiers,* KIT Scientific Publishing, Karlsruhe, Germany, 2010.

24. T. Durhuus, B. Mikkelsen, C. Joergensen, S. L. Danielsen, and K. E. Stubkjaer, "All-optical wavelength conversion by semiconductor optical amplifiers," *J. Lightwave Tech.*, vol. 14, pp. 942–954, June 1996.

25. S. J. B. Yoo, "Wavelength conversion technologies for WDM network applications," *J. Lightwave Tech.*, vol. 14, pp. 955–966, June 1996.

26. M. Asghari, I. H. White, and R. V. Penty, "Wavelength conversion using semiconductor optical amplifiers," *J. Lightwave Tech.*, vol. 15, pp. 1181–1190, July 1997.

27. J. Capmany, S. Sales, D. Pastor, A. Martínez, and B. Ortega, "Wavelength conversion of SCM signals using semiconductor optical amplifiers: Theory, experiments, and applications," *J. Lightwave Tech.*, vol. 21, pp. 961–972, Apr. 2003.

28. K. Obermann, S. Kindt, D. Breuer, K. Petermann, C. Schmidt, S. Diez, and H. G. Weber, "Noise characteristics of semiconductor optical amplifiers used for wavelength conversion via cross-gain and cross-phase modulation," *IEEE Photonics Tech. Lett.*, vol. 9, pp. 312–314, Mar. 1997.

29. S. L. Danielsen, P. B. Hansen, K. E. Stubkjaer, M. Schilling, K. Wünstel, W. Idler, P. Doussiere, and F. Pommerau, "All optical wavelength conversion for increased input power dynamic range," *IEEE Photonics Tech. Lett.*, vol. 10, pp. 60–62, Jan. 1998.

30. K. Obermann, I. Koltchanov, K. Petermann, S. Diez, R. Ludwig, and H. G. Weber, "Noise analysis of frequency converters utilizing semiconductor-laser amplifiers," *IEEE J. Quantum Electron.*, vol. 33, pp. 81–88, Jan. 1997.

31. J. T. Hsieh, P. M. Gong, S. L. Lee, and J. Wu, "Improved dynamic characteristics on SOA-based FWM wavelength conversion in light-holding SOAs," *IEEE J. Selected Topics in Quantum Electronics*, vol. 10, pp. 1187–1196, Sept./Oct. 2004.

32. T. F. Morgen, J.P.R. Lacey, and R. S. Tucker, "Widely tunable FWM in semiconductor optical amplifiers with constant conversion frequency," *IEEE Photonics Tech. Lett.*, vol. 10, pp. 1401–1403, Oct. 1998.

33. H. Jang, S. Hur, Y. Kim, and J. Jeong, "Theoretical investigation of optical wavelength conversion techniques for DPSK modulation formats using FWM in SOAs and frequency comb in 10 Gb/s transmission systems,"

J. Lightwave Tech., vol. 23, pp. 2638–2646, Sept. 2005.

34. C. Politi, D. Klonidis, and M. J. O'Mahony, "Waveband converters based on four-wave mixing in SOAs," *J. Lightwave Tech.*, vol. 24, pp. 1203–1217, Mar. 2006.

35. J. S. Russell, *Reports of the Meetings of the British Assoc. for the Advancement of Science*, 1844.

36. Y. Lai and H. A. Haus, "Quantum theory of solitons in optical fibers. II. Exact solution," *Physical Review*, vol. 40, pp. 854–866, July 1989.

37. H. A. Haus, "Optical fiber solitons: Their properties and uses," *Proc. IEEE*, vol. 81, pp. 970–983, July 1993.

38. H. Haus and W. S. Wong, "Solitons in optical communications," *Rev. Mod. Physics*, vol. 68, pp. 432–444, 1996.

39. L. F. Mollenauer, J. P. Gordon, and P. V. Mamyshev, "Solitons in high bit-rate, long-distance transmissions," in I. P. Kaminow and T. L. Koch, eds., *Optical Fiber Telecommunications–III*, Vol. A, Academic, New York, 1997.

40. E. Iannone, F. Matera, A. Mecozzi, and M. Settembre, *Nonlinear Optical Communication Networks*, Wiley, New York, 1998.

41. A. Hasegawa and M. Matsumoto, *Optical Solitons in Fibers*, Springer, New York, 3rd ed., 2003.

42. J. R. Taylor, ed., *Optical Solitons: Theory and Experiment*, Cambridge University Press, 2005.

43. L. F. Mollenauer and J. P. Gordon, *Solitons in Optical Fibers: Fundamentals and Applications*, Academic, 2006.

CHAPTER 13

Optical Networks

This chapter covers performance and implementation issues related to optical fiber links that can be utilized in various types of networks to connect users having a wide range of transmission capacities and speeds.[1–8] The links between these users can range in length from short localized connections within a building or a campus environment to networks that span continents and cross oceans. A major motivation for developing sophisticated communication networks has been the rapid proliferation of information exchange desired by institutions involved in fields such as commerce, finance, education, scientific and medical research, health care, national and international security, and entertainment. The potential for this information exchange arose from the ever-increasing power of computers and data-storage devices, which need to be interconnected by high-speed, high-capacity networks.

Section 13.1 defines basic terminology and general network concepts, discusses the concept of network layering, and describes fiber optic network categories. Five broad transmission-system implementation categories are long-haul, metro, access, campus, and local-area networks. Section 13.2 illustrates bus and star network topologies and examines the design tradeoffs among them. For the terrestrial and undersea long-haul category, the closely coupled SONET and SDH standards provide a mechanism for multiplexing and transmitting optical signals so they can be shared between networks within the global telecommunications infrastructure. Section 13.3 discusses the physical layer aspects of SONET/SDH ring networks as they relate to optical transmission lines and networks. The basic SONET/SDH characteristics covered include the standard data-frame structure, the optical interface specifications, and the fundamental ring architectures. Also shown are possible network architectures that can be constructed from SONET/ SDH rings. Owing to the widespread use of long-haul and metro networks, reliability and availability are of high importance because many millions of users depend on them. Therefore, included in this section are discussions on how various protection-switching features can be built into a SONET/SDH ring network to ensure a high degree of survivability.

To increase the capacity of optical fiber links, research and development engineers and scientists are devising increasingly higher-rate transmission methods and sophisticated data encoding schemes. Section 13.4 gives some examples for links operating at 10, 40, 100, and 160 Gb/s. Of particular interest is the feasibility of sending information at rates up to 160 Gb/s over installed standard G.652 single-mode fibers using special encoding techniques.

Section 13.5 describes the major transmission elements used in high-capacity long-haul and metro WDM networks. Two critical elements for implementing these networks include fixed and reconfigurable optical

add/drop multiplexers (OADMs) and optical crossconnects (OXCs). Sections 13.5 and 13.6, respectively, define these elements and describe how they route signals along wavelength channels or lightwave paths through multiple intermediate nodes from the data source to the network client at the receiving end. Along the transmission path, an OADM or OXC located at an intermediate node may switch wavelengths to alternate paths or convert the wavelength when forwarding the data stream onto another link segment. Section 13.6 also introduces the concepts of wavelength routing, optical packet switching, and optical burst switching.

Section 13.7 illustrates applications of network elements, such as OADMs and OXCs, to wideband long-haul and narrowband metro WDM networks. The architectures and operations of passive optical networks (PON) are examined in Sec. 13.8. The process of exchanging information between communication elements using the Internet Protocol (IP) is discussed in Sec.13.9. The worldwide success in deploying Ethernet in local area networks has resulted in its extension into access and metropolitan area networks. This implementation, which is known as *optical Ethernet*, is the topic of Sec. 13.10. In designing WDM networks, transmission impairments such as chromatic and polarization-mode dispersion, optical amplifier transients, and timing jitter need to be taken into account. Section 13.11 discusses various mitigation techniques for these impairments.

13.1 | Network Concepts

This section provides some background material concerning the concepts of optical networks. The discussion illustrates different network architectures, notes what organizations own and operate various network segments, and defines some network terminology.

13.1.1 Network Terminology

Before getting into networking details, let us define some terminology using Fig. 13.1 as a guideline.

Stations Devices that network subscribers use to communicate are called *stations*. These may be computers, monitoring equipment, telephones, fax machines, or other telecommunication equipment.
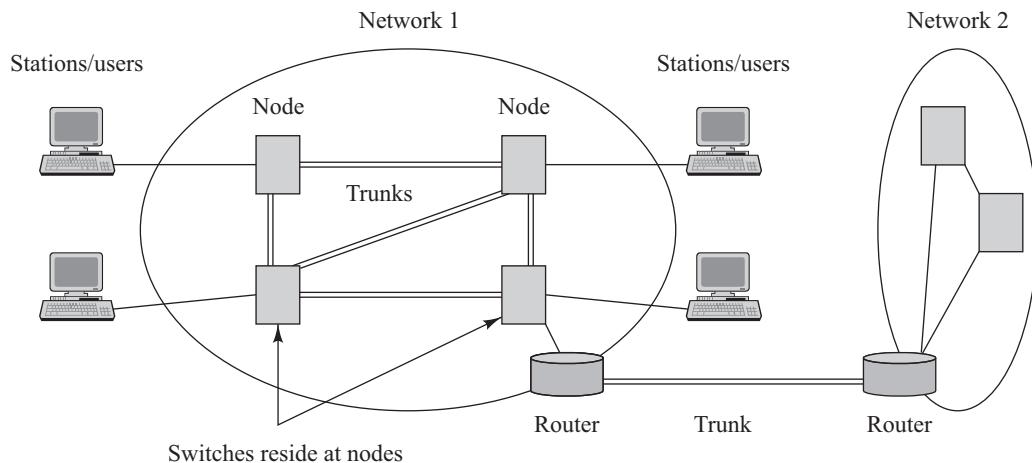
Fig. 13.1 *Definitions of various elements of a network*

Networks To establish connections between these stations, transmission paths run between them to form a collection of interconnected stations called a *network*.

Node Within this network, a *node* is a point where one or more communication lines terminate and/or where stations are connected. Stations also can be connected directly to a transmission line.

Trunk The term *trunk* normally refers to a transmission line that runs between nodes or networks and supports large traffic loads.

Topology The *topology* is the logical manner in which nodes are linked together by information-transmitting channels to form a network.

Switching and routing The transfer of information from source to destination through a series of intermediate nodes is called *switching*, and the selection of a suitable path through a network is referred to as *routing*.

Thus a *switched communication network* consists of an interconnected collection of nodes, in which information streams that enter the network from a station are routed to the destination by being switched from one transmission path to another at a node.

13.1.2 Network Categories

Networks can be divided into the following broad categories, as Fig. 13.2 illustrates:

Local Area Network (LAN)

A *local area network* (LAN) interconnects users in a localized area such as a large room or work area, a department, a home, a building, an office or factory complex, or a small group of buildings. A LAN employs relatively inexpensive hardware that allows users to share common expensive resources such
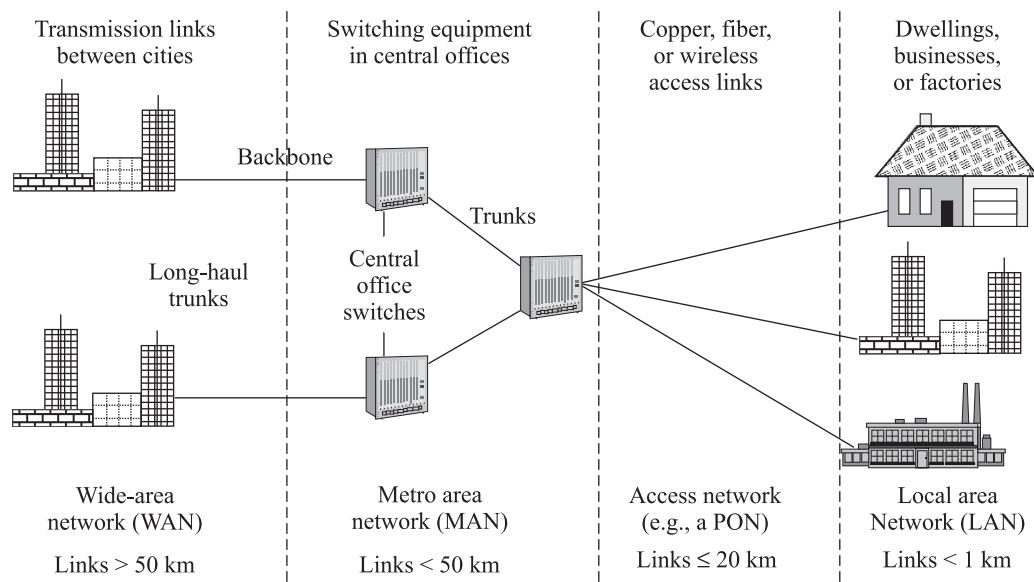
Fig. 13.2 *Definitions of some terms used in describing different segments of a public network*

as servers, high-performance printers, specialized instrumentation, or other equipment. Ethernet is the most popular networking technology used in LANs. Local area networks usually are owned, used, and operated privately by a single organization.

Campus Network

A *campus network* is an extension of a LAN and can be considered as an interconnection of a collection of LANs in a localized area. Similar to a LAN, it is an autonomous network owned and managed by a single organization that exists within a local geographical area. In networking terminology the word *campus* refers to any group of buildings that are within reasonable walking distance of each other. Thus a campus network could be deployed in a university campus, a business park, a government center, a research center, or a medical center. Typically a campus network uses routers to provide an access path into a larger network, such as a metropolitan area network or the Internet.

Metropolitan Area Network (MAN)

A MAN is commonly referred to as a *metro network* and spans a larger area than a LAN or a campus network. The interconnected facilities could range from buildings located in several city blocks to an entire city and the metropolitan area surrounding it. Thus the distances between central switching offices for a metro network range from a few to several tens of kilometers. Metro network resources are owned and operated by many telecommunication organizations.

Access Network

An access network lies between a metro network and a LAN or campus network. This network category encompasses connections that extend from a centralized switching facility to individual businesses, organizations, and homes. One function of an access network is to concentrate the information flows that originate in a local network and send this aggregated traffic to the switching facility. This path is called the *upstream direction*. In the other transmission direction (*downstream* or toward the user), the network provides voice, data, video, and other services to subscribers. The transmission distances tend to be up to 20 km. A particular access network is owned by a single telecommunication service provider.

Wide Area Network (WAN)

A WAN spans a large geographical area. The transmission distances can range from links between switching facilities in neighboring cities to long-haul cross-country terrestrial or intercontinental undersea lines. The large collection of WAN resources are owned and operated by either private organizations or telecommunication service providers.

Enterprise and Public Networks

When a private organization (for example, a company, a government entity, a medical facility, a university, or a commercial enterprise) owns and operates a network, it is called an *enterprise network*. These networks only provide services to the members of the organization. On the other hand, networks that are owned by the telecommunication carriers provide services such as leased lines or real-time telephone connections to the general public. These networks are referred to as *public networks*, since the services are available to any users or organizations that wish to subscribe to them. Public networks have another set of terminologies associated with them. As shown in Fig. 13.2, these terms include the following:

Central office A centralized switching facility in a public network is called a *central office* (CO) or a *point of presence* (POP). The CO houses a series of large telecommunication switches that establish

temporary connections for the duration of a requested service time between subscriber lines or between users and network resources.

Backbone The term *backbone* means a link that connects multiple network segments. For example, a backbone handles internetwork traffic; that is, traffic that originates in one network segment and is transmitted to another segment. Backbones can be short or long links.

Long-haul network A *long-haul network* interconnects distant cities or geographical regions and spans hundreds to thousands of kilometers between central offices. For example, this could be a high-capacity link carrying terabits of information between New York and San Francisco or between India and Singapore.

Passive optical network Many different transmission media can be used in an access network, including twisted-pair copper wires, coaxial cable, optical fibers, and radio links. Optical distribution networks that do not require any active optoelectronic components in the access region offer a number of operating advantages over other media. This implementation is called a *passive optical network* (PON) and is the basis for the fiber-to-the-premises (FTTP) networks described in Sec. 13.8.

13.1.3 Network Layers

When discussing the design and implementation of a telecommunication system, the term *network architecture* is used to describe the general physical arrangement and operational characteristics of communicating equipment together with a common set of communication protocols. A *protocol* is a set of rules and conventions that governs the generation, formatting, control, exchange, and interpretation of information that is transmitted through a telecommunication network or that is stored in a database.[9–12]

A traditional approach to setting up a protocol is to subdivide it into a number of individual pieces or layers of manageable and comprehensible size. The result is a layered structure of services, which is called a *protocol stack*. In this scheme each layer is responsible for providing a set of functions or capabilities to the layer above it by using the functions or capabilities of the layer below. A user at the highest layer is offered all the capabilities of the lower levels for interacting with other users and peripheral equipment distributed on the network.

As an example of a structured approach for simplifying the complexity of modern networks, in the early 1980s the International Standards Organization (ISO) developed an *open system interconnect (OSI) reference model* for dividing the functions of a network into seven operational layers.[10–13] As Fig. 13.3 illustrates, by convention these layers are viewed as a vertical sequence with the numbering starting at the bottom layer. Each layer performs specific functions using a standard set of protocols. A given layer is responsible for providing a service to the layer above it by using the services of the layer below it. Thereby an increasingly larger number of functions are provided when moving up the protocol stack, and the level of functional abstraction increases. The lower layers govern the *communication facilities*. These deal with the physical connections, data link control, and routing and relaying functions that support the actual transmission of data. The upper layers support *user applications* by structuring and organizing data for the needs of the user.

In the classical OSI model, the various layers carry out the following functions:

Physical layer The *physical layer* refers to a physical transmission medium, such as a wire or an optical fiber, which can handle a certain amount of bandwidth. It provides different types of physical interfaces to equipment, and its functions are responsible for actual transmission of bits across an optical fiber or metallic wire.

Data link layer The purpose of the *data link layer* is to establish, maintain, and release links that directly connect two nodes. Its functions include framing (defining how data is structured for transport), multiplexing, and demultiplexing of data. Examples of data link protocols include the *point-to-point protocol* (PPP) and the *high-level data link control* (HDLC) protocol.

User application program

Layer 7	Application	Provides general services (e.g., file transfer, user access)
Layer 6	Presentation	Formats data (encodes, encrypts, compresses)
Layer 5	Session	Maintains dialog between communicating devices
Layer 4	Transport	Provides reliable end-to-end data transmission
Layer 3	Network	Switches and routes information units
Layer 2	Data link	Provides data exchange between peer devices
Layer 1	Physical	Transmits bit stream to physical medium

Responsibility of host system

Responsibility of network

Fig. 13.3 *General structure and functions of the seven-layer OSI reference model*

Network layer

The function of the *network layer* is to deliver data packets from source to destination across multiple network links. Typically, the network layer must find a path through a series of connected nodes, and the nodes along this path must forward the packets to the appropriate destination. The dominant network layer protocol is the *Internet Protocol* (IP).

Transport Layer

The *transport layer* is responsible for reliably delivering the complete message from the source to the destination to satisfy a *quality of service* (QoS) requested by the upper layer. The QoS parameters include throughput, transit delay, bit-error rate, delay time to establish a connection, cost, information security, and message priority. The *transmission control protocol* (TCP) used in the Internet is an example of a transport layer protocol.

Higher Layers

The higher layers (session, presentation, application) support user applications, which are not covered here.

Note that there is nothing inherently unique about using seven layers or about the specific functionality in each layer. In actual applications some of the layers may be omitted and other layers can be subdivided into further sublayers. Thus the layering mechanism should be viewed as a framework for discussions of implementation and not as an absolute requirement. As an example, consider the widely used IP over SONET/SDH architecture (see Sec. 13.3). In this case, IP operates at the network and data link layers to format packets in such a way that they can be routed from end to end across a packet-switched network. Here the IP network views the SONET/SDH transport network merely as a set of physical point-to-point links between IP routers. However, as described in Sec. 13.3, the internal switching and routing operations of the SONET itself encompass physical, data link, and network layer functions.
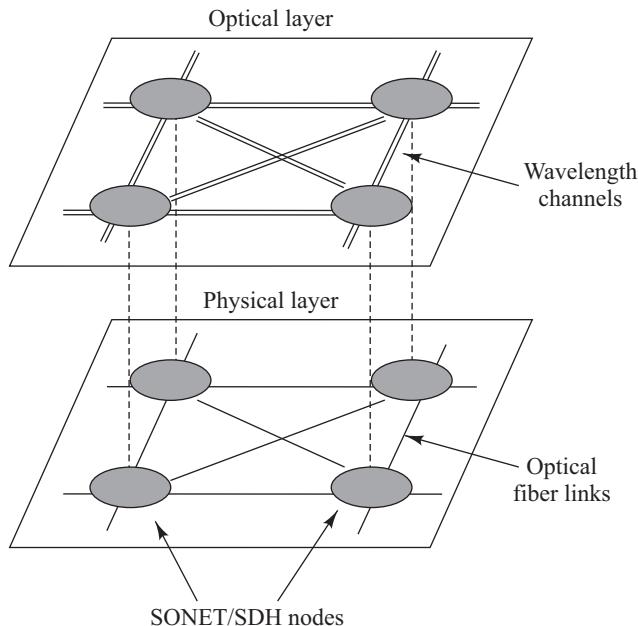
Fig. 13.4   *The optical layer describes the wavelength connections that lie on top of the physical layer.*

13.1.4   Optical Layer
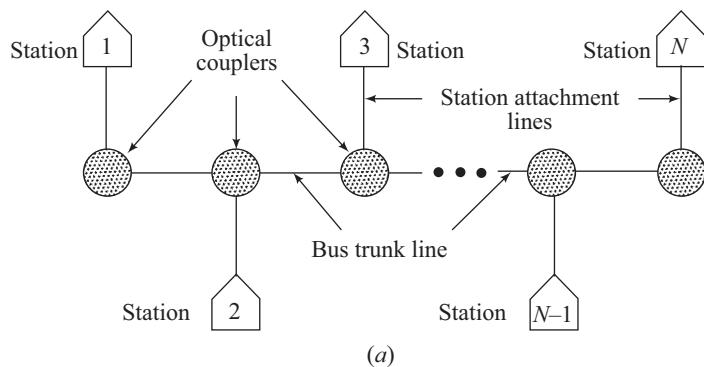
When dealing with optical network concepts, the term *optical layer* is used to describe various network functions or services. The optical layer is a wavelength-based concept and lies just above the physical layer, as shown in Fig. 13.4. This means that whereas the physical layer provides a physical connection between two nodes, the optical layer provides *lightpath services* over that link. A *lightpath* is an end-to-end optical connection that may pass through one or more intermediate nodes. For example, in an eight-channel WDM link there are eight lightpaths, which may go over a single physical line. Note that for a specific multiple-segment lightpath the wavelengths between various node pairs in the overall link may be different.[14–16]

The optical layer may carry out processes such as wavelength multiplexing, adding and dropping wavelengths, and support of optical crossconnects or wavelength switching. Networks that have these optical-layer functions are referred to as *wavelength-routed networks*. Section 13.6 has further details on these types of networks.
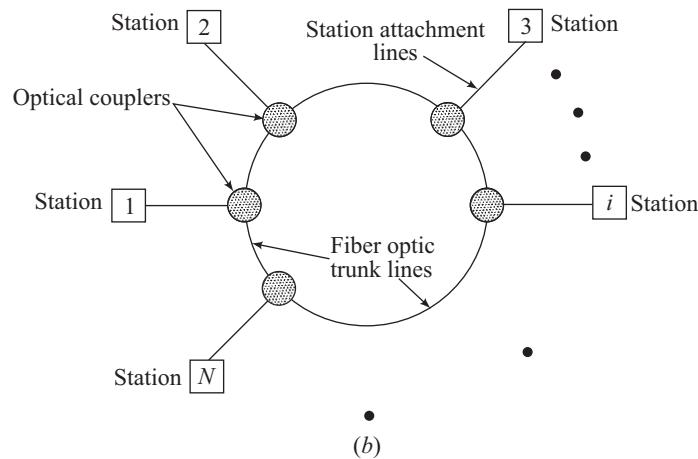
13.2   Network Topologies

Figure 13.5 shows the four common topologies used for fiber optic networks. These are the *linear bus, ring, star,* and *mesh configurations*. Each has its own particular advantages and limitations in terms of reliability, expandability, and performance characteristics.
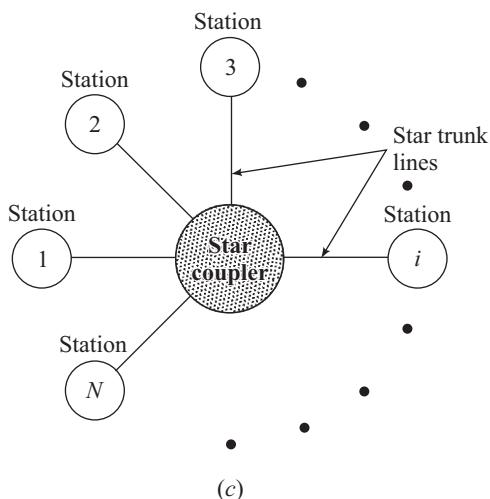
Nonoptical bus networks, such as standard Ethernet, employ coaxial cable as the transmission medium. The primary advantages of such a network are the passive nature of the transmission medium and the ability to easily install low-perturbation (high-impedance) taps on the coaxial line without disrupting
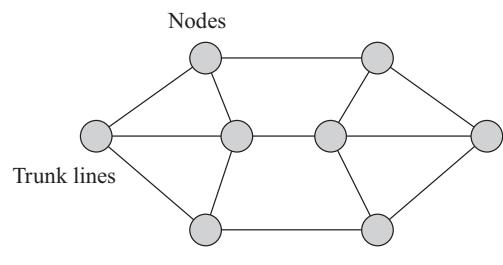
Fig. 13.5 *Basic optical fiber network topologies: (a) bus, (b) ring, (c) star, and (d) mesh configurations*

the operating network. In contrast to the coaxial bus, a fiber-optic-based bus network is more difficult to implement. The impediment is that there are no low-perturbation optical-tap equivalents to coax taps for efficiently coupling optical signals into and out of the main optical fiber trunk line. Access to an optical data bus is achieved by means of a coupling element, which can be either active or passive. An *active coupler* converts the optical signal on the data bus to its electric baseband counterpart before any data processing (such as injecting additional data into the signal stream or merely passing on the received data) is carried out. A *passive coupler* employs no electronic elements. It is used passively to tap off a portion of the optical power from the bus. Examples of these are the 2×2 couplers described in Chapter 10.

In a ring topology, consecutive nodes are connected by point-to-point links that are arranged to form a single closed path. Information in the form of data packets (a group of information bits plus overhead bits) is transmitted from node to node around the ring. The interface at each node is an active device that has the ability to recognize its own address in a data packet in order to accept messages. The active node forwards those messages that do not contain its own address on to its next neighbor.

In a star architecture, all nodes are joined at a single point called the *central node* or *hub*. The central node can be an active or a passive device. Using an active hub, one can control all routing of messages in the network from the central node. This is useful when most of the communications are between the central and the outlying nodes, as opposed to information exchange between the attached stations. If there is a great deal of message traffic between the outlying nodes, then a heavy switching burden is placed on an active central node. In a star network with a passive central node, a power splitter is used at the hub to divide the incoming optical signals among all the outgoing lines to the attached stations.

In a mesh network as shown in Fig. 13.5d, point-to-point links connect the nodes in an arbitrary fashion that can vary greatly from one application to another. This topology allows significant network configuration flexibility and offers connection protection in case there are multiple link or node failures. Link protection in mesh networks is accomplished by means of a mechanism that first determines where a failure has occurred and then restores the interrupted services by redirecting the traffic from a failed link or node to travel over another link in the mesh.[17,18]

13.2.1   Performance of Passive Linear Buses

To evaluate the performance of a passive linear bus, let us examine the various locations of power loss along the transmission path. We consider this in terms of the fraction of power lost at a particular interface or within a particular component. First, as described in Sec. 3.1, over an optical fiber of length x (in kilometers), the ratio A_0 of received power $P(x)$ to transmitted power $P(0)$ is given by

$$A_0 = \frac{P(x)}{P(0)} = 10^{-\alpha x/10} \tag{13.1}$$

where α is the fiber attenuation in units of dB/km.

The losses encountered in a passive coupler used for a linear bus are shown schematically in Fig. 13.6. This is nominally a cascaded combination of two directional couplers where two ports (one in each directional coupler) are not used. For simplicity, we do not show these unused ports here. The coupler thus has four functioning ports: two for connecting the coupler onto the fiber bus, one for receiving tapped-off light, and one for inserting an optical signal onto the line after the tap-off to keep the signal out of the local receiver. If a fraction F_c of optical power is lost at each port of the coupler, then the *connecting loss* L_c is

$$L_c = -10 \log (1 - F_c) \tag{13.2}$$
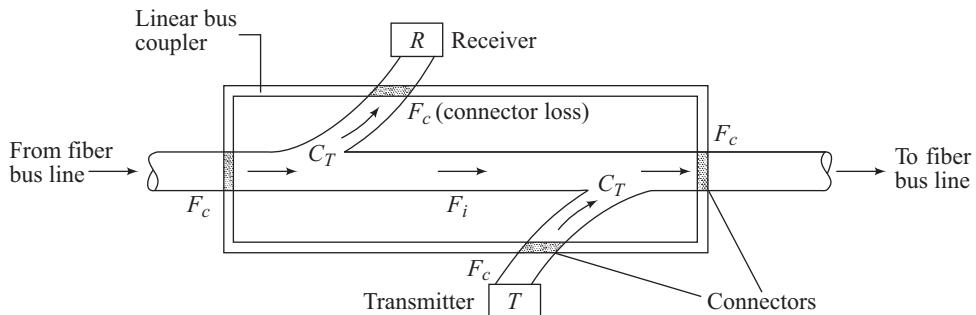
Fig. 13.6 *Losses encountered in a passive linear bus coupler consisting of a cascade of two directional couplers*

For example, if we take this fraction to be 20 percent, then L_c is about 1 dB; that is, the optical power gets reduced by 1 dB at any coupling junction.

Let C_T represent the fraction of power that is removed from the bus and delivered to the detector port. The power extracted from the bus is called a *tap loss* and is given by

$$L_{tap} = -10 \log C_T \tag{13.3}$$

For a symmetric coupler, C_T also is the fraction of power that is coupled from the transmitting input port to the bus. Thus, if P_0 is the optical power launched from a source flylead, the power coupled to the bus is $C_T P_0$. Note that, in general, when calculating the throughput power for intermediate stations, the transmission path through the bus coupler passes two tap points, since optical power is extracted at both the receiving and the transmitting taps of the device. The power removed at the transmitting tap goes out of the unused port and thus is lost from the system. The *throughput coupling loss* L_{thru} in decibels is then given by

$$L_{thru} = -10 \log (1 - C_T)^2 = -20 \log (1 - C_T) \tag{13.4}$$

In addition to connection and tapping losses, there is an intrinsic transmission loss L_i associated with each bus coupler. If the fraction of power lost in the coupler is F_i, then the *intrinsic transmission loss* L_i in decibels is

$$L_i = -10 \log (1 - F_i) \tag{13.5}$$

Generally, a linear bus will consist of a number of stations separated by various lengths of bus line. However, for analytical simplicity, here we consider a simplex linear bus of N stations uniformly separated by a distance L, as shown in Fig. 13.7. Thus from Eq. (13.1) the fiber attenuation between any two adjacent stations in decibels is

$$L_{fiber} = -10 \log A_0 = \alpha L \tag{13.6}$$

Example 13.1 Consider a specific bus coupler that extracts 5 percent of the light to an optical power monitor. Furthermore, assume 2 percent of the optical power inserted into the coupler is lost internally. What are the throughput coupling and intrinsic losses of this device?

Solution: (a) From Eq. (13.4) with $C_T = 0.05$, the throughput coupling loss is

$$L_{thru} = -20 \log (1 - C_T) = -20 \log 0.95 = 0.45 \text{ dB}$$

(b) From Eq. (13.5) the intrinsic loss is

$$L_i = -10 \log (1 - F_i) = -10 \log 0.98 = 0.09 \text{ dB}$$
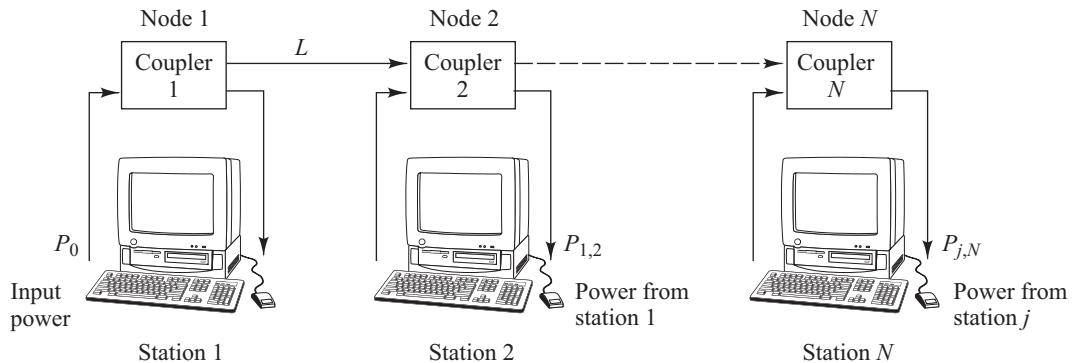
Fig. 13.7 *Topology of a linear bus consisting of N uniformly spaced stations*

The term *simplex* means that, in this configuration, information flows only from left to right. To have *full-duplex* communication, in which stations can communicate in either direction, one needs an analogous configuration with a separate parallel line and another set of N couplers. The information flow in the second line would then be from right to left.

To determine the power budget, we first examine the link in terms of fractional power losses at each link element. The examples will then show the power-budget calculation (in decibels) using a tabular form. For the fractional-loss method, we use the notation $P_{j,k}$ to denote the optical power received at the detector of the kth station from the transmitter of the jth station. For simplicity, we assume that a bus coupler exists at every terminal of the bus, including the two end stations.

Nearest-Neighbor Power Budget

The smallest distance in transmitted and received power occurs for adjacent stations, such as between stations 1 and 2 in Fig. 13.7. If P_0 is the optical power launched from a source into a fiber at station 1, then the power detected at station 2 is

$$P_{1,2} = A_0\, C_T^2\, (1 - F_c)^4\, (1 - F_i)^2\, P_0 \tag{13.7}$$

since the optical power flow encounters the following loss-inducing mechanisms:
- One fiber path with attenuation A_0.
- Tap points at both the transmitter and the receiver, each with coupling efficiencies C_T.
- Four connecting points, each of which passes a fraction $(1 - F_c)$ of the power entering them.
- Two couplers which pass only the fraction $(1 - F_i)$ of the incident power owing to intrinsic losses.

Using Eqs. (13.2) through (13.4) and Eq. (13.6), the expression for the losses between stations 1 and 2 can be expressed in logarithmic form as

$$10\ \log\left(\frac{P_0}{P_{1,2}}\right) = \alpha L + 2L_{\text{tap}} + 4L_c + 2L_i \tag{13.8}$$

Example 13.2   A short linear bus network consists of six computer stations that are interconnected with 10-m lengths of polymer optical fiber having a loss of 0.1 dB/m. Assume the bus couplers each have the following characteristics: an optical power tap off of 10 percent, a 10 percent connection loss, and a 2 percent intrinsic loss. What is the optical power loss in the link between each computer?

Solution: From Eq. (13.2):
$$L_c = -10 \log (1 - F_c) = -10 \log 0.90 = 0.46 \text{ dB}$$

From Eq. (13.3):
$$L_{\text{tap}} = -10 \log C_T = -10 \log 0.10 = 10.0 \text{ dB}$$

From Eq. (13.8) we have
Loss between adjacent stations $= (0.1 \text{ dB/m}) \times (10 \text{ m})$
$+ 2 L_{\text{tap}} + 4 L_c + 2 L_i = 23.02 \text{ dB}$

Largest-Distance Power Budget

The largest distance for transmitted and received power occurs between stations 1 and N. At the transmitting end the fractional power level coupled into the first length of cable from the transmitter through the bus coupler at station 1 is

$$F_1 = (1 - F_c)^2 \ C_T (1 - F_i) \tag{13.9a}$$

Similarly, at station N the fraction of power from the bus coupler input port that emerges from the detector port is

$$F_N = (1 - F_c)^2 \ C_T (1 - F_i) \tag{13.9b}$$

For each of the $(N - 2)$ intermediate stations, the fraction of power passing through each coupling module (shown in Fig. 13.6) is

$$F_{\text{coup}} = (1 - F_c)^2 \ (1 - C_T)^2 \ (1 - F_i) \tag{13.10}$$

since, from the input to the output of each coupler, the power flow encounters two connector losses, two tap losses, and one intrinsic loss. Combining the expressions from Eqs. (13.9a), (13.9b), and (13.10), and the transmission losses of the $N - 1$ intervening fibers, we find that the power received at station N from station 1 is

$$\begin{aligned}
P_{1,N} &= A_0^{N-1} \ F_1 \ F_{\text{coup}}^{N-2} \ F_N \ P_0 \\
&= A_0^{N-1} (1 - F_c)^{2N} (1 - C_T)^{2(N-2)} C_T^2 (1 - F_i)^N \ P_0
\end{aligned} \tag{13.11}$$

Using Eqs. (13.2) through (13.6), the power budget for this link then is

$$\begin{aligned}
10 \log \left(\frac{P_0}{P_{1,N}} \right) &= (N - 1) \, \alpha L + 2N L_c + (N - 2) \, L_{\text{thru}} + 2 L_{\text{tap}} + N L_i \\
&= [\text{fiber} + \text{connector} + \text{coupler throughput} \\
&\quad + \text{ingress/egress} + \text{coupler intrinsic}] \text{ losses} \\
&= N \, (\alpha L + 2 L_c + L_{\text{thru}} + L_i) - \alpha L - 2 L_{\text{thru}} + 2 L_{\text{tap}}
\end{aligned} \tag{13.12}$$

The last expression shows that the losses (in decibels) of the linear bus increase linearly with the number of stations N.

Example 13.3 Let us compare the power budgets of three linear buses, having 5, 10, and 50 stations, respectively. Assume that $C_T = 10$ percent, so that $L_{\text{tap}} = 10$ dB and $L_{\text{thru}} = 0.9$ dB. Let $L_i = 0.5$ dB and $L_c = 1.0$ dB. If the stations are relatively close together, say 500 m, as they might be in a LAN, then for an attenuation of 0.4 dB/km at 1300 nm the fiber loss is 0.2 dB. Using Eq. (13.12), the power budgets for these three cases can be calculated using the tabular method shown in Table 13.1. For actual calculations, it is best to use a spreadsheet on a computer.

Example 13.4 For the applications given in Example 13.3 suppose that for implementing a 10-Mb/s bus we have a choice of an LED that emits −10 dBm or a laser diode capable of emitting +3 dBm of optical power from a fiber flylead. At the destination, assume we have an APD receiver with a sensitivity of −48 dBm at 10 Mb/s in the 1300-nm window. In the LED case, the power loss allowed between the source and the receiver is 38 dB. As shown in Fig. 13.8 this allows up to 5 stations on the bus. For the laser diode source, we have an additional 13 dB of margin, so for this example we can have a maximum of 8 stations connected to the bus.

The total loss values given in Table 13.1 are plotted in Fig. 13.8, which shows that the loss (in dB) increases linearly with the number of stations.

Dynamic Range

Owing to the serial nature of the linear bus, the optical power available at a particular node decreases with increasing distance from the source. Thus a performance quantity of interest is the system dynamic range. This is the maximum optical power range to which any detector must be able to respond. The worst-case dynamic range (DR) is found from the ratio of Eq. (13.7) to Eq. (13.11)

$$\text{DR} = 10 \log \left(\frac{P_{1,2}}{P_{1,N}} \right) = 10 \log \left\{ \frac{1}{[A_0 (1 - F_c)^2 (1 - C_T)^2 (1 - F_i)]^{N-2}} \right\}$$

$$= (N - 2) (\alpha L + 2L_c + L_{\text{thru}} + L_i) \tag{13.13}$$

For example, this could be the difference in power levels received at station N from station $(N - 1)$ and from station 1 (i.e., $P_{1,2} = P_{N-1, N}$).

Table 13.1 *Comparison of the power budgets of three linear buses that have 5, 10, and 50 stations, respectively*

Coupling/loss factor	Loss expression	Loss (dB)	Losses for 5 stations	Losses for 10 stations	Losses for 50 stations
Source connector	Eq. (13.2)	1.0	1.0	1.0	1.0
Coupling (tap) loss	Eq. (13.3)	2×10.0	20.0	20.0	20.0
Coupler-to-fiber loss	Eq. (13.2)	$2 (N - 1) \times 1.0$	8.0	18.0	98.0
Fiber loss (500 m)	Eq. (13.6)	$(N - 1) \times 0.2$	0.8	1.8	9.8
Coupler throughput	Eq. (13.4)	$(N - 2) \times 0.9$	2.7	7.2	43.2
Intrinsic coupler loss	Eq. (13.5)	$N \times 0.5$	2.5	5.0	25.0
Receiver connector	Eq. (13.2)	1.0	1.0	1.0	1.0
Total loss (dB)	Eq. (13.12)	–	36.0	54.0	198.0
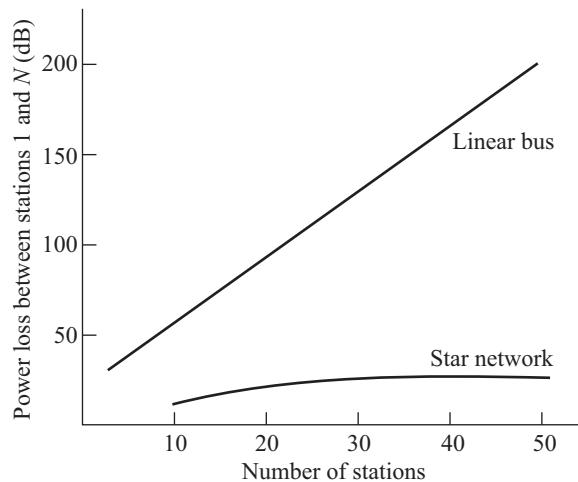
Fig. 13.8 *Total optical power loss as a function of the number of attached stations for linear bus and star architectures*

Example 13.5 Consider the linear buses described in Example 13.3. What are the dynamic ranges for 5 and 10 stations?

Solution: For $N = 5$, from Eq. (13.13) the dynamic range is

$$DR = 3[0.2 + 2(1.0) + 0.9 + 0.5]\ dB = 10.8\ dB$$

For $N = 10$ stations,

$$DR = 8[0.2 + 2(1.0) + 0.9 + 0.5]\ dB = 28.8\ dB$$

13.2.2 Performance of Star Architectures

To see how a star coupler can be applied to a given network, let us examine the various optical power losses associated with the coupler. Section 10.2.4 gives the details of how an individual star coupler works. As a quick review, the excess loss is defined as the ratio of the input power to the total output power. That is, it is the fraction of power lost in the process of coupling light from the input port to all the output ports. From Eq. (10.25), for a single input power P_{in} and N output powers, the excess loss in decibels is given by

$$\text{Fiber star excess loss} = L_{\text{excess}} = 10\ \log\left(\frac{P_{in}}{\sum\limits_{i=1}^{N} P_{\text{out, i}}}\right) \tag{13.14}$$

In an ideal star coupler the optical power from any input is evenly divided among the output ports. The total loss of the device consists of its splitting loss plus the excess loss in each path through the star. The *splitting loss* is given in decibels by

$$\text{Splitting loss} = L_{\text{split}} = -10\log\left(\frac{1}{N}\right) = 10\log N \tag{13.15}$$

To find the power-balance equation, we use the following parameters:

- P_S is the fiber-coupled output power from a source in dBm.
- P_R is the minimum optical power in dBm required at the receiver to achieve a specific bit-error rate.
- α is the fiber attenuation.
- All stations are located at the same distance L from the star coupler.
- L_c is the connector loss in decibels.

Then, the power-balance equation for a particular link between two stations in a star network is

$$P_S - P_R = L_{\text{excess}} + \alpha(2L) + 2L_c + L_{\text{split}}$$
$$= L_{\text{excess}} + \alpha(2L) + 2L_c + 10 \log N \qquad (13.16)$$

Here, we have assumed connector losses at the transmitter and the receiver. This equation shows that, in contrast to a passive linear bus, for a star network the loss increases much slower as log N. Figure 13.8 compares the performance of the two architectures.

Example 13.6   Consider two star networks that have 10 and 50 stations, respectively. Assume each station is located 500 m from the star coupler and that the fiber attenuation is 0.4 dB/km. Assume the excess loss is 0.75 dB for the 10-station network and 1.25 dB for the 50-station network. Let the connector loss be 1.0 dB. What is the power margin for 10 and for 50 stations?

Solution: For $N = 10$, from Eq. (13.16) the power margin between the transmitter and the receiver is

$$P_S - P_R = [0.75 + 0.4(1.0) + 2(1.0) + l0 \log 10]\text{dB}$$
$$= 13.2 \text{ dB}$$

For $N = 50$,

$$P_S - P_R = [1.25 + 0.4(1.0) + 2(1.0) + 10 \text{ log } 50] \text{ dB}$$
$$= 20.6 \text{ dB}$$

Using the transmitter output and receiver sensitivity values given in Example 13.4, we see that an LED transmitter can easily accommodate the losses in this 50-station star network. In comparison, a laser transmitter could not even meet the 10-station design in a passive linear bus.

13.3   SONET/SDH

With the advent of fiber optic transmission lines, the next step in the evolution of the digital time-division-multiplexing (TDM) scheme was a standard signal format called *synchronous optical network* (SONET) in North America and *synchronous digital hierarchy* (SDH) in other parts of the world. This section addresses the basic concepts of SONET/SDH, its optical interfaces, and fundamental network implementations. The aim here is to discuss only the physical-layer aspects of SONET/SDH as they relate to optical transmission lines and optical networks. Topics such as the detailed data format structure, SONET/SDH operating specifications, and the relationships of switching methodologies such as carrier Ethernet service aggregation with SONET/SDH are beyond the scope of this text. These can be found in numerous sources.[9,10,19–21]

13.3.1   Transmission Formats and Speeds

In the mid-1980s, several service providers in the USA started efforts to develop a standard that would allow network engineers to interconnect fiber optic transmission equipment from various vendors
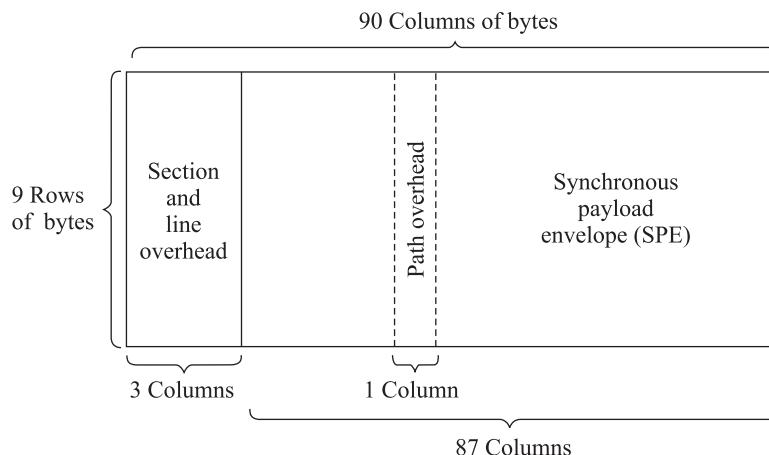
Fig. 13.9 *Basic structure of an STS-1 SONET frame*

through multiple-owner trunking networks. This soon grew into an international activity, which, after many differences of opinion of implementation philosophy were resolved, resulted in a series of ANSI T1.105 standards[22] for SONET and a series of ITU-T recommendations for SDH.[23] Of particular interest here are the ANSI TI.105 standard and the ITU-T G.957 recommendation. Although there are some implementation differences between SONET and SDH, all SONET specifications conform to the SDH recommendations.

Figure 13.9 shows the basic structure of a SONET frame. This is a two-dimensional structure consisting of 90 columns by 9 rows of bytes, where one byte is eight bits. Here, in standard SONET terminology, a *section* connects adjacent pieces of equipment, a *line* is a longer link that connects two SONET devices, and a *path* is a complete end-to-end connection. The fundamental SONET frame has a 125-μs duration. Thus, the transmission bit rate of the basic SONET signal is

$$\text{STS-1} = (90 \text{ bytes/row})(9 \text{ rows/frame})(8 \text{ bits/byte})/(125 \text{ } \mu\text{s/frame}) = 51.84 \text{ Mb/s}$$

This is called an STS-1 signal, where STS stands for *synchronous transport signal.* All other SONET signals are integer multiples of this rate, so that an STS-*N* signal has a bit rate equal to *N* times 51.84 Mb/s. When an STS-*N* signal is used to modulate an optical source, the *logical STS-N signal* is first scrambled to avoid long strings of ones and zeros and to allow easier clock recovery at the receiver. After undergoing electrical-to-optical conversion, the resultant *physical-layer optical signal* is called OC-*N,* where OC stands for *optical carrier.* In practice, it has become common to refer to SONET links as OC-*N* links. Algorithms have been developed for values of *N* ranging between 1 and 768.

In SDH the basic rate is equivalent to STS-3, or 155.52 Mb/s. This is called the *synchronous transport module—level* 1 (STM-1). Higher rates are designated by STM-*M*. (Note: Although the SDH standard uses the notation "STM-*N*," here we use the designation "STM-*M*," to avoid confusion when comparing SDH and SONET rates.) Values of *M* supported by the ITU-T recommendations are *M* = 1, 4, 16, and 64. These are equivalent to SONET OC-*N* signals, where *N* = 3*M* (i.e., *N* = 3, 12, 48, and 192). This shows that, in practice, to maintain compatibility between SONET and SDH, *N* is a multiple of three. Analogous to SONET, SDH first scrambles the logical signal. In contrast to SONET,

Table 13.2 *Commonly used SONET and SDH transmission rates*

SONET level	Electrical level	SDH level	Line rate (Mb/s)	Common rate name
OC-N	STS-N	—	$N \times 51.84$	—
OC-1	STS-1	—	51.84	—
OC-3	STS-3	STM-1	155.52	155 Mb/s
OC-12	STS-12	STM-4	622.08	622 Mb/s
OC-48	STS-48	STM-16	2488.32	2.5 Gb/s
OC-192	STS-192	STM-64	9953.28	10 Gb/s
OC-768	STS-768	STM-256	39813.12	40 Gb/s

SDH does not distinguish between a logical electrical signal (e.g., STS-*N* in SONET) and an optical signal (e.g., OC-*N*)*,* so that both signal types are designated by STM-*M*. Table 13.2 lists commonly used values of OC-*N* and STM-*M*.

Referring to Fig. 13.9, the first three columns comprise transport overhead bytes that carry network management information. The remaining field of 87 columns is called the *synchronous payload envelope* (SPE) and carries user data plus nine bytes of *path overhead* (POH). The POH supports performance monitoring by the end equipment, status, signal labeling, a tracing function, and a user channel. The nine path-overhead bytes are always in a column and can be located anywhere in the SPE. An important point to note is that the synchronous byte-interleaved multiplexing in SONET/SDH (unlike the asynchronous bit interleaving used in earlier TDM standards) facilitates add/drop multiplexing of individual information channels in optical networks.

For values of *N* greater than 1, the columns of the frame become *N* times wider, with the number of rows remaining at nine, as shown in Fig. 13.10*a*. Thus, an STS-3 (or STM-1) frame is 270 columns wide with the first nine columns containing overhead information and the next 261 columns being payload data. The basic frame structure for SDH is shown in Fig. 13.10*b*. An STM-*N* frame has a 125-μs duration and consists of nine rows, each of which has a length of $270 \times N$ bytes. Since the line and section overhead bytes differ somewhat between SONET and SDH, a translation mechanism is needed to interconnect SONET and SDH equipment.

13.3.2 Optical Interfaces

To ensure interconnection compatibility between equipment from different manufacturers, the SONET and SDH specifications provide details for the optical source characteristics, the receiver sensitivity, and transmission distances for various types of fibers. Six transmission ranges and the associated fiber types are defined with different terminology for SONET and SDH, as Table 13.3 indicates. The transmission distances are specified for G.652, G.653, and G.655 fibers. The ITU-T G.957 Recommendation also designates the SDH categories by codes such as I-1, S-1.1, L-.1, and so on as indicated in the table.

The optical fibers specified in ANSI T1.105.06 and ITU-T G.957 fall into the following three categories and operational windows:

1. Graded-index multimode in the 1310-nm window (O-band)
2. Conventional nondispersion-shifted single-mode in the 1310-nm and 1550-nm windows (O-band and C-band)
3. Dispersion-shifted single-mode in the 1550-nm window (C-band)

$90 \times N$ Columns

9 Rows

Section and line overhead

Path overhead

SPE

$3 \times N$ Columns

N Columns

$87 \times N$ Columns

(*a*)

$270 \times N$ Columns

9 Rows

Section and line overhead

Path overhead

SPE

$9 \times N$ Columns

N Columns

$261 \times N$ Columns

(*b*)
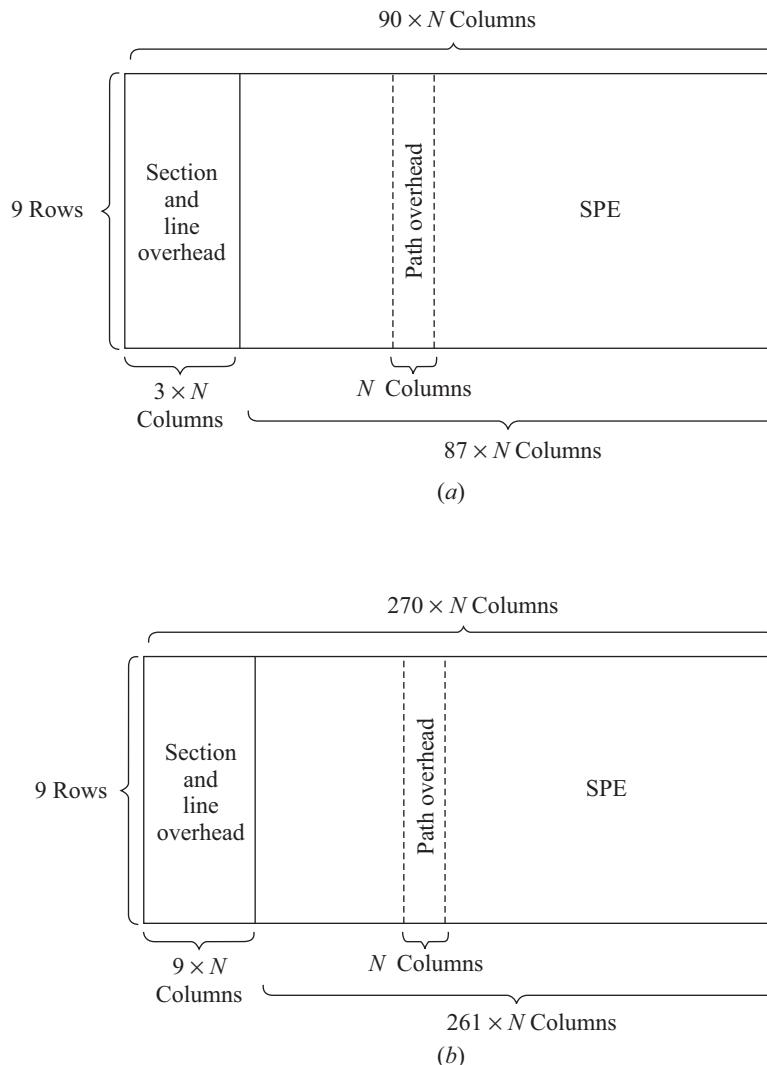
Fig. 13.10 *Basic formats of (a) an STS-N SONET frame and (b) an STM-N SDH frame*

Table 13.4 shows the wavelength and attenuation ranges specified in these fibers for transmission distances up to 80 km.

Depending on the attenuation and dispersion characteristics for each hierarchical level shown in Table 13.3, feasible optical sources include light-emitting diodes (LEDs), multimode lasers, and various single-mode lasers. The system objective in ANSI T1.105 and ITU-T G.957 is to achieve a bit-error rate (BER) of less than 10^{-10} for rates less than 1 Gb/s and 10^{-12} for higher rates and/or higher-performance systems.

The receiver sensitivities in G.957 are the worst-case, end-of-life values. They are defined as the minimum-acceptable, average, received power needed to achieve a 10^{-10} BER. The values take into account extinction ratio, pulse rise and fall times, optical return loss at the source, receiver connector

Table 13.3 *Transmission distances and their SONET and SDH designations, where x denotes the STM-x level.*

Transmission distance	Fiber type	SONET terminology	SDH terminology
≤ 2 km	G.652	Short-reach (SR)	Intraoffice (I-1)
15 km at 1310 nm	G.653	Intermediate-reach (IR-1)	Short-haul (S-*x*.1)
15 km at 1550 nm	G.653	Intermediate-reach (IR-2)	Short-haul (S-*x*.2)
40 km at 1310 nm	G.655	Long-reach (LR-1)	Long-haul (L-*x*.1)
80 km at 1550 nm	G.655	Long-reach (LR-2)	Long-haul (L-*x*.3)
120 km at 1550 nm	G.655	Very long-reach (VR-1)	Very long (V-*x*.3)
160 km at 1550 nm	G.655	Very long-reach (VR-2)	Ultra long (U-*x*.3)

degradations, and measurement tolerances. The receiver sensitivity does not include power penalties associated with dispersion, jitter, or reflections from the optical path; these are included in the maximum optical path penalty. Table 13.5 lists the receiver sensitivities for various link configurations up through long-haul distances (80 km). Note that the ANSI and ITU-T recommendations are updated periodically, so the reader should refer to the latest version of the documents for specific details.

Longer transmission distances are possible using higher-power lasers. To comply with eye-safety standards, an upper limit is imposed on fiber-coupled powers. If the maximum total output power (including ASE) is set at the Class-3*A* laser limit of $P_{3A} = +17$ dBm, then for ITU-T G.655 fiber this allows transmission distances of 160 km for a single-channel link. Using this condition, then for *M* operational WDM channels the maximum nominal channel power P_{chmax} should be limited to $P_{chmax} = P_{3A} - 10 \log M$. Table 13.6 lists the maximum nominal optical powers per channel for $M = 8$.

13.3.3 SONET/SDH Rings

A key characteristic of SONET and SDH is that they are configured as either a ring or mesh architecture.[24,25] This is done to create *loop diversity* for uninterrupted service protection purposes in case of link or equipment failures. The SONET/SDH rings are commonly called *self-healing rings* because the traffic flowing along a certain path can automatically be switched to an alternate or standby path following failure or degradation of the link segment.

Three main features, each with two alternatives, classify all SONET/SDH rings, thus yielding eight possible combinations of ring types. First, there can be either two or four fibers running between the nodes on a ring. Second, the operating signals can travel either clockwise only (which is termed a *unidirectional ring*) or in both directions around the ring (which is called a *bidirectional ring*). Third, protection switching can be performed either via a line-switching or a path-switching scheme.[18,19,26–28]

Table 13.4 *Wavelength ranges and attenuation for transmission distances up to 80 km*

Distance	Wavelength range at 1310 nm	Wavelength range at 1550 nm	Attenuation at 1310 nm (dB/km)	Attenuation at 1550 nm (dB/km)
≤ 15 km	1260 – 1360 nm	1430 – 1580 nm	3.5	Not specified
≤ 40 km	1260 – 1360 nm	1430 – 1580 nm	0.8	0.5
≤ 80 km	1280 – 1335 nm	1480 – 1580 nm	0.5	0.3

Table 13.5 *Source output, attenuation, and receiver sensitivity ranges for various rates and distances up to 80 km (see ITU-T G.957)*

Parameter	Intraoffice	Short-haul (1)	Short-haul (2)	Long-haul (1)	Long-haul (3)
Wavelength	1310 nm	1310 nm	1550 nm	1310 nm	1550 nm
Fiber		SM	SM	SM	SM
Distance (km)	≤ 2 km	15 km	15 km	40 km	80 km
Designation	I - 1	S - 1.1	S - 1.2	L - 1.1	L - 1.3
Source range (dBm)					
155 Mb/s	− 15 to − 8	− 15 to − 8	− 15 to − 8	0 to 5	0 to 5
622 Mb/s	− 15 to − 8	− 15 to − 8	− 15 to − 8	− 3 to +2	− 3 to +2
2.5 Gb/s	− 10 to − 3	− 5 to 0	− 5 to 0	− 2 to +3	− 2 to +3
Attenuation range (dB)					
155 Mb/s	0 to 7	0 to 12	0 to 12	10 to 28	10 to 28
622 Mb/s	0 to 7	0 to 12	0 to 12	10 to 24	10 to 24
2.5 Gb/s	0 to 7	0 to 12	0 to 12	10 to 24	10 to 24
Receiver sensitivity					
155 Mb/s	− 23	− 28	− 28	− 34	− 34
622 Mb/s	− 23	− 28	− 28	− 28	− 28
2.5 Gb/s	− 18	− 18	− 18	− 27	− 27

Upon link failure or degradation, *line switching* moves all signal channels of an entire OC-*N* channel to a protection fiber. Conversely, *path switching* can move individual payload channels within an OC-*N* channel (e.g., an STS-1 subchannel in an OC-12 channel) to another path.

Of the eight possible combinations of ring types, the following two architectures have become popular for SONET and SDH networks:

- two-fiber, unidirectional, path-switched ring (two-fiber UPSR).
- two-fiber or four-fiber, bidirectional, line-switched ring (two-fiber or four-fiber BLSR).

The common abbreviations of these configurations are given in parentheses. They are also referred to as unidirectional or bidirectional self-healing rings (USHRs or BSHRs), respectively.

Table 13.6 *Maximum nominal optical power per wavelength channel based on the total optical power being +17 dBm (see ITU-T Recommendation G.692, Ref. 23a)*

Number of wavelengths (channels)	Nominal power per channel (dBm)
1	17.0
2	14.0
3	12.2
4	11.0
5	10.0
6	9.2
7	8.5
8	8.0
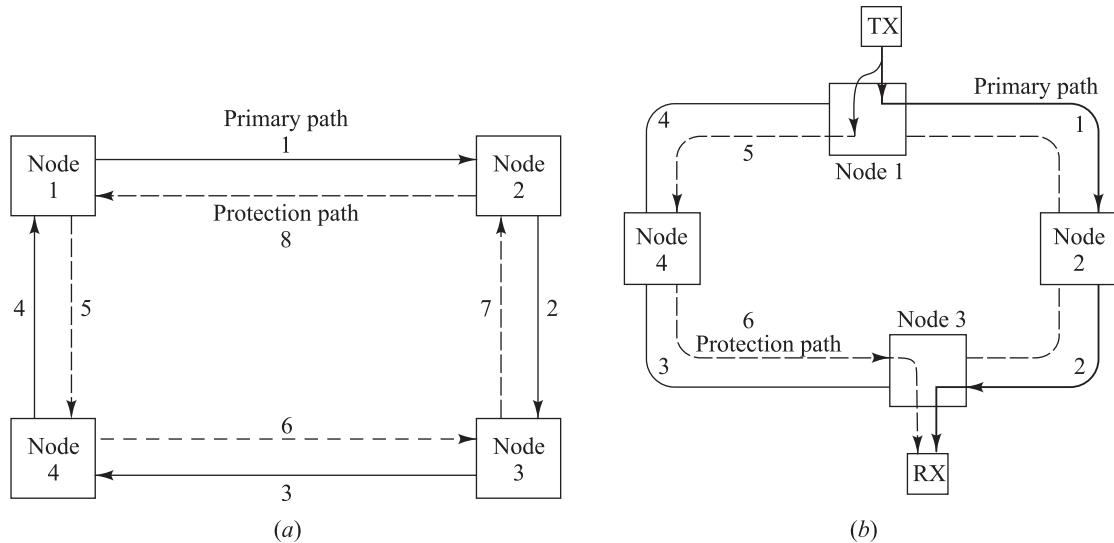
Fig. 13.11 *(a) Generic two-fiber unidirectional path-switched ring (UPSR) with a counter-rotating protection path. (b) Flow of primary and protection traffic from node 1 to node 3.*

Figure 13.11 shows a two-fiber unidirectional path-switched ring network. By convention, in a unidirectional ring the normal working traffic travels clockwise around the ring, on the *primary path*. For example, the connection from node 1 to node 3 uses links 1 and 2, whereas the traffic from node 3 to node 1 traverses links 3 and 4. Thus, two communicating nodes use a specific bandwidth capacity around the entire perimeter of the ring. If nodes 1 and 3 exchange information at an OC-3 rate in an OC-12 ring, then they use one-quarter of the capacity around the ring on all the primary links. In a unidirectional ring the counterclockwise path is used as an alternate route for protection against link or node failures. This *protection path* (using links 5 through 8) is indicated by dashed lines. To achieve protection, the signal from a transmitting node is dual-fed into both the primary and protection fibers. This establishes a designated protection path on which traffic flows counterclockwise: namely, from node 1 to node 3 via links 5 and 6, as shown in Fig. 13.11.

Consequently, two identical signals from a particular node arrive at their destination from opposite directions, usually with different delays, as denoted in Fig. 13.11*b*. The receiver normally selects the signal from the primary path. However, it continuously compares the fidelity of each signal and chooses the alternate signal in case of severe degradation or loss of the primary signal. Thus each path is individually switched based on the quality
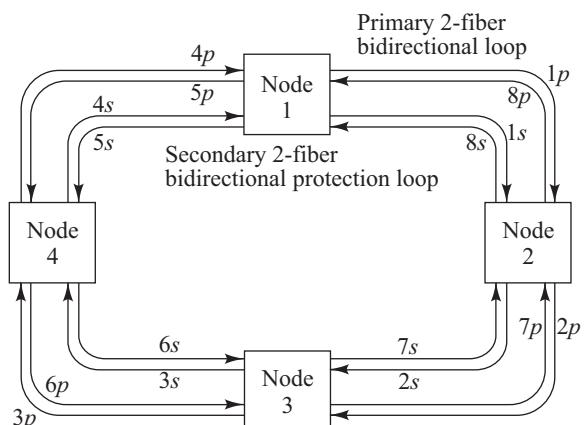
Fig. 13.12 *Architecture of a four-fiber bidirectional line-switched ring (BLSR)*
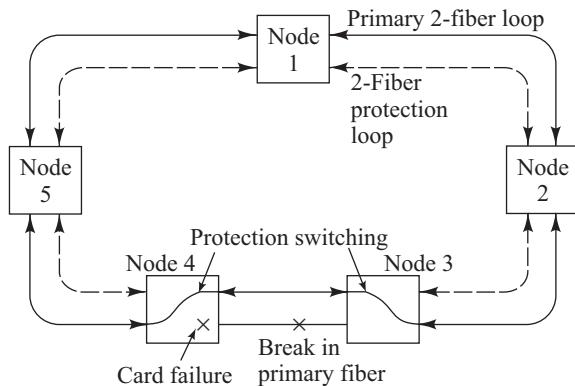
Fig. 13.13 *Reconfiguration of a four-fiber BLSR under transceiver or line failure*

of the received signal. For example, if path 2 breaks or equipment in node 2 fails, then node 3 will switch to the protection channel to receive signals from node 1.

Figure 13.12 illustrates the architecture of a four-fiber bidirectional line-switched ring. Here, two primary fiber loops (with fiber segments labeled $1p$ through $8p$) are used for normal bidirectional communication, and the other two secondary fiber loops are standby links for protection purposes (with fiber segments labeled $1s$ through $8s$). In contrast to the two-fiber UPSR, the four-fiber BLSR has a capacity advantage since it uses twice as much fiber cabling and because traffic between two nodes is sent only partially around the ring. To see this,

consider the connection between nodes 1 and 3. The traffic from node 1 to node 3 flows in a clockwise direction along links $1p$ and $2p$. Now, however, in the return path the traffic flows counterclockwise from node 3 to node 1 along links $7p$ and $8p$. Thus the information exchange between nodes 1 and 3 does not tie up any of the primary channel bandwidth in the other half of the ring.

To see the function and versatility of the standby links in the four-fiber BLSR, consider first the case where a transmitter or receiver circuit card used on the primary ring fails in either node 3 or 4. In this situation, the affected nodes detect a loss-of-signal condition and switch both primary fibers connecting these nodes to the secondary protection pair, as shown in Figure 13.13. The protection segment between nodes 3 and 4 now becomes part of the primary bidirectional loop. The exact same reconfiguration scenario will occur when the primary fiber connecting nodes 3 and 4 breaks. Note that in either case the other links remain unaffected.

Now suppose an entire node fails, or both the primary and the protection fibers in a given span are severed, which could happen if they are in the same cable duct between two nodes. In this case, the nodes on either side of the failed internodal span internally switch the primary-path connections from their receivers and transmitters

Fig. 13.14 *Reconfiguration of a four-fiber BLSR under node or fiber-cable failure*

Fig. 13.15 *Generic configuration of a large SONET or SDH network consisting of linear chains and various types of interconnected rings*

to the protection fibers, in order to loop traffic back to the previous node. This process again forms a closed ring, but now with all of the primary and protection fibers in use around the entire ring, as shown in Fig. 13.14.

13.3.4 SONET/SDH Networks

Commercially available SONET/SDH equipment allows the configuration of a variety of network architectures, as shown in Fig. 13.15. For example, one can build point-to-point links, linear chains, unidirectional path-switched rings (UPSR), bidirectional line-switched rings (BLSR), and interconnected rings. The OC-192 four-fiber BLSR could be a large national backbone network with a number of OC-48 rings attached in different cities. The OC-48 rings can have lower-capacity localized OC-12 or OC-3 rings or chains attached to them, thereby providing the possibility of attaching equipment that has an extremely wide range of rates and sizes. Each of the individual rings has its own failure-recovery mechanism and SONET/SDH network management procedures.

A fundamental SONET/SDH network element is the *add/drop multiplexer* (ADM). This piece of equipment is a fully synchronous, byte-oriented multiplexer that is used to add and drop subchannels within an OC-*N* signal. Figure 13.16 shows the functional concept of an ADM. Here, various OC-12s
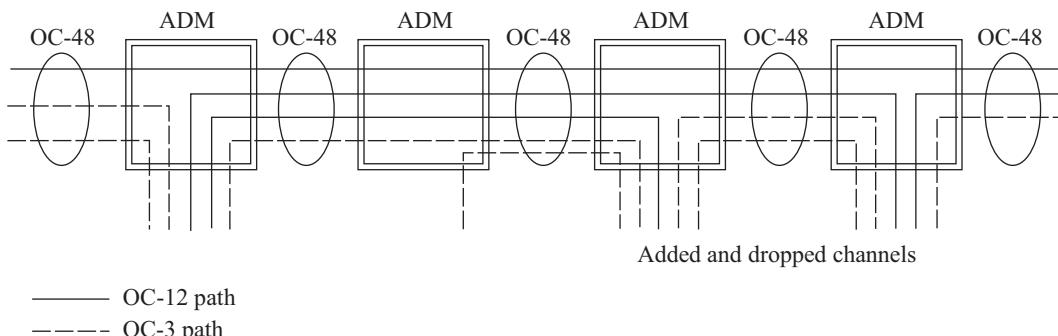
Fig. 13.16 *Functional concept of an electronic add/drop multiplexer for SONET/SDH applications*

TX: Optical transmitter
TX: Optical receiver
VA: Variable attenuator
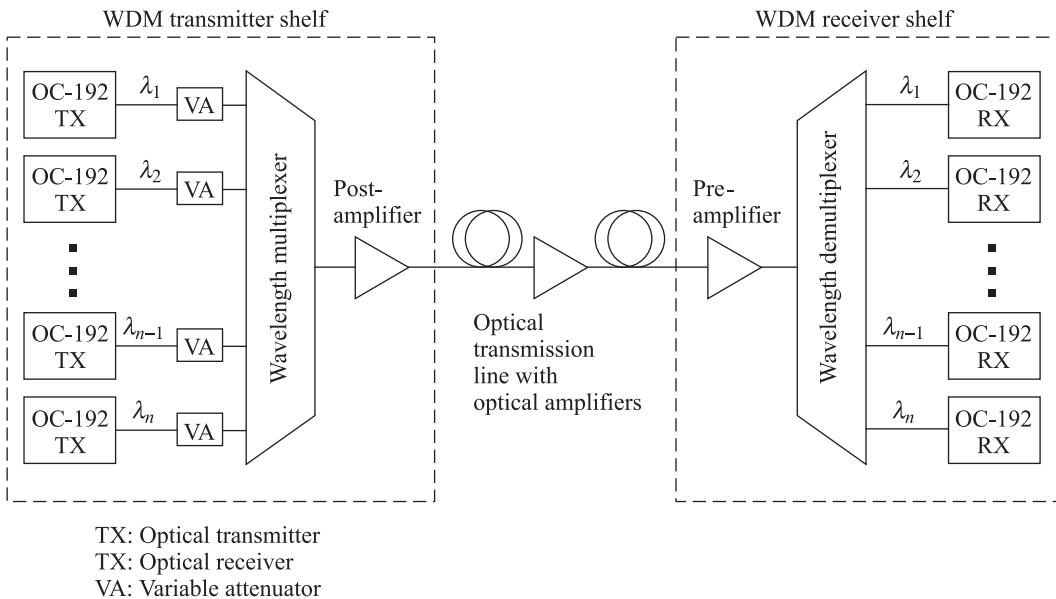
Fig. 13.17 *DWDM deployment of n wavelengths in an OC-192/STM-64 trunk ring*

and OC-3s are multiplexed into an OC-48 stream. Upon entering an ADM, these subchannels can be individually dropped by the ADM and others can be added. For example, in Fig. 13.16, one OC-12 and two OC-3 channels enter the left-most ADM as part of an OC-48 channel. The OC-12 is passed through and the two OC-3s are dropped by the first ADM. Then, two more OC-12s and one OC-3 are multiplexed together with the OC-12 channel that is passing through and the aggregate (partially filled) OC-48 is sent to another ADM node downstream.

The SONET/SDH architectures can also be implemented with multiple wavelengths. For example, Fig. 13.17 shows a dense WDM deployment on an OC-192 trunk ring for *n* wavelengths (e.g., one could have *n* = 16). The different wavelength outputs from each OC-192 transmitter are passed first through a variable attenuator to equalize the output powers. These are then fed into a wavelength multiplexer, possibly amplified by a post-transmitter optical amplifier, and sent out over the transmission fiber. Additional optical amplifiers might be located at intermediate points and/or at the receiving end.

13.4 High-Speed Lightwave Links

A challenge to creating efficient and reliable optical networks that satisfy an ever-growing demand for bandwidth is the development of high-speed optical fiber transceivers. As shown in Fig. 4.43, a variety of transceivers exist that incorporate both an optical transmitter and receiver in the same miniaturized package. For example, the *small form factor pluggable* (SFP) transceiver shown in Fig. 13.18 can be used for DWDM applications. Of interest for such devices are a *hot-pluggable* capability (meaning they can be inserted and removed from transmission equipment line cards without turning off the electric power) and the incorporation of sophisticated wavelength control into the package. Such transceivers operating at 2.5 Gb/s for DWDM applications with 100-GHz wavelength spacing are in wide use.[29] Recall from Chapter 4 that laser diodes can be modulated directly up to 2.5 Gb/s (in some cases up to 10 Gb/s), but usually need an external modulator beyond that point. Therefore,

new challenges emerge for transceivers operating at higher rates, such as 10, 40, and 160 Gb/s.

13.4.1 Links Operating at 10 Gb/s

Many 10-Gb/s optical fiber transmission systems have been studied and installed worldwide.[8,30–33] These include Fibre Channel connections for storage area networks, 10-gigabit Ethernet (known as *10-GbE* or *10GigE*) lines for local-area and metro networks, and SONET/SDH OC192/STM64 terrestrial and undersea long-haul lines and metro links. A wide selection of industry-standardized transceiver packages are on the market for these applications.[34,35]

Fig. 13.18 *A standard SFP transceiver package. (Photo courtesy of Finisar Corporation; www. finisar.com.)*

As a result of product improvement efforts, several multimode fibers with different bandwidth grades exist for 10–Gb/s use. To identify these different fiber categories, the ISO/IEC 11801 Structured Cabling Standard gives four classifications of multimode fiber in terms of bandwidth. As Table 13.7 shows, these fibers are referred to as grades OM1 through OM4. Note that the values of the bandwidth depend on the measurement standard used, and the maximum transmission distances may vary depending on the exact fiber design from different manufacturers. To summarize this table, note that

- **OM1 grade fiber** is the original (often called *legacy*) multimode fiber that was designed to be used with LEDs. Most of these legacy fibers have a 62.5-μm-diameter core, although in some cases the early installed fibers had a 50-μm core diameter. The bandwidth of these fibers is 200 MHz-km at 850 nm and 500 MHz-km at 1310 nm. The data rates with LEDs are limited to about 100 Mb/s.
- **OM2 grade fiber** has an improved bandwidth and can be used to expand networks that contain existing legacy 50-μm core diameter fiber. If only this fiber is used, then at 850 nm it is possible to send 1-Gb/s signals over about 750 m distances and 10 Gb/s over 82 m.
- **OM3 grade fiber** has a higher bandwidth and can support 10-Gb/s data rates over distances up to 300 m.
- **OM4 grade fiber** has a bandwidth of 4700 MHz-km and increases the transmission distance to 550 m when using cost-saving 850-nm VCSELs for existing 1-Gb/s and 10-Gb/s applications as well as for future 40- and 100-Gb/s Ethernet systems.

Table 13.7 *Multimode fiber classifications and their use with 1- and 10-Gb/s Ethernet*

Fiber class and size	BW @850 nm (MHz-km)	BW @1300 nm (MHz-km)	Max distance for 1 Gb/s @ 850 nm	Max distance for 1 Gb/s @ 1300 nm	Max distance for 10 Gb/s @ 850 nm
OM1 62.5/125	200	500	300 m	550 m	33 m
OM2 50/125	500	500	750 m	200 m	82 m
OM3 50/125	2000	500	950 m	600 m	300 m
OM4 50/125	4700	500	1040 m	600 m	550 m

For a short-reach 10-Gb/s network to be compliant with installation standards, all segments of the network should use the same grade of multimode fiber. However, in some situations it may be impractical or too expensive to replace existing legacy multimode fiber with a higher grade fiber. In this case most likely a link will contain a mixture of, for example, OM2 and OM3 fibers that are spliced together. If the link is to transport 10-Gb/s data, then the bandwidths of the fibers will determine the resulting *effective maximum link length*. If all geometric parameters of the interconnected OM2 and OM3 fibers are the same, then a general expression for determining the effective maximum length L_{max} is given by

$$L_{\text{max}} = L_{\text{OM2}} \, \frac{BW_{\text{OM3}}}{BW_{\text{OM2}}} + L_{\text{OM3}} \qquad (13.17)$$

where L_{OMx} and BW_{OMx} are the length and bandwidth, respectively, of the OMx grade fiber. To be compliant with the maximum allowed dispersion, the effective maximum link length calculated by Eq. (13.17) must be less than the achievable link length if only OM3 fiber is used. For standard OM3 fiber this length is 300 m. With OM4 fibers, the maximum length is 550 m at an 850-nm wavelength for operation at 10 Gb/s.

Example 13.7   An engineer wants to create a link consisting of 40 m of OM2 fiber that has a 500-MHz bandwidth and 100 m of OM3 fiber that has a 2000-MHz bandwidth. What is the effective maximum link length?

Solution: From Eq. (13.17), the maximum link length is

$L_{\text{max}} = (40 \text{ m})(2000/500) + 100 \text{ m} = 260 \text{ m}$

Since the calculated length is less than 300 m, this link would comply with the installation standard.

If 10-Gb/s transmission is attempted over legacy OM1 multimode fibers, the signal is restricted to distances of only about 33 m. To achieve 300-m transmission distances at 10-Gb/s over legacy multimode fibers, an alternative scheme is to send four 3.125-Gb/s data streams over four different wavelengths. A number of manufacturers offer 850-nm VCSEL sources that can be modulated directly at 10 Gb/s. Both the transmitter and receiver can be housed in the same package.

For access network applications ranging from 7 to 20 km, which the 10-GbE specification calls *long-reach* (LR), the links need to use InGaAsP-based distributed-feedback (DFB) lasers operating at 1310 nm over single-mode fibers. These links operate near the 1310-nm dispersion minimum of G.652 single-mode fibers and the light sources can be modulated directly.

For metro network applications ranging from 40 to 80 km, which the 10-GbE specification calls *extended-reach* (ER), the links need to use externally modulated distributed-feedback (DFB) lasers operating at 1550 nm over single-mode fibers. A number of vendors offer a variety of transceiver packages for both LR and ER applications (see Fig. 4.43). Three of the several configurations include the 300-pin, XFP, and SFP modules (also see Fig. 4.43). Figure 13.19 shows the relative sizes of these specific transceivers. The larger modules, such as the 300-pin device, contain other internal electronics besides the normal circuitry used for the light source and photodetection operations. These electronics include clocking functions for transmit and receive signals, a line loopback capability for diagnostic signals, and circuitry for multiplexing sixteen 622-Mb/s SONET/SDH inputs into a 10-Gb/s stream, which becomes the electric signal input to the laser. In the receive direction, a photodetector changes the incoming 10-Gb/s optical signal to an electrical format, which then is demultiplexed back into sixteen 622-Mb/s electrical

signals. External electronics need to provide these functions for the smaller transceiver modules.

13.4.2 Links Operating at 40 Gb/s

New challenges in terms of transceiver response characteristics, chromatic dispersion control, and polarization-mode dispersion compensation arise when transitioning to higher capacity links, such as 40 Gb/s data rates.[36–40] For example, compared to a 10-Gb/s system, when using a conventional on-off keying (OOK) modulation format a link operating at 40 Gb/s is sixteen times more sensitive to chromatic dispersion, is four times more sensitive to polarization-mode dispersion, and needs an optical signal-to-noise ratio (OSNR) that is at least 6 dB higher to reach an equivalent bit-error rate (BER).

Fig. 13.19 *Top-view size comparison of 300-pin, XFP, and SFP optical transceivers*

Therefore, alternate modulation schemes besides OOK have been considered. One method is *differential binary phase-shift keying*, which is known as DBPSK or simply DPSK.[41,42] One advantage of DPSK is that when using a balanced receiver the OSNR needed to reach a specific BER is about 3 dB lower than that required for OOK. (A *balanced receiver* uses a pair of matched photodiodes to achieve a higher sensitivity.) Since a factor of 3 dB means only half the amount of optical power is needed for an equivalent BER, the lower OSNR that is needed for DPSK can be allocated to reducing the optical power level at the receiver, to relaxing the loss specifications for various link components, or to extending the transmission distance. For example, the transmission distance could be doubled if there are no other signal impairments, such as additional nonlinear effects. In addition, DPSK is quite resilient to nonlinear effects, which are a problem when using OOK for data rates greater than 10 Gb/s. This resilience is due to the fact that compared to OOK the optical power is more evenly distributed in DPSK (that is, there is optical power in each bit slot, as the next paragraph describes) and the optical peak power is 3 dB lower for the same average optical power. These factors reduce nonlinear effects that are dependent on the bit pattern and on optical power levels.

In contrast to OOK where the information is conveyed through amplitude changes, DPSK carries the information in the optical phase. Optical power appears in each DPSK bit slot and can occupy the entire slot for NRZ-DPSK, or it can occupy part of the slot in the form of a pulse in a RZ-DPSK format. Binary data is encoded as either a 0 or a π optical phase shift between adjacent slots. For example, the information bit 1 may be transmitted by a 180° carrier phase shift relative to the carrier phase in the previous slot, while the information bit 0 is transmitted by no phase shift relative to the carrier phase in the previous signaling interval.

Various long-haul terrestrial and undersea field trials have examined the feasibility of using different types of DPSK modulation formats to enable 40-Gb/s transmission over conventional G.652 nonzero-dispersion-shifted fibers (NZDSF). Among these are RZ-DPSK, chirped return-to-zero (CRZ) DPSK, carrier-suppressed return-to-zero (CSRZ) DPSK, and RZ differential quadrature phase-shift keying (RZ-DQPSK). The most widely accepted format has been RZ-DPSK, for which transceiver modules that can interface to SONET-678/SDH-256 equipment are available.

13.4.3 Standards for 40-Gigabit and 100-Gigabit Ethernet

Ethernet has become a worldwide technology for enabling telecommunications via the Internet and other packet-switched networks using the Internet Protocol (IP). Owing to Ethernet's low cost, proven reliability,
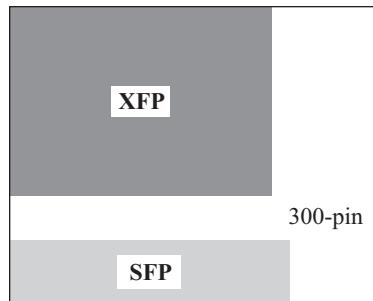
and operational simplicity, currently most Internet traffic starts or ends on an Ethernet connection. By 2006 Ethernet rates of 10 Gb/s were already being used. This technology was referred to as *10 Gigabit Ethernet*. However, the constantly growing demand for more bandwidth was starting to challenge the capabilities of networks employing 10 Gigabit Ethernet links. Starting in 2006 preliminary studies determined that two new Ethernet rates were needed. These were one specification of 40 Gb/s for server and computing applications and another specification of 100 Gb/s for network aggregation applications. Consequently, in January 2008 the IEEE P802.3ba 40 Gb/s and 100 Gb/s Task Force was formed to develop a 40 Gigabit Ethernet standard and a 100 Gigabit Ethernet standard. Included in the standards effort was the development of physical layer specifications for communication across backplanes, copper cabling, multimode fiber, and single-mode fiber. The target date for the ratification of the standards is June 2010.

13.4.4 OTDM Links Operating at 160 Gb/s

To determine the possibility of sending information at data rates even higher than 40 Gb/s, initiatives were started for testing systems that can operate at 160 Gb/s over a single wavelength using installed standard G.652 single-mode fibers. These test links used the concept of *optical time-division multiplexing* (OTDM) to form a 160-Gb/s data stream, since electronic devices that are needed for carrying out signal processing at these rates were not available.

One option is to use *bit-interleaved* OTDM.[43–46] This multiplexing technique is similar to WDM in that the access nodes share many small channels operating at a peak rate that is a fraction of the media rate. For example, the individual channel rates could vary from 2.5 to 40 Gb/s, whereas the time-multiplexed media rate can be up to 160 Gb/s, depending on the pulse widths of the multiplexed data streams. Figure 13.20 illustrates the basic concept of point-to-point transmission using bit-interleaved optical TDM. A laser source produces a regular stream of very narrow return-to-zero optical pulses at a repetition rate *B*. This rate can range from 2.5 to 40 Gb/s, which corresponds to the bit rate of the electronic data tributaries feeding the system. An optical splitter divides the pulse train into *N* separate streams. In the example in Fig. 13.20, the pulse stream is 10 Gb/s and $N = 4$. Each of these channels can be modulated independently by an electrical tributary data source at a bit rate *B*. The modulated outputs are delayed individually by
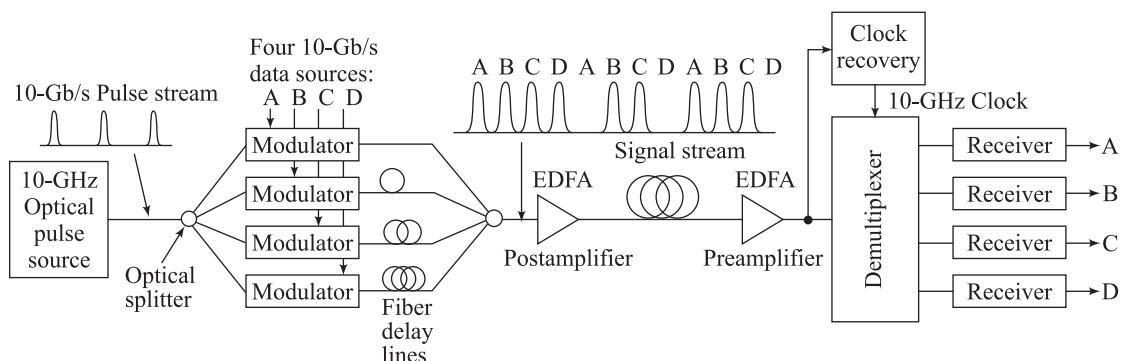
Fig. 13.20 *Example of an ultra-fast point-to-point transmission system using optical TDM. (Adapted with permission from Cotter, Lucek, and Marcenak,[44] © IEEE, 1997.)*

different fractions of the clock period, and are then interleaved through an optical combiner to produce an aggregate bit rate of $N \times B$.

Optical postamplifiers and preamplifiers generally are included in the link to compensate for splitting and attenuation losses. At the receiving end, the aggregate pulse stream is demultiplexed into the original N independent data channels for subsequent electronic signal processing. In this technique, a clock-recovery mechanism operating at the base bit rate B is required at the receiver to drive and synchronize the demultiplexer.

Several field trials have demonstrated the feasibility of long-haul 160-Gb/s transmission systems. An interesting point to note about these 160-Gb/s experiments is that good performance was obtained using installed standard G.652 single-mode fibers.

- Researchers in Germany achieved repeaterless error-free transmissions of 1×170-Gb/s signals over 185 km and 8×170-Gb/s signals (that is, eight WDM signals) over 140 km of *standard single-mode fiber* (SSMF) using return-to-zero DPSK (RZ-DPSK) modulation.[47] The 170-Gb/s signal was created by time-interleaving four channels operating at 42.7 Gb/s each. The system consisted of a cascade of links through seven cities and used hybrid EDFA and Raman optical amplification at intermediate nodes.

- To cope with transmission impairments from CD and PMD, researchers in Japan investigated the use of 2-bit/symbol encoding techniques such as differential quadrature phase shift keying (DQPSK) and simultaneous amplitude shift keying (ASK) and DPSK. The 160-Gb/s signal was composed of eight 20-Gb/s channels. Relatively stable BER characteristics were obtained after transmissions over 200 km of installed G.652 single-mode fiber.[48]

- In a 160-Gb/s experiment carried out in the United Kingdom by European researchers, the impact of chromatic and polarization-mode dispersions were examined on 275- and 550-km links of installed SSMF. The 160-Gb/s signal was created by time-interleaving sixteen channels operating at 10 Gb/s each. The experiments showed excellent operation of clock recovery, bit-error rate, and the functions of dropping and adding wavelength channels (see Sec. 13.5).[49]

13.5 Optical Add/Drop Multiplexing

An *optical add/drop multiplexer* (OADM) is a device that allows the insertion or extraction of one or more wavelengths from a fiber at a network node. For example, an OADM might have the capability to drop and insert three wavelengths from a set of N being transported over a single fiber. The remaining $N - 3$ wavelengths pass unaffected through (also called *express through*) the OADM to the next node. Without an OADM, if only three wavelengths out of $N \gg 3$ are needed at the local node, then all the other wavelengths also would need to be processed there by means of optoelectronic transceivers. This will greatly increase the equipment costs at that node. Thus the advantage of the add/drop function is that no signal processing is required for the set of wavelengths that passes straight through the OADM.

The OADM can reside at an optical amplifier site in a long-haul network or it can be located at a node in a metro network. Depending on how it is designed, an OADM can operate either statically in a fixed add/drop configuration or it can be reconfigured dynamically from a remote network management site. A fixed optical add/drop multiplexer is simply referred to as an OADM. A dynamic device is called a *reconfigurable OADM* (ROADM).[50–54] A fixed OADM obviously is not as flexible as a ROADM and may require a change of hardware components if a different set of wavelengths needs to be dropped or added.

Depending on whether an engineer is designing a metro or a long-haul network, different performance specifications need to be addressed when implementing an optical add/drop capability in the network. In general, because of the nature of the services provided, changes in the add/drop configuration for a long-haul network occur less frequently than in a metro network where there tends to be a high turnover rate in service requests and in wavelengths being transported. In addition, compared to a metro network the wavelength spacing usually is much narrower in a long-haul network and the optical amplifiers used must cover a wider spectral band. Section 13.7 looks at some examples of this.

13.5.1 OADM Configurations

A number of different OADM configurations are possible.[14,15,50] Most OADMs are constructed using the WDM elements described in Chapter 10. These can be a series of dielectric thin-film filters, an arrayed waveguide grating (AWG), a set of liquid crystal devices, or a series of fiber Bragg gratings used in conjunction with optical circulators. The architecture selected for a particular application depends on implementation factors such as the number of wavelengths to be dropped and added at a node, the desired modularity of the OADM (e.g., how easy is it to upgrade this device), and the constraint of whether random individual wavelengths or a neighboring group of wavelengths should be processed. Here we assume that N wavelengths traveling on a single fiber enter the OADM and a subset M of them is processed at the node. After processing, these M wavelengths are reinserted onto the fiber line to rejoin the through-going $N - M$ channels.

In the configuration illustrated in Fig. 13.21, all N incoming wavelengths are separated into individual channels at the OADM input by means of a wavelength demultiplexer. This gives a very versatile architecture, since any of the N wavelengths can be dropped, processed at the node, and then reinserted onto the outgoing fiber by means of a wavelength multiplexer. In the illustration given here, M wavelengths are dropped and the remaining $N - M$ channels individually pass through the OADM. The M dropped wavelengths are labeled λ_i through λ_k to indicate that any combination of M wavelengths selected from the N incoming light channels can be dropped. Such a configuration is useful if a large fraction of the N wavelengths needs to be dropped and added. However, it is not cost-effective for dropping and adding a small fraction of the incoming wavelengths.

Figure 13.22 shows a more modular OADM architecture. Here the N incoming light channels are split into several *bands of wavelengths*. This function can be achieved using either a set of thin-film
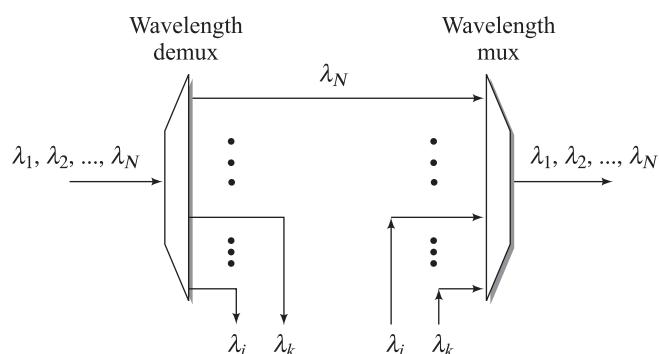
Fig. 13.21 *Simple passive optical add/drop multiplexer*
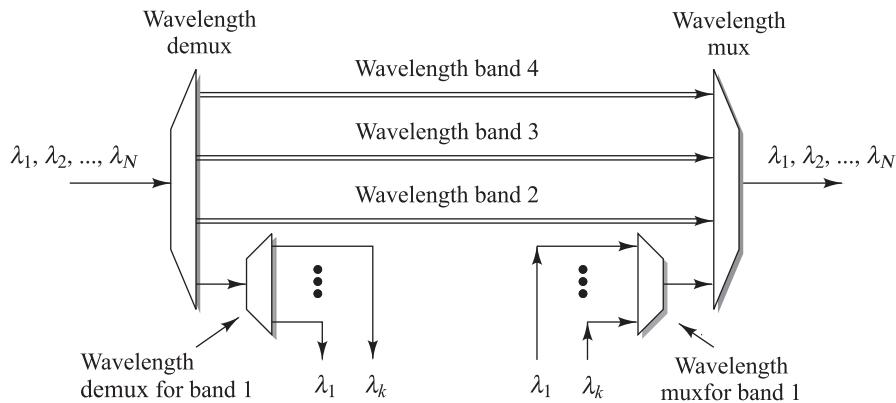
Fig. 13.22 *Modular expandable passive OADM architecture*

filters or an AWG. The band to be dropped can be sent through a second demultiplexer followed by processing of the individual wavelengths. As an example, if $N = 12$ and three wavelengths need to be dropped, the incoming wavelengths can be divided into four bands of three wavelengths each. Added lightwave channels go through two stages of multiplexing to rejoin the pass-through bands. An advantage of this approach is that for future OADM upgrades to drop another wavelength band at this node, network engineers can incorporate another second-stage demultiplexer to handle the next desired wavelength band.

13.5.2   Reconfigurable OADM

The task of reconfiguring a fixed OADM manually may require several days of planning and implementation of hardware changes. In contrast, the use of ROADMs gives service providers the ability to reconfigure a network within minutes from a remote network-management console. The dynamic flexibility to drop and add selected wavelengths quickly at a particular node upon customer demand for new or expanded services is known popularly as *service provisioning on the fly*. This is particularly important in metro networks where changes in service requests tend to be significantly more dynamic than in long-haul networks and clients expect very rapid responses to their requests. Such service requests can have fairly diverse origins such as time-varying business applications, on-demand entertainment, and emergency or disaster-response communications.

A variety of ROADM architectures are possible. In this section we look at three types. These are based on *wavelength blockers, arrays of small switches*, and *wavelength-selective switches*. Section 13.6 describes the concept of an optical crossconnect, which uses a more complex architecture of multiple wavelength-selective switches. Each ROADM type can have a number of alternative configurations and different operational characteristics. The design of a particular ROADM depends on factors such as cost, reliability, technology maturity, desired network operating flexibility, and upgrade capability of the equipment. Before going into architecture details, let us first look at some desired features and terminology. These include:

- **Wavelength dependence.** If the architectures are wavelength-dependent they are said to be *colored* or have *colored ports*. When the ROADM operation is independent of wavelength, it is referred to as being *colorless* or having *colorless ports*.

- **ROADM degree.** The *degree* of a ROADM refers to the number of bidirectional multiwavelength interfaces the device supports. Thus a degree-2 ROADM has two bidirectional WDM interfaces and a degree-4 ROADM can support up to four bidirectional WDM interfaces, for example, in north, south, east and west directions.

- **Remote reconfiguration.** The ability to change the ROADM configuration from a remotely located network-management workstation is an important feature, since it greatly reduces operations costs by eliminating the need for a service person to visit the ROADM site in order to manually upgrade the device.

- **Express channels.** The ability to have *express channels* that allow a selected set of wavelengths to pass through the node without the need for optical-to-electrical-to-optical conversion saves the expense of having optical transceivers for those wavelengths.

- **Modular expansion.** To avoid an intial high setup cost to connect transmitters and receivers to each add/drop port, service providers usually first activate the minimum number of ports needed to support current traffic and then later add more channels as the service demand increases. This is known popularly as the *pay-as-you-grow approach*.

- **Minimum optical impairment.** Having an express wavelength feature requires a careful engineering design to avoid the accumulation of optical signal impairments as a particular wavelength set passes through several sequential ROADMs. These include effects such as crosstalk between channels, wavelength-dependent attenuation, ASE noise, and polarization-dependent loss.

Wavelength Blocker Configuration

Figure 13.23 shows the simplest ROADM configuration, which uses a *broadcast-and-select approach*.[50,51] In this degree-2 architecture a passive optical coupler splits the incoming lightwave signal power into two paths. One branch is an express path and the other branch is diverted to a drop site. Located in the express path is a device called a *wavelength blocker* that can be configured to block those wavelengths that are to be received at the node. The drop segment contains a $1 \times N$ optical power splitter, which divides and directs the light signal into N tunable filters that allow selection of any desired wavelength. The add segment contains N tunable laser sources that allow
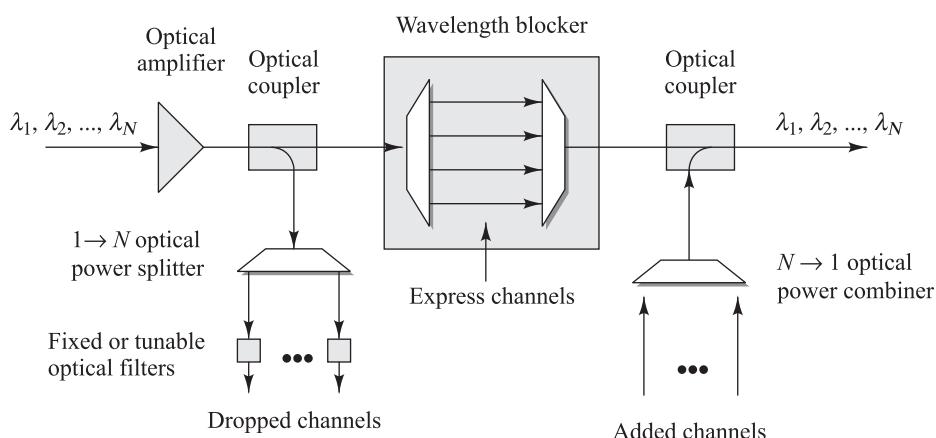
Fig. 13.23 *Example of a ROADM based on the use of a wavelength blocker*
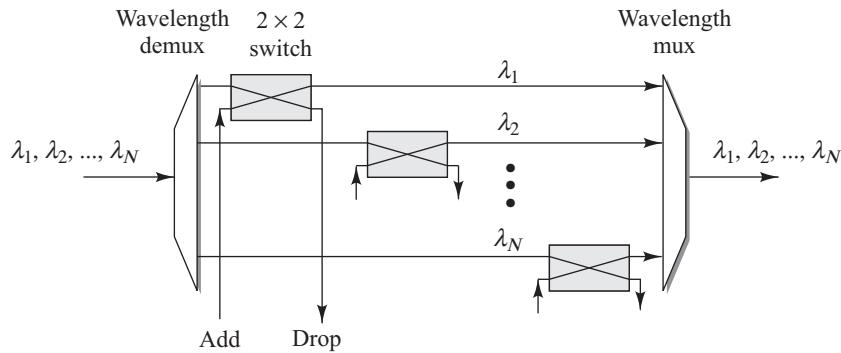
Fig. 13.24 *A colored (wavelength-dependent) switching-based ROADM*

the inserted wavelengths to join the express wavelengths on the main fiber line by means of an $N \times 1$ optical power combiner and another passive optical coupler. An attractive feature of this ROADM is that initially only the two passive couplers and the wavelength blocker are needed. Later different sizes of optical power splitters and combiners can be selected to drop and add other desired wavelengths, which now need to be prevented from traversing the express path by proper settings of the wavelength blocker. One drawback of this ROADM is that the losses of the optical power splitters and combiners increase as their size grows, that is, as the number of wavelengths to be dropped and added increases.

Switch Array Configuration

Figure 13.24 shows a colored *switching-based ROADM* configuration, which is also known as a *demux-switch-mux approach*. Here the N incoming wavelengths first pass through a demultiplexer. Individual 2×2 or 1×2 switches then allow each wavelength to either bypass the node or be dropped. As Fig. 13.25 shows, this wavelength can be added back onto the outgoing fiber using either the same 2×2 switch as used by the incoming wavelength or a 2×1 switch analogous to the drop device. After the N lightwaves emerge from the 2×2 or 2×1 switches, a wavelength multiplexer combines them onto the outgoing fiber.

In a less versatile variation of this colored architecture, the ROADM may be designed to express through a certain set of K wavelengths, say λ_1 through λ_K, without using switches for them. One

(*a*) Add-and-drop setting   (*b*) Express setting
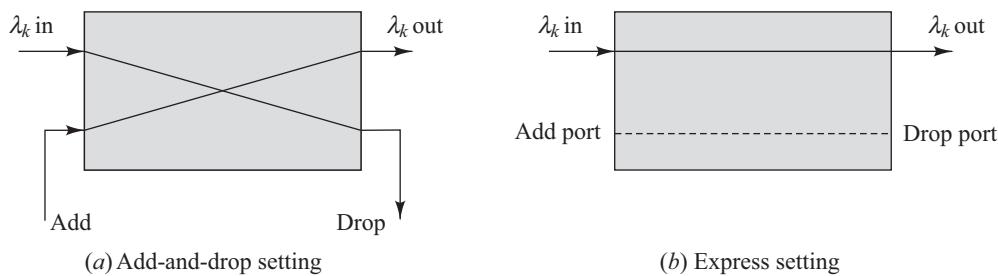
Fig. 13.25 *Internal connections of a 2 × 2 switch for (a) add/drop and (b) express functions*
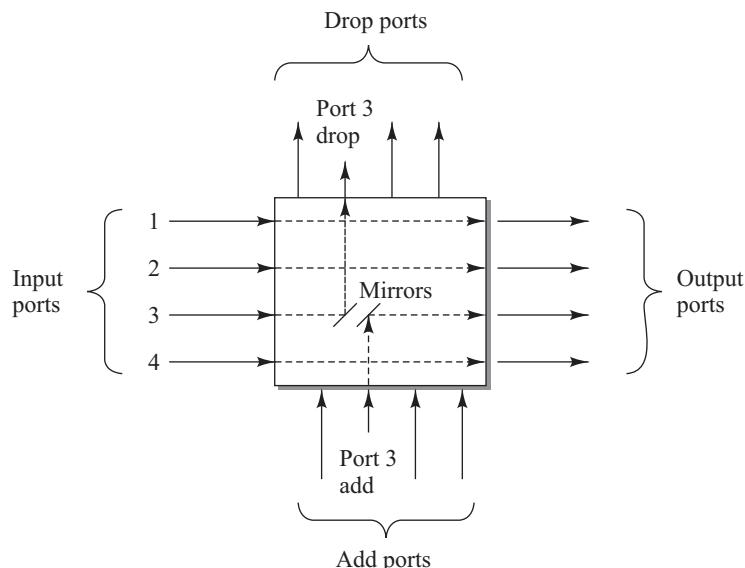
Fig. 13.26 *Example of ROADM based on a 4 × 4 array of MEMS mirrors*

implementation alternative for this is to replace the set of 2×2 switches with an $N \times M$ switch that has N ports for N incoming wavelengths and M add and drop ports. This could be accomplished with an $N \times M$ array of MEMS (micro-electro-mechanical system) mirrors. Figure 13.26 gives an example for a 4×4 configuration. To drop and add a wavelength, a miniature mirror in the switch path of that channel is set at an angle to divert the incoming light to a drop port and to insert the same wavelength via an add port. Figure 13.26 shows this for port 3.

In either case (full or partial add/drop capability), a limitation of the switching-based ROADM architecture is that it can be costly and complex to implement. If the ROADM is designed to have a full add/drop capability of all wavelengths, then the transceivers of all express wavelengths will not be used.
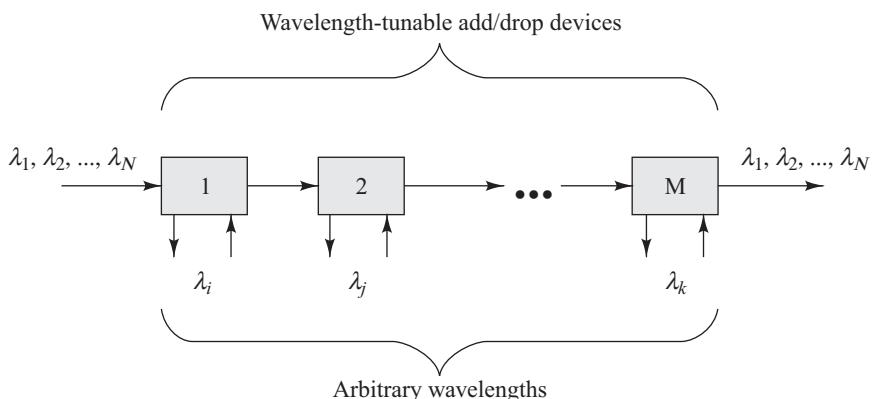
Fig. 13.27 *Wavelength-tunable ROADM for adding and dropping small channel counts*

In case of a partial add/drop capability, a smaller set of transceivers is needed, but careful planning is required to determine what the $K = N - M$ express wavelengths should be and the ROADM implementation becomes constrained.

Figure 13.27 shows a variation of the switching-based ROADM architecture, which is colorless and has more design and implementation flexibility. In contrast to the parallel architecture shown in Fig. 13.24, this is a serial configuration. This configuration has a series of $M \leq N$ wavelength-tunable devices that allow any single wavelength to be added or dropped by any device. Here M is the maximum number of wavelengths that can be used at the node. Note that in this *tuning-based ROADM* architecture, there is no constraint on which particular wavelengths are express or add/drop channels, since the tunable elements are colorless; that is, they can be set to any wavelength value. However, this configuration usually is limited to a small number of drop channels; otherwise the accumulated optical loss through the series of switching elements may be fairly large.

Wavelength-Selective Switch Configuration

In the evolution of add/drop multiplexer technology to enable more sophisticated mesh network applications, *wavelength-selective switch* (WSS) technology was introduced.[54–56] A key feature of a WSS is that it can direct each wavelength entering a common input port to any one of a multiple number of output ports, as indicated generically in Fig. 13.28. A basic ROADM is formed using one WSS module for dropping wavelengths and another for adding them. Each module contains a set of wavelength selective switches. To designate the size of such a switch, it is labeled as a $1 \times M$ switch, where M indicates the number of output ports to which a wavelength can be directed. If there are N incoming wavelengths that are to be switched to any of M output ports, then one WSS module contains N wavelength selective switches of size $1 \times M$. Available WSS configurations can send wavelengths to between four and ten ports.

Figure 13.29 illustrates the switching concept using 16 input wavelengths (λ_1 through λ_{16}), which can be directed to any of four output lines with sixteen 1×4 wavelength selective switches. These four output lines include one express port for through-going wavelengths and three drop ports. In Fig. 13.29 the incoming wavelengths are divided into individual channels by a demultiplexer and then pass through a variable optical attenuator (VOA). The function of the VOAs is to ensure that all wavelengths leaving the ROADM have the same optical power level. The 1×4 WSS following a VOA either sends a wavelength to the express port or switches it to any of the three drop ports. A versatile feature is that any collection of wavelengths can be switched to a given drop port simultaneously, thereby creating an *any-to-any switching capability*. Since each ROADM port has a multiple-wavelength capability, this increases the degree of a WSS-based ROADM to more than 2.
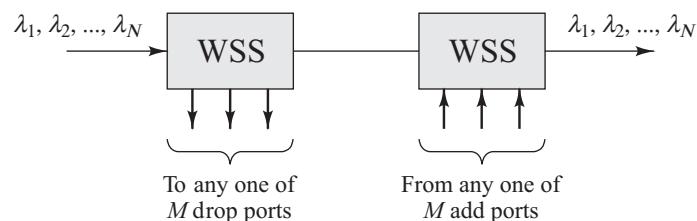
Fig. 13.28 *Flexible ROADM architecture based on a pair of wavelength-selective switches*
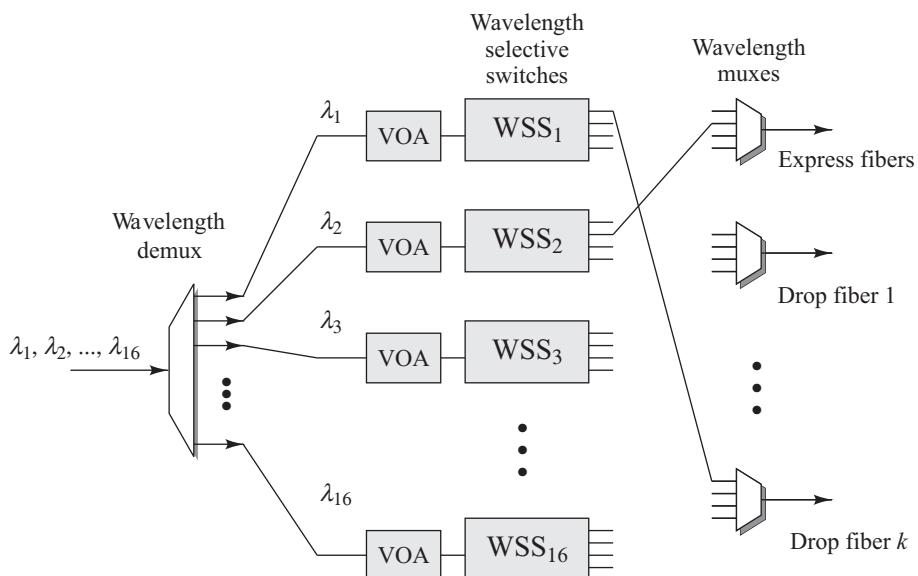
Fig. 13.29 *WSS-based ROADM that can switch any combination of sixteen input wavelengths simultaneously to any output port*

13.6 Optical Switching

The appearance of enhanced multimedia services requiring huge bandwidths, such as broadcast high-definition television (HDTV) and video-on-demand, created a need for transitioning from basic interconnected rings to extremely high-capacity rings adjoined to mesh networks that can support clusters of up to 50 nodes in a metropolitan area. Therefore, another network element with more sophisticated switching capabilities than a ROADM is needed. This element, which is called an *optical crossconnect* (OXC), provides switched pass-through paths for express traffic that does not terminate at the node and an interface for dropping and adding optical signals at the node. The express traffic can be switched from any input to any output fiber. Internally such a device can have either an electrical or optical switching fabric. To understand basic optical switching technology, Sec. 13.6.1 looks at general OXC configurations, Sec. 13.6.2 considers the performance impact when wavelength conversion is used, and Sec. 13.6.3 describes the standard implementations of wavelength routing or optical circuit switching.

As network traffic volume rises, particularly from the desire to have high-bandwidth multimedia services, the number of wavelengths per fiber will increase. This means that changes are needed for the earlier switching methods in which optical signals are converted to electrical signals, switched electronically, and then converted back to an optical format. The main reason is that the performance of the electronic equipment used in this *optical-to-electrical-to-optical* (O/E/O) conversion process is strongly dependent on the data rate and protocol. To overcome these limitations, the concept of all-optical switching was explored. Two approaches to this concept are optical burst switching (OBS) and optical packet switching (OPS). Sections 13.6.4 and 13.6.5, respectively, discuss these concepts.

13.6.1 Optical Crossconnect

A high degree of path-configuration modularity, capacity scaling, and flexibility in adding or dropping channels at a client site can be achieved by introducing the concept of an optical crossconnect switch in the physical path structure of an optical network. These *optical crossconnect* (OXC) switches sit at junctions in ring and mesh networks that are interconnected through many hundreds of optical fibers, each carrying dozens of wavelength channels. At such a junction an OXC can dynamically route, set up, and take down very high-capacity lightpaths.[57–60]

To visualize the operations of an OXC, consider first the OXC architecture shown in Fig. 13.30 that uses a switching matrix for directing the incoming wavelengths that arrive on a series of *M* input fibers. The switch matrix can operate electrically or optically; that is, it can switch the incoming signals in either the electric or the optical domain. Each of the *M* input fibers carries *N* wavelengths, any or all of which could be added or dropped at a node. For illustration simplicity, here we let $M = 2$ and $N = 4$. At each fiber input, the arriving aggregate of *N* signal wavelengths is demultiplexed and then enters the switch matrix. The switch matrix directs the channels either to one of the eight output lines if it is a through-traveling signal, or to a particular receiver attached to the OXC at output ports 9 through 12 if it has to be dropped to a user at that node. Signals that are generated locally by a user get connected electrically via the *digital crossconnect matrix* (DXC) to an optical transmitter. From here they enter the switch matrix, which directs them to the appropriate output line. The *M* output
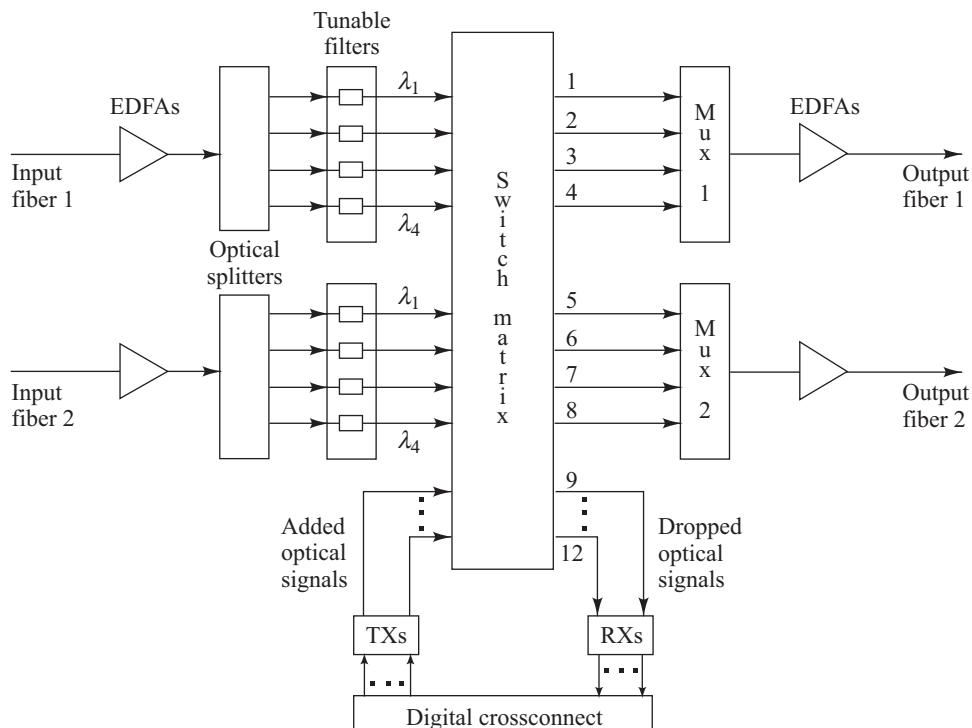
Fig. 13.30 *Optical crossconnect architecture using optical space switches and no wavelength converters*

Example 13.8   Consider the 4 × 4 OXC shown in Fig. 13.31. Here two input fibers are each carrying two wavelengths. Either wavelength can be switched to any of the four output ports. The OXC consists of three 2 × 2 switch elements. Suppose that λ_2 on input fiber 1 needs to be switched to output fiber 2 and that λ_1 on input fiber 2 needs to be switched to output fiber 1. This is achieved by having the first two switch elements set in the bar state (the straight-through configuration) and the third element set in the cross state, as indicated in Fig. 13.31. Without wavelength conversion there would be wavelength contention at both output ports. By using wavelength converters, the cross-connected lightwaves can be prevented from contending for the same output fiber.
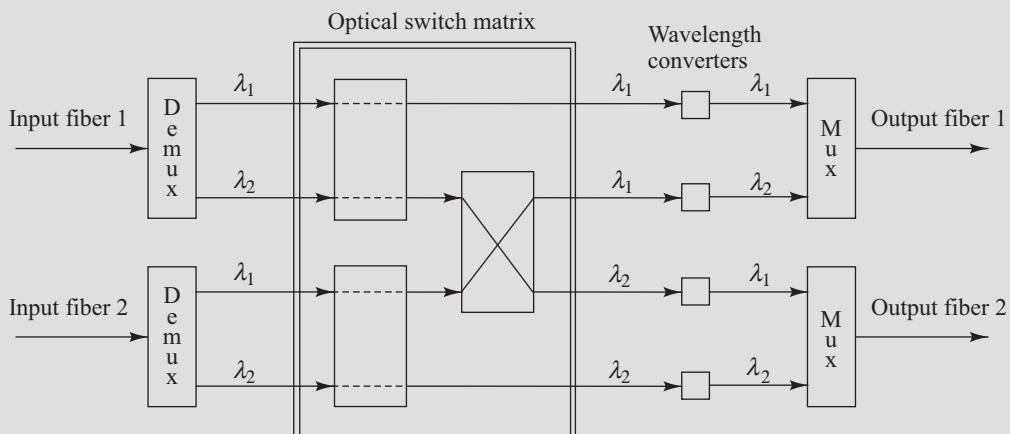
Fig. 13.31   *Example of a simple 4 × 4 optical crossconnect architecture using optical space switches and wavelength converters*

lines, each carrying separate wavelengths, are fed into a wavelength multiplexer to form a single aggregate output stream. An optical amplifier to boost the signal level for transmission over the trunk fiber normally follows this.

Contentions could arise in the architecture shown in Fig. 13.30 when channels having the same wavelength but traveling on different input fibers enter the OXC and need to be switched simultaneously to the same output fiber. This could be resolved by assigning a fixed wavelength to each optical path throughout the network, or by dropping one of the incoming channels at the node and retransmitting it at another wavelength. However, in the first case, wavelength reuse and network scalability (expandability) are reduced, and in the second case the add/drop flexibility of the OXC is lost. These blocking characteristics can be eliminated by using wavelength conversion at any output of the OXC, as shown in the example above.

13.6.2   Wavelength Conversion

Numerous studies have been carried out to quantify the benefits of wavelength conversion.[61–63] These efforts employed either probabilistic models or used deterministic algorithms on specific network topologies. The studies indicate that the benefits are greater in a mesh network than in a ring or fully connected network.

To illustrate the effect of wavelength conversion, we show a simple model that is based on standard series independent-link assumptions commonly used in circuit-switched networks.[58,61] In this simplified example, during a request for establishing a lightpath connection between two stations, the usage of a

wavelength on a fiber is statistically independent of other fiber links and other wavelengths. Although this model tends to overestimate the probability that a wavelength is blocked along a path, it provides insight into the network performance improvement when using wavelength conversion.

Assume that there are H links (or hops) between two nodes that need to be connected, which we will call nodes A and B. Take the number of available wavelengths per fiber link to be F, and let ρ be the probability that a wavelength is used on any fiber link. Then, since ρF is the expected number of busy wavelengths on any link, ρ is a measure of the *fiber utilization* along the path.

First, consider a network with wavelength conversion. In this case, a connection request between nodes A and B is blocked if one of the H intervening fibers is full; that is, the fiber is already supporting F independent sessions on different wavelengths. Thus, the probability P_b' that the connection request from A to B is blocked is the probability that there is a fiber link in this path with all F wavelengths in use, so that

$$P_b' = 1 - (1 - \rho^F)^H \tag{13.18}$$

If q is the *achievable utilization* for a given blocking probability in a network *with* wavelength conversion, then

$$q = [1 - (1 - P_b')^{1/H}]^{1/F} \approx \left(\frac{P_b'}{H} \right)^{1/F} \tag{13.19}$$

where the approximation holds for small values of P_b'/H. Figure 13.32 shows the achievable utilization q for $P_b' = 10^{-3}$ as a function of the number of wavelengths for $H = 5$, 10, and 20 hops. The effect of path length is small, and q rapidly approaches 1 as F becomes large.

Now consider a network without wavelength conversion. Here, a connection request between A and B can be honored only if there is a free wavelength; that is, if there is a wavelength that is unused on each
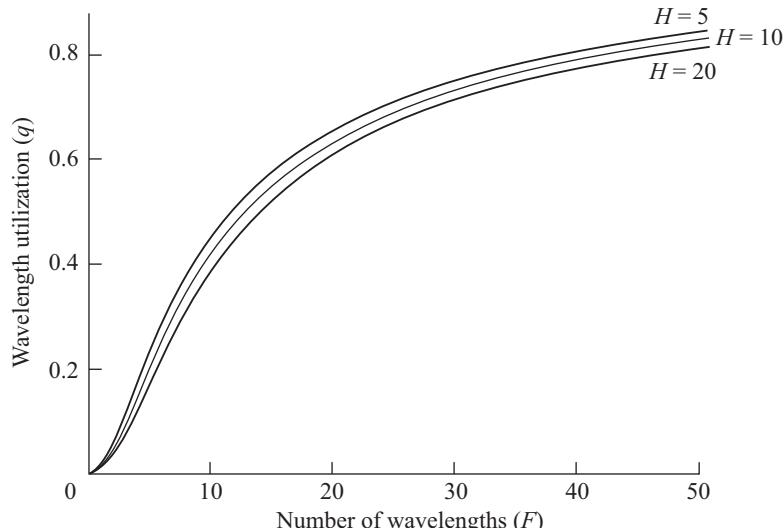
Fig. 13.32 *Achievable wavelength utilization as a function of the number of wavelengths for a 10^{-3} blocking probability in a network using wavelength conversion. (Reproduced with permission from Barry and Humblet,[61] © IEEE, 1996.)*

of the H intervening fibers. Thus, the probability P_b that the connection request from A to B is blocked is the probability that each wavelength is used on at least one of the H links, so that

$$P_b = [1 - (1 - \rho)^H]^F \tag{13.20}$$

Letting p be the *achievable utilization* for a given blocking probability in a network without wavelength conversion, then

$$p = 1 - (1 - P_b^{1/F})^{1/H} \approx -\frac{1}{H} \ln (1 - P_b^{1/F}) \tag{13.21}$$

where the approximation holds for large values of H and for $P_b^{1/F}$ not too close to unity. In this case, the achievable utilization is inversely proportional to the number of hops H between A and B, as one would expect. Figure 13.33 shows this effect. Analogous to Fig. 13.32, this depicts the achievable utilization p for $P_b = 10^{-3}$ as a function of the number of wavelengths for $H = 5$, 10, and 20 hops. In contrast to the previous case, here the effect of path length (i.e., the number of links) is dramatic.

Example 13.9 Consider two optical networks, each of which uses 30 wavelengths. For a 10^{-3} blocking probability in a 10-hop implementation, what is the wavelength utilization when (*a*) there is wavelength conversion; (*b*) there is no wavelength conversion?

Solution: (*a*) From Eq. (13.19) with $P_b' = 10^{-3}$, $H = 10$, and $F = 30$, we find that the achievable wavelength utilization is

$$q \approx (P_b'/H)^{1/F} = (10^{-4})^{0.033} = 0.74$$

(*b*) From Eq. (13.21) we find

$$p \approx -\frac{1}{H} \ln\left(1 - P_b^{1/F}\right) = -\frac{1}{10} \ln[1 - (.001)^{0.033}] = 0.16$$
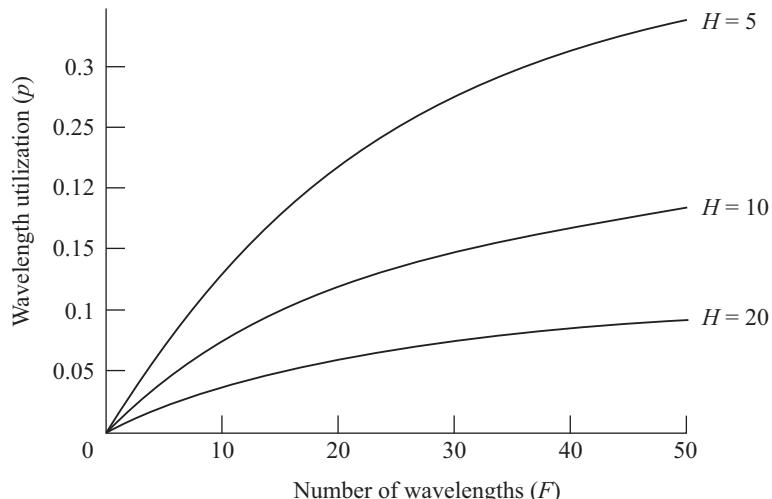
Fig. 13.33 *Achievable wavelength utilization as a function of the number of wavelengths for a 10^{-3} blocking probability in a network not using wavelength conversion. (Reproduced with permission from Barry and Humblet,[61] © IEEE, 1996.)*

To measure the benefit of wavelength conversion, define the gain $G = q/p$ to be the increase in fiber or wavelength utilization for the same blocking probability. Setting $P_b' = P_b$ in Eqs. (13.18) and (13.20), we have

$$G \equiv \frac{q}{p} = \frac{[1 - (1 - P_b)^{1/H}]^{1/F}}{1 - (1 - P_b^{1/F})^{1/H}}$$

$$\approx H^{1-1/F} \frac{P_b^{1/F}}{-\ln(1 - P_b^{1/F})} \qquad (13.22)$$

As an example, Fig.13.34 shows G as a function of F for $H = 5$, 10, and 20 links for a blocking probability of $P_b = 10^{-3}$. This figure shows that as F increases, the gain increases and peaks at about $F = H/2$. The gain then slowly decreases, since large trunking networks are more efficient than small ones.

13.6.3 Wavelength Routing

To send information quickly and reliably across a network, service providers use various techniques to establish a circuit-switched *lightpath* (that is, a temporary point-to-point optical connection) between communicating end equipment. An OXC is a key element to set up express paths through intermediate nodes for this process. Since an OXC is a large complex switch, it is used in extended mesh backbone networks, where there is a heavy volume of traffic between nodes, to connect equipment such as SONET/SDH terminals, IP routers, and ROADMs. In such a network the lightpath normally is set up for long periods of time. Depending on the desired service running between distant nodes, this time connection can range from minutes to months and even longer.

Lightpaths running from a source node to a destination node may traverse many fiber link segments along the route. At intermediate points along the connection route, the lightpaths may be switched between
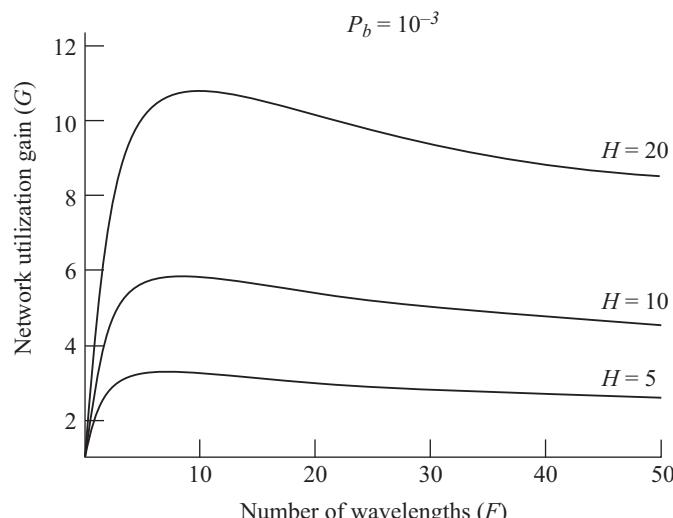
Fig. 13.34 *Increase in network utilization as a function of the number of wavelengths for a 10⁻³ blocking probability when wavelength conversion is used (Reproduced with permission from Barry and Humblet,[61] © IEEE, 1996)*

different links, and sometimes the lightpath wavelength may need to be changed when entering another link segment. As noted in Sec. 13.6.2, this wavelength conversion is necessary if two lightpaths entering some segment happen to have the same wavelength.

The process of establishing a lightpath is called by various names, such as wavelength routing, optical circuit switching, or lightpath switching. The more popular terms are *wavelength routing* and *wavelength routed network* (WRN). Many different static and dynamic approaches have been proposed and implemented for establishing a lightpath. Since a method for setting up a lightpath requires deciding which path to traverse and what wavelength to use, it involves a *routing and wavelength assignment* (RWA) procedure. In general the RWA problem is fairly complex and special software algorithms have been developed for solving it.[14,64–69]

13.6.4 Optical Packet Switching

The success of electronic packet-switched networks lies in their ability to achieve reliable high packet throughputs and to adapt easily to traffic congestion and transmission link or node failures. Various studies have been undertaken to extend this capability to all-optical networks in which no *O/E/O* conversion takes place along a lightpath. In the concept of an *optical packet-switched* (OPS) network, user traffic is routed and transmitted through the network in the form of optical packets along with in-band control information that is contained in a specially formatted header or label.[70] For OPS systems examined so far, the header processing and routing functions are carried out electronically, and the switching of the optical payloads is done in the optical domain for each individual packet. This decoupling between header or label processing and payload switching allows the packets to be routed independent of payload bit rate, coding format, and packet length.

Optical label swapping (OLS) is a technique for realizing a practical OPS implementation.[70–74] In this procedure, optically formatted packets (which contain a standard IP header and an information payload as shown in Fig. 13.35) first have an optical label or control packet attached to them before they enter the OPS network. Note that in some OPS schemes the wavelength used to transmit the label may be different from that used by the packet. When the payload-plus-label packet travels through an OPS network, the optical packet switches at intermediate nodes process only the optical label electronically. The offset time is needed to allow time for the optical packet switch to be set after the optical label is processed. This is done to extract routing information for the packet and to determine other factors such as the wavelength on which the packet is being transmitted and the bit rate of the encapsulated payload. Since the payload remains in an optical format as it moves through the network, it may use any modulation scheme and may be encoded at a very high bit rate.

Several methods have been examined in different network testbeds for creating and attaching a label to an optical packet, and experiments have demonstrated label swapping at data rates of up to 40 and 160 Gb/s. The limitation of an OPS network is that the technology for creating practical optical buffers needs further development. Similar to other switching methodologies, these buffers are needed to store the optical packets temporarily during the time it takes to set up an output path through an intermediate
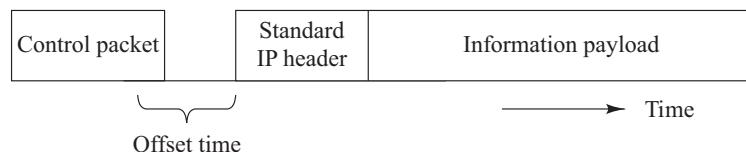
Fig. 13.35 *Optically formatted packet used for optical label swapping*

optical packet switch and to resolve any port contentions that may arise between two or more incoming packets destined for the same output port. This technology limitation can be circumvented through the optical burst switching concept described in the next section.

13.6.5 Optical Burst Switching

Optical burst switching was conceived to provide an efficient solution for sending high-speed bursty traffic over WDM networks.[75–82] Traffic is considered as being *bursty* if there are long idle times between the busy periods in which a large number of packets arrive from users. This format is typical of data traffic in contrast to voice traffic, which is characterized by a more continuous bit stream nature. There are two advantages to OBS. First, it offers the high bandwidth and packet-sized granularity of optical packet-switched networks without the need for complex optical buffering. Second, it provides the low packet-processing overhead that is characteristic of wavelength-routed networks. Thus, the performance characteristics of OBS lie between those of a wavelength-routed network and an optical packet-switched network.

In the OBS network concept, a collection of optical burst switches are interconnected with WDM links to form the central core of the network. Devices called *edge routers* collect traffic flows from various sources at the periphery of a WDM network, as illustrated in Fig. 13.36. The flows then are sorted into different classes according to their destination address and grouped into variable-sized elementary switching units called *bursts*. The characteristics of an edge router play a critical role in an OBS system, since the overall network performance depends on how a burst is assembled based on particular types of traffic statistics.

Before a burst is transmitted, the edge router generates a *control packet* and sends it to the destination to set up a lightpath for this burst. As the control packet travels toward the destination, each optical burst switch along the lightpath reads the burst size and arrival time from the control packet. Then, in advance of the burst arrival, the burst switch schedules an appropriate time period on a wavelength that the next lightpath segment will use to carry the burst. This reservation for the arriving burst is called *burst scheduling*.
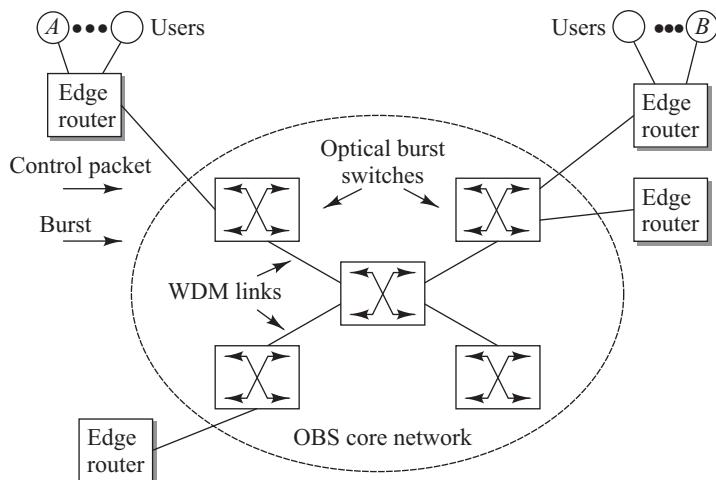
Fig. 13.36 *Generic structure of an optical burst switching (OBS) network*
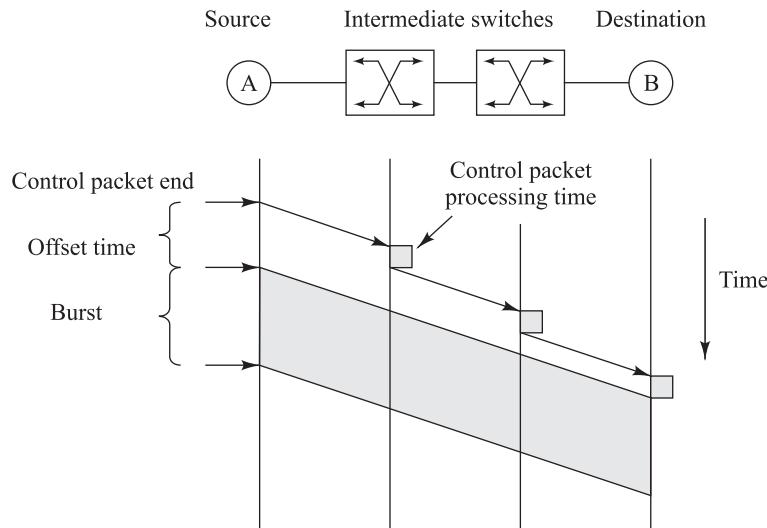
Fig. 13.37 *Timing diagram for the progression of a control packet and a burst through an OBS network*

Figure 13.37 shows an example of a *timing diagram* for the connection setup through two intermediate burst switches from source node *A* to destination node *B*. This figure shows that following the transmission of the control packet, the burst itself is sent after a specified delay called an *offset time*. The time it takes to set up a connection depends on the end-to-end propagation time of the control packet, the sum of all the processing delays t_{proc} of the control packet at intermediate nodes, and the time t_{conf} it takes to configure the link. Once the link is configured, the travel time for the burst to reach the destination node is equal to the propagation time, since no further processing is required for the burst at any intermediate node. Therefore, the start of the burst has to be delayed by an offset time following transmission of the control packet in order to give the link enough time to be configured. If there are *N* intermediate nodes in the link, then the offset time must be at least $Nt_{proc} + t_{conf}$. Note that the edge router sends the burst immediately after expiration of the offset time without waiting for an acknowledgment from the destination to indicate that the link is complete. Furthermore, the burst transmission may start before the control packet reaches its destination.

Proposed solutions for burst-assembly algorithms include formatting a burst based on a fixed assembly time, a fixed burst length, or a hybrid time-/burst-length method. The main parameters are a maximum assembly time threshold *T*, a maximum burst length *B*, and a minimum burst length B_{min}. When the incoming packet flow is low, the threshold *T* time ensures that a packet is not delayed too long in the assembly queue. When the incoming packet flow is high, the upper limit *B* on the burst size also limits packet delay times by restricting the time needed to assemble a burst. Otherwise, if the burst size is not limited, long assembly times could result during heavy incoming traffic loads. If the assembled burst length is smaller than B_{min} by the time the burst is scheduled to be transmitted, padding bits are added to bring the length to B_{min}.

Studies to determine the performance of various proposed OBS approaches include factors such as burst assembly methodologies, schemes for reserving bandwidth over a lightpath, burst scheduling procedures, and network-resource contention resolution when two or more bursts from different edge

routers try to leave the same switch output port simultaneously. In addition, the idea of implementing a hybrid optical network has been proposed as an alternative to a pure OBS network.[82] In this network architecture concept, hybrid optical switches that can accept both wavelength-routed and OBS traffic are used as the crossconnect switches in the network core.

13.7 │ WDM Network Examples

This section presents some WDM network implementation examples for wideband long-haul and narrowband metro networks.

13.7.1 Wideband Long-Haul WDM Networks

Wideband long-haul DWDM networks consist of a collection of ROADMs, switches, and optical crossconnects interconnected with point-to-point high-capacity trunk lines. This collection of equipment can be configured through any combination of ring or mesh networks, as Fig. 13.38 illustrates.

Each long-haul trunk cable contains a large number of single-mode fibers, as Fig. 13.39 illustrates. The cables often are based on fiber-ribbon configurations, which are available with up to 864 optical fibers housed in a cable structure that is less than one inch (24.4 mm) in diameter (see Fig. 2.44 for a cable example). The individual fibers can transport many closely spaced wavelengths simultaneously, and each of these independent lightwave channels can support multigigabit data rates. For example, depending on the service needs, a single fiber in a standard long-haul link can carry up to 160 channels of 2.5-Gb/s, 10-Gb/s, or 40-Gb/s of traffic using 160 individual wavelengths. These data rate numbers correspond to the SONET/SDH designations OC-48/STM-16, OC-192/STM-64, and OC-768/STM-256, respectively. In addition, as described in Sec. 13.4, transmission systems capable of sending data rates of 160 Gb/s over installed standard G.652 single-mode fibers have been demonstrated.
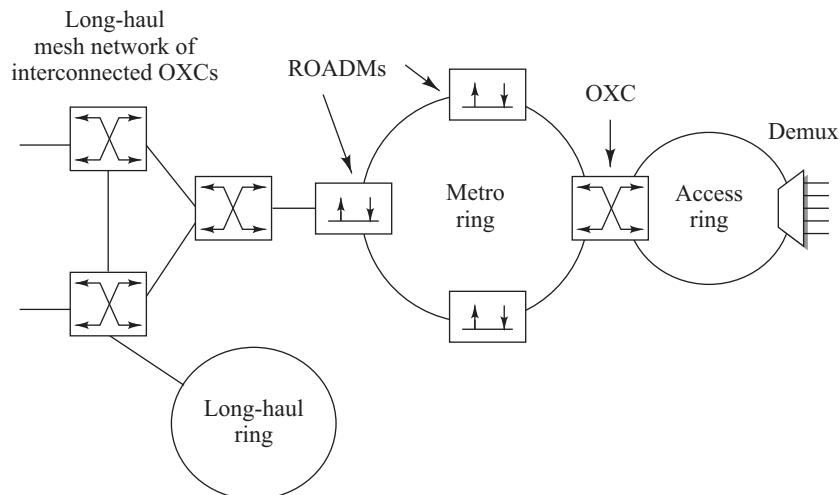
Fig. 13.38 *Wideband DWDM networks can be configured through any combination of ring or mesh networks*
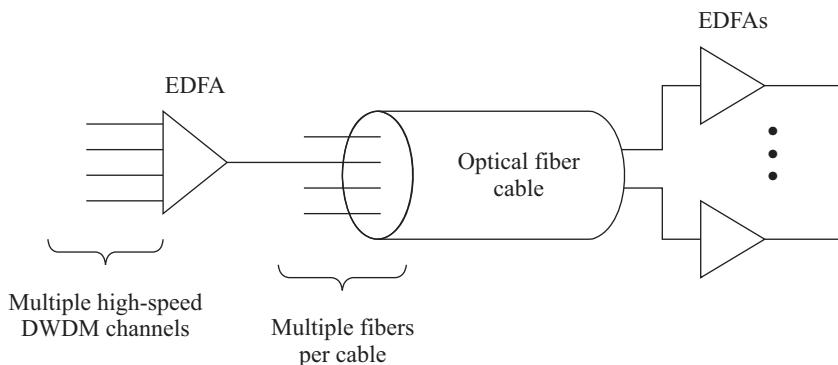
Fig. 13.39 *Long-haul trunk cables contain a large number of high-capacity single-mode fibers*

A long-haul trunk line also can have various ROADMs of degree-*N* for inserting and extracting traffic at remote intermediate points. Standard transmission distances in long-haul terrestrial DWDM links are 600 km with an 80-km nominal spacing between optical amplifiers.

If 160 DWDM channels are separated by 50 GHz, the required frequency span is 8 THz (8000 GHz), which is equivalent to a wavelength spectral band of about 65 nm. This requires operation over both the C-band and either the S- or L-band simultaneously. Consequently the active and passive components intended for long-haul applications must meet high performance requirements, such as the following:

- The optical amplifiers must operate over a wide spectral band.
- The optical amplifiers need high-power pump lasers to amplify a large number of channels.
- Each wavelength has to exit an optical amplifier with the same power level to prevent an increasing skew in power levels from one wavelength to another as the signals pass through successive amplifiers.
- Strict temperature stabilization and optical frequency controls are required of laser transmitters to prevent crosstalk between lightwave channels.
- High-rate transmission over long distances requires optical signal-conditioning techniques such as chromatic-dispersion and polarization-mode-dispersion compensations.

A typical erbium-doped fiber amplifier (EDFA) is limited to the 1530-to-1560-nm C-band in which a standard erbium-doped fiber has a high gain response. By adding a Raman amplification mechanism, the gain response can be extended into both the S-band and the L-band. Figure 13.40 illustrates one amplification concept for operation in the C- and S-bands. Here a multiple-wavelength distributed Raman amplifier pump unit is added ahead of a band-splitting device, which separates the S- and C-bands. The Raman amplification boosts the power levels in both the S-band and the C-band. After passing through the band splitter, the gains of the S-band wavelengths traveling in the bottom path can be enhanced further with an S-band amplifier (e.g., by the TDFA described in Chapter 11). An EDFA boosts the power levels of the C-band wavelengths in the top path in Fig. 13.40. Following the amplification processes, all the wavelengths are recombined with a wideband multiplexing unit. The gain-flattening filters (GFFs) that are used for equalizing the final output powers of all the wavelengths can be either passive or active devices.

A network operator can use remotely controllable degree-4 ROADMs to reconfigure the number of added or dropped wavelengths in different parts of the network when necessary. This function can supply another wavelength to a node for additional capacity or can shut down specific wavelength
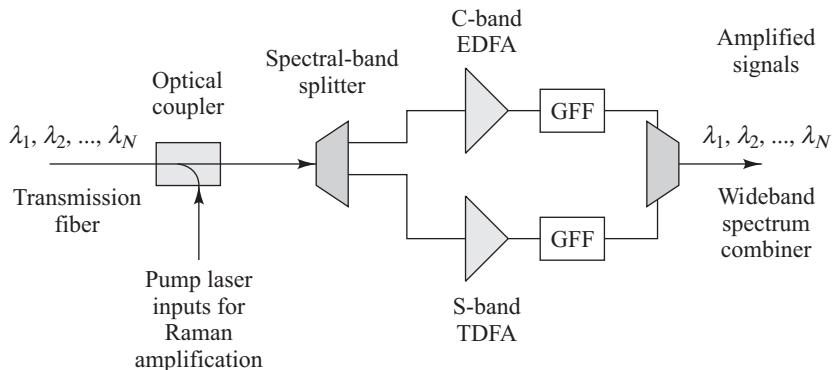
Fig. 13.40 *An amplification concept for operation in the C- and S-bands*

connections that no longer need service. The degree-4 OXCs can set up new routes for specific wavelengths, can close discontinued connections, or can reroute high-capacity lightpaths in case of link congestion or faults.

13.7.2 Narrowband Metro WDM Networks

Nominally a *metro network* is configured as a mesh network or a SONET/SDH bidirectional, line-switched ring (BLSR), as Fig. 13.12 illustrates.[83–89] A ring consists of 2.5- or 10-Gb/s (OC-48/STM-16 or OC-192/STM-64) point-to-point connections between central offices that are spaced 10 to 20 km apart. A metro network may comprise from three to eight nodes and the ring circumference nominally is less than 80 km. At switching nodes within the metro ring, the degree-2 ROADMs allow multiple selected wavelengths to be extracted and inserted at the metro central offices. A degree-2 ROADM or a degree-4 OXC provides inter-connections to a long-haul network. Within the metro central offices, the SONET/SDH equipment has STS-1 grooming capabilities. The term *grooming* means that low-speed SONET/SDH rates, such as 51.84-Mb/s STS-1 or 155.52-Mb/s STM-1 rates, are packaged into higher-rate circuits, such as 2.5-Gb/s or 10-Gb/s.

Attached to a metro network is the *access network*, which consists of links between end users and a central office. The access network could be either a ring or a star-based passive optical network or PON (see Sec. 13.8). An *access ring* configuration ranges from 10 to 40 km in circumference and typically contains three or four nodes, whereas a maximum of 32 users located up to 20 km away can be attached to a single PON. Within the access network ring, either a fixed OADM or a ROADM provides the capability to add or drop multiple wavelengths to local users or to other regional networks. A degree-4 OXC provides the connection from the access ring to a metro network.

In contrast to the stringent performance specifications imposed on wideband long-haul DWDM systems, the shorter spans in metro and access networks relax some of the requirements. In particular, if CWDM technology is employed, the 20-nm spectral band tolerances allow the use of optoelectronic devices that are not temperature-controlled. However, other requirements unique to metro applications arise, such as the following:

- A *high degree of connectivity* is required to support meshed traffic in which various wavelengths are inserted and extracted at different points along the path.
- A *modular* and *flexible switching device* such as a ROADM is needed since the wavelength add/drop patterns and link capacities vary dynamically with various levels of new and completed service demands from different nodes.

Table 13.8 *Data formats and protocols that need to be accommodated by a metro WDM network*

Format or protocol	Data rate
OC-3/OC-3c and STM-1/STM-1c	155 Mb/s
OC-12/OC-12c and STM-4/STM-4c	622 Mb/s
OC-48/OC-48c and STM-16/STM-16c	2.488 Gb/s
OC-192/OC-192c and STM-64/STM-64c	9.953 Gb/s
Fast Ethernet	125 Mb/s
Gigabit Ethernet (GigE)	1.25 Gb/s
10-Gigabit Ethernet (10GigE)	10 Gb/s
ESCON	200 Mb/s
Fibre Channel	133 Mb/s to 1.06 Gb/s
Digital video	270 Mb/s

- Since the add/drop functions can change dynamically from node to node, *special components*, such as a *variable optical attenuator* (VOA), are needed to equalize the power levels of newly added wavelengths with those that already are on the fiber.
- *Optical amplifiers* optimized for metro network use are needed, since interconnection losses can be fairly high and the power level of express wavelengths can change as the light signal passes through successive nodes.

Metro WDM networks must support a wide variety of transmission formats, protocols, and bit rates. As listed in Table 13.8, these include SONET/SDH traffic ranging from OC-3/STM-1 to OC-192/STM-64, ESCON (Enterprise Systems Connection from IBM), Fibre Channel, Fast Ethernet, Gigabit Ethernet, and digital video. Emerging protocols that need to be supported include 40 and 100 Gigabit Ethernet.

13.8 | Passive Optical Networks

A *passive optical network* (PON) is based on using CWDM wavelengths and bidirectional transmission on a single optical fiber.[90–94] In a PON there are no active components between the central office and the customer premises. Instead, only passive optical components are placed in the network transmission path to guide the traffic signals contained within specific optical wavelengths to the user endpoints and back to the central office.

13.8.1 Basic PON Architectures

Figure 13.41 illustrates the architecture of a typical PON in which a fiber optic network connects switching equipment in a central office with a number of service subscribers. In the central office, data and digitized voice are combined and sent downstream to customers over an optical link by using a 1490-nm wavelength. The upstream (customer to central office) return path for the data and voice uses a 1310-nm wavelength. Video services are sent downstream using a 1550-nm wavelength. There is no video service in the upstream direction. Note that all users attached to a particular PON share the same 1310-nm upstream wavelength. This requires carefully timed assignments of transmission permits to the users, so that upstream traffic streams from different users do not interfere with each other.

Starting at the central office, one single-mode optical fiber strand runs to a passive *optical power splitter* near a housing complex, an office park, or some other campus environment. At this point the passive
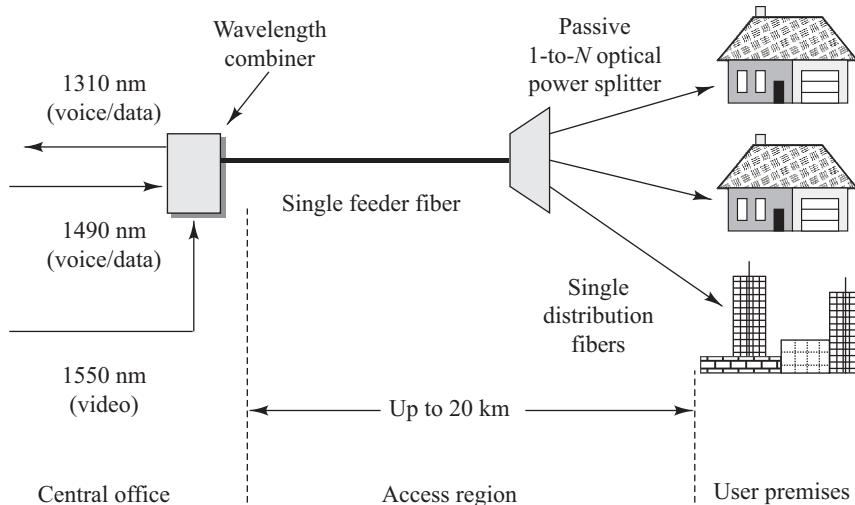
Fig. 13.41 *Architecture of a typical passive optical network*

splitting device simply divides the optical power into N separate paths to the subscribers. If the splitter is designed to divide the incident optical power evenly and if P is the optical power entering the splitter, then the power level going to each subscriber is P/N. Designs of power dividers with other splitting ratios are also possible depending on the application. The number of splitting paths can vary from 2 to 64, but in a PON they typically are 8, 16, or 32. From the optical splitter, individual single-mode fibers run to each building or serving equipment. The optical fiber transmission span from the central office to the user can be up to 20 km. Thus active devices exist only in the central office and at the user end.

There are several alternative PON implementation schemes.[91,94–106] The main ones are as follows:

- *Broadband* PON (BPON) is based on the G.983.1 to G.983.5 series of ITU-T Recommendations that specify the asynchronous transport mode (ATM) technique as the transport and signaling protocol. This PON type is being phased out due to its relatively high cost compared to Ethernet and its incompatibility with emerging IP over DWDM technology (see Sec. 13.9).
- *Ethernet* PON (EPON) or *gigabit Ethernet* PON (GE-PON) uses 1-Gb/s Ethernet as the underlying protocol. The GE-PON is based on the IEEE 802.3ah EPON standard and is a dominant PON technology in Asia. The higher-speed 10G EPON is based on the IEEE 802.3av EPON standard and on ITU-T Recommendation G.987 for 10-Gb/s upstream and downstream transmission.
- *Gigabit* PON (GPON) combines the features of both ATM and Ethernet to provide a more efficient and flexible network usage. GPON provides downstream speeds of 2.5 Gb/s and upstream speeds of 1.25 Gb/s. It is based on the G.984.1 to G.984.6 series of ITU-T Recommendations. GPON provides broadband Internet services and natively delivers ATM, TDM, and Ethernet traffic. Typically the bandwidth is shared by 32 users.
- *WDM* PON uses a different wavelength for each user to greatly enhance network capacity. In this architecture a wavelength multiplexer (usually an AWG) is used in place of the power splitter shown in Fig. 13.41. Its service-offering flexibility is a major advantage of a WDM PON compared to other PON types. Since each user has a dedicated wavelength that is not shared with others, a customer with very high bandwidth demands can easily be accommodated without affecting other lower-usage customers. The dedicated wavelength also provides for a higher level of information security compared to other PONs.

The application of PON technology for providing broadband connectivity in the access network to homes, multiple-occupancy units, and small businesses commonly is called *fiber-to-the-x* (FTTx).[90–92] Here *x* is a letter indicating how close the fiber endpoint comes to the actual user. FTTP for *fiber-to-the-premises* has become the prevailing term that encompasses the various FTTx concepts. FTTP networks can use any of the various PON technologies.

13.8.2 Active PON Modules

This section gives a snapshot of the basic functions and compositions of the optoelectronic equipment located in the central office and at or near the endpoint users.

Optical Line Termination (OLT)

The *optical line termination* (OLT) is located in a central office and controls the bidirectional flow of information across the network. An OLT must be able to support transmission distances of up to 20 km. In the downstream direction the function of an OLT is to take in voice, data, and video traffic from a long-haul or metro network and broadcast it to all the users on the PON. In the reverse direction (upstream), an OLT accepts and distributes multiple types of voice and data traffic from the network users.

A typical OLT is designed to control more than one PON. Figure 13.42 gives an example of an OLT that is capable of serving four independent passive optical networks. In this case, if there are 32 connections to each PON, then the OLT can distribute information to 128 users.

Simultaneous transmission of separate service types on the same PON fiber is enabled by using different wavelengths for each direction. For downstream transmissions, a PON uses a 1490-nm wavelength for combined voice and data traffic and a 1550-nm wavelength for video distribution. Upstream voice and data traffic use a 1310-nm wavelength. Passive WDM couplers perform the wavelength combining and separation functions. Depending on the particular PON standard being used, the downstream and upstream transmission equipment operates at 155 Mb/s, 622 Mb/s, 1.25 Gb/s, or 2.5 Gb/s. In some cases the transmission rates are the same in either direction (a *symmetric* network). In other PON standards the downstream rate may be higher than the upstream rate, which is called an *asymmetric* implementation. A number of different transmission formats can be used for the downstream video transmission at 1550 nm.

Optical Network Termination (ONT)

An *optical network termination* (ONT) is located directly at the customer premises. Its purpose is to provide an optical connection to the PON on the upstream side and to interface electrically to the local customer equipment. Depending on the communication requirements of the customer or block of users,
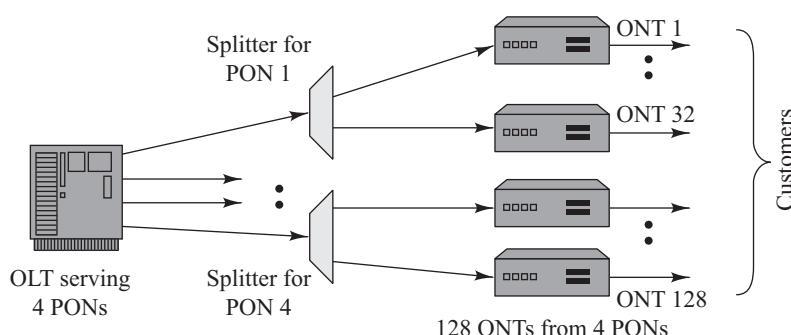
Fig. 13.42 *An OLT that is capable of serving four independent passive optical networks*

the ONT typically supports a mix of telecommunication services, including various Ethernet rates, T1 or E1 (1.544 or 2.048 Mb/s) and DS3 or E3 (44.736 or 34.368 Mb/s) telephone connections, ATM interfaces (155 Mb/s), and digital and analog video formats.

A wide variety of ONT functional designs and chassis configurations are available to accommodate the needs of different levels of demand. The size of an ONT can range from a simple box that may be attached to the outside of a house to a fairly sophisticated unit mounted in a standard indoor electronics rack for use in large apartment complexes or office buildings. At the high-performance end, an ONT can aggregate, groom, and transport various types of information traffic coming from the user site and send it upstream over a single-fiber PON infrastructure. The term *grooming* means that the switching equipment looks inside of a time-division multiplexed data stream, identifies the destinations of the individual multiplexed channels, and then reorganizes the channels so that they can be delivered efficiently to their destinations. In conjunction with the OLT, an ONT allows dynamic bandwidth allocation based on varying demands of users to enable smooth delivery of data traffic that typically arrives in bursts from the users.

Figure13.43 shows a compact, rugged outdoor cabinet that serves as an interface between the termination of a multiple-fiber PON cable and multiple ONT equipment units located in facilities such as an apartment building or an office building. By means of such a cabinet, individual incoming fibers can be activated, moved, or disconnected based on whether ONTs of specific users are being turned on, moved to a different part of the building, or turned off.

Fig. 13.43 *A compact, rugged outdoor cabinet used to interface a multiple-fiber PON cable to multiple indoor ONTs. (Photo courtesy of Charles Industries; www .charlesindustries.com.)*

Optical Network Unit (ONU)

An optical network unit (ONU) normally is housed in an outdoor equipment shelter. These installations would include shelters located at a curb or in a centralized place within an office park. Thus the ONU equipment must be environmentally rugged to withstand wide temperature variations. The shelter for the outdoor ONU must be water-resistant, vandal-proof, and able to withstand high winds. In addition, there has to a local power source to run the equipment, together with emergency battery backup. The link from the ONU to the customer premises can be a twisted-pair copper wire, a coaxial cable, an independent optical fiber link, or a wireless connection.

13.8.3   Traffic Flows

Two key network functions of an OLT are to control user traffic and to assign bandwidth dynamically to the ONT modules. Up to 32 ONTs use the same wavelength and share a common optical fiber transmission line, so some type of transmission synchronization must be used to avoid collisions between traffic coming from different ONTs. The simplest method is to use *time-division multiple access* (TDMA) wherein each user transmits information within a specific assigned time slot at a prearranged data rate. However, this does not make efficient use of the available bandwidth because many time slots will be empty if network users do not have information to be sent back to the central office.

A more efficient process is *dynamic-bandwidth allocation* (DBA), wherein time slots of an idle or low-utilization user are assigned to a more active customer. The exact DBA scheme implemented through an OLT in a particular network depends on factors such as user priorities, the quality of service guaranteed to specific customers, the desired response time for bandwidth allocation, and the amount of bandwidth requested (and paid for) by a customer.

As shown in Fig. 13.44, the OLT uses *time-division* multiplexing (TDM) to combine incoming voice and data streams that are destined for users on the PON. As a simple example of this, if there are N independent information streams coming into the OLT, each of which is running at a data rate of R b/s, then the TDM scheme interleaves them electrically into a single information stream operating at a higher rate of $N \times R$ b/s. The resulting multiplexed downstream signal is broadcast to all the ONTs. Each ONT discards or accepts the incoming information packets depending on the packet header address. Encryption may be necessary to maintain privacy because the downstream signal is broadcast, and every ONT receives all the information destined for each end terminal.

Sending traffic in the upstream direction is more complicated because all users have to time-share the same wavelength. To avoid collisions between the transmissions of different users, the system uses a *time-division multiple access* (TDMA) protocol. Figure 13.45 gives a simple example of this. The OLT controls and coordinates the traffic from each ONT by sending permissions to them to transmit during a specific time slot. The time slots must be synchronized because transit times vary between ONTs (see Example 13.10).

13.8.4   GPON Characteristics

Table 13.9 highlights the major GPON service requirements (GSR). First a GPON must be a *full-service network*, which means that it should be able to carry all service types. These include 10- and 100-Mb/s Ethernet, legacy analog telephone, digital T1/E1 traffic, ATM packets, and higher-speed leased-line traffic. The data rates can be either symmetrical or asymmetrical with higher rates being sent downstream from the OLT to the ONTs. A service provider can offer a lower upstream rate to those GPONs in which the downstream traffic is much larger than in the upstream direction, as is the case when subscribers use the IP data service mainly for applications such as lower-rate upstream Internet surfing or e-mail and higher-rate downstream downloads of large files.
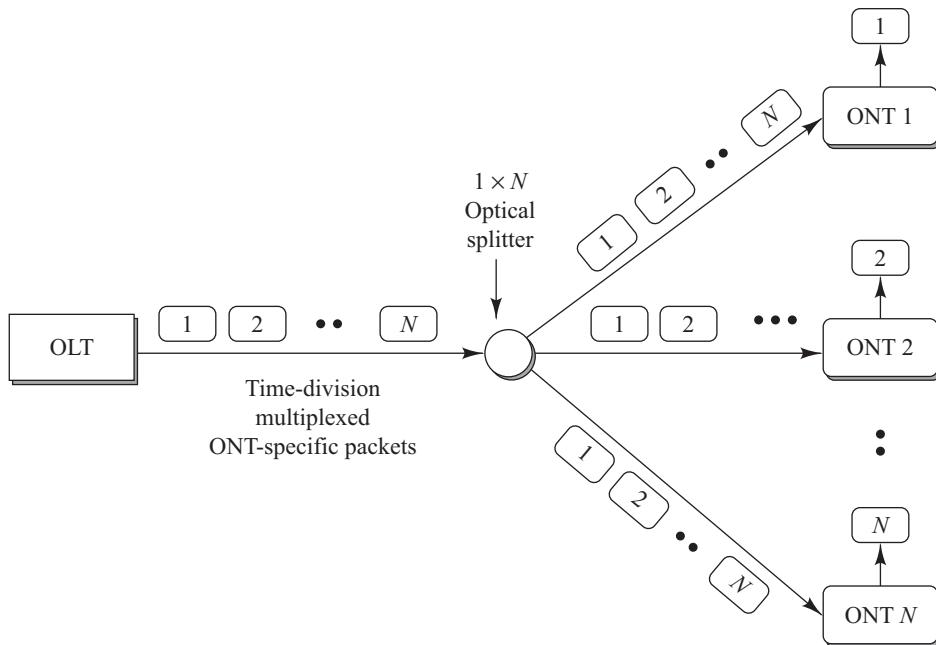
Fig. 13.44 *Operation of a time-division multiplexing process*

Example 13.10 Consider two ONTs located 3 km and 20 km from the OLT, respectively. For these two links, what are the roundtrip propagation times for messages coming from and returning to the OLT? Let the speed of light in the fiber be $s_{glass} = 2 \times 10^8$ m/s.

Solution: The roundtrip propagation times are

$$t(3 \text{ km}) = 2(3 \text{ km})/(2 \times 10^8 \text{ m/s}) = 30 \text{ } \mu s$$

$$t(20 \text{ km}) = 2(20 \text{ km})/(2 \times 10^8 \text{ m/s}) = 200 \text{ } \mu s$$
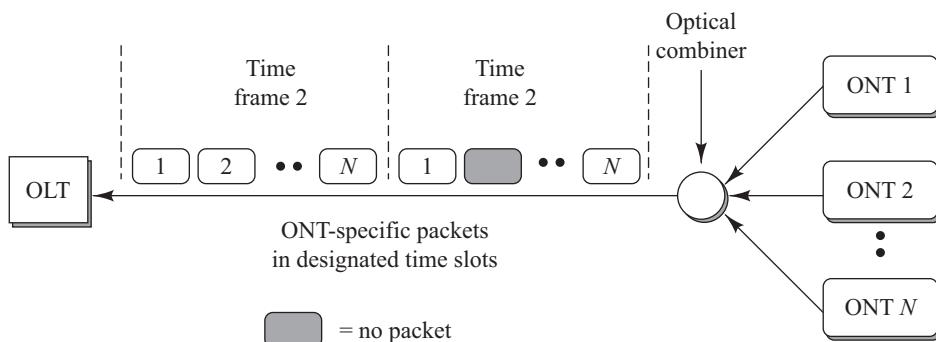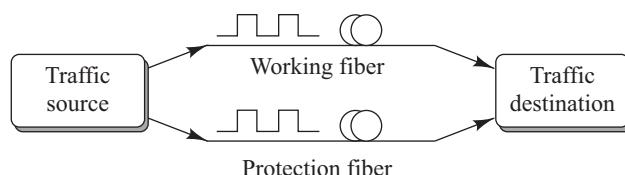
Fig. 13.45 *Operation of a time-division multiple access protocol*

Table 13.9   *Summary of GPON service requirements*

Parameter	GSR specification
Service	Full service; for example, 10/100 BASE-T Ethernet, analog telephony, SONET/SDH TDM, ATM
Access data rate	Downstream: 1.244 and 2.488 Gb/s
	Upstream: 155 Mb/s, 622 Mb/s, 1.244 Gb/s, 2.488 Gb/s
Distance	10 km or 20 km maximum
Splitting number	Maximum of 64
Wavelengths	Downstream voice/data: 1480 to 1500 nm
	Upstream voice/data: 1260 to 1360 nm
	Downstream video distribution: 1550 to 1560 nm
Protection switching	Fully redundant 1 + 1 protection
	Partially redundant 1:N protection
Security	Information security at the protocol level for downstream traffic; for example, the use of the *Advanced Encryption Standard* (AES)

The ITU-T Recommendation G.984.1 describes the use of a protection switching mechanism. This allows several different types of PON configurations that include redundancy of links and equipment for network protection. Among these are a fully redundant 1 + 1 protection and a partially redundant 1:N protection. Figure 13.46 shows that in 1 + 1 protection the traffic is transmitted simultaneously over two separate fiber lines from the source to the destination. Typically these two paths do not overlap at any point, so that a cable cut would affect only one fiber transmission path. In the 1 + 1 protection scheme, the receiving equipment selects one of the links as the *working fiber* for reception of information. If a fiber in that link is cut or the transmission equipment on that link fails, then the receiver switches over to the *protection fiber* and continues to receive information. This protection method provides rapid switchover during failures and does not require a protection signaling protocol between the source and destination. However, it requires duplicate fibers and redundant transmission equipment for each link.

The 1:N protection procedure offers a more economical use of fibers and equipment. As shown in Fig. 13.47, here one protection fiber is shared among N working fibers. This arrangement offers protection in the event that one of the working fibers fails. For most operational networks this level of protection is adequate as failures of multiple fibers are rare (unless all the fibers are in the same cable). In contrast to the 1 + 1 protection method, in the 1:N protection scheme traffic is transmitted only over the working fiber in normal operation. When there is a failure on a particular link, the source and destination both switch over to the protection fiber. This requires an automatic switching protocol between the endpoints to enable use of the protection link.

Fig. 13.46   *Fully redundant 1 + 1 protection of links*
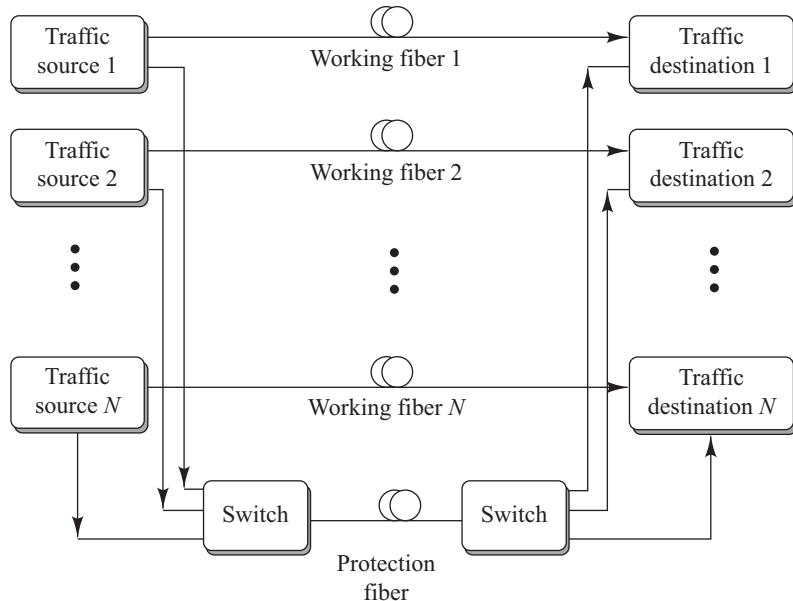
Fig. 13.47   *The 1:N protection procedure*

Since the downstream data from the OLT is broadcast to all ONTs, every transmitted message can be seen by all the users attached to the GPON. Thus the GPON standard describes the use of an information security mechanism to ensure that users are allowed to access only the data intended for them. In addition, such a security mechanism ensures that no malicious eavesdropping threat is probable. One example of a point-to-point encryption mechanism is the *Advanced Encryption Standard* (AES), which is used to protect the information payload of the data field in the GPON frame. *Encryption* is a technique whereby data is transformed into an unintelligible format at the sending end to protect it against unauthorized disclosure, modification, utilization, or destruction as it travels through the network.

The AES encryption algorithm encrypts and decrypts 128-bit data blocks from their original format called *plaintext* to an unintelligible form called *ciphertext*. The cipher keys can have lengths of 128, 192, or 256 bits, which makes the encryption extremely difficult to compromise. A key can be changed periodically (for example, once per hour) without disturbing the information flow. Decrypting the ciphertext at the destination converts the data back into its original form.

The GPON encapsulation method (GEM) provides a generic means to send different services over a GPON. The encapsulated payload can be up to 1500 bytes long. If an ONT has a packet to send larger than 1500 bytes, the ONT must break the packet into smaller *fragments* that fit into the allowed payload length. The destination equipment is responsible for reassembling the fragments into the original packet format. A key advantage of the GEM scheme is that it provides an efficient means to encapsulate and fragment user information packets. Encapsulation on a GPON allows proper management of the multiple service flows from different ONTs that share a common optical fiber transmission link. The purpose of fragmentation is to send packets from a user efficiently regardless of their size and to recover the original packet format reliably from the physical layer transmission windows on the GPON.

13.8.5 WDM PON Architectures

The ever-increasing demand for higher-capacity triple-play services (such as interactive video, virtualization, and cloud computing) can quickly lead to demands of 100 Mb/s per subscriber. Since a standard three-wavelength FTTP network will not be able to satisfy such demands, a possible enhancement is to use more wavelengths to create a WDM PON. This method uses a separate wavelength for each transmitting ONT, so that an ONT can send its information continuously over the shared upstream fiber without having to wait for a specific assigned transmission time slot.

In this architecture a wavelength multiplexer (usually an AWG) is used in place of the power splitter shown in Fig. 13.41. Its service-offering flexibility is a major advantage of a WDM PON compared to other PON types. Since each user has a dedicated wavelength that is not shared with others, a customer with very high bandwidth demands can easily be accommodated without affecting other lower-usage customers. The dedicated wavelength also provides for a higher level of information security compared to other PON types.

In such a WDM PON it is desirable to have a *colorless* ONT, which means that no ONT should have a fixed transmission wavelength assigned to it. An obvious but extremely expensive solution is to use a tunable laser at each ONT, but low-cost end equipment is a driving factor in PON implementations, so this is not a viable solution. Thus a major challenge for a WDM PON implementation is to have a low-cost, high-output optical source.

One proposed method is to use spectral splicing of a single broadband relatively inexpensive light source. Various techniques are being explored to achieve this. One idea is to have each ONT contain a source with a broad optical output spectrum within the transmitter, such as a *superluminescent light emitting diode* (SLED). The broad spectral output of the source at the ONT is connected to one port on a local WDM device, such as a thin-film filter or an AWG. Only those optical spectral components from the SLED that can pass through the WDM channel are transmitted through to the central office. Although all the ONTs have identical SLEDs, each user is connected to a different port on the local WDM device, so it is possible to slice a different part of the available optical spectrum for each ONT. This scheme thereby provides each ONT with a different transmitting wavelength.

Another concept is to use a broadband source (for example, a superluminescent laser diode or a broadband EDFA source) at the central office and send the source output downstream through an AWG. Unique spectrally sliced wavelengths travel to individual ONTs and seed a relatively inexpensive source at each ONT, such as a Fabry-Perot (FP) laser diode. The seeding action forces the FP laser to operate in a quasi single mode. Since a different seed wavelength arrives at each ONT, the FP laser located at that ONT can transmit data upstream at its uniquely locked wavelength.

Other techniques use a reflective semiconductor optical amplifier (ROSA) or a reflective electro-absorption modulator to create unique wavelength sources at an ONT for upstream transmission. Although by 2010 several field trials on WDM PONs were under way, including providing service to 150,000 customers in Korea, further research is needed to find a cost-effective method for establishing transmitters with unique wavelengths at a customer site.

13.9 IP over DWDM

The main purpose of any communication network is to exchange information between elements such as telephones, computers, office equipment, or various types of instrumentation. As shown in Fig. 13.3, this information exchange is accomplished by invoking a set of protocols that govern the generation, formatting, control, exchange, and interpretation of information that is transmitted through a telecommunication

Has a high degree of framing overhead and four layers of management

| IP/MPLS |
| ATM |
| SONET/SDH |
| Optical adaptation layer |
| Physical layer |

Manages WDM channel setups and takedowns; provides some level of link protection and recovery

Performs functions such as optical amplification, wavelength switching and conversion, and add/drop multiplexing
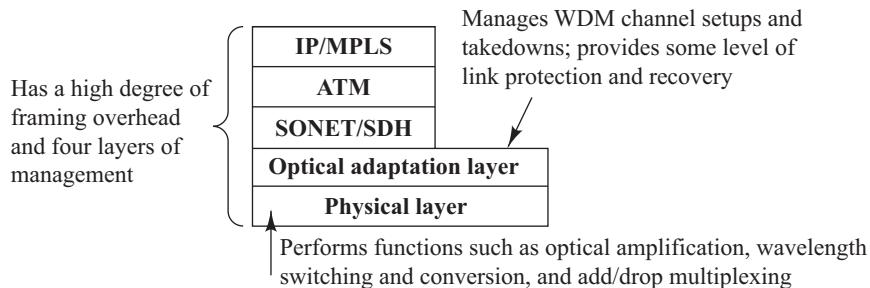
Fig. 13.48 *The layering architecture of a first-generation routed IP network*

network or that is stored in a database. In the protocol stack shown in Fig. 13.3, the function of the network layer is to deliver data packets from source to destination across multiple network links. Typically, the network layer must find a path through a series of connected nodes, and the nodes along this path must forward the packets to the appropriate destination. The dominant network layer protocol is the Internet Protocol (IP).

There are multiple ways to deliver IP packets over a physical optical transport network.[100,107–108] Figure 13.48 shows the layering architecture of a first-generation routed IP network used in the early 2000s. User information is encapsulated into IP packets, which contain a data field and a series of both addressing and control bits to allow information to be routed within a network. For transportation over an optical network, the IP packets are segmented into fixed-length *asynchronous-transfer mode* (ATM) packets. A SONET or SDH network then is used to carry the ATM cells over a WDM network to the information destination. The role of the WDM network is to manage the setups and takedowns of WDM channels, to provide some level of network protection and recovery, and to provide physical layer functions such as optical amplification, add/drop multiplexing, and wavelength conversion.

ATM is a high-performance switching and multiplexing technology that utilizes fixed-length 53-byte packets to carry different types of traffic.[9–12] The ATM packets, which are called *cells*, consist of 48 bytes of information payload and a 5-byte header. The advantage of ATM is that it enables carriers to offer multiple classes of service, to connect devices that operate at different speeds, and to mix traffic types having different transmission requirements, such as voice, video, and data traffic.

The drawback of the architecture shown in Fig. 13.48 is that redundant management functions are incurred in each layer, so this layering method is not efficient for transporting IP traffic. Consequently, the legacy ATM started being phased out in favor of creating an IP over SONET architecture, as illustrated in Fig. 13.49. Within this architecture a technique called *Multiprotocol Label Switching* (MPLS) is implemented to provide for the efficient designation, routing, forwarding, and switching of traffic flows through the network.[109–110]

IP → **SONET/SDH** → **DWDM**

Three management layers

Fig. 13.49 *Concept of an IP-over-SONET-over-DWDM architecture*
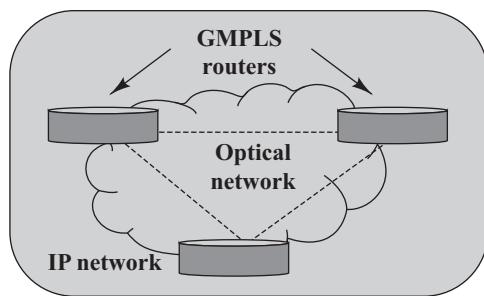
Fig. 13.50 *GMPLS routers receive information from the IP network and establish routing paths through the optical network*

The fundamental concept of MPLS is to forward IP traffic using a packet label. In a traditional IP network every router along the transmission path makes an independent packet-forwarding decision based on the network layer header of each packet. With MPLS, when a packet initially enters a network, it is assigned a specific *forwarding equivalence class* label based on preestablished IP routing information, which designates the path the packet will follow through the network. This label instructs the routers and the switches in the network where to forward the packets. Consequently, once a packet has entered a network, intermediate routers do not need to perform a header analysis. This also means that real-time traffic that cannot afford time delays, such as voice and video, can be mapped easily to low-latency routes across a network.

An extension of MPLS called *Generalized Multiprotocol Label Switching* (GMPLS) enhances the MPLS architecture by having a complete separation of the control and data planes of various networking layers.[111-113] This concept is illustrated in Fig. 13.50. GMPLS gives routers the ability to intelligently interface to the optical layer to establish, change, or take down optical links in real time. In addition, it allows end-to-end provisioning, control, and traffic engineering even when the start and the end nodes belong to heterogeneous networks.

13.10 Optical Ethernet

Optical Ethernet is designed to send standard well-known 10/100-Mb/s, 1-Gb/s, and 10-Gb/s Ethernet frames directly over optical fibers.[114-117] Since the Ethernet protocol has a lower complexity than traditional telecom protocols, such as SONET/SDH, optical Ethernet offers reduced network complexity. In addition and importantly, optical Ethernet enables direct interfaces to the enormous ubiquitous worldwide suite of Ethernet-based equipment. Thus there is no need to translate Ethernet frames to and from another communications protocol between the sender and receiver.

13.10.1 Basic Optical Ethernet Implementations

Ethernet has become the most widely used protocol for local area and campus networks and has been extended into both metro and wide area networks. There is an installed base of over a billion Ethernet ports, which has resulted in low-cost and highly reliable components that are widely available worldwide. Using Ethernet as a first-mile transport technology in the access network allows engineers to build access networks with IP and Ethernet, since data is carried in standard Ethernet frames. This avoids the cost and complexity of protocol conversion, as is done with BPON implementations. In addition, the deployment of

Ethernet in the first mile enables network managers to take advantage of existing and familiar management and analysis tools that have been designed for monitoring and control of an Ethernet network.

In June 2004 the IEEE approved the 802.3ah *Ethernet in the First Mile* (EFM) standard. The *first mile* is defined as the network infrastructure that connects business or residential subscribers to the central office of a *telecom carrier* (a company that is authorized by regulatory agencies to operate a telecommunications system) or a *service provider* (a company that provides access to telephone and related communications services). Sometimes the first mile can be a distribution network in a building, which is connected through a router or layer-3 switch to a service provider via some high-speed connection. Such buildings can include apartments or multidwelling units (MDUs), office buildings or multitenant units (MTUs), and hotels or multihospitality units (MHUs).

Three different EFM physical transport schemes are possible. As illustrated in Fig. 13.51, one scheme uses an *Ethernet passive optical network* (EPON) methodology; the other two methods employ point-to-point (P2P) links over either copper wires or optical fibers that directly connect users and the central office. Thus there can be either many lines running to individual premises from the central office or a single fiber or copper link can connect the central office to a local distribution switch or an optical splitter. Table 13.10 highlights the physical layer characteristics of these three options. They are spelled out in more detail in the IEEE EFM standard. Note that in Japan and other parts of Asia, the 1-Gb/s EPON is widely known as the GE-PON (*Gigabit Ethernet passive optical network*).
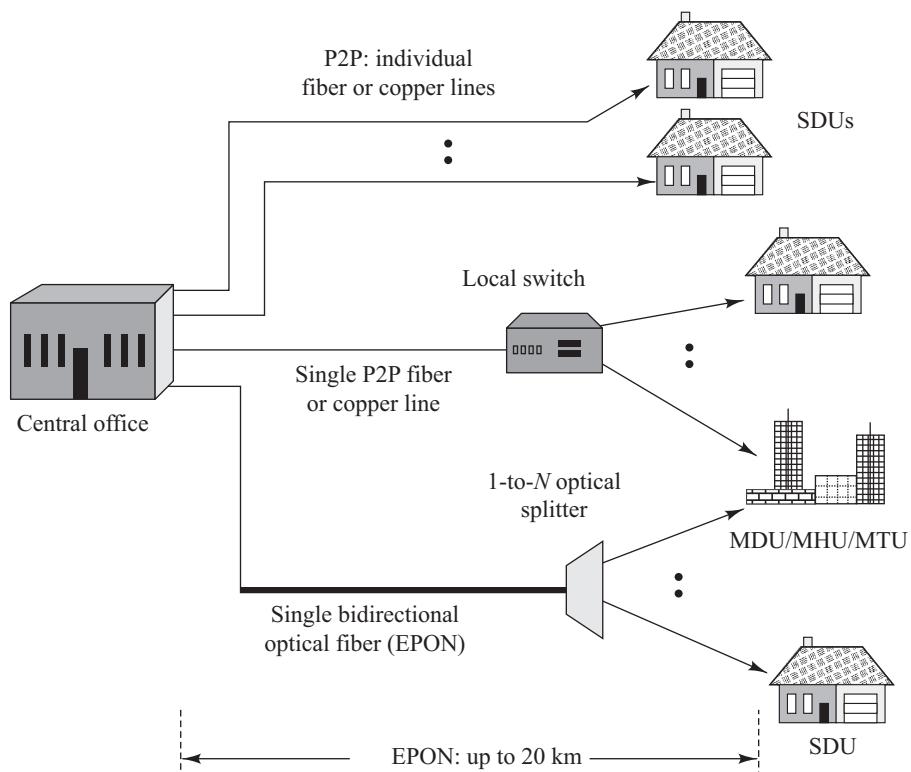
Fig. 13.51 *Three different EFM physical transport schemes. The SDU is a single-dwelling unit.*

	Table 13.10 *The main physical layer characteristics of three EFM options*
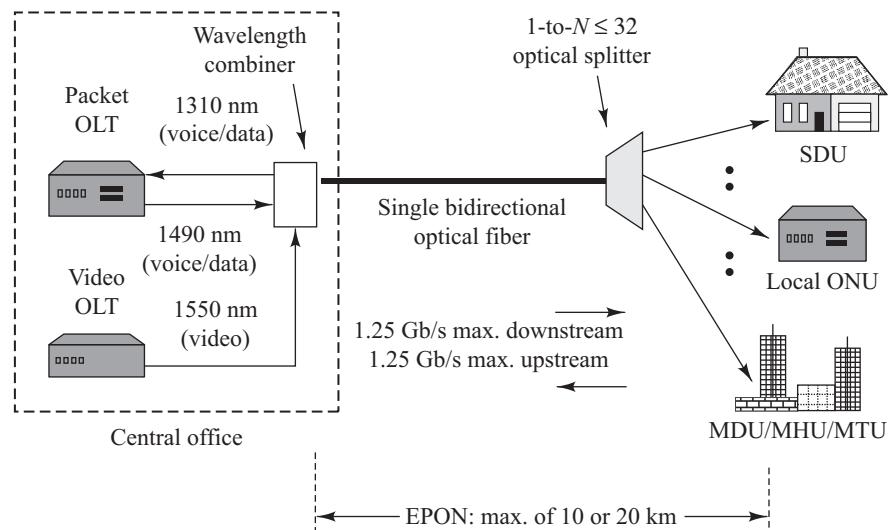
EFM scheme	Physical layer options
EPON/GE-PON	• 10-km distance; 1 Gb/s; 1 × 32 splitter; one bidirectional single-mode fiber • 20-km distance; 1 Gb/s; 1 × 16 or 1 × 32 splitter; one bidirectional single-mode fiber
P2P over fiber	• 1000BASE-LX: extended temperature range optics for 1-Gb/s links • 1000BASE-X: 10 km over one bidirectional single-mode fiber for 1-Gb/s links • 100BASE-X: 10 km over one bidirectional single-mode fiber for 100-Mb/s links
P2P over copper	• 750-m distance; 10 Mb/s full-duplex transmission over single pair nonloaded voice-grade copper wires

13.10.2 EPON/GE-PON Architecture

Figure 13.52 shows the basic EPON/GE-PON architecture and operational concept. This architecture follows the standard PON layout, which has one main feeder line going to an optical splitter. Up to 32 distribution branches leave the splitter and interface to ONTs. As noted in Table 13.10, the IEEE 802.3ah EFM standard specifies operational conditions for a maximum transmission distance of either 10 or 20 km between the OLT and an ONT. The achievable transmission distance depends on the optical splitter size (16 or 32 subscriber ports) and on whether short-reach (class B) or long-reach (class C) optics are selected.

The EPON implementation uses commercially available Ethernet *media access control* (MAC) and *physical layer* (PHY) chipsets. This results in a significant economic benefit because of the wide availability and proven reliability of these components.

An EPON uses a 1490-nm wavelength for downstream transmission of voice and data to the ONTs and a 1310-nm wavelength for the upstream return path from an ONT to the OLT. Thus the 1550-nm window is available for other services, such as multichannel video transmission from the OLT to the users. Since it is based on the standard gigabit Ethernet protocol, an EPON has a nominal bit rate of

Fig. 13.52 *EPON architecture and operational concept*
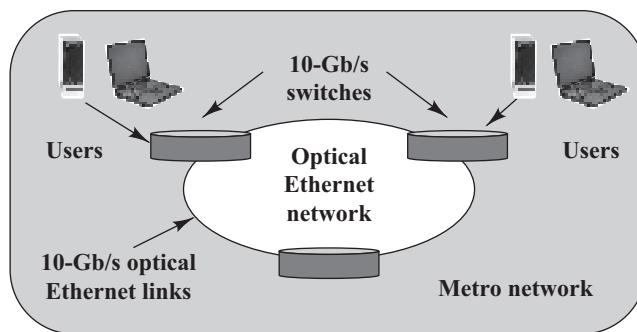
Fig. 13.53 *Optical Ethernet delivering traffic over a metro network backbone based on the 10G Ethernet standard (Photos © Getty RF)*

1250 Mb/s, which is sent using 8B10B encoding. In this 8B10B encoding scheme, two extra redundant bits are incorporated into every eight-bit block of data to provide adequate timing for signal recovery and to have error-monitoring features. In addition to the 256 eight-bit data characters in an 8B10B code (since $2^8 = 256$), there are 13 special ten-bit control characters. These control characters are used, for example, to indicate whether the information bits are for idle data, test messages, or frame delineation.

13.10.3 Metro Optical Ethernet

Metro optical Ethernet is based on using advanced switching methodologies and optical fiber telecom technologies that allow Ethernet to extend over greater distances than the original campus network implementation. Using 1550-nm wavelength systems, optical Ethernet enables transmission links to be up to 70 km (about 43 miles) long. Metro optical Ethernet allows business customers to connect multiple locations within a service area using native Ethernet protocol. In a typical service offering, transmission speeds can range from 3 Mb/s to 1 Gb/s. The service granularity can be in increments of 10 Mb/s in the 10-to-100 Mb/s service range and in increments of 100 Mb/s in the 100-to-1000 Mb/s service range. This service allocation allows for flexible growth and permits customers to choose a bandwidth profile that fits their actual traffic needs.

Figure 13.53 shows the general network operation of a metro optical Ethernet. Instead of the usual SONET/SDH ring, the transmission backbone is based on the 10-Gb/s Ethernet standard. The interface to the metro backbone is a 10-Gb/s Ethernet switch, which an enterprise can use to aggregate and exchange traffic from entities such as local telecom equipment, remote users, branch offices, and business partners. This process is known as *metropolitan aggregation*. The same Ethernet switch also provides high-performance, IP/MPLS, 10-Gb/s access to a wide area SONET/SDH network. Typical metro Ethernet switches have a variety of interface capabilities ranging from 10 Mb/s to 1 Gb/s. Such switches offer a high degree of flexibility because they integrate metro core and edge network functions for high-capacity aggregation, service provisioning, and network management in the metro core network.

13.11 Mitigation of Transmission Impairments

When designing WDM networks, transmission impairments such as chromatic and polarization-mode dispersion, optical amplifier transients, and timing jitter need to be taken into account. This section discusses various mitigation techniques for these impairments.

13.11.1 Chromatic Dispersion Compensating Fiber

A large base of more than 50 million kilometers of G.652 single-mode fiber was installed throughout the world in the early 1980s for use in single-wavelength transmission systems. Such fibers work very well for long-distance links of several hundred kilometers if the transmission wavelength is close to the 1310-nm zero-dispersion point. However, the large values of accumulated dispersion over even moderate distances of tens of kilometers at other wavelengths pose a problem. For example, the chromatic dispersion of a G.652 fiber is about 17 ps/(nm-km) in the C-band. Newer installations use G.655 non-zero dispersion-shifted fiber (NZDSF). Since these fibers have a chromatic dispersion of about 4 ps/(nm-km) in the C-band, for such links the chromatic dispersion accumulates slower than in G.652 fibers.

One approach to negate the accumulated dispersion of the transmission fiber is to use *dispersion compensation* techniques.[118–121] Two possible methods are the insertion of a dispersion-compensating fiber into the link or the use of a chirped fiber Bragg grating. A *dispersion-compensating fiber* (DCF) has a dispersion characteristic opposite that of the transmission fiber. Dispersion compensation is achieved by inserting a loop of DCF into the transmission path. The total dispersion in the DCF loop needs to be equal and opposite to the accumulated dispersion in the transmission fiber. If the transmission fiber has a low positive dispersion [say, 4 ps/(nm-km)], then the DCF will have a large negative dispersion [e.g., −90 ps/(nm-km)]. With this technique, the total accumulated dispersion is zero after some distance, but the absolute dispersion per length is nonzero at all points along the fiber. The nonzero absolute dispersion value causes a phase mismatch between wavelength channels, thereby destroying the possibility of effective FWM production.

Figure 13.54 shows that the DCF can be inserted at either the beginning or the end of an installed fiber span between two optical amplifiers. A third option is to have a DCF at both ends of an amplified link. In *precompensation* schemes the DCF is located right after the optical amplifier and just before the transmission fiber. Conversely, in *postcompensation* schemes the DCF is placed right after the
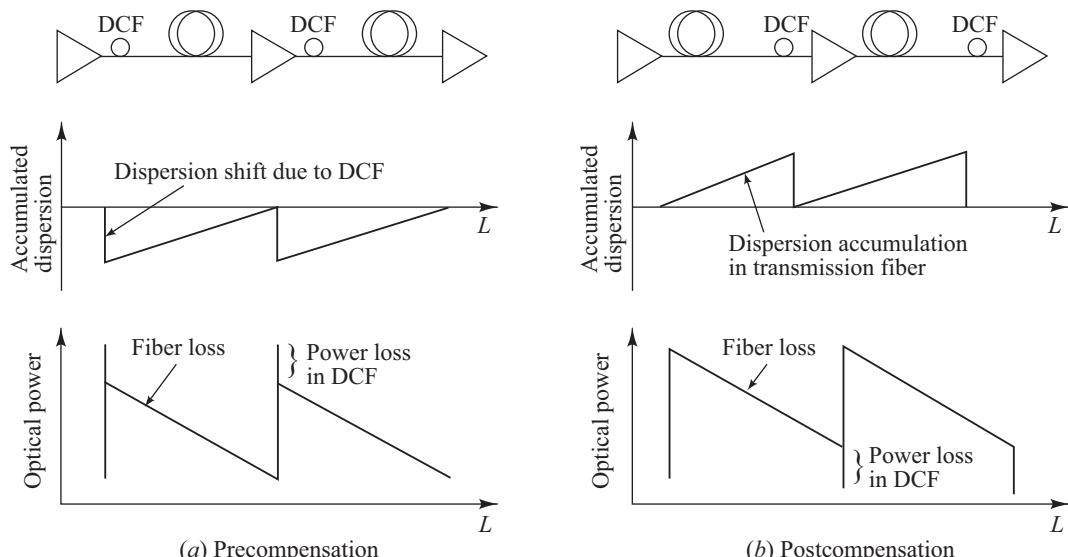
Fig. 13.54 *Dispersion maps and power maps for (a) precompensation and (b) postcompensation methods*

transmission fiber and just before the optical amplifier. Figure 13.54 also shows plots of the accumulated dispersion and the power level as functions of distance along the fiber. These plots are called *dispersion maps* and *power maps*, respectively. As Fig. 13.54a illustrates, in precompensation the DCF causes the dispersion to drop quickly to a low negative level from which it slowly rises toward zero (at the next optical amplifier) with increasing distance along the trunk fiber. This process repeats itself following amplification. The power map shows that the optical amplifier first boosts the power level to a high value. Since the DCF is a loop of fiber, there is a drop in power level before the signal enters the actual transmission path, in which it decays exponentially before being amplified once more.

Similar processes occur in postcompensation, as shown in Fig. 13.54b. In either case the accumulated dispersion is near zero after some distance to minimize the effects of pulse spreading, but the absolute dispersion per length is nonzero at all points, thereby causing a phase mismatch between different wavelengths, which mitigates FWM effects.

In actual systems, both experiments and simulations have shown that a *combination* of postcompensation and precompensation provides the best solution for dispersion compensation. When implementing such a compensation technique, the length of the DCF should be as short as possible since the special fiber used has a higher loss than the transmission fiber. The loss is around 0.5 dB/km at 1550 nm compared to 0.21 dB/km for G.655 fiber. Since around 1 km of DCF is needed for every 10 to 15 km of operational fiber, the additional DCF loss needs to be taken into account when designing a link. The required length L_{DCF} of the DCF fiber can be calculated using the expression

$$L_{DCF} = |\, D_{TX}/D_{DCF}\,| \times L \tag{13.23}$$

Here L is the length of the operational fiber, D_{TX} is the dispersion of the operational fiber, and D_{DCF} is the dispersion of the DCF.

13.11.2 Bragg Grating Dispersion Compensators

Another way of viewing dispersion is to consider the propagation speed of the different wavelength constituents of an optical pulse. When an optical pulse travels along a fiber in the anomalous-dispersion region (where $D_{TX} > 0$), the shorter-wavelength (higher-frequency) components of the pulse travel faster than the longer-wavelength (shorter-frequency) components. This is a dispersive effect that broadens the pulse.

To compensate for the difference in arrival times of the various frequency components, one can use a chirped fiber Bragg grating. As shown in Fig. 13.55, in such a dispersion compensator the grating spacing varies linearly over the length of the grating. This results in a range of wavelengths (or frequencies) that satisfy the Bragg condition for reflection. In the configuration shown, the spacings decrease along the fiber, which means that the Bragg wavelength decreases with distance along the grating length. Consequently the shorter-wavelength components of a pulse travel farther into the fiber before being reflected. Thereby they experience more delay in going through the grating than the longer-wavelength components. The relative delays induced by the grating on the different frequency components of the pulse are the opposite of the delays caused by the fiber. This results in dispersion compensation because the grating compresses the pulse.

Prior to the year 2000, manufacturing difficulties limited gratings to lengths of about 10 cm. Since the round-trip time T_R inside the grating of length L_G is given by $T_R = 2n_G L_G c$, where n_G is the refractive index of the grating fiber, the maximum round-trip delay time of light through a 10-cm long grating is 1 ns. The delay per unit length corresponds to the product of the dispersion D_G introduced by the grating and the spectral bandwidth $\Delta\lambda$ of the light being delayed, that is,

$$T_R/L_G = D_G \Delta\lambda \tag{13.24}$$
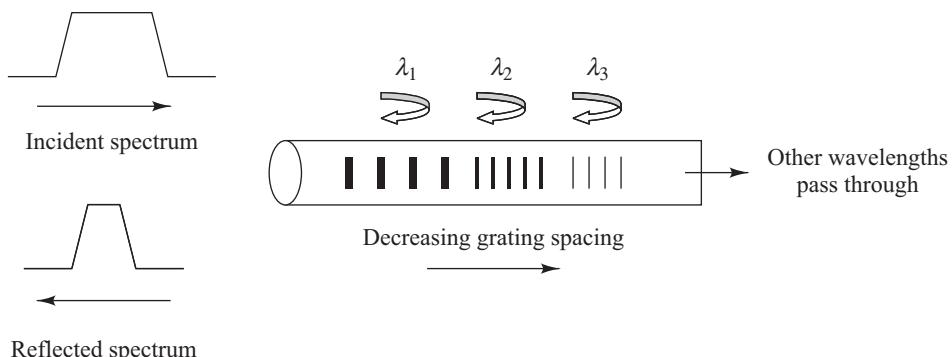
Fig. 13.55 *A chirped fiber Bragg grating can be used to compensate chromatic dispersion. Here $\lambda_1 > \lambda_2 > \lambda_3$.*

Thus one can make some application tradeoffs when using the chirped Bragg grating for dispersion compensation, such as 1000 ps/nm over a 1-nm bandwidth. Then, for example, since for G.655 fiber the total dispersion that arises over a 100-km length is about 500 ps/nm, in order to use a chirped Bragg grating to compensate for dispersion caused by several km of fiber, the gratings must have a very narrow bandwidth. That means the short gratings made prior to 2000 only could be used to compensate for individual wavelengths in a WDM system.

Improved manufacturing techniques now enable the fabrication of fiber gratings that are several meters long. For example, it is possible to fabricate a chirped fiber Bragg grating that is about 2 m long, has a bandwidth of over 30 nm, an insertion loss of about 1 dB, and a delay slope of -1.1 ps/nm^2. These types of chirped Bragg gratings are multichannel components that are capable of compensating for chromatic dispersion across the entire C-band or L-band in individual fibers.

13.11.3 Polarization-Mode Dispersion Compensation

The other dispersion factor in fibers that needs to be taken into account when designing a WDM network is *polarization-mode dispersion* (PMD). This dispersion results from the fact that light-signal energy at a given wavelength in a single-mode fiber actually occupies two orthogonal polarization states or modes (see Fig. 3.17). At the start of the fiber the two polarization states are aligned. However, since fiber material is not perfectly uniform throughout its length, each polarization mode will encounter a slightly different refractive index and will travel at a slightly different velocity. The resulting difference in propagation times between the two orthogonal polarization modes commonly is referred to as *differential group delay* (DGD) and will result in pulse spreading. This differential time delay between the polarization modes is called *first-order polarization-mode dispersion*. Furthermore, another dispersion factor arises since this delay depends on frequency and varies over the bandwidth of the pulse, which is similar to chromatic dispersion. This effect, called *second-order polarization-mode dispersion*, is the frequency derivative of PMD and can lead to further pulse spreading.

Currently manufactured low-PMD fibers have PMD values between 0.05 and 0.1 ps/$\sqrt{\text{km}}$. However, for a large percentage of the fielded fiber systems, the PMD values vary between 0.5 and 1.0 ps/$\sqrt{\text{km}}$. An availability of better than 99.999 percent is a standard specification for telecommunication networks.[122] This number translates to network outage times (total time during which the system is unavailable) of

less than 5 min/yr. Comparing this number to PMD-induced pulse spreading results in a mean PMD of 15 percent of a bit period. For a 40-Gb/s system the pulse period is 25 ps, which yields a maximum spread of 3.75 ps. This number restricts the transmission distance to about 55 km for a fiber with an average PMD of 0.5 ps/$\sqrt{\text{km}}$.

Whereas several methods exist for mitigating the effects of chromatic dispersion, it is more difficult to compensate for polarization-mode dispersion. This is because PMD varies with wavelength and slowly drifts randomly with time on the order of milliseconds. This random drift factor requires any compensation technique to adapt dynamically to polarization state changes while the system is running.

PMD compensation can be done either electronically or optically.[123–131] Electronic PMD compensation using adaptive electronics has been demonstrated at 10 Gb/s, and optical compensation techniques were used in field trails that transmitted data rates of 10, 40, and 160 Gb/s over standard G.652 single-mode fiber.

In an optical PMD-compensation process the basic procedure is to divide the received signal into its fast and slow polarization components. The fast polarization mode then is delayed in order to allow the slow mode to catch up. Ideally, the resulting polarization-mode dispersion should be zero. However, in practice, since the differential group delay changes with time, the separation and delay of both polarization modes can be estimated with only a limited accuracy and within a limited bandwidth.

Electronic compensation is based on the fact that, analogous to other forms of dispersion, PMD-induced pulse spreading causes *intersymbol interference* (ISI). Since *equalization* is the standard technique for mitigating ISI in electronic digital systems, a similar ISI-equalization process can be used to compensate for PMD. Equalization is done after the photodetection process by using linear and nonlinear adaptive electronic filters, which have transfer functions that are close to the inverse of the photodiode transfer function. The selection of the filters is quite complex due to noise effects and the statistical variation of PMD.

13.11.4   Optical Amplifier Gain Transients

Devices such as EDFAs, ROADMs, and OXCs are among the key components in WDM telecommunication networks. The ROADMs and OXCs allow rapid service provisioning, lightpath switching, congestion control, and fault recovery by dynamically adding and dropping wavelength channels. Since EDFAs usually operate in their saturation region, the rapid insertion and deletion of channels can create significant power transients in the remaining channels that the EDFA accommodates. As these abrupt power changes propagate through a series of optical amplifiers in a network, they can induce serious degradations in performance parameters such as the bit-error rate and the signal-to-noise ratio.

To avoid these performance degradations, it is essential to maintain a constant optical power gain for each amplified WDM channel. Several optical, electronic, and hybrid optoelectronic *automatic gain control* (AGC) techniques are in use to mitigate gain transients in EDFAs.[132–135] This process is referred to as *gain clamping* or *gain stabilization*.

Optical methods clamp the EDFA gain through the use of a separate control channel that is outside the spectral band of the signal wavelengths but inside the spectral response band of the EDFA. The two common optical control methods use optical power feedback mechanisms to stabilize the EDFA saturation level. In the first mechanism, distributed Bragg reflector gratings located at both ends of the EDFA create a negative feedback at the Bragg wavelength. Another optical control method uses a ring laser and wavelength-selective optical filters that allow only the lasing wavelength to be fed back into the EDFA. The settling times for the gain are tens of microseconds. However, the technique requires relatively high optical pump powers for the gain-clamping channel, and the technique produces some unwanted laser relaxation oscillations.

Electronic gain control is achieved by rapidly tuning the EDFA pump power. The first of two basic electronic gain control techniques uses a feedforward method that adjusts the pump power according to variations in the total EDFA optical input power. The second electronic gain method uses a negative feedback loop to adjust the pump power. Settling times for these two methods vary from 1 to 10 μs. In an actual implementation, a gain-controlled EDFA might consist of several amplification stages to optimize various performance functions. The functions of these stages could include achieving a low noise figure, generating a high per channel optical gain, and compensating for power loss in EDFA-associated components such as gain-equalizing filters and dispersion-compensating fibers.

Gain transients in Raman amplifiers also need to be considered when using a hybrid Raman-plus-EDF amplifier. With such an amplifier the dynamic adding and dropping of WDM channels affects not only the EDFA gain but also the Raman gain. Although the Raman amplification mechanisms may involve the use of several Raman pump lasers, investigations have shown that variations in the spectral distribution of each amplified WDM channel are negligible during gain transients. Thus for practical applications the variations in the combined EDFA-plus-Raman gains can be compensated for either by employing fast clamping of only the EDFA gain or by means of a simple feedforward gain-control method that uses the total input signal power as the feed signal to adjust the Raman pump powers.

PROBLEMS

13.1 An engineer plans to construct an in-line optical fiber data bus operating at 10 Mb/s. The stations are to be separated by 100 m, for which optical fibers with a 3-dB/km attenuation are used. The optical sources are laser diodes having a 500-μW (–3 dBm) output from a fiber flylead, and the detectors are avalanche photodiodes with a 1.6-nW (–58-dBm) sensitivity. The couplers have a power tap-off factor of $C_T = 5$ percent and a 10-percent fractional intrinsic loss F_i. The power loss at the connectors is 20 percent (1 dB).

(*a*) Make a plot of P_{1N} in dBm as a function of the number of stations N for $2 \leq N \leq 12$.

(*b*) What is the operating margin of the system for eight stations?

(*c*) What is the worst-case dynamic range for the maximum allowable number of stations if a 6-dB power margin is required?

13.2 Consider an N-node star network in which 0 dBm of optical power is coupled from any given transmitter into the star. Let the fiber loss be 0.3 dB/km. Assume the stations are located 2 km from the star, the receiver sensitivity is –38 dBm, each connector has a 1-dB loss, the excess loss in the star coupler is 3 dB, and the link margin is 3 dB.

(*a*) Determine the maximum number of stations N that can be incorporated on this network.

(*b*) How many stations can be attached if the receiver sensitivity is -32 dBm?

13.3 A two-story office building has two 10-feet-wide hallways per floor that connect four rows of offices with eight offices per row as is shown in Fig. 13.56. Each office is a 15 feet $\times$ 15 feet square. The office ceiling height is 9 feet with a false ceiling hung 1 foot below the actual ceiling. Also, as shown in Fig. 13.56, there is a wiring room for LAN interconnection and control equipment in one corner of each floor. Every office has a local-area network socket on each of the two walls that are perpendicular to the hallway wall. If we assume that cables can be run only in the walls and in the ceilings, estimate the length of cable (in feet) that is required for the following configurations:

(*a*) A coaxial cable bus with a twisted-pair wire drop from the ceiling to each outlet.

(*b*) A fiber optic star that connects each outlet to the wiring room on the corresponding floor and a vertical fiber optic riser that connects the stars in each wiring room.
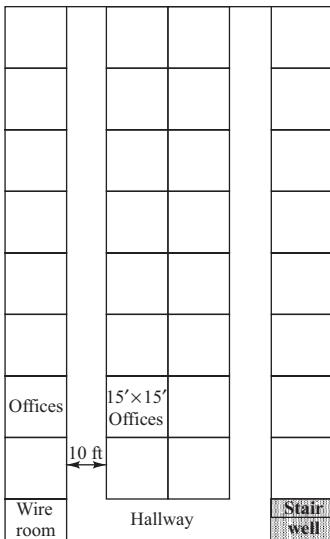
Fig. 13.56

13.4 Consider the $M \times N$ grid of stations shown in Fig. 13.57 that are to be connected by a local-area network. Let the stations be spaced a distance d apart and assume that interconnection cables will be run in ducts that connect nearest-neighbor stations (i.e., ducts are not run diagonally in Fig. 13.57). Show that for the following configurations, the cable length for interconnecting the stations is as stated:

(*a*) $(MN - 1)\, d$ for a bus configuration.

(*b*) MNd for a ring topology.
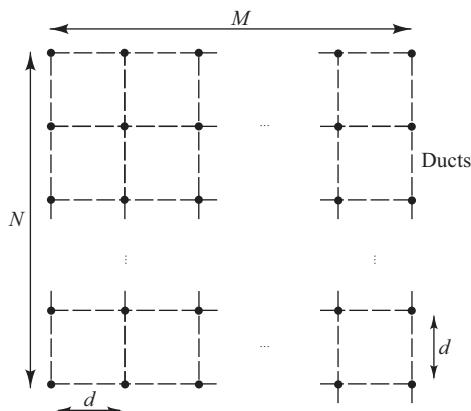
Fig. 13.57

(*c*) $MN\,(M + N - 2)\, d/2$ for a star topology where each subscriber is connected *individually* to the network hub located *in one corner* of the grid.

13.5 Consider the $M \times N$ rectangular grid of computer stations shown in Fig. 13.57, where the spacing between stations is d. Assume these stations are to be connected by a star-configured LAN using the duct network shown in the figure. Furthermore, assume that each station is connected to the central star by means of its own dedicated cable.

(*a*) If m and n denote the relative position of the star, show that the total cable length L needed to connect the stations is given by

$$L = [\, MN(M + N + 2)/2 - Nm \\ \times (M - m + 1) - Mn\,(N - n + 1)]\, d$$

(*b*) Show that if the star is located in one corner of the grid, then this expression becomes

$$L = MN\,(M + N - 2)d/2$$

(*c*) Show that the shortest cable length is obtained when the star is at the center of the grid.

13.6 In this problem do not resort to wavelength conversion but assume that wavelengths can be reused in different parts of a network.

(*a*) Show that the minimum number of wavelengths required to connect N nodes in a WDM network is as follows:
- $N - 1$ for a star network.
- $(N/2)^2$ for N even or $(N - 1)\,(N + 1)/4$ for N odd in a bus network.
- $N(N - 1)/2$ for a ring network.

(*b*) Draw example networks and their wavelength assignments for $N = 3$ and 4.

(*c*) Plot the number of wavelengths versus the number of nodes for star, bus, and ring networks for $2 \le N \le 20$.

13.7 Using information from the web, find some vendors of tap couplers and from their data sheets list representative performance specifications for several different types.

Consider characteristics such as tap ratio (use 1 to 5 percent), insertion loss, return loss, optical power handling limit, package configuration, and physical size.

13.8 (*a*) Calculate how many 64-kb/s voice channels can be carried by an STS-3, STS-48, and STS-192 system.

(*b*) How many 20-Mb/s digitized video channels can be transported over these systems?

13.9 Compare the system margins for 40-km and 80-km long-haul OC-48 (STM-16) links at 1550 nm for the minimum and maximum source output ranges. Assume there is a 1.5-dB coupling loss at each end of the link. Use Tables 13.4 and 13.5.

13.10 Verify that the maximum optical powers per wavelength channel given in Table 13.6 yield a total power level of + 17 dBm in an optical fiber.

13.11 Two SONET rings need to be interconnected at two mutual nodes in order to ensure redundant paths under failure conditions. Draw the interconnection between two bidirectional line-switched (BLSR) SONET rings showing the primary and secondary path setups that designate the signal flows under normal and failure conditions. In designing the interface, consider the following possible failure conditions of the two mutual nodes:

(*a*) Failure of a transmitter or receiver in one of the nodes.

(*b*) Failure of an entire node.

(*c*) A fiber break in the link between the two nodes.

13.12 Similar to Prob. 13.11 draw the interconnection between a UPSR and a BLSR.

13.13 Consider the four-node network shown in Fig.13.58. Each node uses a different combination of three wavelengths to communicate with the other nodes, so that there are six different wavelengths in the network. Given that node 1 uses λ_2, λ_4, and λ_6 for information exchange with the other nodes (i.e., these wavelengths are added and dropped at node 1, and the remaining wavelengthsw from other nodes pass through), establish wavelength assignments for the other nodes.
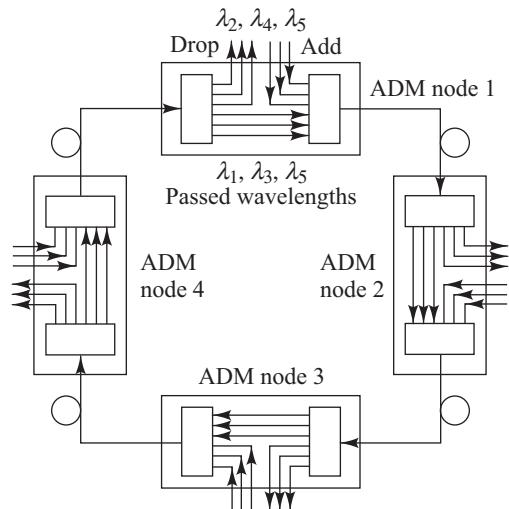
Fig. 13.58

13.14 Figure 13.59 shows two architectures in which switches share wavelength converters for cost-saving purposes. Find a set of wavelength connections that can be set up with the share-per-node architecture in Fig. 13.59*a* but not with the share-per-link architecture in Fig. 13.59*b*, and vice versa.
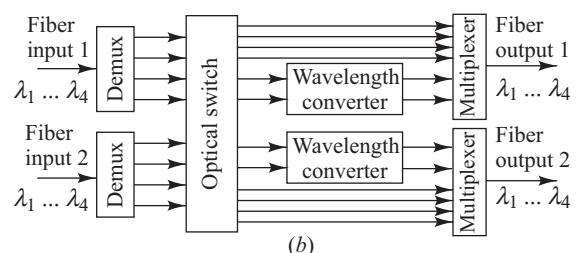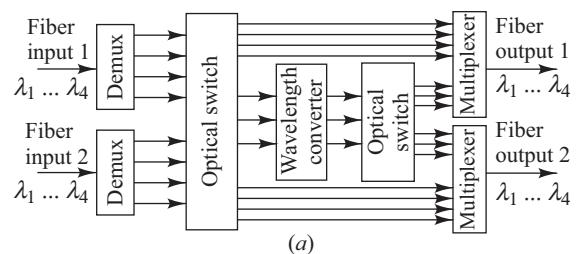
Fig. 13.59

13.15 Consider the network consisting of three interconnected rings, shown in Fig. 13.60. Here, the circles represent nodes that contain optical switches and wavelength converters. These nodes can receive two wavelengths from any direction and can transmit them out over any line. The squares are access stations that have tunable optical transmitters and receivers (both wavelengths can be transmitted and received at any access station). Suppose the network has two wavelengths available to set up the following paths:

(*a*) A-1-2-5-6-F

(*b*) B-2-3-C

(*c*) B-2-5-8-H

(*d*) G-7-8-5-6-F

(*e*) A-1-4-7-G

Assign the two wavelengths to these paths and show at which nodes wavelength conversion is needed.

13.16 Analogous to Fig.13.34, use Eq. (13.22) to plot the gain $G = q/p$ as a function of the number of wavelengths F for blocking probabilities $P_b = 10^{-4}$ and 10^{-5}. Use the same hopping number values $H = 5$, 10, and 20. For $H = 20$, plot G versus F on the same graph for blocking probabilities $P_b = 10^{-3}$, 10^{-4}, and 10^{-5}. What is the effect on the gain as the blocking probability increases?

13.17 Suppose an engineer wants to attach a 50-m long standard 50-μm fiber that has a bandwidth of 500 MHz-km to a 100-m higher-grade 50-μm fiber that has a bandwidth of 2000 MHz-km. A 10-m length of the lower grade fiber is used at each end in order to connect to transmission equipment. Will this link work at 10 Gb/s?
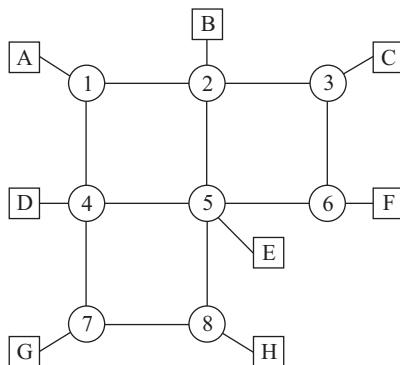
13.18 Consider an OBS system in which the processing and configuration times are both 1 ms. Similar to Fig. 13.37, suppose the burst goes between edge routers A and B through four intermediate OXCs and that all nodes are 10 km apart. If the OBS control packet is sent by A at time $t = 0$ and the propagation delay in an optical fiber is 5 μs/km, calculate the following:

(*a*) The offset time

(*b*) The times at which the control packet will arrive at and leave from each intermediate OXC

(*c*) The times at which the burst will arrive at each intermediate OXC.

13.19 Consider the passive optical network shown in Fig. 13.41. Assume the following:

(*a*) Laser sources at 1310 and 1490 nm can launch optical powers of 2.0 and 3.0 dBm, respectively, into a fiber

(*b*) The passive optical splitter is 10 km from the central office

(*c*) The users are 5 km from the power splitter

(*d*) The insertion losses are 13.5 and 16.6 dB for 1×16 and 1×32 splitters, respectively.

(*e*) The fiber attenuation is 0.6 dB/km at 1310 nm and 0.3 dB/km at 1550 nm

(*f*) For simplicity assume there are no connector, splice, or other losses in the link.

Find the total link loss for the following four situations:

(i) A 1310-nm laser transmits upstream (user to central office) through a 1×16 splitter

(ii) A 1310-nm laser transmits upstream through a 1×32 splitter

(iii) A 1550-nm laser transmits downstream through a 1×16 splitter

(iv) A 1550-nm laser transmits downstream through a 1×32 splitter

13.20 Suppose an $N \times N$ star coupler is constructed of a set of M 2×2 couplers, each of which

Fig. 13.60

has a 0.1-dB insertion loss. Find the maximum value of M and the maximum size N if the power budget for the star coupler is 30 dB.

13.21 Suppose the optical power emerging from a feeder fiber is to be distributed among eight individual houses. Assume that they are separated by 100 m and lie along a straight line going out from the end of the feeder fiber. One way to distribute the power is to use a 1×8 star coupler and run individual fibers to each house. Why is this preferable to the configuration where a single fiber runs along the line to the homes and individual tap couplers in the fiber line each extract 10 percent of the power from the line for each house that the line passes?

13.22 (*a*) Why may it be more important to encrypt information in the downstream direction in a PON and not necessarily in the upstream path?

(*b*) How might information be compromised either maliciously or unintentionally in either direction?

(*c*) How might encryption be implemented in the upstream and the downstream directions?

13.23 In terms of network configuration flexibility and ease of maintenance and troubleshooting, at what cable interfaces to units such as optical splitters, fiber termination boxes, patch panels, and WDM couplers would it be better to use optical connectors instead of splices?

13.24 Draw network diagrams of how a 1+1 protective method and a 1:1 arrangement could be implemented for a GPON. Describe how each configuration would function.

13.25 The basic unit for processing the Advanced Encryption Standard (AES) algorithm is the byte. For convenience, byte values are denoted using hexadecimal notation with each of two groups of four bits denoted by a single character, as shown in Table 13.11. For example, the element {01100011} can be represented by {63}. Using Table 13.11, find the binary equivalent of the following words:

(*a*) 32 43 f6 a8 88 5a 30 8d 31 31 98 a2 e0 37 07 34

Table 13.11 *Hexadecimal representation of bit patterns*

Bit pattern	Character
0000	0
0001	1
0010	2
0011	3
0100	4
0101	5
0110	6
0111	7
1000	8
1001	9
1010	a
1011	b
1100	c
1101	d
1110	e
1111	f

(*b*) 2b 7e 15 16 28 ae d2 a6 ab f7 15 88 09 cf 4f 3c

13.26 Since the ranging procedure in a PON has limited accuracy, a guard time is placed between consecutive bursts from the ONTs to avoid collisions of the independent packets, as Fig. 7.18 illustrates. Verify that a 25.6-ns guard time consumes 16 bits at 622 Mb/s, 32 bits at 1244 Mb/s, and 64 bits at 2.488 Gb/s.

13.27 An ONT intended for indoor installation requires a 12-V dc power supply and consumes 15 W of electric power during normal operation. If there is an electric power outage, the ONT disables nonessential services and now consumes 7 W of power supplied by backup batteries. Using vendor data sheets, select an indoor uninterruptible power supply (UPS) that satisfies these requirements.

13.28 An ONT intended for outdoor installation requires a 12-V dc power supply and consumes 15 W of electric power during normal operation. If there is an electric power outage, the ONT disables nonessential services and now consumes 7 W of power supplied by backup batteries. Using vendor data sheets, select an

outdoor uninterruptible power supply (UPS) that satisfies these requirements. Assume the

UPS must operate over temperature extremes of −25°C to +35°C.

REFERENCES

1. A. Gladisch, R.-F. Braun, D. Breuer, A. Ehrhardt, H.-M. Foisel, M. Jaeger, R. Leppla, M. Schneiders, S. Vorbeck, W. Weiershausen, and F.-J. Westphal, "Evolution of terrestrial optical system and core network architecture," *Proc. IEEE*, vol. 94, pp. 869–891, May 2006.

2. N. Taesombut, F. Uyeda, A. A. Chien, L. Smarr, T. A. DeFanti, P. Papadopoulos, J. Leigh, M. Ellisman, and J. Orcutt, "OptIPuter: High-performance, QoS-guaranteed network service for emerging e-science applications," *IEEE Commun. Mag.*, vol. 44, pp. 38–45, May 2006.

3. C. H. Cox III, E. I. Ackerman, G. E. Betts, and J. L. Prince, "Limits on the performance of RF-over-fiber links and their impact on device design," *IEEE Trans. Microw. Theory Tech.*, vol. 54, pp. 906–920, Feb. 2006.

4. R. Gaudino, D. Cárdenas, M. Bellec, B. Charbonnier, N. Evanno, P. Guignard, S. Meyer, A. Pizzinat, I. Möllers, and D. Jäger, "Perspective in next-generation home networks: Toward optical solutions?," *IEEE Commun. Mag.*, vol. 48, pp. 39–47, Feb. 2010.

5. T. Plevyak and V. Sahin, *Next Generation Telecommunications Networks, Services, and Management*, Wiley, Hoboken, NJ, 2010.

6. B. St. Arnaud, J. Wu, and B. Kalali, "Customer-controlled and –managed optical networks," *J. Lightwave Tech.*, vol. 21, pp. 2804–2810, Nov. 2003.

7. H. Al-Raweshidy and S. Komaki, eds., *Radio over Fiber Technologies*, Artech House, Boston, 2002.

8. N. S. Bergano, "Wavelength division multi-plexing in long-haul transoceanic transmission systems," *J. Lightwave Tech.*, vol. 23, pp. 4125–4139, Dec. 2005.

9. B. A. Forouzan, *Data Communications Networking*, McGraw-Hill, Burr Ridge, IL, 4th ed., 2007.

10. A. Leon-Garcia and I. Widjaja, *Communication Networks*, McGraw-Hill, Burr Ridge, IL, 2nd ed., 2004.

11. L. Peterson and B. Davie, *Computer Networks*, Morgan Kauffman, 2004.

12. G. Keiser, *Local Area Networks*, McGraw-Hill, Burr Ridge, IL, 2nd ed., 2002.

13. ITU-T Recommendation X.210

14. B. Mukherjee, *Optical WDM Networks*, Springer, New York, 2006.

15. R. Ramaswami and K. N. Sivarajan, *Optical Networks*, Morgan Kaufmann, San Francisco, 3rd ed., 2009.

16. H. Tsushima, S. Hanatani, T. Kanetake, J. A. Fee, and S.-K. Liu, "Optical cross-connect system for survivable optical layer networks," *Hitachi Review*, vol. 47, pp. 85–90, no. 2, 1998.

17. W. D. Grover, *Mesh-Based Survivable Networks: Options and Strategies for Optical, MPLS, SONET, and ATM Networking*, Prentice Hall, Upper Saddle River, NJ, 2004.

18. C. Assi, W. Huo, A. Shami, and N. Ghani, "Analysis of capacity re-provisioning in optical mesh networks," *IEEE Commun. Lett.*, vol. 9, pp. 658–660, July 2005.

19. H. G. Perros, *Connection-Oriented Networks: SONET/SDH, ATM, MPLS and Optical Networks*, Wiley, Hoboken, NJ, 2005.

20. W. Goralski, *SONET/SDH*, McGraw-Hill, New York, 3rd ed., 2003.

21. D. Minoli, P. Johnson, and E. Minoli, *SONET-Based Metro Area Networks*, McGraw-Hill, New York, 2002.

22. American National Standards Institute (ANSI). Sample SONET standards:
   (*a*) T1.105, *Basic Description Including Multiplex Structures, Rates, and Formats*
   (*b*) T1.105.06, *SONET-Physical Layer Specification*
   (*c*) T1.119, *Operations, Administration, Maintenance, and Provisioning Communications*
   (*d*) T1.514, *SONET Bit Rates*

23. International Telecommunication Union-Telecommunication Standardization Sector (ITU-T). Sample SDH recommendations:

(*a*) G.692, *Optical Interfaces for Multichannel Systems with Optical Amplifiers*

(*b*) G.707, *Network Node Interface for the Synchronous Digital Hierarchy*

(*c*) G.841, *Types and Characteristics of SDH Network Protection Architectures*

(*d*) G.957, *Optical Interfaces for Equipments and Systems Relating to the Synchronous Digital Hierarchy*

24. H. C. Ji, J. H. Lee, and Y. C. Chung, "Evaluation of system outage probability due to temperature variation and statistically distributed chromatic dispersion of optical fiber," *J. Lightwave Tech.* vol. 22, pp. 1893–1898, Aug. 2004.

25. S. Wright, H. Fahmy, and A. J. Vernon, "Deployment challenges for access/metro optical networks and services," *J. Lightwave Tech.*, vol. 22, pp. 2606–2616, Nov. 2004.

26. H. T. Mouftah and P.-H. Ho, *Optical Networks: Architecture and Survivability*, Springer, New York, 2003.

27. W.-P. Lin, M.-S. Kao, and S. Chi, "A DWDM/SCM self-healing architecture for broad-band subscriber networks," *J. Lightwave Tech.*, vol. 21, pp. 319–328, Feb. 2003.

28. M.-J. Li, M. J. Soulliere, D. J. Tebben, L. Nederlof, M. D. Vaughn, and R. E. Wagner, "Transparent optical protection ring architectures and applications," *J. Lightwave Tech.*, vol. 23, pp. 3388–3403, Oct. 2005.

29. S. Priyadarshi, S. Z. Zhang, M. Tokhmakhian, H. Yang, and N. Margalit, "The first hot pluggable 2.5 Gb/s DWDM transceiver in an SFP form factor," *IEEE Commun. Mag.*, vol. 43, pp. *S*29–*S*31, Feb. 2005.

30. I. Zacharapoulos, A. Tzanakaki, D. Parcharidou, and I. Tomkos, "Optimization study of advanced modulation formats for 10-Gb/s metropolitan networks," *J. Lightwave Tech.*, vol. 23, pp. 321–329, Jan. 2005.

31. C. M. DeCusatis, "Fiber optic cable infrastructure and dispersion compensation for storage area networks," *IEEE Commun. Mag.*, vol. 43, pp. 86–92, Mar. 2005.

32. P. Herve and S. Ovadia, "Optical network technologies for enterprise networks," *Intel Tech. J.*, vol. 8, no. 2, pp. 73–82, May 2004.

33. C. Babla, "Addressing challenges in serial 10 Gb/s multimode fiber enterprise networks," *IEEE Commun. Mag.*, vol. 43, pp. S22–S28, Feb. 2005.

34. P. Kirkpatrick, W.-C. Fang, H. Johansen, B. Christensen, J. Hanberg, M. Lobel, T. B. Mader, S. Q. Shang, C. Schulz, D. Sprock, and J.-M. Verdiel, "10 Gb/s optical transceivers: Fundamentals and emerging technologies," *Intel Tech. J.*, vol. 8, no. 2, pp. 83–99, May 2004.

35. J. Hecht, "Optical modules ease design of fiber optic systems," *Laser Focus World*, vol. 40, pp. 114–118, Aug. 2004.

36. R. DeSalvo, A. G.Wilson, J. Rollman, D. F. Schneider, L. M. Lunardi, S. Lumish, N. Agrawal, A. H. Steinbach, W. Baun, T. Wall, R. Ben-Michael, M. A. Itzler, A. Fejzuli, R. A. Chipman, G. T. Kiehne, and K. M. Kissa, "Advanced components and sub-system solutions for 40 Gb/s transmission," *J. Lightwave Tech.*, vol. 20, pp. 2154–2181, Dec. 2002.

37. O. Leclerc, B. Lavigne, E. Balmefrezol, P. Brindel, L. Pierre, D. Rouvillain, and F. Seguineau, "Optical regeneration at 40 Gb/s and beyond," *J. Lightwave Tech.*, vol. 21, pp. 2779–2790, Nov. 2003.

38. M. Daikoku, N. Yoshikane, T. Otani, and H. Tanaka, "Optical 40-Gb/s 3R regenerator with a combination of the SPM and XAM effects for all-optical networks," *J. Lightwave Tech.*, vol. 24, pp. 1142–1148, Mar. 2006.

39. A. Scavennec and O. Leclerc, "Toward high-speed 40-Gb/s transponders," *Proc. IEEE*, vol. 94, pp. 986–996, May 2006.

40. J.-X. Cai, C. R. Davidson, M. Nissov, H. Li, W. T. Anderson, Y. Cai, L. Liu, A. N. Pilipetskii, D. G. Foursa, W. W. Patterson, P. C. Corbett, A. J. Lucero, and N. S. Bergano, "Transmission of 40-Gb/s WDM signals over transoceanic distance using conventional NZ-DSF with receiver dispersion slope compensation," *J. Lightwave Tech.*, vol. 24, pp. 191–200, Jan. 2006.

41. J. G. Proakis and M. Salehi, *Digital Communications*, McGraw-Hill, New York, 5th ed., 2008.

42. L. W. Couch II, *Digital and Analog Communication Systems*, Prentice Hall, Upper Saddle River, NJ, 7th ed., 2007.

43. C.-K. Chan, L.-K. Chen, and K.-W. Cheung, "A fast channel-tunable optical transmitter for ultra-high-speed all-optical time-division multiaccess networks," *IEEE J. Sel. Areas Commun.*, vol. 14, pp. 1052–1056, June 1996.

44. D. Cotter, J. K. Lucek, and D. D. Marcenac, "Ultra-high-bit-rate networking: From the transcontinental backbone to the desktop," *IEEE Commun. Mag.*, vol. 35, pp. 90–95, Apr. 1997.

45. A. R. Chraplyvy and R. W. Tkach, "Terabit/second transmission experiments," *IEEE J. Quantum Electron.*, vol. 34, pp. 2103–2108, Nov. 1998.

46. A. Bogoni, L. Poti, P. Ghelfi, M. Scaffardi, C. Porzi, F. Ponzini, G. Meloni, G. Berrettini, A. Malacarne, G. Prati, "OTDM-based optical communications networks at 160 Gbit/s and beyond," *Optical Fiber Technol.*, vol. 13, pp. 1–12, Jan. 2007.

47. M. Schneiders, S. Vorbeck, R. Leppla, E. Lach, M. Schmidt, S. B. Papernyi, and K. Sanapi, "Field transmission of 8×170 Gb/s over high-loss SSMF link using third-order distributed Raman amplification," *J. Lightwave Tech.*, vol. 24, pp. 175–182, Jan. 2006.

48. T. Miyazaki, Y. Awaji, Y. Kamio, and F. Kubota, "Field demonstration of 160-Gb/s OTDM signal using eight 20-Gb/s 2-bit/symbol channels over 200 km," *OFC/NFOEC Technical Digest*, paper OFF1, March 2005.

49. J. P. Turkkiewicz, E. Tangiongga, G. Lehmann, H. Rohde, W. Schairer, Y. R. Zhou, E.S.R. Sikora, A. Lord, D. B. Payne, G.-D. Khoe, and H. de Waardt, "160 Gb/s OTDM networking using deployed fiber," *J. Lightwave Tech.*, vol. 23, pp. 225–235, Jan. 2005.

50. C. F. Lam, N. F. Frigo, and M. D. Feuer, "A taxonomical consideration of optical add/drop multiplexers," *Photonic Network Commun.*, vol. 3, no. 4, pp. 327–334, Oct. 2001.

51. M. Vasilyev, I. Tomkos, M. Mehendale, J.-K. Rhee, A. Kobyakov, M. Ajgaonkar, S. Tsuda, and M. Sharma, "Transparent ultra-long-haul DWDM networks with broadcast-and-select OADM/OXC architecture," *J. Lightwave Tech.*, vol. 21, pp. 2661–2672, Nov. 2003.

52. N. Kataoka, N. Wada, K. Sone, Y. Aoki, H. Miyata, H. Onaka, and K. Kitayma, "Field trial of data-granularity-flexible reconfigurable OADM with wavelength-packet-selective switch," *J. Lightwave Tech.*, vol. 24, pp. 88–94, Jan. 2006.

53. H. Zhu and B. Mukherjee, "Online connection provisioning in metro optical WDM networks using reconfigurable OADMs," *J. Lightwave Tech.*, vol. 23, pp. 2893–2901, Oct. 2005.

54. D. M. Marom, D. T. Neilson, D. S. Greywall, C.-S. Pai, N. R. Basavanhally, V. A. Aksyuk, D. O. López, F. Pardo, M. E. Simon, Y. Low, P. Kolodner, and C. A. Bolle, "Wavelength-selective $1 \times K$ switches using free-space optics and MEMS micromirrors: Theory, design, and implementation," *J. Lightwave Tech.*, vol. 23, pp. 1620–1630, Apr. 2005.

55. D. J. Blumenthal, P. R. Prucnal, and J. R. Sauer, "Photonic packet switches: Architectures and experimental implementations," *Proc. IEEE*, vol. 82, pp. 1650–1667, Nov. 1994.

56. G. I. Papadimitriou, C. Papazoglou, and A. S. Pomportsis, "Optical switching: Switch fabrics, techniques, and architectures," *J. Lightwave Tech.*, vol. 21, pp. 384–405, Feb. 2003.

57. E. Iannone and R. Sabella, "Optical path technologies: A comparison among different cross-connect architectures," *J. Lightwave Tech.*, vol. 14, pp. 2184–2196, Oct. 1996.

58. H. Yuan, W.-D. Zhong, and W. Hu, "FBG-based bidirectional optical crossconnects for bidirectional WDM ring networks," *J. Lightwave Tech.*, vol. 22, pp. 2710–2721, Dec. 2004.

59. J. Gripp, M. Duelk, J. E. Simsarian, A. Bhardwaj, P. Bernasconi, O. Laznicka, and M. Zirngibl, "Optical switch fabrics for ultra-high capacity IP routers," *J. Lightwave Tech.*, vol. 21, pp. 2839–2850, Nov. 2003.

60. A. Watanabe, K. Noguchi, K. Shimano, T. Kawai, E. Yoshida, A. Sahara, T. Takahashi, S. Okamoto, T. Goh, Y. Takigawa, M. Koga, and K.-I. Sato, "Photonic MPLS router to create bandwidth-abundant IP routers," *J. Lightwave Tech.*, vol. 21, pp. 2851–2862, Nov. 2003.

61. R. A. Barry and P. Humblet, "Models of blocking probability in all-optical networks with and without wavelength conversion," *IEEE J. Select Areas Commun.*, vol. 14, pp. 858–867, June 1996.

62. D. Marcenac, "Benefits of wavelength conversion in optical ring-based networks," *Optical Networks Mag.*, vol. 1, no. 2, pp. 29–35, Apr. 2000.

63. J. Strand, R. Doverspike, and G. Li, "Importance of wavelength conversion in an optical network," *Optical Networks Mag.*, vol. 2, no. 3, pp. 33–44, May/June 2001.

64. X. Chu, B. Li, and I. Chlamtac, "Wavelength converter placement under different RWA algorithms in wavelength-routed all-optical networks," *IEEE Trans. Commun.*, vol. 51, pp. 607–617, Apr. 2003.

65. X. Cao, V. Anand, Y. Xiong, and C. Qiao, "A study of waveband switching with multilayer multigranular optical cross-connects," *IEEE J. Select Areas Commun.*, vol. 21, pp. 1081–1095, Sept. 2003.

66. C. Nuzman, J. Leuthold, R. Ryf, S. Chandrasekhar, C. R. Giles, and D. T. Neilson, "Design and implementation of wavelength-flexible network nodes," *J. Lightwave Tech.*, vol. 21, pp. 648–663, Mar. 2003.

67. X. Chu and B. Li, "Dynamic routing and wavelength assignment in the presence of wavelength conversion for all-optical networks," *IEEE/ACM Trans. Networking*, vol. 13. pp. 704–715, June 2005.

68. J. M. Tang and K. A. Shore, "Wavelength-routing capability of reconfigurable optical add/drop multiplexers in dynamic optical networks," *J. Lightwave Tech.*, vol. 24, pp. 4296–4303, Nov. 2006.

69. H.-Y. Jeong and S.-W. Seo, "Blocking in wavelength-routed optical networks with heterogeneous traffic," *IEEE J. Sel. Areas Commun.*, vol. 23, pp. 1643–1657, Aug. 2005.

70. A. Pattavina, "Architectures and performance of optical packet switching nodes for IP networks," *J. Lightwave Tech.*, vol. 23, pp. 1023–1032, Mar. 2005.

71. D. J. Blumenthal, B.-E. Olsson, G. Rossi, T. E. Dimmick, L. Rau, M. Masanovic, O. Lavrova, R. Doshi, O. Jerphagnon, J. E. Bowers, V. Kaman, L. A. Coldren, and J. Barton, "All-optical label swapping networks and technologies," *J. Lightwave Tech.*, vol. 18, pp. 2058–2075, Dec. 2000.

72. D. Gurkan, S. Kumar, A. E. Willner, K. R. Parameswaran, and M. M. Fejer, "Simultaneous label swapping and wavelength conversion of multiple independent WDM channels in an all-optical MPLS network using PPLN waveguides as wavelength converters," *J. Lightwave Tech.*, vol. 21, pp. 2739–2745, Nov. 2003.

73. W. Wang, L. G. Rau, and D. J. Blumenthal, "160 Gb/s variable length packet/10 Gb/s-label all-optical label switching with wavelength conversion and unicast/multicast operation," *J. Lightwave Tech.*, vol. 23, pp. 211–218, Jan. 2005.

74. Y.-M. Lin, M. C. Yuang, S.-L. Lee, and W.I. Way, "Using superimposed ASK label in a 10 Gbps multi-hop all-optical label swapping system," *J. Lightwave Tech.*, vol. 22, pp. 351–361, 2004.

75. J. P. Jue and V. M. Vokkarane, *Optical Burst Switched Networks*, Springer, New York, 2005.

76. H. Yang and S. J. B. Yoo, "All-optical variable buffering strategies and switch fabric architectures for future all-optical data routers," *J. Lightwave Tech.*, vol. 23, pp. 3321–3330, Oct. 2005.

77. G. I. Papadimitriou, C. Papazoglou, and A. S. Pomportsis, *Optical Switching*, Wiley, Hoboken, NJ, 2007.

78. T. S. El-Bawab, *Optical Switching*, Springer, New York, 2010.

79. R. Rajaduray, S. Ovadia, and D. J. Blumenthal, "Analysis of an edge router for span-constrained optical burst switched (OBS) networks," *J. Lightwave Tech.*, vol. 22, pp. 2693–2705, Nov. 2004.

80. X. Yu, J. Li, X. Cao, Y. Chen, and C. Qiao, "Traffic statistics and performance evaluation in optical burst switched (OBS) networks," *J. Lightwave Tech.*, vol. 22, pp. 2722–2738, Dec. 2004.

81. M. Izal, J. Aracil, D. Morató, and E. Magaña, "Delay-throughput curves for timer-based OBS burstifiers with light load," *J. Lightwave Tech.*, vol. 24, pp. 277–285, Jan. 2006.

82. H. Le Vu, A. Zalesky, E.W.M. Wong, Z. Rosberg, S.M.H. Bilgrami, M. Zukerman, and R. S. Tucker, "Scalable performance evaluation of a hybrid optical switch," *J. Lightwave Tech.*, vol. 23, pp. 2961–2973, Oct. 2005.

83. D. Simeonidou, R. Nejabati, G. Zervas, D. Klonidis, A. Tzanakaki, and M. J. O'Mahony, "Dynamic optical-network architectures and technologies for existing and emerging grid services," *J. Lightwave Tech.*, vol. 23, pp. 3347–3357, Oct. 2005.

84. D. F. Grosz, A. Agarwal, S. Banerjee, D. N. Maywar, and A. P. Küng, "All-Raman ultra-long-haul single-waveband DWDM transmission system with OADM capability," *J. Lightwave Tech.*, vol. 22, pp. 423–432, Feb. 2004.

85. K. Iwatsuki, J.-I. Kani, H. Suzuki, and M. Fujiwara, "Access and metro networks based on WDM technologies," *J. Lightwave Tech.*, vol. 22, pp. 2623–2630, Nov. 2004.

86. N. Ghani, J.-Y. Pan, and X. Cheng, "Metropolitan optical networks," in *Optical Fiber Telecommunications IV-B: Systems and Impairments*, I. P. Kaminow and T. Li, eds., pp. 329–403, Academic, San Diego, 2002.

87. A. Zapata, M. Düser, J. Spencer, P. Bayvel, I. de Miguel, D. Breuer, N. Hanik, and A. Gladisch, "Next-generation 100-Gigabit Metro Ethernet (100 GbME) using multiwavelength optical rings," *J. Lightwave Tech.*, vol. 22, pp. 2420–2434, Nov. 2004.

88. P. Bedell, *Gigabit Ethernet for Metro Area Networks*, McGraw-Hill, New York, 2002.

89. S. Bottacchi, *Optical Fibre Transmission Theory, Technology, and Design, Multi-Gigabit Transmission over Multimode Optical Fibre: Theory and Design Methods for 10GbE Systems*, Wiley, Hoboken, NJ, 2006.

90. M. Abrams, P. C. Becker, Y. Fujimoto, V. O'Byrne, and D. Piehler, "FTTP deployments in the United States and Japan—Equipment choices and service provider imperatives," *J. Lightwave Tech.*, vol. 23, pp. 236–246, Jan. 2005.

91. G. Keiser, *FTTX Concepts and Applications*, Wiley, Hoboken, NJ, 2006.

92. P. E. Green Jr., *Fiber to the Home: The New Empowerment*, Wiley, Hoboken, NJ, 2006.

93. C. Lin, ed., *Broadband Optical Access Networks and Fiber-to-the-Home: Systems Technologies and Deployment Strategies*, Wiley, Hoboken, NJ, 2006.

94. N. Nadarajah, E. Wong, and A. Nirmalathas, "Packet labeling technique using electronic code-division multiple-access for WDM packet-based access networks," *IEEE Photonics Technol. Lett.*, vol. 18, pp. 607–609, Feb. 2006.

95. ITU-T Recommendation G.983.1, "Broadband optical access systems based on passive optical networks (PON)," Jan 2005.

96. ITU-T Recommendation G.984.1, "Gigabit-capable passive optical networks (GPON): General characteristics," Mar. 2008.

97. ITU-T Recommendation G.987.1, "10-Gigabit-capable passive optical networks (XG-PON): General requirements," Jan. 2010.

98. C. H. Yeh, C. W. Chow, C. H. Wang, F. Y. Shih, H. C. Chien, and S. Chi, "A self-protected colorless WDM-PON with 2.5 Gb/s upstream signal based on RSOA," *Opt. Express,* vol. 16, pp. 12296–12301, 2008.

99. G. Kramer, *Ethernet Passive Optical Networks*, McGraw-Hill, New York, 2005.

100. E. Desurvire, *Global Telecommunications: Signaling Principles, Network Protocols, and Wireless Systems*, Wiley, Hoboken, NJ, 2004.

101. S.-G. Mun, J.-H. Moon, H.-K. Lee, J.-Y. Kim, and C.-H. Lee, "A WDM-PON with a 40 Gb/s (32×1.25 Gb/s) capacity based on wavelength-locked Fabry-Perot laser diodes," *Opt. Express*, vol. 16, pp. 11361–11368, 2008.

102. M. Presi, R. Proietti, K. Prince, G. Contestabile, and E. Ciaramella, "A 80 km reach fully passive WDM-PON based on reflective ONUs," *Opt. Express*, vol. 16, pp. 19043–19048, 2008.

103. S.-C. Lin, S.-L. Lee, and C.-K. Liu, "Simple approach for bidirectional performance enhancement on WDM-PONs with direct modulation lasers and RSOAs," *Opt. Express*, vol. 16, pp. 3636–3643, 2008.

104. P. J. Urban, A.M.J. Koonen, G. D. Khoe, and H. de Waardt, "Interferometric crosstalk reduction in an RSOA-based WDM passive optical network," *J. Lightwave Tech.*, vol. 27, pp. 4943–4953, 2009.

105. C.-L. Tseng, C.-K. Liu, J.-J. Jou, W.-Y. Lin, C.-W. Shih, S.-C. Lin, S.-L. Lee, and G. Keiser, "Bidirectional transmission using tunable fiber lasers and injection-locked Fabry–Pérot laser diodes for WDM access networks," *Photonics Technol. Lett.*, vol. 20, pp. 794–796, May 2008.

106. G.-R. Lin, H.-L. Wang, G.-C. Lin, Y.-H. Huang, Y.-H. Lin, and T.-K. Cheng, "Comparison on injection-locked Fabry–Perot laser diode with front-facet reflectivity of 1% and 30% for optical data transmission in WDM-PON system," *J. Lightwave Tech.*, vol. 27, pp. 2779–2785, 2009.

107. R. Van Caenegem, J. M. Martínez, D. Colle, M. Pickavet, P. Demeester, F. Ramos, and J. Martí, "From IP over WDM to all-optical packet switching: Economical view," *J. Lightwave Tech.*, vol. 24, pp. 1638–1645, Apr. 2006.

108. M. Kodialam, V. Lakshman, J. B. Orlin, and S. Sengupta, "Preconfiguring IP-over-optical networks to handle router failures and unpredictable traffic," *IEEE J. Sel. Areas Commun.*, vol. 25, pp. 934–948, June 2007.

109. L. D. Ghein, *MPLS Fundamentals*, Cisco Press, Indianapolis, IN, 2007.

110. I. Minei and J. Lucek, *MPLS-Enabled Applications*, Wiley, Hoboken, NJ, 2008.

111. IETF RFC 3945, *Generalized Multi-Protocol Label Switching (GMPLS) Architecture*, Oct. 2004.

112. A. Farrel and I. Bryskin, *GMPLS: Architecture and Applications*, Morgan Kaufmann, 2006.

113. N. Sambo, I. Cerutti, A. Giorgetti, and P. Castoldi, "Impact of routing and wavelength selection strategies on GMPLS-controlled distributed restoration," *J. Opt. Network.*, vol. 7, pp. 388–399, 2008.

114. W. Zhong, Z. Lian, S. Kumar Bose, and Y. Wang, "Optical resilient Ethernet rings for high-speed MAN networks [Invited]," *J. Opt. Network.*, vol. 4, pp. 784–806, 2005.

115. H. Nakamura, H. Suzuki, J.-I. Kani, and K. Iwatsuki, "Reliable wide-area wavelength division multiplexing passive optical network accommodating Gigabit Ethernet and 10-Gb Ethernet services," *J. Lightwave Tech.*, vol. 24, pp. 2045–2051, Apr. 2006.

116. C. F. Lam and W. I. Way, "Optical Ethernet: Protocols, management, and 1-100 G technologies," in I. P. Kaminow, T. Li, and A. E. Willner, eds., *Optical Fiber Telecommunications–V*, vol. B, pp. 345–400, Academic, New York, 2008.

117. Y. Ishii and H. Yamashita, "Optical access transport system: GE-PON platform," *Fujitsu Sci Tech. J.*, vol. 45, pp. 346–354, Oct. 2009.

118. B.-H. Choi, M. Attygalle, Y. J. Wen, and S. D. Dods, "Dispersion map optimization and dispersion slope mismatch issue on 40 channel × 10 Gbit/s transmission over 3000 km using standard SMF and EDFA amplification," *Optics Comm.*, vol. 242, pp. 525–532, Dec. 2004.

119. L. Grüner-Nielsen, S. N. Knudsen, B. Edvold, T. Veng, D. Magnussen, C. C. Larsen, and H. Damsgaard, "Dispersion-compensating fibers," *Opt. Fiber Technol.*, vol. 6, no. 2, pp. 164–180, Apr. 2000.

120. M. Suzuki and N. Edagawa, "Dispersion-managed high-capacity ultra-long-haul transmission," *J. Lightwave Tech.*, vol. 21, pp. 916–929, Apr. 2003.

121. L. Grüner-Nielsen, M. Wandel, P. Kristensen, C. Jørgensen, L. V. Jørgensen, B. Edvold, B. Pálsdóttir, and D. Jakobsen, "Dispersion-compensating fibers," *J. Lightwave Tech.*, vol. 23, pp. 3566–3579, Nov. 2005.

122. H. Kogelnik, R. M. Jopson, and L. E. Nelson, "Polarization mode dispersion," in *Optical Fiber Telecommunications IV-B*, I. P. Kaminow and T. Li, eds., Academic, San Diego, CA, 2002.

123. P. M. Nellen, R. Brönnimann, M. Held, and U. Sennhauser, "Long-term monitoring of polarization-mode dispersion of aerial optical

cables with respect to line availability," *J. Lightwave Tech.*, vol. 22, pp. 1848–1855, Aug. 2004.

124. D. A. Nolan, X. Chen, and M. J. Li, "Fibers with low polarization-mode dispersion," *J. Lightwave Tech.*, vol. 22, pp. 1066–1077, Apr. 2004.

125. L. Yan, X. S. Yao, M. C. Hauer, and A. E. Willner, "Practical solutions to polarization-mode-dispersion emulation and compensation," *J. Lightwave Technol.*, vol. 24, pp. 3992–4005, Nov. 2006.

126. F. Buchali and H. Bülow, "Adaptive PMD compensation by electrical and optical techniques," *J. Lightwave Technol.*, vol. 22, pp. 1116–1126, Apr. 2004.

127. H. Bülow, F. Buchali, and A. Klekamp, "Electronic dispersion compensation," *J. Lightwave Tech.*, vol. 26, pp. 158–167, Jan. 2008.

128. M. E. McCarthy, J. Zhao, A. D. Ellis, and P. Gunning, "Full-field electronic dispersion compensation of a 10 Gbit/s OOK signal over 4 124 km field-installed single-mode fibre," *J. Lightwave Tech.*, vol. 27, pp. 5327–5334, Dec. 2009.

129. M. S. Alfiad, D. van den Borne, S. L. Jansen, T. Wuth, M. Kuschnerov, G. Grosso, A. Napoli, and H. de Waardt, "A comparison of electrical and optical dispersion compensation for 111-Gb/s POLMUX–RZ–DQPSK," *J. Lightwave Tech.*, vol. 27, no. 16, pp. 3590–3598, Aug. 2009.

130. B.J.C. Schmidt, A. J. Lowery, and J. Armstrong, "Experimental demonstrations of electronic dispersion compensation for long-haul transmission using direct-detection optical OFDM," *J. Lightwave Tech.*, vol. 26, no. 1, pp. 196–203, Jan. 2008.

131. A. C. Singer, N. R. Shanbhag, and H.-M. Bae, "Electronic dispersion compensation," *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 110–130, Nov. 2008.

132. K. Motoshima, N. Suzuki, K. Shimizu, K. Kasahara, T. Kitayama, and T. Yasui, "A channel-number insensitive erbium-doped fiber amplifier with automatic gain and power regulation function," *J. Lightwave Tech.*, vol. 19, pp. 1759–1766, Nov. 2001.

133. Y. Ben-Ezra, M. Haridim, and B. I. Lembrikov, "All-optical AGC of EDFA based on SOA," *IEEE J. Quantum Electron.*, vol. 42, pp. 1209–1214, Dec. 2006.

134. X. Zhou, M. Feuer, and M. Birk, "A simple feed-forward control algorithm for fast dynamic gain profile control in a multiwavelength forward-pumped Raman fiber amplifier," *IEEE Photonics Technol. Lett.*, vol. 18, pp. 1004–1006, May 2006.

135. S. H. Chang, H. S. Chung, H. J. Lee, and K. Kim, "Suppression of transient phenomena in hybrid Raman/EDF amplifier," *IEEE Photonics Technol. Lett.,* vol. 17, pp. 1004–1006, May 2005.

CHAPTER 14

Performance Measurement and Monitoring

Engineers from many diverse disciplines need to make performance measurements at all stages of the design, installation, and operation of an optical fiber communication network. Numerous levels of measurement techniques have been developed for characterizing the operational behavior of devices and fibers, for ensuring that the correct components have been selected for a particular application, and for verifying that the network is configured properly.[1–6] In addition, various operational performance monitoring methods are needed to verify that all the design and operating specifications of a link are met when it is running. A wide selection of sophisticated test equipment exists for each of these measurement categories.

During the link design phase an engineer can find the operational parameters of many passive and active components on vendor data sheets. These include fixed parameters for fibers, passive optical devices, and optoelectronic components such as light sources, photodetectors, and optical amplifiers. For example, fixed fiber parameters include core and cladding diameters, refractive-index profile, mode-field diameter, and cut-off wavelength. Once such fixed parameters are known, generally there is no need for a link design engineer to measure them again.

However, variable parameters of communication system elements, such as optoelectronic components, may change with operational conditions and need to be measured before, during, and after a link is fielded. Of particular importance are accurate and comprehensive measurements of the optical fiber, since this component cannot be replaced readily once it has been installed. Although many physical properties of fiber remain constant, the attenuation and dispersions of a fiber can change during fiber cabling and cable installation. In single-mode fibers, chromatic and polarization-mode dispersions are important factors that can limit the transmission distance or data rate. Chromatic dispersion effects are of particular importance in high-speed DWDM links, since their behavior depends on the link configuration. Measurement and monitoring of polarization-mode dispersion is important for data rates at and above 10 Gb/s, since its statistical behavior ultimately can limit the highest achievable data rate.

When a link is being installed and tested, the operational parameters of interest include bit-error rate, timing jitter, and signal-to-noise ratio as indicated by the eye pattern. During actual operation, measurements are needed for maintenance and monitoring functions to determine factors such as fault locations in fibers and the status of remotely located optical amplifiers and other active devices.
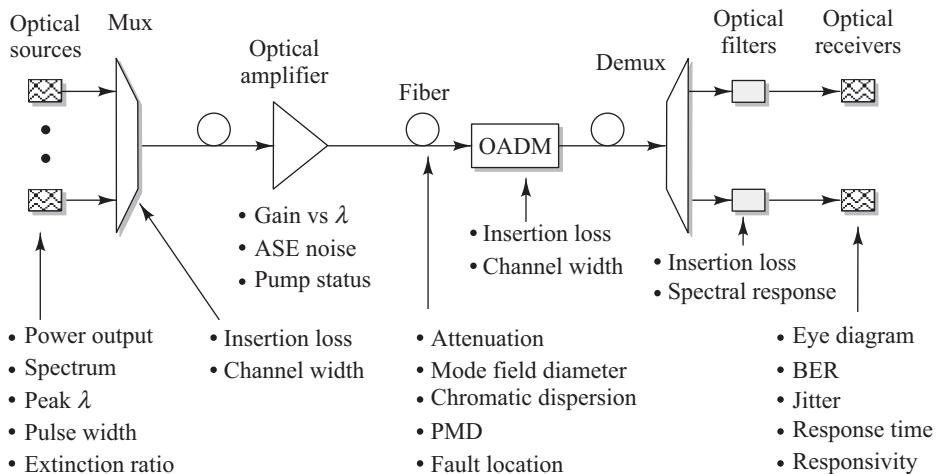
Fig. 14.1 *Components of a typical WDM link and some performance-measurement parameters of user interest*

This chapter discusses measurements and performance monitoring tests of interest to designers, installers, and operators of fiber optic links and networks. Of particular interest here are measurements for WDM links. Figure 14.1 shows some of the relevant test parameters and at what points in a WDM link they are of importance. The operational impact or impairment for many of these factors can be accounted for and controlled through careful network design. Other parameters may need to be monitored and possibly compensated for dynamically during network operation. In either case, all of these parameters must be measured at some point during the time period ranging from network design concept to service provisioning.

First, Sec. 14.1 addresses internationally recognized measurement standards for component and system evaluations. Next, Sec. 14.2 lists basic test instruments for optical fiber communication link characterizations. A fundamental unit in lightwave communications is optical power and its measurement with optical power meters, which is the topic of Sec. 14.3. Turning to measurement techniques, Sec. 14.4 gives an overview of methods and specialized equipment for characterizing optical fiber parameters. In addition to determining geometric parameters, this equipment also can measure attenuation and chromatic dispersion. During and after link installation, several basic parameters need to be checked. For example, the error performance can be estimated through eye pattern measurements, which is the topic of Sec. 14.5. The physical integrity of the fielded link typically is checked with an optical time-domain reflectometer, as Sec. 14.6 explains. Section 14.7 discusses optical performance monitoring, which is essential for managing high-capacity lightwave transmission networks. Network functions that need such monitoring include amplifier control, channel identification, and assessment of the integrity of optical signals. Section 14.8 describes some basic performance measurements procedures.

14.1 Measurement Standards

Before examining measurement techniques, let us first look at what standards exist for fiber optics. As summarized in Table 14.1, the three basic classes are primary, component testing, and system standards.

Table 14.1 *Three standards classes, the involved organizations, and their functions*

Standards class	Involved organizations	Functions of the organizations
Primary	• NIST (USA) • NPL (UK) • PTB (Germany)	• Characterize physical parameters • Support and accelerate development of emerging technologies (NIST)
Component testing	• TIA/EIA • ITU-T • IEC	• Define component evaluation tests • Establish equipment-calibration procedures
System testing	• ANSI • IEEE • ITU-T	• Define physical-layer test methods • Establish measurement procedures for links and networks

Primary standards refer to measuring and characterizing fundamental physical parameters such as attenuation, bandwidth, mode-field diameter for single-mode fibers, and optical power. In the United States the main group involved in primary standards is the *National Institute of Standards and Technology* (NIST).[6,7] This organization carries out fiber optic and laser standardization work, and it sponsors a biennial conference on optical fiber measurements. Another goal is to support and accelerate the development of emerging technologies. Other national organizations include the *National Physical Laboratory* (NPL) in the United Kingdom[8] and the *Physikalisch-Technische Bundesanstalt* (PTB) in Germany.[9]

As Table 14.2 summarizes, several international organizations are involved in formulating component and system testing standards. The major organizations that deal with measurement methods for links and networks are the *Institute for Electrical and Electronic Engineers* (IEEE) and the *Telecommunication Standardization Sector of the International Telecommunication Union* (ITU-T). *Telcordia Technologies* provides a wide range of generic requirements for telecommunications network components, installations, and equipment.

Component testing standards define relevant tests for fiber optic component performance, and they establish equipment-calibration procedures. A key organization for component testing is the *Telecommunication Industry Association* (TIA) in association with the *Electronic Industries Alliance* (EIA). The TIA has a list of over 200 fiber optic test standards and specifications under the general designation TIA/EIA-455-XX-YY, where XX refers to a specific measurement technique and YY refers to the publication year. These standards are also called *Fiber Optic Test Procedures* (FOTP), so that TIA/EIA-455-XX becomes FOTP-XX. These include a wide selection of recommended methods for testing the response of fibers, cables, passive devices, and electro-optic components to environmental factors and operational conditions. For example, TIA/EIA-455-20-B-2004, or FOTP-20, is a method published in 2004 that gives two methods for monitoring the changes in transmittance of optical fibers or cables that may occur during mechanical or environmental testing.

System standards refer to measurement methods for links and networks. The major organizations involved here are the *American National Standards Institute* (ANSI), the *Institute for Electrical and Electronic Engineers* (IEEE), and the ITU-T. Of special interest for fiber optics systems are test standards and recommendations from the ITU-T that are aimed at all aspects of optical networking. Within the TIA, the FO-2 Committee develops physical-layer test procedures, system design guides, and system specifications to assist both suppliers and users of fiber optic communications technology. Interoperability and compatibility between different vendor equipment are the important concerns. The committee

Table 14.2 *Summary of major standards organizations and their functions related to testing*

Organization	Internet address	Testing-related activities
IEEE	www.ieee.org	Establish and publish measurement procedures for links and networks
		• Define physical-layer test methods
		• IEEE 802.3ah Ethernet in First Mile (EFM)
ITU-T	www.itu.int/ITU-T	Create and publish standards in all areas of telecommunications
		• Series *G* for telecommunication transmission systems and media, digital systems, and networks
		• Series *L* for construction, installation, and protection of cables and other outside plant elements
		• Series *O* for specifications of measuring equipment
Telcordia	www.telcordia.com	Provide generic requirements for network elements
		• Fiber optic connectors
		• Indoor and outdoor cabinets
		• Underground, aboveground enclosures
		• Field-deployed products
TIA/EIA	www.tiaonline.org www.eia.org	Created over 200 test specifications under the designation Fiber Optic Test Procedures (FOTP)
		• Define physical-layer test methods
		• TIA/EIA-455-XX or FOTP-XX documents

addresses the performance and reliability of active components and systems, such as transmitters, receivers, amplifiers, and modulators. Systems include single-mode digital and analog systems, optically amplified systems with dense wavelength division multiplexing (DWDM), point-to-point multimode systems, and local-area network (LAN) applications.

14.2 | Basic Test Equipment

As optical signals pass through the various parts of an optical link, they need to be measured and characterized in terms of the three fundamental areas of optical power, polarization, and spectral content. The basic instruments for carrying out such measurements on optical fiber components and systems include optical power meters, attenuators, tunable laser sources, spectrum analyzers, and optical time-domain reflectometers. These come in a variety of capabilities, with sizes ranging from portable, handheld units for field use to sophisticated briefcase-sized bench-top or rack-mountable instruments for laboratory and manufacturing applications. In general, the field units do not need to have the extremely high precision of laboratory instruments, but they need to be more rugged to maintain reliable and accurate measurements under extreme environmental conditions of temperature, humidity, dust, and mechanical stress. However,

Table 14.3 *Some widely used optical system test instruments and their functions*

Test instrument	Function
Test-support lasers (multiple-wavelength or broadband)	Assist in tests that measure the wavelength-dependent response of an optical component or link
Optical spectrum analyzer	Measures optical power as a function of wavelength
Multifunction optical test system	Factory or field instruments with exchangeable modules for performing a variety of measurements
Optical power attenuator	Reduces power level to prevent instrument damage or to avoid overload distortion in the measurements
Conformance analyzer	Measures optical receiver performance in accordance with standards-based specifications
Visual fault indicator	Uses visible light to give a quick indication of a break in an optical fiber
Optical power meter	Measures optical power over a selected wavelength band
BER test equipment	Uses standard eye-pattern masks to evaluate the data-handling ability of an optical link
OTDR (field instrument)	Measures attenuation, length, connector/splice losses, and reflectance levels; helps locate fiber breaks
Optical return loss tester	Measures total reverse power in relation to total forward power at a particular point

even the handheld instruments for field use have reached a high degree of sophistication with automated microprocessor-controlled test features and computer-interface capabilities.

More sophisticated instruments, such as polarization analyzers and optical communication analyzers, are available for measuring and analyzing polarization-mode dispersion (PMD), eye diagrams and pulse waveforms. These instruments enable a variety of statistical measurements to be made at the push of a button, after the user has keyed in the parameters to be tested and the desired measurement range.

Table 14.3 lists some essential test equipment and their functions for installation and operation of optical communication systems. This section defines a selection of the first six instruments in this table. Later sections give more details concerning optical power meters, bit-error rate testers, and optical time-domain reflectometers.

14.2.1 Test-Support Lasers

Specialized light sources are desirable for testing optical components. Table 14.4 lists the characteristics of two such laser-source instruments used for test support.

Tunable laser sources are important instruments for measurements of the wavelength-dependent response of an optical component or link. A number of vendors offer such light sources that generate a true single-mode laser line for every selected wavelength point. Typically the source is an external-cavity semiconductor laser. A movable diffraction grating may be used as a tunable filter for wavelength selection. Depending on the source and grating combination, an instrument may be tunable over (for example) the 1280-to-1330-nm, the

Table 14.4   *Characteristics of laser source instruments used for test support*

Parameter	Tunable source	Broadband source
Spectral output range	Selectable: e.g., 1370–1495 nm or 1460–1640 nm	Peak wavelength ±25 nm
Total optical output power	Up to 8 dBm	> 3.5 mW(5.5 dBm) over a 50-nm range
Power stability	< ±0.02 dB	< ±0.05 dB
Wavelength accuracy	< ±10 pm	(Not applicable)

1370-to-1495-nm, or the 1460-to-1640-nm band. Wavelength scans, with an output power that is flat across the scanned spectral band, can be done automatically. The minimum output power of such an instrument usually is –10 dBm, and the absolute wavelength accuracy is typically ±0.01 nm (±10 pm).

A *broadband incoherent light source* with a high output power coupled into a single-mode fiber is desirable to evaluate passive DWDM components. Such an instrument can be realized by using the amplified spontaneous emission (ASE) of an erbium-doped fiber amplifier. The power spectral density of the output is up to one hundred times (20 dB) greater than that of edge-emitting LEDs and up to 100,000 times (50 dB) greater than white-light tungsten lamp sources. The instrument can be specified to have a total output power of greater than 3.5 mW (5.5 dBm) over a 50-nm range with a spectral density of –13 dBm/nm (50 μW/nm). The relatively high-power spectral density allows test personnel to characterize devices with medium or high insertion loss. Peak wavelengths might be 1200, 1310, 1430, 1550, or 1650 nm.

14.2.2   Optical Spectrum Analyzer

The widespread implementation of WDM systems calls for making optical spectrum analyses to characterize the behavior of various telecommunication network elements as a function of wavelength. One widely used instrument for doing this is an *optical spectrum analyzer* (OSA), which measures optical power as a function of wavelength. The most common implementation uses a diffraction-grating-based optical filter, which yields wavelength resolutions to less than 0.1 nm. Higher wavelength accuracy (±0.001 nm) is achieved with wavelength meters based on Michelson interferometry.

Figure 14.2 illustrates the operation of a grating-based optical spectrum analyzer. Light emerging from a fiber is collimated by a lens and is directed onto a diffraction grating that can be rotated. The exit slit selects or filters the spectrum of the light from the grating. Thus it determines the *spectral resolution* of the OSA. The term *resolution bandwidth* describes the width of this optical filter. Typical OSAs have selectable filters ranging from 10 nm down to 0.1 nm. The optical filter characteristics determine the *dynamic range,* which is the ability of the OSA to simultaneously view large and small signals in the same sweep. The bandwidth of the amplifier is a major factor affecting the sensitivity and sweep time of the OSA. In the O-band through L-band the photodiode is usually an InGaAs device.
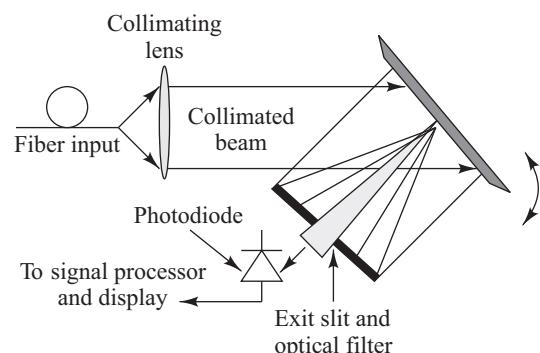
Fig. 14.2   *Operation of a grating-based optical spectrum analyzer*

Fig. 14.3 *Compact, portable multifunction test instrument for use in field environments. (Model FOT-930, photo provided courtesy of EXFO; www.exfo.com. © 2004–2010 EXFO Inc.)*

The OSA normally sweeps (scans in a certain time interval) across a spectral band making measurements at discretely spaced wavelength points. This spacing depends on the bandwidth-resolution of the instrument and is known as the *trace-point spacing.*

14.2.3 Multiple-Function Testers

To reduce the number of individual pieces of test gear, manufacturers are producing multifunction test equipment. Figure 14.3 shows one example of a handheld instrument that has been engineered for FTTP network use. This single portable instrument has the following functions: a power meter, a bidirectional dual-wavelength loss tester, an optical return-loss tester, a visual fault indicator for locating breaks and failures in a fiber cable, and a talk set for full-duplex communications between field personnel. To carry out these functions, the unit contains optical sources which emit at the 1310-, 1490-, and 1550-nm wavelengths used in FTTP networks.

14.2.4 Optical Power Attenuators

In many laboratory or production tests, the characteristics of a high optical signal level may need to be measured. For example, if the level is a strong output from an optical amplifier, the signal may need to be attenuated precisely before being measured. This is done to prevent instrument damage or to avoid overload distortion in the measurements. An *optical attenuator* allows a test engineer to reduce an optical signal level up to, for example, 60 dB (a factor of 10^6) in precise steps at a specified wavelength, which is usually 1310 or 1550 nm. At the lower performance end, these attenuators are small devices (approximately $2 \times 5 \times 10$ cm) intended for quick field measurements that may only need to be accurate to 0.5 dB. Laboratory instruments may have an attenuation precision of 0.001 dB.

14.2.5 Optical Transport Network (OTN) Tester

A major concern in the implementation of an optical transport network (OTN) is how to verify that the various network elements, such as SONET/SDH and Ethernet circuit cards, are functioning properly. In relation to this, the ITU-T has published the following guidelines:
- G.709, Interfaces for the OTN
- G.798, Characteristics of OTN Hierarchy Functional Blocks

Of particular importance is verification that the OTN elements that may have been designed, produced, and installed by different vendors conform to these ITU-T recommendations. Such tests, which are known as *conformance tests*, include the following:
- Verifying the correctness of the interface specifications of the elements
- Checking the correct responses of the device under test (DUT)
- Verifying the correct behavior of forward-error-correcting (FEC) modules
- Examining that mapping and demapping of client signals are performed correctly

Figure 14.4 shows an example of a multifunction, multiport, multiuser portable instrument for field conformance testing. Such an instrument can characterize physical-layer jitter, wander, and bit-error rates up to 43 Gb/s in long-haul, metro, and FTTP access networks.

14.2.6 Visual Fault Indicator

A *visual fault locator* (VFL) is a handheld pen-sized instrument that uses a visible laser light source to locate events such as fiber breaks, overly tight bends in a fiber, or poorly mated connectors. The source emits a bright beam of red light (e.g., 650 nm) into a fiber, thereby allowing the user to see a fiber fault or a high-loss point as a glowing or blinking red light. The VFL is particularly useful for identifying fiber faults within the initial dead zone of an OTDR. In using such a device, the events must occur where the fiber or connector is in the open so that visual observation of the emitted red light is possible.

Fig. 14.4 *A portable multifunction optical network tester. (Photo courtesy of JDSU: Model ONT-503; www.jdsu.com.)*

The nominal light output is 1 mW, so the light will be visible through a fiber jacket at a fault point. This power level allows a user to detect a fiber fault visually, for up to 5 km. The device generally is powered by one 1.5-V size AA battery and operates in either a continuously on or a blinking mode.

14.3 Optical Power Measurements

Optical power measurement is the most basic function in fiber optic metrology. However, this parameter is not a fixed quantity and can vary as a function of other parameters such as time, distance along a link, wavelength, phase, and polarization.

14.3.1 Definition of Optical Power

To get an understanding optical power, let us look at its physical basis and how it relates to other optical quantities such as energy, intensity, and radiance.

- Light particles, called *photons*, have a certain energy associated with them, which changes with wavelength. The relationship between the energy E of a photon and its wavelength λ is given by the equation $E = hc/\lambda$, which is known as *Planck's Law.* In terms of wavelength (measured in units of μm), the energy in electron volts is given by the expression $E(\text{eV}) = 1.2406/\lambda(\mu\text{m})$. Note that $1\ e\text{V} = 1.60218 \times 10^{-19}$ J.
- *Optical power P* measures the rate at which photons arrive at a detector. Thus it is a measure of energy transfer per time. Since the rate of energy transfer varies with time, the optical power is a function of time. It is measured in *watts* or joules per second (J/s).
- As noted in Chapter 4, *radiance* (or *brightness*) is a measure, in watts, of how much optical power radiates into a designated solid angle per unit of emitting surface.

Since optical power can vary with time, its measurement also changes with time. Figure 14.5 shows a plot of the power level in a signal pulse stream as a function of time. Different instantaneous power-level readings are obtained depending on the exact instant when the measurement is made. Therefore, two standard classes of power measurements can be specified in an optical system. These are the peak
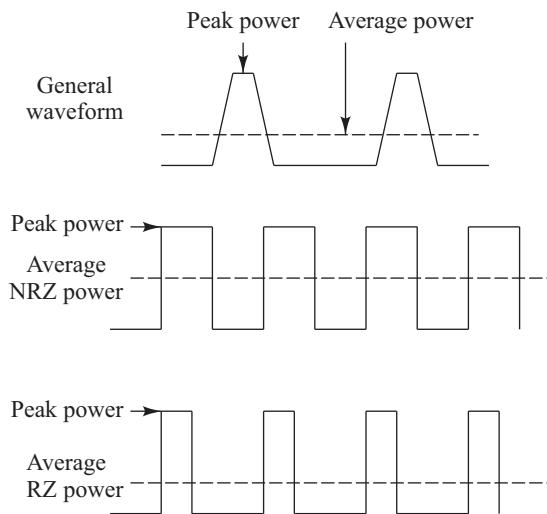
Fig. 14.5 *Peak and average powers in a series of general, NRZ, and RZ optical pulses*

power and the average power. The *peak power* is the maximum power level in a pulse, which might be sustained for only a very short time.

The *average power* is a measure of the power level averaged over a relatively long time period compared to the duration of an individual pulse. For example, the measurement time period could be one second, which contains many signal pulses. As a simple example, in a nonreturn-to-zero (NRZ) data stream there will be an equal probability of 1 and 0 pulses over a long time period. In this case, as shown in Fig. 14.5, the average power is half the peak power. If a return-to-zero (RZ) modulation format is used, the average power over a long sequence of pulses will be one-fourth the peak power because there is no pulse in a 0 time slot and a 1 time slot is only half filled.

The sensitivity of a photodetector normally is expressed in terms of the average power level impinging on it, since the measurements in an actual fiber optic system are done over many pulses. However, the output level for an optical transmitter normally is specified as the peak power. This means that the average power coupled into a fiber, which is the power level that a photodetector measures, is at least 3 dB lower than if the link designer incorrectly uses the peak source output in power-budget calculations as the light level entering the fiber.

14.3.2 Optical Power Meters

The function of an *optical power meter* is to measure total power over a selected wavelength band. Some form of optical power detection is in almost every piece of lightwave test equipment. Handheld instruments come in a wide variety of types with different levels of capabilities. Multiwavelength optical power meters using several photodetectors are the most common instrument for measuring optical signal power levels. Usually the meter outputs are given in dBm (where 0 dBm = 1 mW) or dBμ (where 0 dBμ = l μW).

For example, using a Ge photodetector typically allows a measuring range of +18 to – 60 dBm in the 780-to-1600-nm wavelength band, whereas an InGaAs photodetector allows a measuring range of +3 to –73 dBm in the 840-to-1650-nm wavelength band. In each case, the power measurements can be made at a number of calibrated wavelengths. User-selectable threshold settings can let the instrument show a pass/fail on a built-in display. An RS-232 or USB interface together with application software allows a user to download measurements and view, export, or print them in either tabular or graphic form.

14.4 Optical Fiber Characterization

Many millions of kilometers of optical fibers have been fabricated and installed worldwide. Various types of equipment for factory use have been developed to characterize the physical and performance parameters of these fibers. Whereas early equipment tended to specialize on measuring only one or two parameters, modern sophisticated instruments require only one simple fiber preparation to characterize optical fibers

accurately during the manufacturing process. These parameters include mode-field diameter, attenuation, cut-off wavelength, chromatic dispersion versus wavelength, refractive-index profile, effective area, and geometric properties such as core and cladding diameters, core-to-cladding concentricity error, and fiber noncircularity. The two basic measurement methods used by this specialized equipment are the refracted near-field and the transmitted near-field techniques. This section first describes these two techniques and then looks at some standard ways to measure attenuation and three dispersion factors.

14.4.1 Refracted Near-Field Technique

The *refracted near-field* measurement method is recommended by the ITU-T and TIA for determining the refractive-index profile (RIP).[10] This method determines the index profile by moving a focused laser across the fiber end face and examining the distribution of the light that is refracted sideways out of the core as a function of the radial position of the laser spot. The variation in the detected optical signal level is proportional to the index change at the fiber end face. The RIP parameter can be used to calculate the geometrical parameters of a fiber and to estimate all transmission properties (e.g., chromatic dispersion and the cut-off wavelength) except attenuation and polarization-mode dispersion.[11]

14.4.2 Transmitted Near-Field Technique

The *transmitted near-field* measurement method is recommended by the ITU-T and TIA for measuring mode-field characteristics.[12,13] Knowledge of the mode-field diameter (MFD) is important because it describes the radial optical field distribution across the fiber core. Detailed information of the MFD enables one to calculate characteristics such as source-to-fiber coupling efficiency, splice and joint losses, microbending loss, and dispersion. A transmitted near-field scan directly provides the intensity distribution $E^2(r)$ at the fiber exit. From this distribution one then can calculate the MFD using the Petermann II equation.[14] The Petermann II expression is given by Eq. (2.73) in terms of the field intensity distribution as[11,15]

$$\text{MFD} = 2\sqrt{2} \left[\frac{2\int_0^\infty E^2(r)\, r^3 \, dr}{\int_0^\infty E^2(r)\, r \, dr} \right]^{1/2} \tag{14.1}$$

Since it is easy to program this equation, the measurement equipment software can calculate the MFP directly from the near-field data.

14.4.3 Attenuation Measurements

Attenuation of optical power in a fiber waveguide is a result of absorption processes, scattering mechanisms, and waveguide effects. The manufacturer is generally interested in the magnitude of the individual contributions to attenuation, whereas the system engineer who uses the fiber is more concerned with the total transmission loss of the fiber. Here we treat only measurement techniques for total transmission loss.

Three basic methods are available for determining attenuation in fibers. The earliest devised and most common approach involves measuring the optical power transmitted through a long and a short length of the same fiber using identical input couplings. This method is known as the cutback technique. A less accurate but nondestructive method is the insertion-loss method, which is useful for cables with connectors on them. These two methods are described in this section. Section 14.6 describes the third technique, which involves the use of an OTDR.
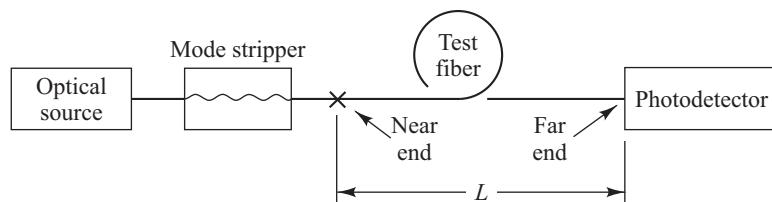
Fig. 14.6   *Schematic experimental setup for determining fiber attenuation by the cutback technique. The optical power is first measured at the far end, then the fiber is cut at the near end, and the power output there is measured*

The Cutback Technique

The *cutback technique*,[16,17] which is a destructive method requiring access to both ends of the fiber, is illustrated in Fig. 14.6. Measurements may be made at one or more specific wavelengths, or, alternatively, a spectral response may be required over a range of wavelengths. To find the transmission loss, the optical power is first measured at the output (or far end) of the fiber. Then, without disturbing the input condition, the fiber is cut off a few meters from the source, and the output power at this near end is measured. If P_F and P_N represent the output powers of the far and near ends of the fiber, respectively, the average loss α in decibels per kilometer is given by

$$\alpha = \frac{10}{L} \log \frac{P_N}{P_F} \tag{14.2}$$

where L (in kilometers) is the separation of the two measurement points. The reason for following these steps is that it is extremely difficult to calculate the exact amount of optical power launched into a fiber. By using the cutback method, the optical power emerging from the short fiber length is the same as the input power to the fiber of length L.

In carrying out this measurement technique, special attention must be paid to how optical power is launched into the fiber because in a multimode fiber different launch conditions can yield different loss

Example 14.1   An engineer wants to find the attenuation at 1310 nm of a 4.95-km long fiber. The only available instrument is a photodetector, which gives an output reading in volts. Using this device in a cutback-attenuation setup, the engineer measures an output of 6.58 V from the photodiode at the far end of the fiber. After cutting the fiber 2 m from the source, the output voltage from the photodetector now reads 2.21V. What is the attenuation of the fiber in dB/km?

Solution: Since the output voltage from the photodetector is proportional to the optical power, we can write Eq. (14.2) as

$$\alpha = \frac{10}{L_1 - L_2} \log \frac{V_2}{V_1}$$

where L_1 is the length of the original fiber, L_2 is the length after cut off, and V_1 and V_2 are the voltage output readings from the long and short lengths, respectively. Then the attenuation in decibels is

$$\alpha = \frac{10}{4.950 - 2} \log \frac{6.58}{2.21} = 0.95 \text{ dB/km}$$
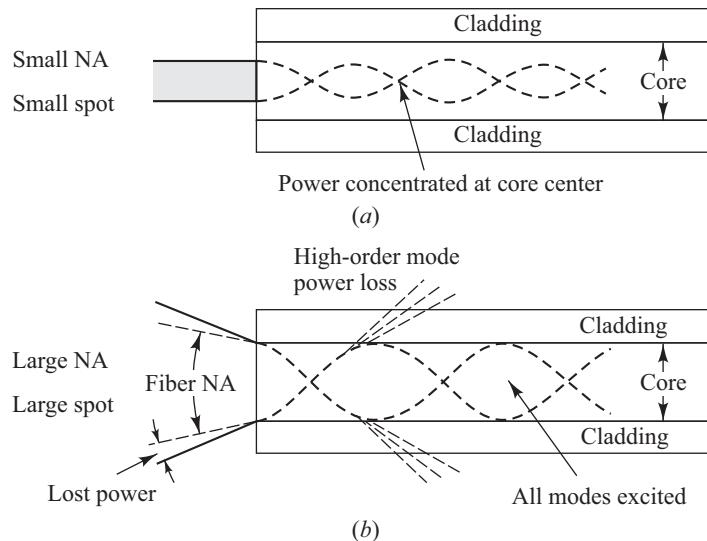
Fig. 14.7 *The effects of launch numerical aperture and spot size on the modal distribution. (a) Underfilling the fiber excites only lower-order modes; (b) an overfilled fiber has excess attenuation from higher-order mode loss*

values. The effects of modal distributions in the multimode fiber that result from different numerical apertures and spot sizes on the launch end of the fiber are shown in Fig. 14.7. If the spot size is small and its NA is less than that of the fiber core, the optical power is concentrated in the center of the core, as Fig. 14.7a shows. In this case, the attenuation contribution arising from higher-order-mode power loss is negligible. In Fig. 14.7b the spot size is larger than the fiber core and the spot NA is larger than that of the fiber. For this overfilled condition, those parts of the incident light beam that fall outside the fiber core and outside the fiber NA are lost. In addition, there is a large contribution to the attenuation arising from higher-mode power loss (see Secs. 5.1 and 5.3).

Steady-state equilibrium-mode distributions are typically achieved by the *mandrel-wrap* method. In this procedure, excess higher-order cladding modes that are launched by initially overexciting the fiber are filtered out by wrapping several turns of fiber around a mandrel, which is about 1.0−1.5 cm in diameter. In single-mode fibers, this type of mode filter is used to eliminate cladding modes from the fiber.

Insertion-Loss Method

For cables with connectors, one cannot use the cutback method. In this case, one commonly uses an *insertion-loss technique*.[16] This is less accurate than the cutback method but is intended for field measurements to give the total attenuation of a cable assembly in decibels.

The basic setup is shown in Fig. 14.8, where the launch and detector couplings are made through connectors. The wavelength-tunable light source is coupled to a short length of fiber that has the same basic characteristics as the fiber to be tested. For multimode fibers, a mode scrambler is used to ensure that the fiber core contains an equilibrium-mode distribution. In single-mode fibers, a cladding-mode stripper is employed so that only the fundamental mode is allowed to propagate along the fiber. A wavelength-selective device, such as an optical filter, is generally included to find the attenuation as a function of wavelength.

(*a*) Reference measurement

(*b*) Cable attenuation measurement
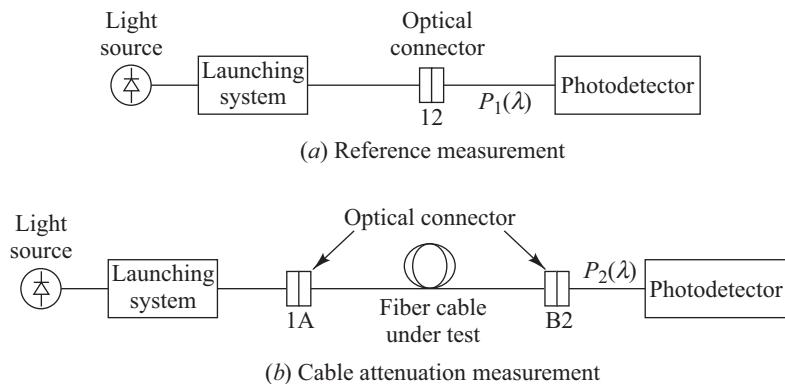
Fig. 14.8 *Test setup for using the insertion-loss technique to measure attenuation of cables that have attached connectors*

To carry out the attenuation tests, the connector of the short-length launching fiber is attached to the connector of the receiving system and the launch-power level $P_1(\lambda)$ is recorded. Next, the cable assembly to be tested is connected between the launching and receiving systems, and the received-power level $P_2(\lambda)$ is recorded. The attenuation of the cable in decibels is then

$$A = 10 \log \frac{P_1(\lambda)}{P_2(\lambda)} \tag{14.3}$$

This attenuation is the sum of the loss of the cabled fiber and the connector between the launch connector and the cable.

Example 14.2 The insertion-loss technique also can be used for measuring the loss through an optical device that has fiber flyleads attached. Suppose an optical filter with attached flyleads is inserted into the link in Fig. 14.8*a*. Consider the case when the power at the photodetector prior to inserting the filter is $P_1 = 0.51$ mW and the power level with the optical filter in the link is $P_2 = 0.43$ mW. What is the insertion loss of the device?

Solution: From Eq. (14.3) we have

Insertion loss = 10 log P_1/P_2 = 10 log 0.51/0.43
= 0.74 dB

14.4.4 Dispersion Measurements

Three basic forms of dispersion produce pulse broadening of lightwave signals in optical fibers, thereby limiting the information-carrying capacity. In multimode fibers, intermodal dispersion arises from the fact that each mode in an optical pulse travels a slightly different distance and thus arrives at the fiber end at slightly offset times. Chromatic dispersion stems from the variation in the propagation speed of the individual wavelength components of an optical signal. Polarization-mode dispersion arises from the splitting of a polarized signal into orthogonal polarization modes, each of which has a different propagation speed.

There are many ways to measure the various dispersion effects. Here, we look at some common methods.

Intermodal Dispersion

For practical purposes in evaluating intermodal dispersion, the fiber can be considered as a filter characterized by an impulse response $h(t)$ or by a power transfer function $H(f)$, which is the Fourier transform of the impulse response.[18] Either of these can be measured to determine the pulse dispersion. The impulse-response measurements are made in the time domain, whereas the power transfer function is measured in the frequency domain.

Both the time-domain and frequency-domain dispersion measurements assume that the fiber behaves quasi-linearly in power; that is, the individual overlapping output pulses from an optical waveguide can be treated as adding linearly. The behavior of such a system in the time domain is described simply as

$$p_{out}(t) = h(t) * p_{in}(t) = \int_{-T/2}^{T/2} p_{in}(t - \tau) h(\tau) \, d\tau \tag{14.4}$$

That is, the output pulse response $p_{out}(t)$ of the fiber can be calculated through the convolution (denoted by *) of the input pulse $p_{in}(t)$ and the power impulse function $h(t)$ of the fiber. The period T between the input pulses should be taken to be wider than the expected time spread of the output pulses.

In the frequency domain, Eq. (14.4) can be expressed as the product

$$P_{out}(f) = H(f) P_{in}(f) \tag{14.5}$$

Here, $H(f)$, the power transfer function of the fiber at the baseband frequency f, is the Fourier transform of $h(t)$:

$$H(f) = \int_{-\infty}^{\infty} h(t) e^{-j2\pi ft} \, dt \tag{14.6}$$

and $P_{out}(f)$ and $P_{in}(f)$ are the Fourier transforms of the output and input pulse responses $p_{out}(t)$ and $p_{in}(t)$, respectively:

$$P(f) = \int_{-\infty}^{\infty} p(t) e^{-j2\pi ft} \, dt \tag{14.7}$$

The transfer function of a fiber optic cable contains the bandwidth information of the system. For pulse dispersion to be negligible in digital systems, one of the following approximately equivalent conditions should be satisfied: (1) the fiber transfer function must not roll off to less than 0.5 of its low-frequency value for frequencies up to half the desired bit rate, or (2) the rms width of the fiber impulse response must be less than one-quarter of the pulse spacing.

Time-Domain Intermodal Dispersion Measurements

The simplest approach for making pulse-dispersion measurements in the time domain is to inject a narrow pulse of optical energy into one end of an optical fiber and detect the broadened output pulse at the other end.[16,19] Figure 14.9 illustrates a setup for this. Here, output pulses from a laser source are coupled through a mode scrambler into a test fiber. The output of the fiber is measured with a sampling oscilloscope that has a built-in optical receiver, or the signal can be detected with an external photodetector and then measured with a regular sampling oscilloscope. Next, the shape of the input pulse is measured the same way by replacing the test fiber with a short reference fiber that has a length less than 1 percent of the test fiber length. This reference fiber can be a short length cut from the test fiber or it can be a fiber segment that has similar properties. The variable delay in the trigger line is used to offset the difference in delay between the test fiber and the shorter reference fiber.
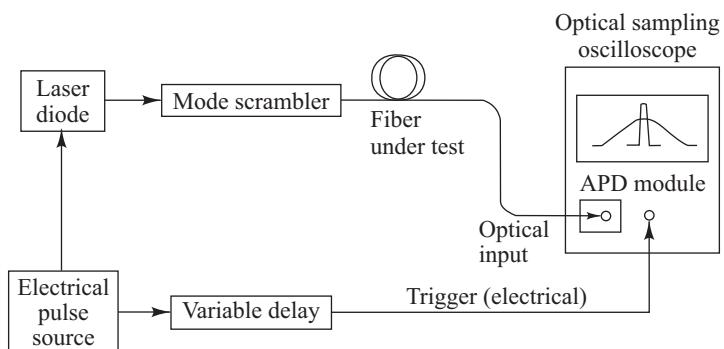
Fig. 14.9 *Test setup for making pulse-dispersion measurements in the time domain*

From the output pulse shape, an rms pulse width σ, as defined in Fig. 14.10, can be calculated by

$$\sigma^2 = \frac{\int_{-\infty}^{\infty} (t - \bar{t})^2 \, p_{\text{out}}(t) \, dt}{\int_{-\infty}^{\infty} p_{\text{out}}(t) \, dt} \tag{14.8}$$

where the center time $\bar{t}$ of the pulse is determined from

$$\bar{t} = \frac{\int_{-\infty}^{\infty} t p_{\text{out}}(t) \, dt}{\int_{-\infty}^{\infty} p_{\text{out}}(t) \, dt} \tag{14.9}$$

The evaluation of Eq. (14.9) requires a numerical integration. An easier method is to assume that the output response of a fiber can be approximated by a gaussian function described by
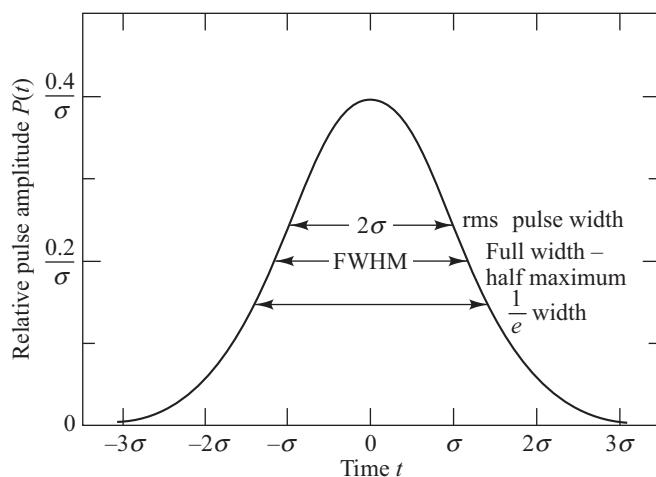
Fig. 14.10 *Definitions of pulse shape parameters*

$$p_{out}(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{t^2}{2\sigma^2}\right) \qquad (14.10)$$

where the parameter σ determines the pulse width, as shown in Fig. 14.10. This figure also shows the parameter t_{FWHM}, which is the full width of the pulse at its half-maximum value. As denoted in Eq. (8.13), this is equal to $2\sigma(2\ln 2)^{1/2} = 2.355\sigma$. As described in Sec. 8.1, the optical bandwidth of the fiber can be defined through a Fourier transform. Normally, this is done in terms of the 3-dB bandwidth, which is the modulation frequency at which the optical power has fallen to one-half the value of the zero frequency modulation (dc value). From Eq. (8.14), this is

$$f_{3-dB\,optical} = \frac{0.440}{t_{FWHM}} = \frac{0.187}{\sigma} \text{ Hz} \qquad (14.11)$$

where "3–dB optical" means a 50 percent optical power reduction. Electrical bandwidths are related to optical bandwidths by $1/\sqrt{2}$, so that

$$f_{3-dB\,electrical} = \frac{1}{\sqrt{2}}\, f_{3-dB\,optical} = \frac{0.311}{t_{FWHM}} = \frac{0.133}{\sigma} \text{ Hz} \qquad (14.12)$$

Frequency-Domain Intermodal Dispersion Measurements

Frequency-domain intermodal dispersion measurements yield information on amplitude-versus-frequency response and phase-versus-frequency response.[19] These data are often more useful for system designers than time-domain pulse-dispersion measurements, especially if equalization techniques are to be performed on the detected signal at the receiver. The dispersion measurements can be made by sinusoidally modulating a narrowband continuous-wave (CW) light signal about a fixed level. The baseband frequency response is then found from the ratio of the sine wave amplitudes at the beginning and end of the fiber.

Figure 14.11 shows an experimental arrangement for finding the fiber baseband frequency response. A swept-frequency RF source or a microwave signal source is used to modulate an optical carrier sinusoidally. The optical signal is coupled through a mode scrambler to the test fiber. At the exit end of the fiber,
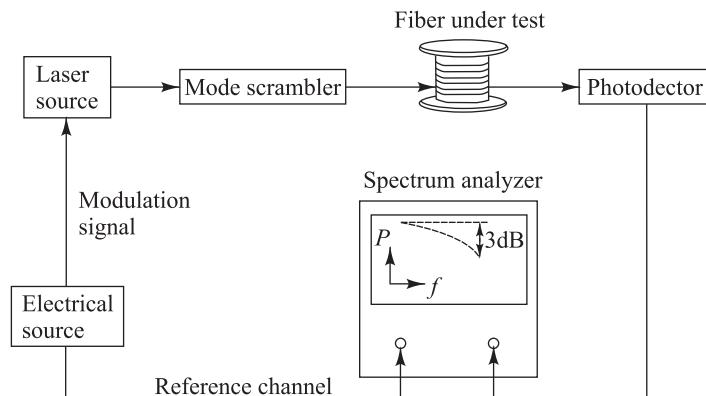
Fig. 14.11   *Test setup for finding the baseband frequency response of a fiber*

a photodetector measures $P_{out}(f)$, the output power as a function of the modulation frequency. The input signal is then measured by substituting a short reference fiber for the test fiber, thereby yielding $P_{in}(f)$.

Comparison of the spectrum at the fiber output with the spectrum at the fiber input provides the baseband frequency response $H(f)$ of the fiber under test:

$$H(f) = \frac{P_{out}(f)}{P_{in}(f)} \tag{14.13}$$

As the modulation frequency is increased, the optical power level at the fiber output will eventually start to decrease. The fiber bandwidth is defined as the lowest frequency at which $H(f)$ has been reduced to 0.5.

Chromatic Dispersion

Chromatic dispersion is a primary dispersive mechanism in single-mode fibers.[19–22] Here, we present one method for its measurement.

Figure 14.12 shows a setup for measuring chromatic dispersion by the *modulation phase-shift method*. An electric signal generator intensity modulates the output of a narrowband, tunable optical source by means of an external modulator. After detecting the transmitted signal with a photodiode receiver, a vector voltmeter is used to measure the phase of the modulation of the received signal relative to the electrical modulation source. The phase measurement is repeated at wavelength intervals $\Delta\lambda$ across the spectral band of interest. Using the measurements at any two adjacent wavelengths, the change in group delay (in ps) over the wavelength interval between them is[19]

$$\Delta\tau_\lambda = \frac{\phi_{\lambda + \Delta\lambda/2} - \phi_{\lambda - \Delta\lambda/2}}{360 f_m} \times 10^6 \tag{14.14}$$

where λ is the wavelength at the center of the interval, f_m is the modulation frequency in MHz, and ϕ is the phase of the measured modulation in degrees.

These data points are then plotted to yield the typical curve shown in Fig. 14.12. The dispersion can be calculated by applying the curve-fitting equations described in Sec. 3.3.3 to the pulse-delay data.
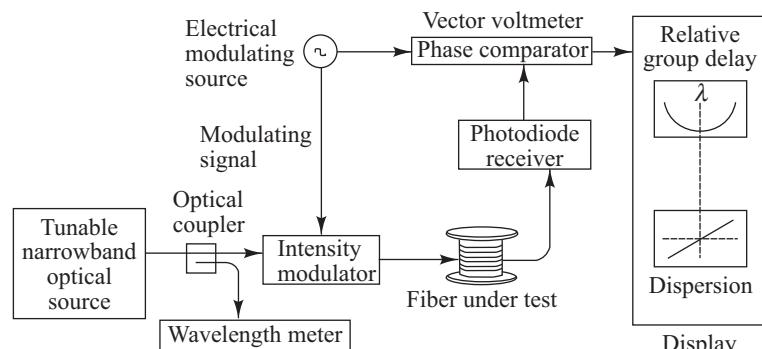
Fig. 14.12   *Test setup and display output for measuring chromatic dispersion by the phase-shift method*

Polarization-Mode Dispersion

As Sec. 3.2.8 describes (see Fig. 3.17), signal energy at a given wavelength occupies two orthogonal polarization modes. Since no fiber is perfectly round and materially symmetrical along its length, it has a varying birefringence along its length. Thus each polarization mode will travel at a slightly different group velocity and the polarization orientation will rotate with distance. The resulting difference in propagation times $\Delta\tau_{\text{pol}}$ between the two orthogonal polarization modes at a given wavelength will result in pulse spreading. This is called *polarization-mode dispersion* (PMD).[23–28] Ultimately, PMD can limit the highest achievable data rates for single-mode systems and its measurement is thus of high importance (see Prob. 14.8).

A special point to note is that, in contrast to chromatic dispersion, which is a relatively stable phenomenon along a fiber, PMD varies randomly along a fiber because of the randomness of the underlying geometric and stress irregularities. Thus statistical predictions are needed to account for its effects. A useful means of characterizing PMD is in terms of the mean or expected value of the differential group delay $\langle\Delta\tau_{\text{PMD}}\rangle$ averaged over time. In contrast to the instantaneous value $\Delta\tau_{\text{PMD}}$, which varies over time and type of source, the expected value does not change from day to day or from source to source. As noted in Sec. 3.2.8, mean values of the polarization-mode dispersion parameter D_{PMD} vary from 0.03 to 1.3 ps/$\sqrt{\text{km}}$, depending on the cable environment.

At least seven different methods have been developed for measuring PMD. Here, we present only the *fixed-analyzer method*.[28] In this technique, the mean differential group delay is evaluated statistically from the number of peaks and valleys appearing in the optical signal power as it is transmitted through a polarizer and scanned as a function of wavelength. Figure 14.13*a* shows a simple setup using a spectrum analyzer. A typical spectrum analyzer trace showing the transmitted power level as a function of wavelength is given in Fig. 14.13*b*. Automatic methods using extrema counting and Fourier analysis are generally used to extract the PMD information from the measurement data. Using extrema counting,

(*a*)

(*b*)
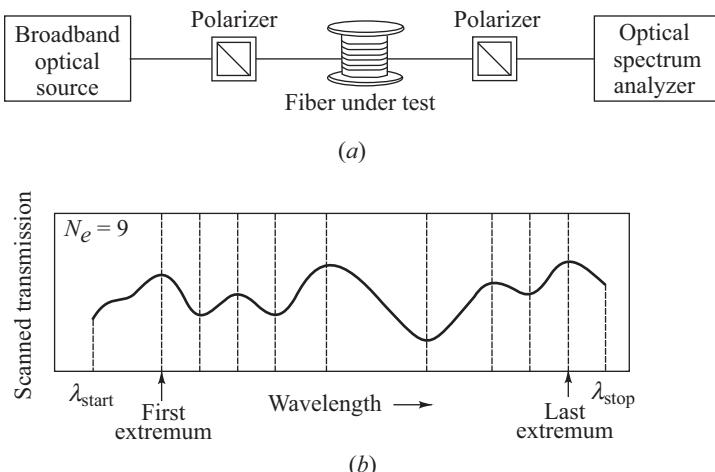
Fig. 14.13 *(a) Setup for measuring polarization-mode dispersion using an optical spectrum analyzer. (b) Typical OSA trace for PMD showing transmitted power level as a function of wavelength.*

the expected value of the differential group delay of the fiber (or of any other device) under test can be calculated from the relationship

$$\left\langle \Delta\tau_{\text{PMD}}\right\rangle_{\lambda} = \frac{kN_e\lambda_{\text{start}}\lambda_{\text{stop}}}{2(\lambda_{\text{start}} - \lambda_{\text{stop}})\,c} \tag{14.15}$$

where λ_{start} and λ_{stop} are the beginning and end, respectively, of the wavelength measurement sweep, N_e represents the number of extrema occurring in the scan, and c is the speed of light. The dimensionless mode-coupling factor k statistically accounts for the wavelength dependence of the polarization states. Its value is 0.84 for randomly mode-coupled fibers and 1.0 for non-mode-coupled fibers and devices.[28] The subscript λ on the $\langle \Delta\tau_{\text{PMD}}\rangle$ term means that the expected value of the differential group delay is determined over a wavelength span.

14.5　Eye Diagram Tests

The use of an *eye diagram* is a traditional technique for quickly and intuitively assessing the quality of a received signal. Modern bit-error rate (also called *bit-error ratio*) measurement instruments construct such eye diagrams by generating a pseudorandom pattern of ones and zeros at a uniform rate but in a random manner. When the pulses in this pattern are superimposed simultaneously, an eye pattern as shown in Fig. 14.14 is formed.[29–35] The word *pseudorandom* means that the generated combination or sequence of ones and zeros will eventually repeat but that it is sufficiently random for test purposes. A *pseudorandom binary sequence* (PRBS) comprises four different 2-bit-long combinations, eight different 3-bit-long combinations, sixteen different 4-bit-long combinations, and so on (that is, sequences of 2^N different N-bit-long combinations) up to a limit set by the instrument. These combinations are randomly selected. The PRBS pattern length is of the form $2^N - 1$, where N is an integer. This choice assures that the pattern-repetition rate is not harmonically related to the data rate. Typical values of N are 7, 10, 15, 20, 23, and 31. After this limit has been reached, the data sequence will repeat.

Ideally, if the signal impairments are small, the received pattern should look like that shown in Fig. 14.14. However, time-varying signal impairments in the transmission path can lead to amplitude variations within the signal and timing skews between the data signal and the associated clock signal. Note that a clock signal, which typically is encoded within a data signal, is used to help the receiver interpret the incoming data correctly. Thus, in an actual link
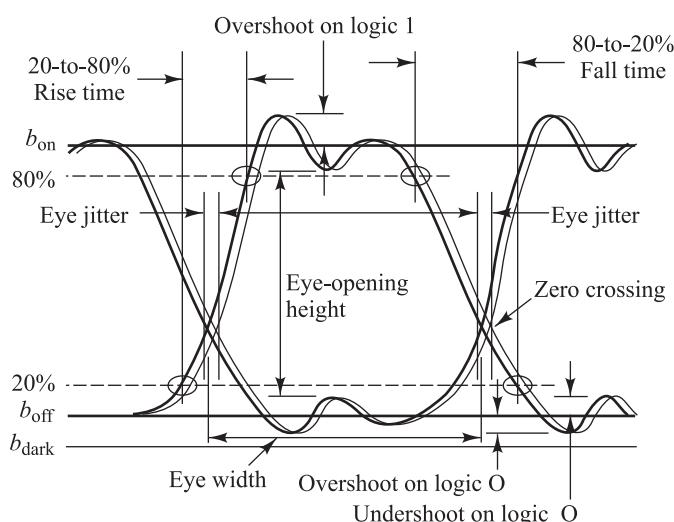
Fig. 14.14　*General configuration of a fairly clean eye diagram showing definitions of fundamental measurement parameters*

the received pattern will become wider or distorted on the sides and on the top and bottom, as shown in Fig. 14.15.

14.5.1 Mask Testing

Interpretation of the characteristics of a distorted eye diagram is done by means *mask testing*.[29,34] Depending on which protocol standard is used, industry-defined masks can take the shape of a polygon or a square, which must fit within the eye diagram opening as shown in Fig. 14.16. In some cases (e.g., 622-Mb/s SONET) the mask is a six-sided polygon sitting in the middle of the eye, whereas for other protocols the mask shape is a rectangle (e.g., OC-48 and OC-192) or a diamond (e.g., Gigabit Ethernet). The *mask height* is sized in proportion to the power level of the signal. This height indicates the minimum separation that is needed

Fig. 14.15 *Signal-distorting effects cause the eye opening to get smaller*

between the logic 1 and 0 levels in order to achieve a specific bit-error rate, as can be derived from the *Q* factor described in Chapter 7. The *slopes* of the polygon edges indicate the allowed 10-to-90 percent rise and fall times. The *mask width* is proportional to the bit rate; that is, the width is narrower for higher bit rates. This is related to the *jitter parameter* shown in Fig. 14.16, which is half the peak-to-peak jitter tolerance associated with the signal. The *overshoot* and *undershoot* parameters bound the amplitudes in terms of the logic 1 and 0 levels, respectively. In Fig. 14.16 the eye measurement parameters are defined as follows:

- P_1 is the mean optical power level associated with a long string of 1 bits
- P_0 is the mean optical power level associated with a long string of 0 bits
- A is the lowest inner upper eye level
- B is the highest inner lower eye level

The operating software of most modern bit-error rate test instruments has a wide selection of built-in masks for different protocols. In addition, the instrument user can key in custom masks for any application or to check the test results differently. Table 14.5 lists the five mask-parameter values of several protocols. The parameter values are given in terms of *unit intervals* (UI), where the pattern height ($P_1 - P_0$) has

Fig. 14.16 *The two bars and the six-sided polygon define the eye-pattern mask.*

Table 14.5 *Standard NRZ eye-mask parameters for several protocols given in terms of unit intervals*

Protocol	Jitter	Rise time	Eye height	Overshoot	Undershoot
OC-3	0.15	0.200	0.60	0.20	0.20
OC-12	0.25	0.150	0.60	0.20	0.20
OC-48/192	0.40	0.000	0.50	0.25	0.25
Gigabit Ethernet	0.22	0.155	0.60	0.30	0.20
Fibre Channel	0.15	0.200	0.60	0.30	0.20
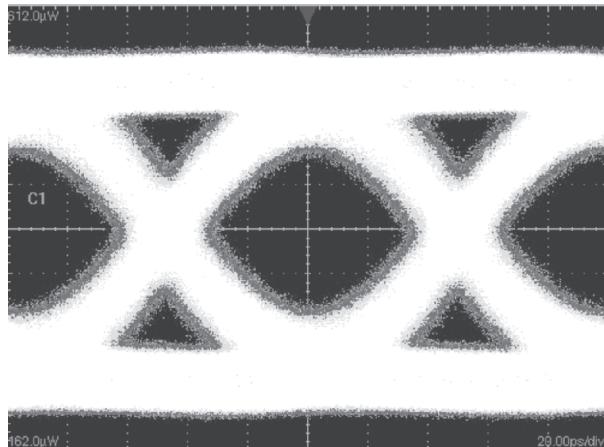
Fig. 14.17 *The inclusion of all possible signal distortion effects results in a stressed eye with only a small opening.*
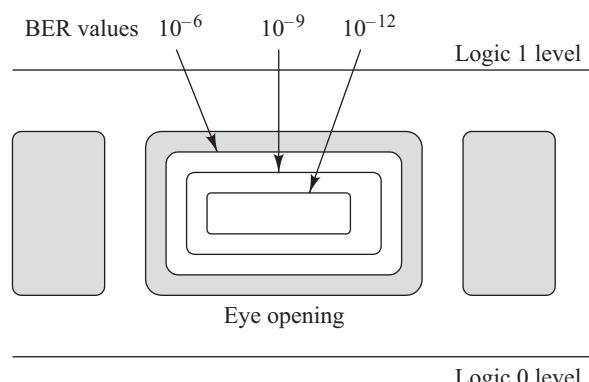
Fig. 14.18 *An eye-contour diagram gives a three-dimensional view of the BER.*

$UI = 1.0$. Note that the rise-time parameter for OC-48 is zero, since the mask is a rectangle.

14.5.2 Stressed Eye

The standards for many high-speed transmission protocols specify a test that uses what is called a *stressed eye*.[35] Among these standards are Gigabit Ethernet, 10-Gigabit Ethernet (l0GigE), Fibre Channel, and SONET OC-48 and OC-192. The concept of this test is to assume that all different possible jitter and intersymbol-interference impairments that might occur to a signal in a fielded link will close the eye down to a diamond shape, as shown in Fig. 14.17. If the eye opening of the optical receiver under test is greater than this diamond-shaped area of assured error-free operation, then it is expected to operate properly in an actual fielded system. The stressed-eye template height typically is between 0.10 and 0.25 UI.

14.5.3 Eye Contour

Associated with the stressed eye is a parameter called the *BER contour*.[35] Basically the BER contours are analogous to geographical contours that indicate the height and steepness profile of a hill. As indicated in Fig. 14.18, the BER contours show different levels of error probability within an eye diagram. From such a plot it can be seen that the different BER contours get closer to each other as the slope becomes steeper. This means that errors are more likely to occur if the receiver operates
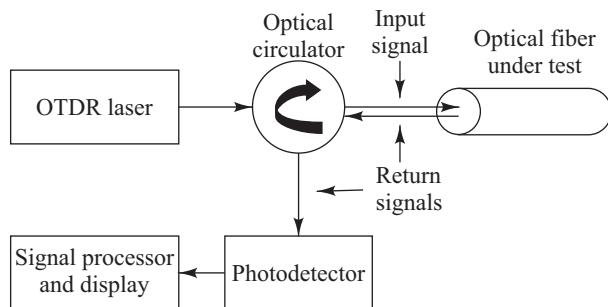
Fig. 14.19 *Operational principle of an OTDR using an optical circulator*

close to the edge of such a steep contour plot. Thus the farther a receiver decision point is within a contour boundary, the better its performance will be. This is referred to as having a *healthier eye*.

14.6 Optical Time-Domain Reflectometer

An *optical time-domain reflectometer* (OTDR) is a versatile portable instrument that is used widely to evaluate the characteristics of an installed optical fiber link. In addition to identifying and locating faults or anomalies within a link, this instrument measures parameters such as fiber attenuation, length, optical connector and splice losses, and light reflectance levels.[36–40]

An OTDR is fundamentally an optical radar. As shown in Fig. 14.19, the OTDR operates by periodically launching narrow laser pulses into one end of a fiber under test by using either a directional coupler or a circulator. The properties of the optical fiber link then are determined by analyzing the amplitude and temporal characteristics of the waveform of the reflected and back-scattered light. A typical OTDR consists of a light source and receiver, data-acquisition and processing modules, an information-storage unit for retaining data either in the internal memory or on an external disk, and a display. Figure 14.20 shows a portable OTDR for making measurements in the field.

Fig. 14.20 *Example of portable universal test sets that can be used for OTDR traces, optical spectrum analysis, dispersion analysis, and other test functions. (Model FTB-500, photo provided courtesy of EXFO; www.exfo.com. © 2004–2010 EXFO Inc.)*
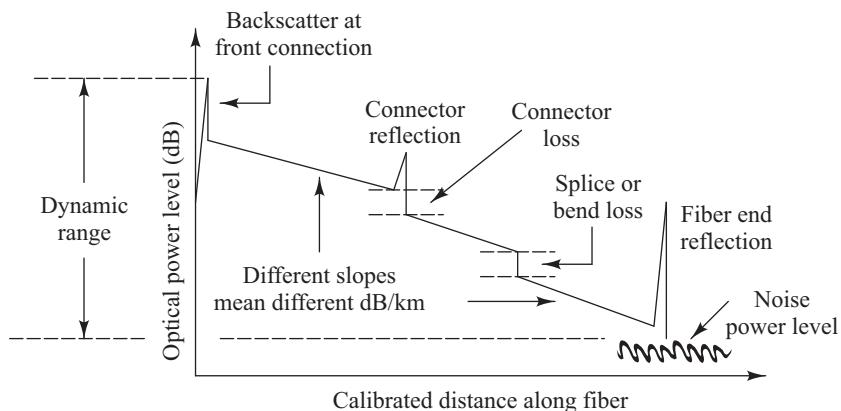
Fig. 14.21 *Representative trace of backscattered and reflected optical power as displayed on an OTDR screen and the meanings of various trace features*

14.6.1 OTDR Trace

Figure 14.21 shows a typical trace as seen on the display screen of an OTDR. The scale of the vertical axis is logarithmic and measures the returning (back-reflected) signal in decibels. The horizontal axis denotes the distance between the instrument and the measurement point in the fiber. In addition to the trace, an OTDR such as the one shown in Fig. 14.20 can place a number next to an event on the display and give a list of these numbers and their corresponding measurement information in a table below the trace.

The backscattered waveform has four distinct features:

- A large initial pulse resulting from Fresnel reflection at the input end of the fiber
- A long decaying tail resulting from Rayleigh scattering in the reverse direction as the input pulse travels along the fiber
- Abrupt shifts in the curve caused by optical loss at joints or connectors in the fiber line
- Positive spikes arising from Fresnel reflection at the far end of the fiber, at fiber joints, and at fiber imperfections

Fresnel reflection and Rayleigh scattering principally produce the backscattered light. *Fresnel reflection* occurs when light enters a medium having a different index of refraction. For a glass–air interface, when light of power P_0 is incident perpendicular to the interface, the reflected power P_{ref} is

$$P_{ref} = P_0 \left(\frac{n_{fiber} - n_{air}}{n_{fiber} + n_{air}} \right)^2 \tag{14.16}$$

where n_{fiber} and n_{air} are the refractive indices of the fiber core and air, respectively. A perfect fiber end reflects about 4 percent of the power incident on it. However, since fiber ends generally are not polished perfectly and perpendicularly to the fiber axis, the reflected power tends to be much lower than the maximum possible value. In particular, this is the case if an *angle-polished connector* (APC) is used.

The detection and measurement accuracy of an event depend on the signal-to-noise ratio (SNR) that an OTDR can achieve at that point. This is defined as the ratio between the back-reflected signal and the noise level. The SNR depends on factors such as the OTDR pulse width, how often the OTDR samples the signal, and the distance to the measurement point.

Two important performance parameters of an OTDR are dynamic range and measurement range. *Dynamic range* is defined as the difference between the initial backscatter power level at the front connector and the noise level peak at the far end of the fiber. It is expressed in decibels of one-way fiber loss. Dynamic range provides information on the maximum fiber loss that can be measured and denotes the time required to measure a given fiber loss. A basic limitation of an OTDR is the tradeoff between dynamic range and *event location resolution.* For high spatial resolution, the pulse width has to be as small as possible. However, this reduces the signal-to-noise ratio and thus lowers the dynamic range. Typical distance resolution values range from 8 cm for a 10-ns pulse to 5 m for a 50 μs pulse.

Measurement range deals with how far away an OTDR can identify events in the link, such as splice points, connection points, or fiber breaks. The maximum range R_{max} depends on the fiber attenuation α and on the pulse width, that is, on the dynamic range D_{OTDR}. If the attenuation is given in dB/km, then the maximum range in km is

$$R_{max} = D_{OTDR} / \alpha \tag{14.17}$$

Example 14.3 Consider an OTDR that has a dynamic range of 36 dB. If a cable installation engineer wants to use this instrument to characterize a fiber with a 0.5-dB/km attenuation, what is the maximum fiber range R_{max} that can be tested?

Solution: From Eq. (14.17) we find that the maximum range is
$$R_{max} = D_{OTDR}/\alpha = 72 \text{ km}$$

14.6.2 Attenuation Measurements

Rayleigh scattering reflects light in all directions throughout the length of the fiber. This factor is the dominant loss mechanism in most high-quality fibers. The optical power that is Rayleigh-scattered in the reverse direction inside the fiber can be used to determine attenuation.

The optical power at a distance x from the input coupler can be written as

$$P(x) = P(0) \exp\left[-\int_0^x \beta(y)\, dy \right] \tag{14.18}$$

Here, $P(0)$ is the fiber input power and $\beta(y)$ is the fiber loss coefficient in km^{-1}, which may be position-dependent; that is, the loss may not be uniform along the fiber. The parameter 2β can be measured in natural units called *nepers*, which are related to the loss $\alpha(y)$ in decibels per kilometer through the relationship (see App. D)

$$\beta(\text{km}^{-1}) = 2\beta(\text{nepers}) = \frac{\alpha(\text{dB})}{10 \log e} = \frac{\alpha(\text{dB})}{4.343} \tag{14.19}$$

Under the assumption that the scattering is the same at all points along the optical waveguide and is independent of the modal distribution, the power $P_R(x)$ scattered in the reverse direction at the point x is

$$P_R(x) = SP(x) \tag{14.20}$$

Here, S is the fraction of the total power that is scattered in the backward direction and trapped in the fiber. Thus the backscattered power from point x that is seen by the photodetector is

$$P_D(x) = P_R(x) \exp\left[-\int_0^x \beta_R(y)\, dy \right] \tag{14.21}$$

where $\beta_R(y)$ is the loss coefficient for the reverse-scattered light. Since the modes in the fiber excited by the backscattered light can be different from those launched in the forward direction, the parameter $\beta_R(y)$ may be different from $\beta(y)$.

Substituting Eqs. (14.18), (14.19), and (14.20) into Eq. (14.21) yields

$$P_D(x) = SP(0) \exp\left[-\frac{2\bar{\alpha}(x)x}{10 \log e} \right] \tag{14.22}$$

where the average attenuation coefficient $\bar{\alpha}(x)$ is defined as

$$\bar{\alpha}(x) = \frac{1}{2x} \int_0^x [\alpha(y) + \alpha_R(y)]\, dy \tag{14.23}$$

Using this equation, the average attenuation coefficient can be found from a data trace such as the one shown in Fig. 14.20. For example, the average attenuation between two points x_1 and x_2, where $x_1 > x_2$, is

$$\bar{\alpha} = -\frac{10[\log P_D(x_2) - \log P_D(x_1)]}{2(x_2 - x_1)} \tag{14.24}$$

Example 14.4   An OTDR is used to measure the attenuation of a long length of fiber. If the optical power level measured by the OTDR at the 8-km point is 0.5 of the measured value at the 3-km point, what is the fiber attenuation?

Solution: We can express Eq. (14.24) as

$$\alpha = \frac{10 \log\left[\dfrac{P_D(x_2)}{P_D(x_1)} \right]}{2(x_2 - x_1)} = \frac{10 \log 0.5}{2(8 - 3)} = 0.3 \text{ dB/km}$$

14.6.3   OTDR Dead Zone

The concept of a dead zone is another important OTDR specification. *Dead zone* is the distance over which the photodetector in an OTDR is saturated momentarily after it measures a strong reflection. As Fig. 14.22 shows, there are two specifications for dead zone. An *event dead zone* specifies the minimum distance over which an OTDR can detect a reflective event that follows another reflective event. Typically vendors specify this as the distance between the start of a reflection and the -1.5-dB point on the falling edge of the reflection. A short pulse width is used when measuring the event dead zone. For example, a 30-ns pulse width would give a 3-m event dead zone.

The *attenuation dead zone* indicates over which distance the photodetector in an OTDR needs to recover following a reflective event before it is again able to detect a splice. This means that the receiver has to recover to within 0.5 dB of the backscatter value. Typical attenuation dead zones range from 10 to 25 m.

Typically an OTDR dead zone is the same length as the distance that the optical pulse covers in a fiber plus a few meters. Thus OTDR vendors have started employing a special length of fiber called an *optical pulse suppressor* (OPS), which is inserted between the OTDR and the fiber. An OPS moves the
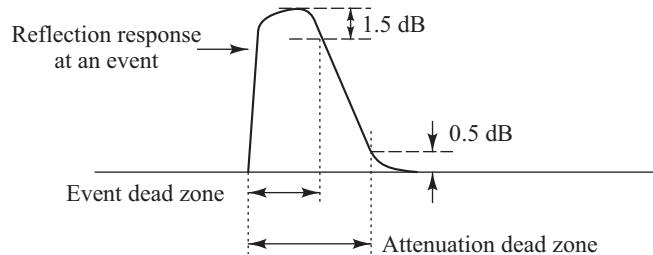
Fig. 14.22 *Two specifications for dead zone are the event dead zone and the attenuation dead zone.*

dead zone from the beginning of the fiber under test to this special fiber. This can reduce the event dead zone to about 1 m, so that anomalies occurring within a short distance, for example, within the cabling system of a central office, may be detected and measured.

14.6.4 Fiber Fault Location

To locate breaks and imperfections in an optical fiber, the fiber length L (and, hence, the position of the break or fault) can be calculated from the time difference between the pulses reflected from the front of the fiber and the event location. If this time difference is t, then the length L is given by

$$L = \frac{ct}{2n_1} \tag{14.25}$$

where n_1 is the core refractive index of the fiber. The number "2" in the denominator accounts for the fact that light travels a length L from the source to the break point and then another length L on the return trip.

Example 14.5 Consider a long optical fiber with a core refractive index $n_1 = 1.460$. Suppose that an engineer uses OTDR to locate a break in the fiber. If the break is located 15 km away, what is the return time of an OTDR test pulse?

Solution: Using Eq. (14.25) we find

$$t = \frac{2n_1 L}{c} = \frac{2(1.460)(15\text{ km})}{3 \times 10^5\text{ km/s}} = 0.146\text{ ms}$$

14.6.5 Optical Return Loss

Reflections of the light in a backward direction occur at various points in optical links that use laser transmitters. This can occur at connectors, fiber ends, optical splitter interfaces, and within the fiber itself due to Rayleigh scattering. The percent of power reflected back from a particular point in a light path is called *back reflection*. If it is not controlled, the back reflections can cause optical resonance in the laser source and result in erratic operation and increased laser noise. In addition, the back reflections can undergo multiple reflections in the transmission line and increase the bit-error rate when they enter the receiver.

Therefore it is desirable to measure the *optical return loss* (ORL), which is the percent of total reverse power in relation to total forward power at a particular point. The ORL is expressed as a ratio of reflected power P_{ref} to incident power P_{inc}

Fig. 14.23 *Compact portable optical return loss (ORL) tester, which also can be used as a power meter. (Model 7340, photo provided courtesy of Kingfisher; www .kingfisher.com.au.)*

$$\text{ORL} = 10\log(P_{\text{ref}}/P_{\text{inc}}) \tag{14.26}$$

One can use either an OTDR or an ORL meter to measure this parameter. Although an OTDR can give precise reflectance values at individual events along a fiber transmission path, it has a limitation in measuring the back reflections near and within the OTDR dead zone. Since such an event can be a major contributor to ORL, it is better to use a return loss meter, such as the one shown in Fig. 14.23.

14.7 Optical Performance Monitoring

Modern communication networks have become an essential part of society with applications ranging anywhere from simple web browsing to high-profile business transactions. Due to the importance of these networks to everyday life, users have come to expect the network to always be available and to function properly. To offer services with an extremely high degree of reliability, operators need to have a means to monitor the health and status of all parts of their network continuously. In a SONET/SDH network this monitoring function is the performance-management subset of a larger set of network management functions. Basically the network health is assessed by means of a continuous in-line BER measurement. The information obtained from this test is used to assure that the quality-of-service (QoS) requirements are met. In addition, another standard network management function is fault monitoring, which checks to see where and why a network failure has occurred or is about to take place.

Optical performance monitoring (OPM) adds to these standard network management concepts by checking the status of elements in the physical layer to examine the temporal behavior of the basic performance factors that affect the signal quality. Depending on the desired network-control complexity and the system cost constraints, optical performance monitoring can range anywhere from simply checking the optical power level of each WDM channel to a highly sophisticated system that identifies the origins of a wide range of signal impairments and assesses their impact on network performance.[41–55]

This section first gives an overview of generic network management functions to show their relationship to OPM. Then Sec. 14.7.2 discusses management functions defined by the ITU-T for the multiple

wavelengths in the optical layer. These functions are an extension of the standard SONET/SDH procedures used for managing a single wavelength. Next Sec. 14.7.3 describes three levels of monitoring functions that can be carried through different categories of OPM. Sections 14.7.4 through 14.7.6 then give some general examples of OPM procedures. These include network maintenance, fault management, and OSNR monitoring. Section 14.8 describes some specific measurement methods.

14.7.1 Management Architecture and Functions

Once the hardware and software elements of an optical network have been installed properly and integrated successfully, they need to be managed to ensure that the required level of network performance is met. In addition, the network devices must be monitored to verify that they are configured properly and to ensure that corporate policies regarding network use and security procedures are being followed. This is carried out through *network management*, which is a service that uses a variety of hardware and software tools, applications, and devices to assist human network managers in monitoring and maintaining networks.

Figure 14.24 shows the components of a typical network management system and their relationships. The *network management console* is a specialized workstation that serves as the interface for the human network manager. There can be several of these workstations that perform different functions in a network. From such a console a network manager can view the health and status of the network to verify that all devices are functioning properly, that they are configured correctly, and that their application software is up to date. A network manager also can see how the network is performing, for example, in terms of traffic loads and fault conditions. In addition, the console allows control of the network resources.

The *managed devices* are network components, such as optical transmitters and receivers, optical amplifiers, optical add/drop multiplexers (OADMs), and optical crossconnects (OXCs). Each such device is monitored and controlled by its *element management system* (EMS). Management software modules, called *agents*, residing in a microprocessor within the elements continuously gather and compile information on the status and performance of the managed devices. The agents store this information in a *management information base* (MIB), and then provide the data to *management entities* within a *network management system* (NMS) that resides in the management workstation. A MIB is a logical base of information that defines
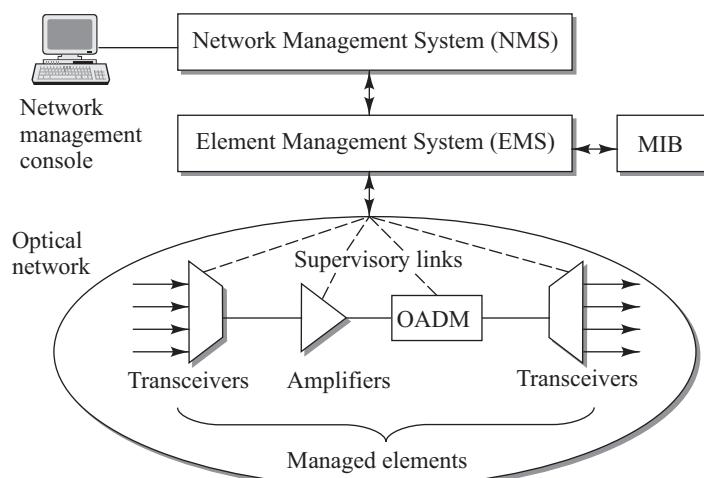
Fig. 14.24 *The components of a typical network management system and their relationships*

Table 14.6 *The purposes of five basic network management functions*

Management function	Purpose
Performance management	Monitor and control parameters that are essential to the proper operation of a network in order to guarantee a specific quality of service to network users.
Configuration management	Monitor network setup information and network device configurations to track and manage the effects on network operation of the various hardware and software elements.
Accounting management	Measure network-utilization parameters so that individuals or groups of users on the network can be regulated and billed for services appropriately.
Fault management	Detect fault or degradation symptoms, determine the origin and possible cause of faults, issue instructions on how to resolve the fault.
Security management	Develop security policies, set up a network security architecture, implement firewall and virus-protection software, establish access-authentication procedures.

data elements and their appropriate syntax and identifier, such the fields in a database. This information may be stored in tables, counters, or switch settings. The MIB does not define how to collect or use data elements. It only specifies what the agent should collect and how to organize these data elements so that other systems can use them. The information transfer from the MIB to the NMS is done via a *network management protocol* such as the widely used *simple network management protocol* (SNMP).

When agents notice problems in the element they are monitoring (for example, link or component faults, wavelength drifts, reduction in optical power levels, or excessive bit-error rates), they send alerts to the management entities. Upon receiving an alert, the management entities can initiate one or more actions such as operator notification, event logging, system shutdown, or automatic attempts at fault isolation or repair. The EMS also can query or poll the agents in the elements to check the status of certain conditions or variables. This polling can be automatic or operator-initiated. In addition, there are *management proxies* that provide management information on behalf of devices that are not able to host an agent.

Network management functions can be classified into the five generic categories listed in Table 14.6. These are performance, configuration, accounting, fault, and security management.

14.7.2 Optical Layer Management

To deal with standardized management functions in the optical layer, the ITU-T defined a three-layer *optical transport network* (OTN) model in ITU-T Rec. G.709, which also is referred to as the *Digital Wrapper* standard. Just as the SONET/SDH standard enabled the management of single-wavelength optical networks using equipment from many different vendors, the G.709 standard enables the broad adoption of technology for managing multiwavelength optical networks. The structure and layers of the OTN closely parallel the path, line, and section sublayers of SONET.

The model is based on a client/server concept. The exchange of information between processes running in two different devices connected through a network may be characterized by a *client/server interaction*. The terms *client* and *server* describe the functional roles of the elements in the network, as Fig. 14.25 illustrates. The process or element that requests or receives information is called the *client*, and the process or element that supplies the information is called the *server*.
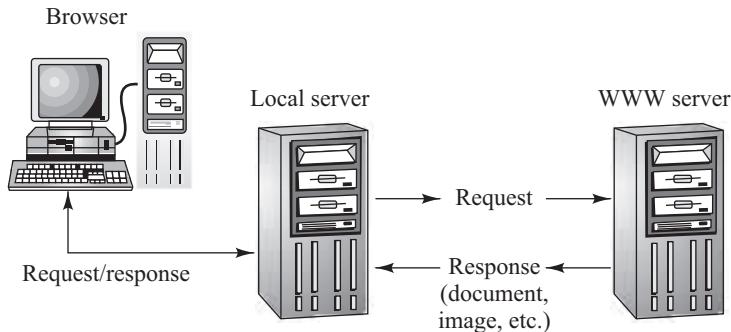
Fig. 14.25   *The terms client and server describe the functional roles of communicating elements in the network. Here the browser is the client.*

Figure 14.26 illustrates the three-layer model for a simple link. Client signals such as IP, Ethernet, or OC-N/STM-M are mapped from an electrical digital format into an optical format in an optical channel (OCh) layer. The OCh deals with single wavelength channels as end-to-end paths or as subnetwork connections between routing nodes. The *optical multiplex section* (OMS) layer represents a link carrying groups of wavelengths between multiplexing equipment or OADMs. The *optical transport section* (OTS) layer relates to a link between two optical amplifiers. Figure 14.27 shows where these sections fit into a link.

The OCh is divided further into three sublayers: the *optical channel transport unit* (OTU), the *optical channel data unit* (ODU), and the *optical channel payload unit* (OPU). Each of these sublayers has its own functions and associated overhead, which are as follows.

Optical channel payload unit   The OPU frame structure contains the client signal payload and the overhead necessary for mapping any client signal into the OPU. Mapping of client signals may include rate adaptation of the client signal to a constant bit-rate signal. Examples of common signals are IP, various forms of Ethernet, ATM, Fibre Channel, and SONET/SDH. The three payload rates associated with the OPU sublayer are 2.5, 10, and 40 Gb/s. These correspond to standard SONET/SDH data rates (OC-48/STM-16, OC-192/STM-64, and OC-768/STM-256, respectively) but may be used for any client signal.

Optical channel data unit   The ODU is the structure used to transport the OPU. The ODU consists of the OPU and the associated ODU overhead and provides path-layer-connection monitoring functions. The ODU overhead contains information that enables maintenance and operation of optical channels. This
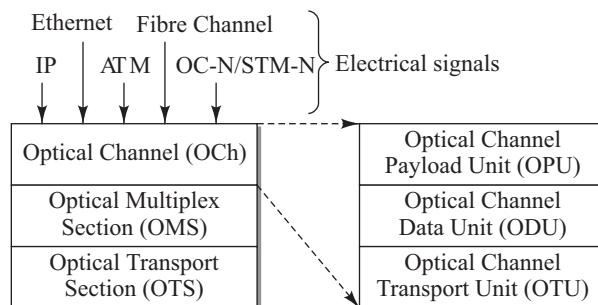
Fig. 14.26   *Three-layer model for a simple link in an OTN. The OCh is divided further into three sublayers.*
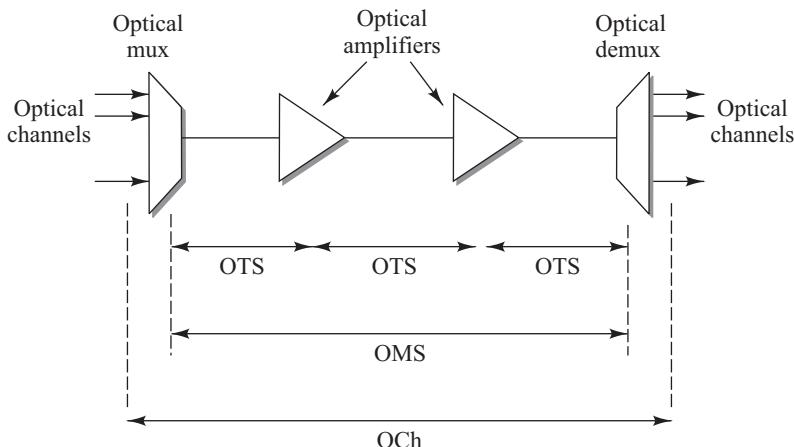
Fig. 14.27 *The OMS layer represents a link carrying wavelengths between multiplexers or OADMs. The OTS layer relates to a link between two optical amplifiers.*

information includes maintenance signals, path monitoring, tandem connection monitoring, automatic protection switching, and designation of fault type and location.

Optical channel transport unit The optical channel transport unit (OTU) contains the ODU frame structure, the OTU overhead, and appended forward error correction (FEC). The OTU changes the digital format of the ODU into a light signal for transport over an optical channel. It also provides error detection and correction and section layer connection monitoring functions.

14.7.3 OPM Functions

The fundamental function of optical performance monitoring is to examine the temporal behavior of performance factors that may affect the health of an optical signal. This process involves checking the operational status of elements in the physical layer and assessing the quality of the optical signals in each WDM channel. OPM can be viewed by means of the following three layers.[41–55]

Transport-layer monitoring deals with optical-domain characteristics that relate to WDM channel management. This involves real-time examinations of factors such as the presence of a channel, whether the wavelength has been registered by the system, and the optical power level, spectral content, and OSNR of each WDM channel.

Optical signal monitoring examines the quality of each WDM channel. This measurement function examines the signal-quality features of an individual channel. Among these features are the Q factor, the electronic SNR, and various eye-diagram statistics such as openness and distortions resulting from dispersion or nonlinear effects.

Protocol performance monitoring deals with digital measurements such as the bit-error rate.

The main factors that OPM is watching for are component faults and signal impairments. *Component faults* can result from malfunctions or degradations of elements, improperly installed or configured equipment, or damage to a network (such as a backhoe digging up a cable or a storm destroying a fiber line). *Signal impairments* can arise from many diverse factors. Among these are noises and transients from optical amplifiers, chromatic and polarization-mode dispersions, nonlinear effects, and timing jitter.

All these factors taken together present a big challenge to devising a comprehensive OPM system. However, a single OPM system does not need to check all possible degradation mechanisms. In fact, cost constraints prevent the deployment of such a super-sophisticated performance-monitoring procedure. In a practical network, the simplest OPM system might monitor only the optical power levels of each channel at a particular point in a WDM network. Advanced OPM systems will include using a miniaturized spectrometer to control the outputs of devices such as optical amplifiers and variable optical attenuators. More complex OPM systems are needed in reconfigurable networks to track the amount of accumulated dispersion on a per channel basis, since the effect on system performance of this impairment will vary as the network configuration changes.

14.7.4 Network Maintenance

An OPM taps off a small portion of the light signals in a fiber and separates the wavelengths or scans them onto a detector or detector array. This enables the measurement of individual channel powers, wavelength, and OSNR. These devices have an important role in controlling DWDM networks. For example, as shown in Fig. 14.28, most long-haul DWDM networks incorporate automated end-to-end power-balancing algorithms that use a high-performance OPM to measure the optical power level of each wavelength at optical amplifiers and at the receiver and to adjust the individual laser outputs at the transmitter. This information is exchanged by means of a separate supervisory channel, which uses a wavelength that lies outside of the signal spectrum but within the response band of the amplifier. In addition, manufacturers may embed an OPM function into dynamic elements such as an EDFA, an OADM, or an OXC to provide feedback for active control of total output power and to balance the power levels between channels. Other functions of an OPM include determining if a particular channel is active, verifying whether wavelengths match the specified channel plan, and checking whether optical power and OSNR levels are sufficient to meet the QoS requirements.

An OPM may have the following operational characteristics:
- Measures absolute channel power to within ±0.5 dBm
- Identifies channels without prior knowledge of the wavelength plan
- Makes full S-, C-, or L-band measurements in less than 0.5 second
- Measures center wavelength accuracy to better than ±50 pm
- Determines OSNR with a 35-dB dynamic range to a ±0.1-dB accuracy
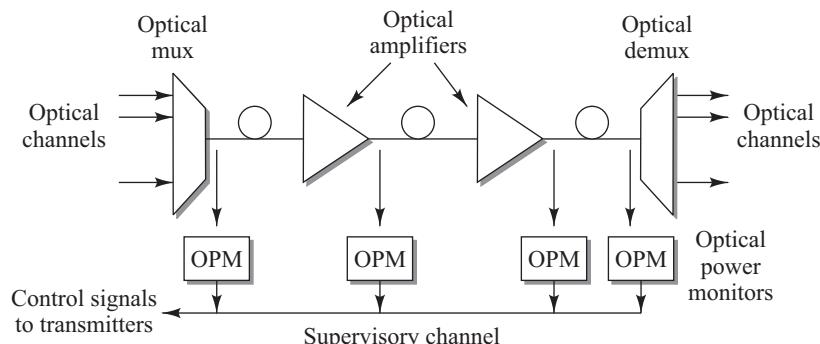
Fig. 14.28 *DWDM networks might use an automated OPM to measure the light level of each wavelength at various network points and to adjust the individual laser outputs at the transmitter.*
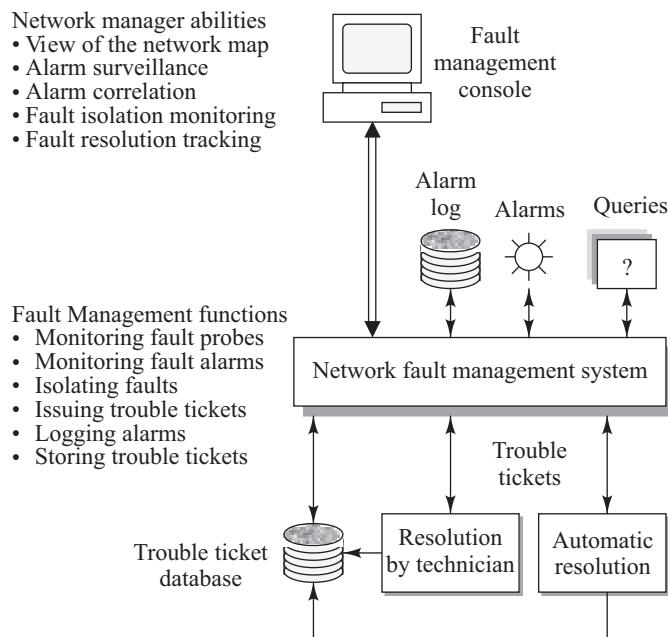
Fig. 14.29   *Functions and interactions of a network fault management system*

14.7.5   Fault Management

Faults in a network, such as physical cuts in a fiber transmission line or failure of a circuit card or optical amplifier, can cause portions of a network to be inoperable. Since network faults can result in system downtime or unacceptable network degradation, *fault management* is one of the most widely implemented and essential network management functions. As Fig. 14.29 illustrates, fault management involves the following processes:

- Detecting fault or degradation symptoms. This can be done with *alarm surveillance*, which involves reporting alarms that may have different levels of severity and indicating possible causes of these alarms. Fault management also provides a summary of unresolved alarms and allows the network manager to retrieve and view the alarm information from an alarm log.
- Determining the origin and possible cause of faults either automatically or through the intervention of a network manager. To determine the location or origin of faults, the management system might use *fault-isolation* techniques such as alarm correlation from different parts of the network and diagnostic testing.
- Once the faults are isolated, the system issues *trouble tickets* that indicate what the problem is and possible means of how to resolve the fault. These tickets go to either a technician for manual intervention or an automatic fault-correction mechanism. When the fault or degradation is corrected, this fact and the resolution method are indicated on the trouble ticket, which then is stored in a database.
- Once the problem has been fixed, the repair is operationally tested on all major subsystems of the network. *Operational testing* involves requesting performance tests, tracking the progress of these tests, and recording the results. The classes of tests that might be performed include echo tests and connectivity examinations.
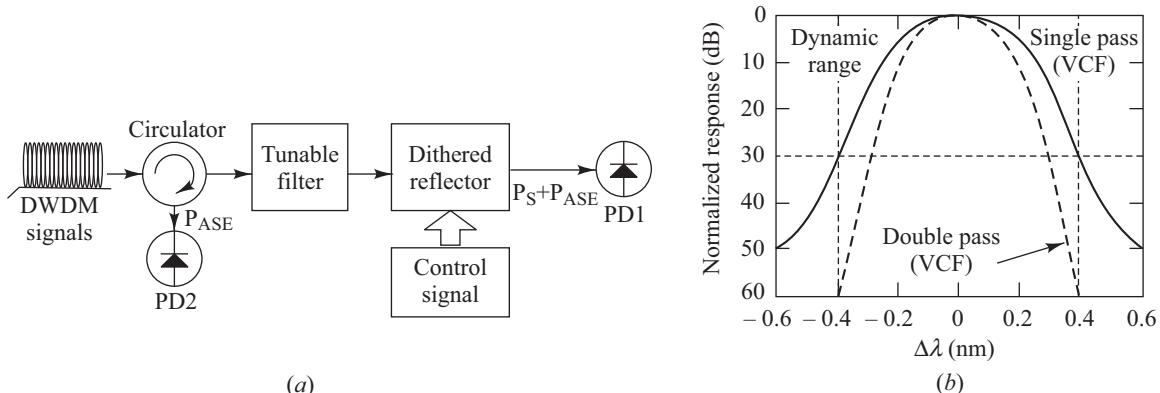
Fig. 14.30 *(a) Schematic diagram of an OSNR monitoring method; (b) response characteristic of the tunable voltage-controlled filter (VCF). (Reproduced with permission from Yang and Lee,[54] © 2004, IEEE.)*

A basic factor in troubleshooting faults is to have a comprehensive physical and logical map of the network. Ideally this map should be part of a software-based management system that can show the network connectivity and the operational status of the constituent elements of the network on a display screen. With such a map, failed or degraded devices can be viewed easily and corrective action can be taken immediately.

14.7.6 OSNR Monitoring

Monitoring of the optical signal-to-noise ratio is a key procedure to implement functions such as gain equalization of optical amplifiers. Many techniques have been devised for OSNR monitoring. The main concept is to use two optical filtering mechanisms: One should be wide enough to accommodate the signal spectrum that is needed to measure the optical power and the other filtering part should be narrow enough to reject power from adjacent channels so that the noise power can be measured. Figure 14.30 shows one method that uses a tunable optical filter and a circulator.[54] In this setup the signal passes through the filter once whereas the noise goes through the filter twice. As shown in the figure, the signal power is measured at the output of the partial reflector, while the noise power is measured at the circulator output. Adding a dithering signal onto the reflector eliminates any residual reflection signals in the system. This setup can measure an OSNR of up to 44 dB with an error of less than 0.4 dB.

14.8 Optical Fiber System Performance Measurements

The evolution of optical fiber communication technology has resulted in highly reliable telecom transmission systems for applications that include high-capacity long-distance links and optical metro, access, and in-building networks. A wide variety of communication protocols, data modulation formats, performance monitoring techniques, and performance testing methods have been devised to keep these systems running smoothly and reliably. In terms of performance testing, the major measurement methods include the bit-error rate (BER), the optical signal-to-noise ratio (OSNR), the *Q* factor, timing jitter, and the optical modulation amplitude. This section provides an overview of these techniques. Further extensive details can be found in the literature.[3,56–63]
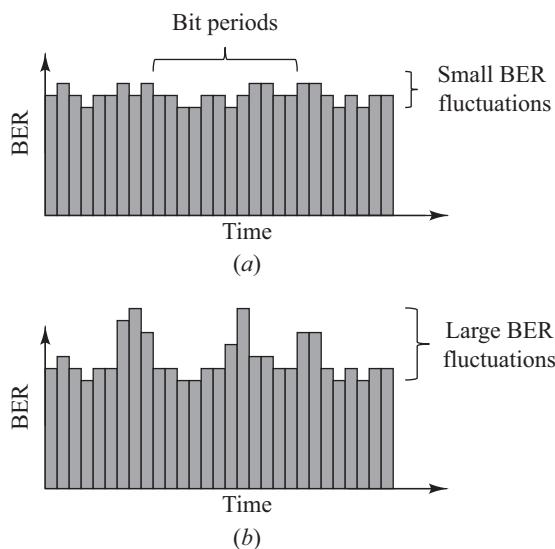
Fig. 14.31 *Sequence of bit periods with (a) relatively stable BER and (b) bursty BER*

14.8.1 Bit-Error Rate Testing

The BER is an important performance quality indicator of a digital communication link. Since BER is a statistical parameter, its value depends on the measurement time and on the factors that cause the errors, such as signal dispersion, accumulated excess noise, and timing jitter. When BER measurements are made, both the number of misinterpreted bits and the total number of received bits are counted in a specific time window ΔT, which is called the *gating time*. If the errors are due to gaussian noise in a relatively stable transmission link, then the BER does not fluctuate significantly over time, as is illustrated in Fig. 14.31a. In this case, a gating time window in which about 100 errors occur is needed to ensure a statistically valid BER. When bursts of errors occur, as shown in Fig. 14.31b, longer measurement times may be needed to accumulate 100 errors in order for the test to be statistically accurate.

From Eq. (7.5) we have that when N_e errors occur in a time window ΔT at a bit rate B, then the BER is given by

$$\text{BER} = \frac{N_e}{B \, \Delta T} \tag{14.26}$$

Thus the gating time window needed to measure $N_e = 100$ errors is $\Delta T = 100/(\text{BER} \times B)$. For high-speed communications greater than 1 Gb/s, the required bit-error rate typically needs to be 10^{-12} or lower.

Example 14.6 A given 10-Gb/s link is designed to operate with a 10^{-12} BER. What is the gating time window for accumulating 100 bit errors?

Solution: From Eq. (14.26) the gating time is

$$T = 100/(\text{BER} \times B) = 100/(10^{-12} \times 10^{10})$$
$$= 10^4 \text{ s} \approx 2.7 \text{ hrs}$$

However, a 10^{-12} BER level may be unacceptable for a 10-Gb/s data rate, so even lower bit-error rates, such as 10^{-15}, may be required to assure customers of a high grade of service. To accumulate 100 errors for such a BER would require over 100 days of measurement time. Since this is not practical, modern BER measuring instruments add an extra precisely calibrated amount of noise into the system, thereby accelerating the occurrence of errors. The extra noise decreases the receiver threshold, which increases the probability of errors and thus greatly reduces the gating time window. Although some accuracy is lost in this method, it reduces the test times to minutes instead of hours or days.

Figure 14.32 shows a typical loopback BER test setup, which consists of a bit pattern generator, a transmission link or device being tested, and a bit-error detector that is colocated with the pattern generator.

Fig. 14.32 *A typical loopback BER test setup. (Photo courtesy of Yokogawa: Multi-Application Test System with BERT module; http://tmi.yokogawa.com.)*

First a pattern generator creates a PRBS that will be sent optically over the transmission link. At the receiving end, the error detector compares the received data stream with a reference signal. In a BER test these two signal streams must be perfectly synchronized in time to be compared correctly. Modern BER test instruments commonly have built-in synchronization features.

To determine the error count, the error detector uses an exclusive OR gate to compare the arriving bits with the bit pattern coming from a reference path. As shown in Fig. 14.33, when the data patterns from the reference and the transmission paths are identical and synchronized in time, the exclusive OR gate gives a zero output. If the patterns do not match in a specific bit period, the output from the exclusive OR gate is a binary 1, which indicates a bit error.

Example 14.7 In a real communication system, such as a SONET/SDH network, network operators also are interested in the frame error rate P_{frame}. If the total number of bits in a frame is k and P_e is the BER, then the probability that no bit in the frame has an error is

$$1 - P_{\text{frame}} = (1 - P_e)^k \approx 1 - P_e k \qquad (14.27)$$

where the approximation results from the condition that $P_e \ll 1$. If $P_e = 10^{-12}$, what is the frame-error rate for an Ethernet frame length of 1518 bytes?

Solution: Since there are 8 bits per byte, this frame has 12,144 bits. From Eq. (14.27) we obtain

$$P_{\text{frame}} = P_e k = (10^{-12}) \times 12,144 = 1.2144 \times 10^{-8}$$

14.8.2 Optical Signal-to-Noise Ratio (OSNR) Estimation

Measuring the SNR and its associated BER is straightforward for unamplified single-wavelength links. However, in multispan optically amplified DWDM networks, the system performance is limited by the
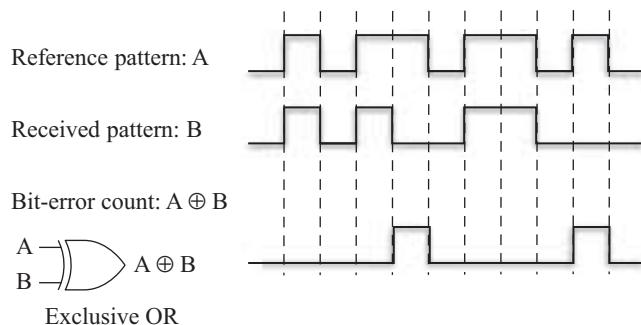
Fig. 14.33 *Basic error detection method using an exclusive OR gate*

optical signal-to-noise ratio (OSNR) rather than by the optical signal power that arrives at the receiver. Although one could demultiplex the incoming DWDM traffic and then do BER estimates on each individual wavelength channel, an optical spectrum measurement can be performed with an *optical spectrum analyzer* (OSA) to derive the OSNR for each individual channel. The OSNR derived from the optical spectrum is an average-power, low-speed measurement, so it does not give information about the effects of temporal impairments on channel performance. However, since it can be correlated to the BER, the OSNR provides indirect BER information for preliminary performance diagnosis of a multichannel system or for giving advance warning of a possible BER degradation on a given DWDM channel.

As described in Sec. 11.5, the OSNR is given by

$$\text{OSNR} = \frac{P_{\text{ave}}}{P_{\text{ASE}}} \tag{14.28}$$

or, in decibels,

$$\text{OSNR(dB)} = 10 \log \frac{P_{\text{ave}}}{P_{\text{ASE}}} \tag{14.29}$$

The OSNR does not depend on factors such as the data format, pulse shape, or optical filter bandwidth, but only on the average optical signal power P_{ave} measured by the OSA and on the average ASE noise power P_{ASE}. OSNR is a metric that can be used for performance verification in the design and installation of networks, as well as to check the health and status of individual optical channels. Sometimes an optical filter is used to significantly reduce the total ASE noise seen by the receiver. Typically such a filter has an optical bandwidth that is large compared to the signal, so that it does not affect the signal, yet that is narrow compared to the bandwidth associated with the ASE background. The ASE noise filter does not change the OSNR. However, it reduces the total power in the ASE noise to avoid overloading the receiver front end.

Example 14.8   Consider an optical signal level of −15 dBm (32 μW) arriving at a *pin* optical receiver in a 10-Gb/s link. If the noise power density is −34.5 dBm (0.35 μW), what is the OSNR?

Solution: From Eq. (14.28) we obtain

$$\text{OSNR} = 32/0.35 = 91$$

or, in decibels,

$$\text{OSNR(dB)} = 10 \log 91 = 19.6 \text{ dB}$$

The IEC Standard 61280-2-9 defines the OSNR as the ratio of the signal power at the peak of a channel to the noise power interpolated at the position of the peak. This document defines OSNR by the following equation

$$\text{OSNR} = 10 \log \frac{P_i}{N_i} + 10 \log \frac{B_m}{B_o} \tag{14.30}$$

where
  - P_i is the optical signal power in watts in the ith channel,
  - N_i is the interpolated value of the noise power in watts measured in the resolution bandwidth B_m at the midchannel spacing point,

- B_m is the resolution bandwidth of the measurement,
- B_o is the reference optical bandwidth, which typically is chosen to be 0.1 nm.

The second term in Eq. (14.30) is used to give an OSNR value that is independent of the resolution bandwidth B_m of the instrument. This allows a comparison of OSNR results that may have been obtained with different OSA instruments. The IEC 61280-2-9 standard also notes that in order to achieve an adequate OSNR measurement, the wavelength measurement range of the OSA must be wide enough to include all DWDM channels plus one-half of an ITU-T grid spacing at each end of the spectral range. In addition, the resolution bandwidth must be wide enough to include the entire signal power spectrum of each modulated channel because this has a direct impact on the accuracy of the noise measurement.

14.8.3   Q Factor Estimation

As noted by Eq. (7.13), the probability of error P_e in a digital communication link is related to the Q factor through the expression

$$P_e = \text{BER} = \frac{1}{2}\text{erfc}\left(\frac{Q}{\sqrt{2}}\right) = \frac{1}{2}\left[1 - \text{erf}\left(\frac{Q}{\sqrt{2}}\right)\right] \approx \frac{1}{\sqrt{2\pi}}\frac{1}{Q}\exp(-Q^2/2) \qquad (14.31)$$

Recall from Eq. (7.14) that Q is proportional to the power difference between the logic 1 and 0 levels. Thus a simple way to examine the error probability versus the Q factor is to vary Q by changing the optical power level at the receiver. For a clear eye diagram, the decision threshold is midway between the 0 and 1 levels, and the noise variance contributed by the receiver remains constant as the input power is varied. These conditions generally hold for *pin* optical receivers where thermal noise in a transimpedance amplifier is the dominant noise.

A different approach is needed when the eye pattern is distorted. In this case, the eye-pattern mask technique described in Sec. 14.5.1 can be used to estimate system performance. This is particularly useful in multispan optically amplified DWDM networks in which the system performance is limited by the OSNR. When optical amplifiers are used in a transmission link, the optical power level at the receiver usually is high enough so that thermal noise and dark current noise can be neglected compared to the signal-ASE noise and the ASE-ASE beat noise. For a distorted eye pattern, the Q factor can be expressed as[61]

$$Q = \frac{2\mathscr{R}(A - B)P_{\text{ave}}}{\sqrt{(G_1 A + G_2)B_e} + \sqrt{(G_1 B + G_2)B_e}} \qquad (14.32)$$

where $G_1 = 4\mathscr{R}(q + \mathscr{R}S_{\text{ASE}})P_{\text{AVE}}$ and $G_2 = S^2_{\text{ASE}}\mathscr{R}^2(2B_o - B_e)$. The dimensionless parameters A and B are the upper and lower eye mask boundaries, respectively, as illustrated in Fig. 14.16. Furthermore, $P_{\text{ave}} = (P_1 + P_2)/2$, $\mathscr{R}$ is the responsivity, S_{ASE} is the power spectral density of the ASE noise (see Sec. 11.4), B_e is the receiver electrical bandwidth, and B_o is the optical bandwidth, which usually is taken to be 0.1 nm.

For OSNR values greater than about 15 dB, the signal-ASE noise is the dominating noise factor, so that the contributions from ASE-ASE beat noise and shot noise can be neglected. In this case, the Q factor for a distorted eye pattern can be expressed by the simplified relationship[61]

$$Q = \frac{(\sqrt{A} - \sqrt{B})\sqrt{P_{ave}}}{\sqrt{S_{ASE}B_e}} = \frac{(\sqrt{A} - \sqrt{B})}{\sqrt{B_e}}\sqrt{OSNR} \qquad (14.33)$$

where the relationship between P_{ave} and OSNR is given by Eqs. (11.24) and (11.36).

Example 14.9 Consider an amplified transmission link for which the receiver Q factor expression for a distorted eye pattern is given by Eq. (14.33) when OSNR > 15 dB. For an OSNR value of 16, compare the values of the Q factor when (*a*) there is no eye closure penalty, that is, $A = 1$ and $B = 0$ and (*b*) with an eye closure penalty that has $A = 0.81$ and $B = 0.25$.

Solution: From Eq. (14.33) we obtain

(*a*) $Q = 1 \times \dfrac{4}{\sqrt{B_e}}$

(*b*) $Q = (0.9 - 0.5) \times \dfrac{4}{\sqrt{B_e}} = 0.4 \times \dfrac{4}{\sqrt{B_e}}$

An experimental setup for measuring OSNR in relation to a bit-error rate is shown in Fig. 14.34. Here the electrical output signal from a pattern generator is used as the input to an optical transmitter. After passing through an optical fiber, an optical coupler (OC) is used to combine the optical signal with additional optical noise from an ASE noise source in order to simulate the effects of ASE noise degradation in an amplified link. The variable optical attenuator (VOA) can be adjusted to simulate different amounts of added ASE noise. The optical bandpass filter (BPF) limits the bandwidth of the ASE noise, thereby limiting the amount of ASE-ASE beat noise. After an erbium-doped fiber amplifier (EDFA) amplifies the combined signal and ASE noise, an optical coupler sends part of the EDFA output to an optical spectrum analyzer (OSA), which monitors the OSNR. The output from the optical coupler is directed to a receiver, which recovers the digital signal. This signal is sent back to the BER test set, which determines the Q factor. This setup thus gives a correlation between the OSNR and the Q factor of the system.
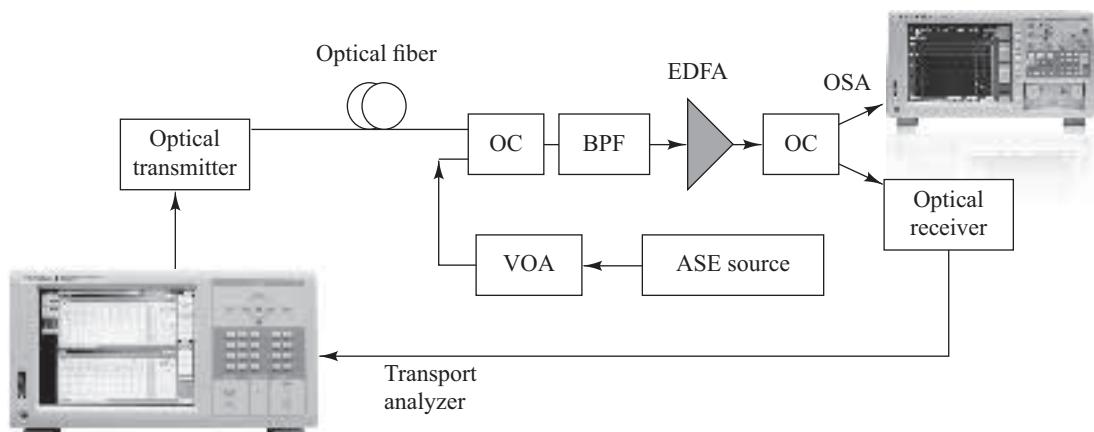
Fig. 14.34 *Setup for measuring receiver Q factor as a function of OSNR. (Photos courtesy of Yokogawa: Transport Analyzer and Optical Spectrum Analyzer; http://tmi.yokogawa.com.)*
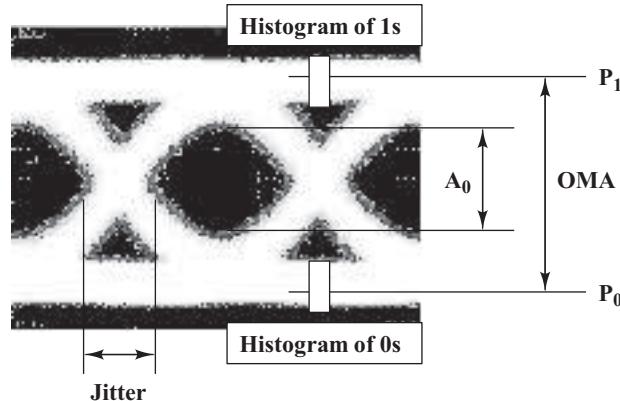
Fig. 14.35 *Definition of measurement parameters for analyzing a stressed eye*

14.8.4 Optical Modulation Amplitude (OMA) Measurement

Compared to long-haul lightwave networks, a different approach must be taken when testing optical Ethernet links that are based on the 10-Gb/s IEEE 802.3ae standard. One reason is that the lasers used for long-haul links usually are high-quality devices that operate with high extinction ratios. Therefore, in the long-haul case, to characterize the sensitivity of an optical receiver to input signals, engineers simply can use a slow-responding power meter to measure the average optical signal power and hence determine the BER.

In contrast, the need to reduce cost in optical Ethernet links means that one needs to use less expensive lasers that have lower extinction ratios but provide adequate 10-Gb/s performance in metro, access, and campus networks. In this case, the low extinction ratio will result in the stressed (partially closed) eye shown in Fig. 14.35, and an average optical power measurement does not give a good indication of receiver performance. Consequently, the traditional tests used to characterize long-haul receivers need to be modified for 10-Gb/s optical Ethernet links. This requirement led to the concept of using an *optical modulation amplitude* (OMA) method.

The parameters for measuring the stressed eye are shown in Fig. 14.35. To measure the OMA, a transmitter puts out a repetitive square-wave pattern of typically five 1s and five 0s (…11111000001111 100000…). The three key parameters derived from this pattern at the receiver are

- A logic 1 amplitude P_1, which is taken from a histogram mean across the middle bit of the run of 1s. Only the middle bit of each sequence of 1 bits is selected so that the measured data points are far from any bit edge.
- A logic 0 amplitude P_0, which is taken from a histogram mean across the middle bit of the run of 0s when using the above repetitive square-wave pattern.
- A_0 is the height of the eye opening.

The *optical modulation amplitude* is defined as the difference between the high and low power levels:

$$OMA = P_1 - P_0 \tag{14.34}$$

The metric spelled out in the IEEE 802.3ae standard for characterizing optical receivers is the *vertical eye closure penalty* (VECP). The VECP measures the vertical opening at the center 1 percent of the eye, which is given by the parameter A_0, and compares that value to the measured OMA value. Thus in decibels,

Table 14.7   *Some IEEE 802.3ae receiver test requirements*

10G Ethernet type	10G-Base-S	10G-Base-L	10G-Base-E
Extinction ratio (dB)	3.0	3.5	3.0
VECP (dB)	3.5	2.2	2.7

$$VECP = 10 \, \log\left(\frac{OMA}{A_0}\right) \tag{14.35}$$

Table 14.7 lists the stressed receiver extinction ratio and VECP requirements for short-, long-, and extended-reach 10-Gb/s Ethernet receivers, which are designated by 10G-Base-S, 10G-Base-L, and 10G-Base-E, respectively. More extensive details on additional test parameters and methods are given in the IEEE802.3ae standard.

14.8.5   Timing Jitter Measurement

In digital communication systems, *time jitter*, also called *timing jitter* or simply *jitter*, is defined as an instantaneous unintentional deviation in the ideal timing between binary symbols. Basically jitter occurs when the transition from one symbol state to the next state occurs earlier or later than the exact end of the bit time interval. Many different factors can contribute to jitter, including random amplitude and noise variations in a signal, periodic noise from switching power supplies, and charge storage mechanisms in circuits and photonic components.

Timing jitter is an especially important issue for high-speed optical fiber transmission systems because pulses are spaced very close together. In this case, incorrect interpretation of the edges of bit periods can lead to high bit-error rates. In digital transmission systems timing jitter can be random or deterministic. Random jitter is caused by noises such as thermal and shot noises in the receiver, and from ASE noise accumulated throughout the transmission link. Deterministic jitter arises from pattern distorting effects due to factors such as chromatic dispersion, self-phase modulation, and interchannel crosstalk.

For a bit sequence with a data rate *B*, a jittered waveform can be expressed as

$$P_{jitter}(t) = P\left[t + \frac{\Delta\varphi(t)}{2\pi B}\right] = P[t + \Delta t(t)] \tag{14.36}$$

Here $\Delta\varphi(t)$ is the phase variation introduced by time jitter, which can be designated in degrees or radians, and $P(t)$ is the waveform in the absence of time jitter. The time deviation Δt, which is given by,

$$\Delta t(t) = \frac{\Delta\varphi(T)}{2\pi B} \tag{14.37}$$

can be expressed in a convenient measure called the *unit interval* (UI). This parameter is the ratio between the time jitter and the bit period $T = 1/B$, and is given by

$$\Delta t_{UI} = \frac{\Delta\varphi(T)}{2\pi} \tag{14.38}$$

A number of different techniques can be used to measure jitter, including the use of BER testers, sampling oscilloscopes, and jitter detectors.[60,61] Figure 14.36 shows one example using a commercial network performance analyzer. Such instruments have built-in high-accuracy jitter measurement capabilities that satisfy the jitter test conditions specified in the ITU-T O.172 Recommendation.[62]
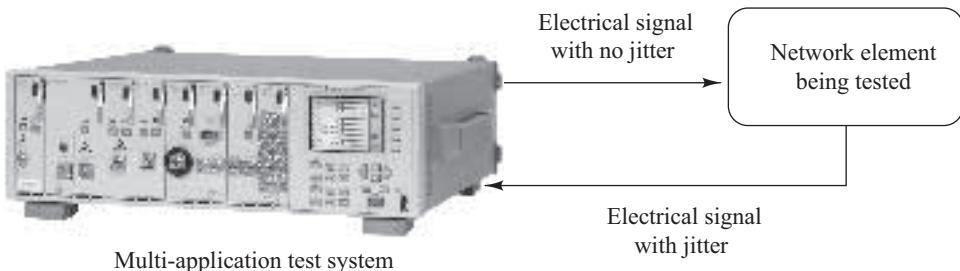
Fig. 14.36 *Setup for measuring jitter with a commercial network performance tester. (Photo courtesy of Yokogawa: http://tmi.yokogawa.com.)*

PROBLEMS

14.1 Optical spectrum analyzers directly measure the wavelength of light in an air environment. However, most wavelength measurements are quoted in terms of wavelengths or optical frequencies in a vacuum. This can lead to errors, particularly in DWDM systems, since the index of refraction of air is a function of wavelength, temperature, pressure, and gas composition. The wavelength dependence of the index of refraction n_{air} of standard dry air at 760 torr and 15°C is[64]

$$n_{air} = 1 + 10^{-8}\left(8342.13 + \frac{2406030}{130 - \frac{1}{\lambda^2}} + \frac{15997}{38.9 - \frac{1}{\lambda^2}}\right)$$

where λ is measured in micrometers.

(a) Given that $\lambda_{vacuum} = \lambda_{air} n_{air}$, what is the error in wavelength measurement at 1550 nm if the effect of the index of refraction of air is ignored? What impact would this have on 0.8-nm-spaced WDM channels in the 1550-nm window?

(b) To compensate for temperature and pressure effects on the value of n_{air}, one can use the relationship

$$n(T,P) = 1 + \frac{(n_{air} - 1)(0.00138823P)}{1 + 0.003671T}$$

where P is measured in torr and T is in Celsius. How much does $n(T, P)$ vary from n_{air} when the pressure is 640 torr and the temperature is 0°C (which would be at a higher elevation and a lower temperature)?

14.2 An engineer wants to find the attenuation at 1310 nm of an 1895-m long fiber. The only available instrument is a photodetector, which gives an output reading in volts. Using this device in a cutback-attenuation setup, the engineer measures an output of 3.31 V from the photodiode at the far end of the fiber. After cutting the fiber 2 m from the source, the output voltage from the photodetector now reads 3.78 V. What is the attenuation of the fiber in dB/km?

14.3 Consider the cutback attenuation measurement technique described by Eq. (14.2) using a photodetector, the power measurements are proportional to the detector output voltage. If the uncertainty in the voltage readings for the two power measurements are ±0.1 percent

each, what is the uncertainty in the attenuation accuracy? How long must the fiber be to get a sensitivity better than ±0.05 dB/km?

14.4 (a) Verify that the full-width at half-maximum of a gaussian pulse is given by $t_{FWHM} = 2\sigma(2 \ln 2)^{1/2}$.

(b) Derive Eq. (14.11), which describes the 3-dB bandwidth of a fiber based on a gaussian output response.

14.5 A gaussian approximation of $|H(f)|$ in the form

$$|H(f)| \approx \exp[-(2\pi f\sigma)^2/2]$$

has been found to be accurate to at least the 0.75-amplitude point in the frequency domain. Using this relationship, plot $P(f)/P(t)$ as a function of frequency from 0 to 1000 MHz for fibers that have impulse responses of full rms pulse widths 2σ equal to 2.0, 1.0, and 0.5 ns. What are the 3-dB bandwidths of these fibers?

14.6 Consider the data shown in Fig. 14.37 of the group delay versus wavelength for a 10-km long fiber. From this data, plot the chromatic dispersion D as a function of wavelength. What is the value of the zero-dispersion slope S_0 in the relationship $D(\lambda) = S_0(\lambda - \lambda_0)$?
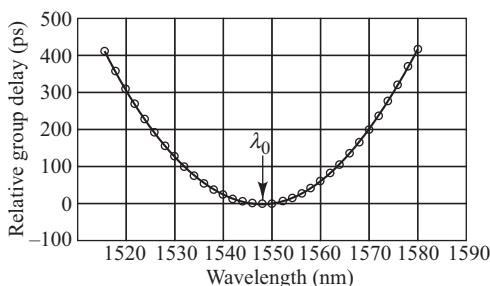
Fig. 14.37 *Chromatic-dispersion measurement of a 10-km fiber*

14.7 Determine the value of the expected differential group delay from the polarization-mode-dispersion measurement response shown in Fig. 14.38 for a non-mode-coupled fiber.
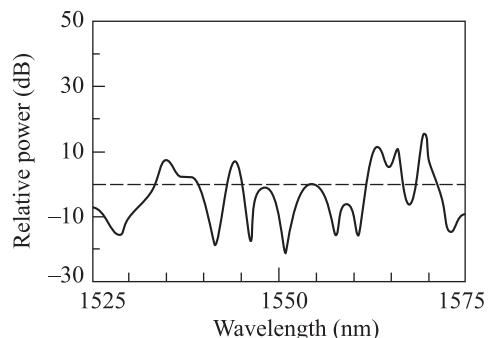
Fig. 14.38 *PMD-induced power excursions about the mean power level (dashed line)*

14.8 The influence of polarization-mode dispersion can be neglected for data rates up to a few Gb/s for intensity-modulated direct-detection systems. However, for higher rates, PMD can cause intersymbol interference (ISI) in long-span links. The ISI power penalty in decibels for PMD is approximately[65]

$$P_{ISI} \approx 26 \frac{\langle \Delta\tau \rangle^2 \gamma(1-\gamma)}{T^2}$$

where T is a bit period (1/bit rate) and γ is the power-splitting ratio between principal polarization states. The maximum power penalty occurs when $\gamma = \frac{1}{2}$. If typical values of the expected value of the differential group delay are 1 ps over a 100-km link and 10 ps over a 1000-km link, find the maximum PMD power penalty for data rates of 10 and 100 Gb/s over these two link distances.

14.9 The optical power in a fiber at a distance x from the input end is given by Eq. (14.18). By assuming that the loss coefficient is uniform along the fiber, use this equation to derive Eq. (14.2).

14.10 Assuming that Rayleight scattering is approximately isotropic (uniform in all directions), show that the fraction S of scattered light trapped in a multimode fiber in the backward direction is given by

$$S \approx \frac{\pi(\text{NA})^2}{4\pi n^2} = \frac{1}{4}\left(\frac{\text{NA}}{n}\right)^2$$

where NA is the fiber numerical aperture, n is the core refractive index, and NA/n represents the half-angle of the cone of captured rays. If NA = 0.20 and n = 1.50, what fraction of the scattered light is recaptured by the fiber in the reverse direction.

14.11 Three 5-km-long fibers have been spliced together in series and an OTDR is used to measure the attenuation of the resultant fiber. The reduced data of the OTDR display is shown in Fig. 14.39. What are the attenuations in decibels per kilometer of the three individual fibers? What are the splice losses in decibels? What are some possible reasons for the large splice loss occurring between the second and third fibers?

14.12 Let α be the attenuation of the forward-propagating light, α_s the attenuation of the backscattered light, and S the fraction of the total output power scattered in the backward direction, as described in Eq. (14.20). Show that the backscatter response of a rectangular pulse of width W from a point a distance L down the fiber is

$$P_s(L) = S\frac{\alpha_s}{\alpha}P_0\,e^{-2\alpha L}(1 - e^{-\alpha W})$$

when $L \geq W/2$, and

$$P_s(L) = S\frac{\alpha_s}{\alpha}P_0\,e^{-\alpha W}(1 - e^{-2\alpha L})$$

for $0 \leq L \leq W/2$
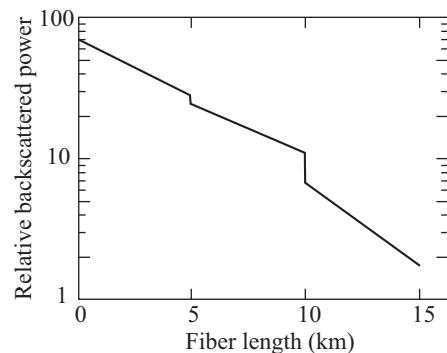
Fig. 14.39 *An OTDR trace of three 5-km spilced fibers*

14.13 Using the expression given in Prob. 14.12 for the backscattered power $P_S(L)$ from a rectangular pulse of width W, show that for very short pulse widths the backscattered power is proportional to the pulse duration. Note: This is the basis of operation of an OTDR.

14.14 The uncertainty U of OTDR loss measurements as a function of the signal-to-noise ratio SNR can be approximated by[36]

$$\log|U| = -0.2\text{SNR} + 0.6$$

Here, U and SNR are given in decibels. If a 0.5-dB splice is located near the far end of a 50-km fiber, what dynamic range must the OTDR have to measure the insertion loss of this splice event with a ±0.05-dB accuracy. Assume the fiber attenuation is 0.33 dB/km.

14.15 Show that, when using an OTDR, an optical pulse width of 5 ns or less is required to locate a fiber fault to within ±0.5 m of its true position.

REFERENCES

1. D. L. Philen and W. T. Anderson, "Optical fiber transmission evaluation," in *Optical Fiber Telecommunications-II*, S. E. Miller and I. P. Kaminow, eds., Academic, New York, 1988.
2. K. J. Gasvik, *Optical Metrology*, Wiley, Hoboken, NJ, 3rd. ed., 2003.
3. D. Derickson, ed., *Fiber Optic Test and Measurement*, Prentice Hall, Upper Saddle River, NJ, 1998.
4. D. Marcuse, *Principles of Optical Fiber Measurements*, Academic, New York, 1981.
5. C. Dorrer, "High-speed measurements for optical telecommunication systems," *IEEE*

J. Sel. Topics Quantum Electronics, vol. 12, pp. 843–858, July/Aug. 2006.

6. *Technical Digest — Symposium on Optical Fiber Measurements (SOFM)*, biennial conference starting in 1980 sponsored by National Institute of Standards and Tech. (NIST).

7. National Institute of Standards and Technology (NIST), Boulder, CO, USA (http://www.nist.gov).

8. National Physical Laboratory (NPL), Teddington, UK (http://www.npl.co.uk).

9. Physikalisch-Technische Bundesanstalt (PTB), Braunschweig, Germany (http://www.ptb.de).

10. ITU-T Recommendation G.650.2, *Definitions and Test Methods for Statistical and Nonlinear Related Attributes of Single-Mode Fibre and Cable*, Jan. 2005.

11. N. Gisin, R. Passy, and B. Perny, "Optical fiber characterization by simultaneous measurement of the transmitted and refracted near field," *J. Lightwave Tech.,* vol. 11, pp. 1875–1883, Nov. 1993.

12. TIA/EIA-455-191 (FOTP-191), *Measurement of the Mode-Field Diameter of Single-Mode Optical Fiber*, 1998.

13. TIA/EIA-455-44B (FOTP-44), *Refractive Index profile, Refracted Ray Method*, 1992.

14. K. Petermann, "Constraints for fundamental mode spot size for broadband dispersion-compensated single-mode fibers," *Electron. Lett.*, vol. 19, pp. 712–714, Sept. 1983.

15. C. Pask, "Physical interpretation of Petermann's strange spot size for single-mode fibers," *Electron. Lett.*, vol. 20, pp. 144, 1984.

16. ITU-T Recommendation G.650.1, *Definitions and Test Methods for Linear, Deterministic Attributes of Single-Mode Fibre and Cable*, June 2004.

17. ANSI/TIA-455-78-B-2002 (FOTP-78), *Measurement Methods and Test Procedures for Attenuation*, Nov. 2002.

18. S. D. Personick, "Receiver design for digital fiber optic communication systems," *Bell Sys. Tech. J.*, vol. 52, pp. 843–874, July-Aug. 1973.

19. P. Hernday, "Dispersion measurements," in D. Derickson, ed., *Fiber Optic Test and Measurement*, Prentice Hall, Upper Saddle River, NJ, 1998.

20. ANSI/TIA-455-175-B-2003 (FOTP-175), *Measurement Methods and Test Procedures— Chromatic Dispersion*, May 2003.

21. J.-N. Maran, R. Slavik, S. LaRochelle, and M. Karasek, "Chromatic dispersion measurement using a multiwavelength frequency-shifted feedback fiber laser," *IEEE Trans. Instrument. Measurement*, vol. 53, pp. 67–71, Feb. 2004.

22. S. G. Murdoch and D. A. Svendsen, "Distributed measurement of the chromatic dispersion of an optical fiber using a wavelength-tunable OTDR," *J. Lightwave Tech.,* vol 24, pp. 1681–1688, Apr. 2006.

23. P. M. Nellen, R. Brönnimann, M. Held, and U. Sennhauser, "Long-term monitoring of polarization-mode dispersion of aerial optical cables with respect to line availability." *J. Lightwave Tech.,* vol. 22, pp. 1848–1855, Aug. 2004.

24. K. E. Cornick, M. Boroditsky, S. Finch, S. D. Dods, and P. M. Farrell, "Experimental comparison of PMD-induced system penalty models," *IEEE Photonics Technol. Lett.*, vol. 18, pp. 1149–1151, May 2006.

25. S. V. Shatalin and A. J. Rogers, "Location of high PMD sections of installed system fiber," *J. Lightwave Tech.,* vol. 24, pp. 3875–3881, Nov. 2006.

26. P. Poggiolini, A. Nespola, S. Abrate, V. Ferrero, and C. Lezzi, "Long-term PMD characterization of a metropolitan G.652 fiber plant," *J. Lightwave Tech.,* vol. 24, pp. 4022–4029, Nov. 2006.

27. H. H. Yaffe and D. L. Peterson Jr., "Experimental determination of system outage probability due to first-order and second-order PMD," *J. Lightwave Tech.,* vol. 24, pp. 4155–4161, Nov. 2006.

28. C. D. Poole and D. L. Favin, "Polarization-mode dispersion measurements based on transmission spectra through a polarizer," *J. Lightwave Tech.*, vol. 12, pp. 917–929, June 1994.

29. S. W. Hinch and C. M. Miller, "Analysis of digital modulation on optical carriers," in D. Derickson, ed., *Fiber Optic Test and Measurement*, Prentice Hall, Upper Saddle River, NJ, 1998.

30. S. Haykin and M. Moher, *An Introduction to Digital and Analog Communications*, Wiley, Hoboken, NJ, 2nd ed. 2006.

31. N. S. Bergano, F. W. Kerfoot. and C. R. Davidson, "Margin measurements in optical amplifier systems," *IEEE Photonics Technol. Lett.*, vol. 5, pp. 304–306, Mar. 1993.

32. ITU-T Recommendation O.201. *Q-factor Test Equipment to Estimate the Transmission Performance of Optical Channels*, July 2003.

33. I. Shake, H. Takara, and S. Kawanishi, "Simple measurement of eye diagram and BER using high-speed asynchronous sampling," *J. Lightwave Tech.*, vol. 22, pp. 1296–1302, Jan. 2004.

34. C. Fernando, J. Coelho, and R. Wong, "Parametric representation of custom masks for compliance tests of high-speed optical transmitters," *Optical Fiber Commun. Conf. 2002 (OFC 2002)*, pp. 163–165, Mar. 2002.

35. G. Foster, "Bridging the gap between BER and eye diagrams: A BER contour tutorial," *Technical Note SR-TN032*, SyntheSys Research, Oct. 2005.

36. J. Beller, "OTDRs and backscatter measurements," in D. Derickson, ed., *Fiber Optic Test and Measurement*, Prentice Hall, Upper Saddle River, NJ, 1998.

37. F. Caviglia, V. C. Di Biase, and A. Gnazzo, "Optical maintenance in PONs," *Optical Fiber Tech.*, vol. 5, pp. 349–362, Oct. 1999.

38. C. Mas and P. Thiran, "A review on fault location methods and their application to optical networks," *Optical Networks Magazine*, vol. 2, pp. 73–87, July/Aug. 2001.

39. A. Girard, *FTTx PON Tech. and Testing*, EXPO, Quebec, Canada, 2006.

40. D. Anderson, L. Johnson, and F. G. Bell, *Troubleshooting Optical Fiber Networks: Understanding and Using Optical Time-Domain Reflectometers*, Academic, New York, 2004.

41. E. Park, "Error monitoring for optical metropolitan network services," *IEEE Commun. Mag.*, vol. 40, pp. 104–109, Feb. 2002.

42. ITU-T Recommendation G.697, *Optical Monitoring for DWDM Systems*, Jun. 2004.

43. D. C. Kilper and W. Weingartner, "Monitoring optical network performance degradation due to amplifier noise," *J. Lightwave Tech.*, vol. 21, pp. 1171–1178, May 2003.

44. S. Wielandy, M. Fishteyn, and B. Zhu, "Optical performance monitoring using nonlinear detection," *J. Lightwave Tech.*, vol. 22, pp. 784–793, Mar. 2004.

45. Z. Li, C. Lu, Y. Wang, and G. Li, "In-service signal quality monitoring and multi-impairment discrimination based on asynchronous amplitude-histogram evaluation for NRZ-DPSK systems," *IEEE Photonics Technol. Lett.*, vol. 17, pp. 1998–2000, Sept. 2005.

46. T. Luo, C. Yu, Z. Pan, Y. Wang, Y. Arieli, and A. E. Willner, "Dispersive effects monitoring for RZ data by adding a frequency-shifted carrier along the orthogonal polarization state," *J. Lightwave Tech.*, vol. 23, pp. 3295–3301, Oct. 2005.

47. M. Dinu, D. C. Kilper, and H. R. Stuart, "Optical performance monitoring using data stream intensity autocorrelation," *J. Lightwave Tech.*, vol. 24, pp. 1194–1202, Mar. 2006.

48. C. Pinart and G. Junyent, "The INIM system: In-service non-intrusive monitoring for QoS-enabled transparent WDM," *IEEE J. Sel. Topics Quantum Electron.*, vol. 12, no. 4, pp. 635–644, July/Aug. 2006.

49. S. B. Jun, H. Kim, P.K.J. Park, J. H. Lee, and Y. C. Chung, "Pilot-tone-based WDM monitoring technique for DPSK systems," *IEEE Photonics Technol. Lett.*, vol. 18, pp. 2171–2173, Oct. 2006.

50. Z. Li and G. Li, "Chromatic dispersion and polarization-mode dispersion monitoring for RZ-DPSK signals based on asynchronous amplitude-histogram evaluation," *J. Lightwave Tech.*, vol. 24, pp. 2859–2866, July 2006.

51. G.-W. Lu, M.-H. Cheung, L.-K. Chen, and C.-K. Chan, "Simultaneous PMD and OSNR monitoring by enhanced RF spectral dip analysis assisted with a local large-DGD element," *IEEE Photonics Technol. Lett.*, vol. 17, pp. 2790–2792, Dec. 2005.

52. X. Yi, W. Chen, and W. Shieh, "An OSNR monitor for optical packet switched networks," *IEEE Photonics Technol. Lett.*, vol. 18, pp. 1448–1450, July 2006.

53. J. H. Lee, H. Y. Choi, S. K. Shin, and Y. C. Chung, "A review of the polarization-nulling technique for monitoring optical-signal-to-noise ratio in dynamic WDM networks," *J. Lightwave Tech.*, vol. 24, pp. 4162–4171, Nov. 2006.

54. C.-L. Yang and S.-L. Lee, "OSNR monitoring using double-pass filtering and dithered tunable reflector," *IEEE Photonics Technol. Lett.*, vol. 16, pp. 1570–1572, June 2004.

55. C.-L. Yang, S.-L. Lee, H.-W. Tsao, and J. Wu, "Simultaneous channel and OSNR monitoring using polarization-selective modulator and a LED," *IEEE Photonics Technol. Lett.*, vol. 16, pp. 945–947, Mar. 2004.

56. A. E. Willner, Z. Pan, and C. Yu, "Optical performance monitoring," in *Optical Fiber Telecommunications V-B*, eds. I. P. Kaminow, T. Li, and A. E. Willner, chap. 7, Academic Press New York, 2008.

57. N. Skorin-Kapov, O. K. Tonguz, and N. Puech, "Toward efficient failure management for reliable transparent optical networks," *IEEE Commun. Mag.*, vol. 47, pp. 72–79, May 2009.

58. T. B. Anderson, A. Kowalczyk, K. Clarke, S. D. Dods, D. Hewitt, and J. C. Li, "Multi impairment monitoring for optical networks," *J. Lightwave Tech.*, vol. 27, no. 16, pp. 3729–3736, Aug. 2009.

59. Z. Pan, C. Yu, and A. E. Willner, "Optical performance monitoring for the next generation optical communication networks," *Optical Fiber Tech.*, vol. 16, no. 1, pp. 20–45, Jan. 2010.

60. D. Derickson and M. Müller, *Digital Communications Test and Measurement*, Prentice Hall, Upper Saddle River, NJ, 2008.

61. R. Hui and M. O'Sullivan, *Fiber Optic Measurement Techniques*, Elsevier Academic Press, Burlington, MA, 2009.

62. ITU-T Recommendation O.172. *Jitter and Wander Measuring Equipment for Digital Systems which Are Based on the Synchronous Digital Hierarchy (SDH)*, Nov. 2002.

63. C.C.K. Chan, *Optical Performance Monitoring: Advanced Techniques for Next-Generation Photonic Networks*, Academic, New York, 2010.

64. J. Vobis and D. Derickson, "Optical spectrum analysis," in D. Derickson, ed., *Fiber Optic Test and Measurement*, Prentice Hall, Upper Saddle River, NJ, 1998.

65. M. Movassaghi, M. K. Jackson, V. M. Smith, and W. J. Hallam, "Noise figure of EDFAs in saturated operation," *J. Lightwave Tech.*, vol. 16, pp. 812–818, May 1988.

APPENDIX A

International System of Units

Quantity	Unit	Symbol	Dimensions
Length	meter	m	
Mass	kilogram	kg	
Time	second	s	
Temperature	kelvin	K	
Current	ampere	A	
Frequency	hertz	Hz	$1/s$
Force	newton	N	$(kg \cdot m)/s^2$
Pressure	pascal	Pa	N/m^2
Energy	joule	J	$N \cdot m$
Power	watt	W	J/s
Electric charge	coulomb	C	$A \cdot s$
Potential	volt	V	J/C
Conductance	siemens	S	A/V
Resistance	ohm	Ω	V/A
Capacitance	farad	F	C/V
Magnetic flux	weber	Wb	$V \cdot s$
Magnetic induction	tesla	T	Wb/m^2
Inductance	henry	H	Wb/A

APPENDIX B

Useful Mathematical Relations

Some of the mathematical relations ecountered in this text are listed for convenient reference. More comprehensive listings are available in various handbooks.[1–5]

B.1 Trigonometric Identities

$$e^{\pm j\theta} = \cos\theta \pm j\sin\theta$$

$$\sin^2\theta + \cos^2\theta = 1$$

$$\cos^2\theta - \sin^2\theta = \cos 2\theta$$

$$4\sin^3\theta = 3\sin\theta - \sin 3\theta$$

$$4\cos^3\theta = 3\cos\theta + \cos 3\theta$$

$$8\sin^4\theta = 3 - 4\cos 2\theta + \cos 4\theta$$

$$8\cos^4\theta = 3 + 4\cos 2\theta + \cos 4\theta$$

$$\sin(\alpha \pm \beta) = \sin\alpha\cos\beta \pm \cos\alpha\sin\beta$$

$$\cos(\alpha \pm \beta) = \cos\alpha\ \cos\beta \mp \sin\alpha\sin\beta$$

$$\tan(\alpha \pm \beta) = \frac{\tan\alpha \pm \tan\beta}{1 \mp \tan\alpha\tan\beta}$$

B.2 Vector Analysis

The symbols $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$ denote units vectors lying parallel to the x, y, and z axes, respectively, of the rectangular coordinate system. Similarly, $\mathbf{e}_r$, $\mathbf{e}_\phi$, and $\mathbf{e}_z$ are unit vectors for cylindrical coordinates. The

unit vectors $\mathbf{e}_r$ and $\mathbf{e}_\phi$ vary in direction as the angle ϕ changes. The conversion from cylindrical to rectangular coordinates is made through the relationships

$$x = r\cos\phi \qquad y = r\sin\phi \qquad z = z$$

B.2.1 Rectangular Coordinates

$$\text{Gradient } \nabla f = \frac{\partial f}{\partial x}\mathbf{e}_x + \frac{\partial f}{\partial y}\mathbf{e}_y + \frac{\partial f}{\partial z}\mathbf{e}_z$$

$$\text{Divergence } \nabla \cdot \mathbf{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$$

$$\text{Curl } \nabla \times \mathbf{A} = \begin{vmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix}$$

$$\text{Laplacian } \nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

B.2.2 Cylindrical Coordinates

$$\text{Gradient } \nabla f = \frac{\partial f}{\partial r}\mathbf{e}_r + \frac{1}{r}\frac{\partial f}{\partial \phi}\mathbf{e}_\phi + \frac{\partial f}{\partial z}\mathbf{e}_z$$

$$\text{Divergence } \nabla \cdot \mathbf{A} = \frac{1}{r}\frac{\partial (rA_r)}{\partial r} + \frac{1}{r}\frac{\partial A_\phi}{\partial \phi} + \frac{\partial A_z}{\partial z}$$

$$\text{Curl } \nabla \times \mathbf{A} = \begin{vmatrix} \dfrac{1}{r}\mathbf{e}_r & \mathbf{e}_\phi & \dfrac{1}{r}\mathbf{e}_z \\ \dfrac{\partial}{\partial r} & \dfrac{\partial}{\partial \phi} & \dfrac{\partial}{\partial z} \\ A_r & rA_\phi & A_z \end{vmatrix}$$

$$\text{Laplacian } \nabla^2 f = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r} \right) + \frac{1}{r^2}\frac{\partial^2 f}{\partial \phi^2} + \frac{\partial^2 f}{\partial z^2}$$

B.2.3 Vector Identities

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$$

$$\nabla^2 \mathbf{A} = \nabla^2 A_x \mathbf{e}_x + \nabla^2 A_y \mathbf{e}_y + \nabla^2 A_z \mathbf{e}_z$$

B.3 Integrals

$$\int \sin x \; dx = -\cos x$$

$$\int \cos x \; dx = \sin x$$

$$\int \sqrt{a^2 - x^2} \, dx = \frac{1}{2}\left(x\sqrt{a^2 - x^2} + a^2 \; \arcsin \frac{x}{a} \right)$$

$$\int x\sqrt{a^2 - x^2} \, dx = -\frac{1}{3}(a^2 - x^2)^{3/2}$$

$$\int x^2 \sin^2 x \; dx = \frac{x^3}{6} - \left(\frac{x^2}{6} - \frac{1}{8} \right) \sin 2x - \frac{x \cos 2x}{4}$$

$$\int \frac{dx}{\cos^n x} = \frac{1}{n-1}\frac{\sin x}{\cos^{n-1} x} + \frac{n-2}{n-1}\int \frac{dx}{\cos^{n-2} x}$$

$$\int u \; dv = uv - \int v \; du$$

$$\int e^{ax} dx = \frac{1}{a} e^{ax}$$

$$\int \sin^2 x \; dx = \frac{x}{2} - \frac{1}{4}\sin 2x$$

$$\int \sin^n x \; dx = -\frac{\sin^{n-1} x \cos x}{n} + \frac{n-1}{n}\int \sin^{n-2} x \; dx$$

$$\int \cos^2 x \; dx = \frac{x}{2} + \frac{1}{4}\sin 2x$$

$$\int \cos^n x \; dx = \frac{1}{n}\cos^{n-1} x \sin x + \frac{n-1}{n}\int \cos^{n-2} x \; dx$$

$$\int_{-\infty}^{\infty} \frac{e^{jpx}}{(\beta + jx)^n} dx = \begin{cases} 0 & \text{if } p < 0 \\ \dfrac{2\pi(p)^{n-1}e^{-\beta p}}{\Gamma(n)} & \text{if } p \geq 0 \end{cases} \quad \text{where } \Gamma(n) = (n-1)!$$

$$\int_{-\infty}^{\infty} e^{-p^2 x^2 + qx} dx = e^{q^2/4p^2} \frac{\sqrt{\pi}}{p}$$

$$\int_{-\infty}^{\infty} \frac{1}{1 + (x/a)^2} dx = \frac{\pi a}{2}$$

$$\frac{2}{\sqrt{\pi}} \int_0^t e^{-x^2} dx = erf(t)$$

B.4 Series Expansions

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2!} x^2 + \frac{n(n-1)(n-2)}{3!} x^3 + \cdots \quad \text{for } |nx| < 1$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots$$

REFERENCES

1. M. Kurtz, *Handbook of Applied Mathematics for Engineers and Scientists*, McGraw-Hill, New York, 1991.

2. D. Zwillinger, ed., *Standard Mathematical Tables and Formulae*, CRC Press, Boca Raton, FL, 31th ed., 2003.

3. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions,* Dover, New York, 10th ed., 1972.

4. I. S. Gradshteyn, A. Jeffrey, I. M. Ryzhik, and D. Zwillinger, *Table of Integrals, Series and Products,* Academic, New York, 7th ed., 2007.

5. A.A.D. Polyanin and A. V. Manzhirov, *Handbook of Mathematics for Engineers and Scientists*, CRC Press, Boca Raton, FL, 2007.

APPENDIX C

Bessel Functions

This appendix lists the definitions and some recurrence relations for integer-order Bessel functions of the first kind $J_v(z)$ and modified Bessel functions $K_v(z)$. Detailed mathematical properties of these and other Bessel functions can be found in Refs. 24 through 26 of Chapter 2. Here, the parameter v is any integer and n is a positive integer or zero. The parameter $z = x + jy$.

C.1 Bessel Functions of the First Kind

C.1.1 Various Definitions

A Bessel function of the first kind of order n and argument z, commonly denoted by $J_n(z)$, is defined by

$$J_n(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{jz \sin \theta - jn\theta} d\theta$$

or, equivalently,

$$J_n(z) = \frac{1}{\pi} \int_0^{\pi} \cos(z \sin \theta - n\theta)\, d\theta$$

Just as the trigonometric functions can be expanded in power series, so can the Bessel functions $J_v(z)$:

$$J_v(z) = \sum_{k=0}^{\infty} \frac{(-1)^k \left(\frac{1}{2} z\right)^{v+2k}}{k!(v+k)!}$$

In particular, for $v = 0$,

$$J_0(z) = 1 - \frac{\frac{1}{4} z^2}{(1!)^2} + \frac{\left(\frac{1}{4} z^2\right)^2}{(2!)^2} - \frac{\left(\frac{1}{4} z^2\right)^3}{(3!)^2} + \cdots$$

For $v = 1$,

$$J_1(z) = \frac{1}{2}z - \frac{\left(\frac{1}{2}z\right)^3}{2!} + \frac{\left(\frac{1}{2}z\right)^5}{2!3!} - \cdots$$

and so on for higher values of v.

C.1.2 Recurrence Relations

$$J_{v-1}(z) + J_{v+1}(z) = \frac{2v}{z}J_v(z)$$

$$J_{v-1}(z) - J_{v+1}(z) = 2J_v'(z)$$

$$J_v'(z) = J_{v-1}(z) - \frac{v}{z}J_v(z)$$

$$J_v'(z) = -J_{v+1}(z) + \frac{v}{z}J_v(z)$$

$$J_0'(z) = -J_1(z)$$

C.2 Modified Bessel Functions

C.2.1 Integral Representations

$$K_0(z) = \frac{-1}{\pi}\int_0^\pi e^{\pm z\cos\theta}[\gamma + \ln(2z\sin^2\theta)]d\theta$$

where Euler's constant $\gamma = 0.57722$.

$$K_v(z) = \frac{\pi^{1/2}\left(\frac{1}{2}z\right)^v}{\Gamma\left(v + \frac{1}{2}\right)}\int_0^\infty e^{-z\cosh t}\sinh^{2v}t\,dt$$

$$K_0(\mathrm{x}) = \int_0^\infty \cos(x\sinh t)\,dt = \int_0^\infty \frac{\cos(xt)}{\sqrt{t^2 + 1}}dt \qquad (x > 0)$$

$$K_v(\mathrm{x}) = \sec\left(\frac{1}{2}v\pi\right)\int_0^\infty \cos(x\sinh t)\cosh(vt)\,dt \quad (x > 0)$$

C.2.2   Recurrence Relations

If $L_v = e^{j\pi v} K_v$, then

$$L_{v-1}(z) - L_{v+1}(z) = \frac{2v}{z} L_v(z)$$

$$L_v'(z) = L_{v-1}(z) - \frac{v}{z} L_v(z)$$

$$L_{v-1}(z) + L_{v+1}(z) = 2 L_v'(z)$$

$$L_v'(z) = L_{v+1}(z) + \frac{v}{z} L_v(z)$$

C.3   Asymptotic Expansions

For fixed v ($\neq -1, -2, -3, \cdots$) and $z \to 0$,

$$J_v(z) \simeq \frac{\left(\frac{1}{2} z\right)^v}{\Gamma(v + 1)}$$

For fixed v and $|z| \to \infty$,

$$J_v(z) \simeq \left(\frac{2}{\pi z}\right)^{1/2} \cos\left(z - \frac{v\pi}{2} - \frac{\pi}{4}\right)$$

For fixed v and large $|z|$,

$$K_v(z) \simeq \left(\frac{\pi}{2z}\right)^{1/2} e^{-z}\left[1 - \frac{\mu - 1}{8z} + \frac{(\mu - 1)(\mu - 9)}{2!(8z)^2} + \cdots\right]$$

where $\mu = 4v^2$.

C.4   Gamma Function

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} \, dt$$

For integer n,

$$\Gamma(z + 1) = n!$$

For fractional values,

$$\Gamma\left(\frac{1}{2}\right) = \pi^{1/2} = \left(-\frac{1}{2}\right)! \simeq 1.77245$$

$$\Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\pi^{1/2} = \left(\frac{1}{2}\right)! \simeq 0.88623$$

APPENDIX D

Decibels

D.1   Definition

In designing and implementing an optical fiber link, it is of interest to establish, measure, and/or inter-relate the signal levels at the transmitter, at the receiver, at the cable connection and splice points, at the input and output of a link component, and in the cable. A convenient method for this is to reference the signal level either to some absolute value or to a noise level. This is normally done in terms of a power ratio measured in *decibels* (dB) defined as

$$\text{Power ratio in dB} = 10 \log \frac{P_2}{P_1} \tag{D.1}$$

where P_1 and P_2 are electric or optical powers.

    The logarithmic nature of the decibel allows a large ratio to be expressed in a fairly simple manner. Power levels differing by many orders of magnitude can be compared easily when they are in decibel form. Some very helpful figures to remember are given in Table D-1. For example, doubling the power means a 3-dB gain (the power level increases by 3 dB), halving the power means a 3-dB loss (the power level decreases by 3 dB), and power levels differing by factors of 10^N or 10^{-N} have decibel differences of $+10N$ dB and $-10N$ dB, respectively.

Table D-1   *Examples of decibel measures of power ratios*

Power ratio	10^N	10	2	1	0.5	0.1	10^{-N}
dB	$+10N$	$+10$	$+3$	0	-3	-10	$-10N$

D.2   The dBm

The decibel is used to refer to ratios or relative units. For example, we can say that a certain optical fiber has a 6-dB loss (the power level gets reduced by 75 percent in going through the fiber) or that a particular connector has a 1-dB loss (the power level gets reduced by 20 percent at the connector). However, the

decibel gives no indication of the absolute power level. One of the most common derived units for doing this in optical fiber communications is the *dBm*. This is the decibel power level referred to 1 mW. In this case, the power in dBm is an absolute value defined by

$$\text{Power level} = 10 \log \frac{P}{1 \text{ mW}} \qquad \text{(D.2)}$$

A useful relationship to remember is that 0 dBm = 1 mW. Negative dBm numbers designate power levels less than 1 mW, whereas positive dBm values indicate power levels greater than 1 mW. Some examples are shown in Table D-2.

Table D-2 *Examples of dBm units (decibel measure of power relative to 1 mW)*

Power (mW)	100	10	2	1	0.5	0.1	0.01	0.001
Value (dBm)	+20	+10	+3	0	−3	−10	−20	−30

D.3 The Neper

The *neper* (N) is an alternative unit that is sometimes used instead of the decibel. If P_1 and P_2 are two power levels, with $P_2 > P_1$, then the power ratio in nepers is given as the natural (or naperian) logarithm of the power ratio:

$$\text{Power ratio in nepers} = \frac{1}{2} \ln \frac{P_2}{P_1} \qquad \text{(D.3)}$$

where

$$\ln e = \ln 2.71828 = 1$$

To convert nepers to decibels, multiply the number of nepers by

$$20 \log e = 8.686$$

APPENDIX E

Acronyms

AES	Advanced encryption standard
AGC	Automatic gain control
AM	Amplitude modulation
ANSI	American National Standards Institute
APC	Angle-polished connector
APD	Avalanche photodiode
ARQ	Automatic repeat request
ASE	Amplified spontaneous emission
ASK	Amplitude shift keying
ATM	Asynchronous transfer mode
AWG	Arrayed waveguide grating
BER	Bit error rate
BH	Buried heterostructure
BLSR	Bidirectional line-switched ring
BPON	Broadband PON
BS	Base station
CAD	Computer-aided design
CATV	Cable TV
CNR	Carrier-to-noise ratio
CO	Central office
CRC	Cyclic redundancy check
CRZ	Chirped return-to-zero
CS	Control station
CSO	Composite second order
CTB	Composite triple beat
CW	Continuous wave
CWDM	Course wavelength division multiplexing
DBA	Dynamic bandwidth assignment
DBR	Distributed Bragg reflector
DCE	Dynamic channel equalizer

DCF   Dispersion compensating fiber
DCM   Dispersion compensating module
DFA   Doped-fiber amplifier
DFB   Distributed feedback (laser)
DGD   Differential group delay
DGE   Dynamic gain equalizer
DI   Delay interferometer
DPSK   Differential phase-shift keying
DQPSK   Differential quadrature phase-shift keying
DR   Dynamic range
DS   Digital system
DSF   Dispersion-shifted fiber
DUT   Device under test
DWDM   Dense wavelength division multiplexing
DXC   Digital cross-connect matrix
EAM   Electro-absorption modulator
EDFA   Erbium-doped fiber amplifier
EDWA   Erbium-doped wave guide amplifier
EFM   Ethernet in the first mile
EH   Hybrid electric-magnetic mode
EHF   Extremely high frequency (30-to-300 GHz)
EIA   Electronics Industries Alliance
EM   Electromagnetic
EMS   Element management system
EO   Electro-optical
EPON   Ethernet PON
ER   Extended reach
FBG   Fiber Bragg grating
FDM   Frequency division multiplexing
FEC   Forward error correction
FM   Frequency modulation
FOTP   Fiber Optic Test Procedure
FP   Fabry-Perot
FSK   Frequency shift keying
FSR   Free spectral range
FTTH   Fiber to the home
FTTP   Fiber to the premises
FTTx   Fiber to the x
FWHM   Full-width half-maximum
FWM   Four-wave mixing
GEM   GPON encapsulation method
GE-PON   Gigabit Ethernet PON
GFF   Gain-flattening filter
GMPLS   Generalized multiprotocol label switching
GPON   Gigabit PON

GR	Generic Requirement
GUI	Graphical user interface
GVD	Group velocity dispersion
HDLC	High-Level Data Link Control
HE	Hybrid magnetic-electric mode
HFC	Hybrid fiber/coax
IEC	International Electrotechnical Commission
IEEE	Institute for Electrical and Electronic Engineers
ILD	Injection laser diode
IM	Intermodulation
IMD	Intermodulation distortion
IM-DD	Intensity-modulated direct-detection
IP	Internet Protocol
ISI	Intersymbol interference
ISO	International Standards Organization
ITU	International Telecommunications Union
ITU-T	Telecommunication Sector of the ITU
LAN	Local area network
LEA	Large effective area
LED	Light-emitting diode
LO	Local oscillator
LP	Linearly polarized
MAC	Media access control
MAN	Metro area network
MCVD	Modified chemical vapor deposition
MDU	Multiple dwelling unit
MEMS	Micro electro-mechanical system
MFD	Mode-field diameter
MHU	Multiple hospitality unit
MIB	Management information base
MPLS	Multiprotocol label switching
MQW	Multiple quantum well
MTU	Multiple tenant unit
MZI	Mach-Zehnder interferometer
MZM	Mach-Zehnder modulator
NA	Numerical aperture
NCF	Nanostructure core fiber
NEP	Noise equivalent power
NF	Noise figure
NIC	Network interface card
NIST	National Institute of Standards and Technology
NMS	Network management system
NPL	National Physical Laboratory
NRZ	Nonreturn-to-zero
NZDSF	Non-zero dispersion-shifted fiber

O/E/O	Optical-to-electrical-to-optical
OADM	Optical add/drop multiplexer
OBS	Optical burst switching
OC	Optical carrier
OC	Optical coupler
ODF	Optical distribution frame
ODU	Optical channel data unit
OFGW	Optical fiber ground wire
OLS	Optical label swapping
OLT	Optical line terminal
OMA	Optical modulation amplitude
OMI	Optical modulation index
OMS	Optical multiplex section
ONT	Optical network terminal
ONU	Optical network unit
OOK	On-off keying
OPM	Optical performance monitor
OPS	Optical pulse suppressor
OPS	Optical packet switching
OPU	Optical channel payload unit
ORL	Optical return loss
OSA	Optical spectrum analyzer
OSI	Open system interconnect
OSNR	Optical signal-to-noise ratio
OST	Optical standards tester
OTDM	Optical time-division multiplexing
OTDR	Optical time domain reflectometer
OTN	Optical transport network
OTS	Optical transport section
OTU	Optical channel transport unit
OVPO	Outside vapor-phase oxidation
OXC	Optical crossconnect
P2P	Point-to-point
PBG	Photonic bandgap fiber
PC	Personal computer
PCE	Power conversion efficiency
PCF	Photonic crystal fiber
PCVD	Plasma-activated chemical vapor deposition
PDF	Probability density function
PDH	Plesiochronous digital hierarchy
PDL	Polarization-dependent loss
PHY	Physical layer
pin	(p-type)-intrinsic-(n-type)
PLL	Phase-locked loop
PM	Phase-modulation

PMD	Polarization mode dispersion
PMMA	Polymethylmethacrylate
POF	Polymer (plastic) optical fiber
POH	Path overhead
PON	Passive optical network
POP	Point of presence
PPP	Point-to-point protocol
PRBS	Pseudorandom binary sequence
PSK	Phase shift keying
PTB	Physikalisch-Technische Bundesanstalt
PVC	Polyvinyl chloride
QCE	Quantum conversion efficiency
QoS	Quality of service
RAPD	Reach-through avalanche photodiode
RC	Resistance-capacitance
RF	Radio-frequency
RFA	Raman fiber amplifier
RIN	Relative intensity noise
RIP	Refractive index profile
rms	Root mean square
ROADM	Reconfigurable OADM
ROF	Radio-over-fiber
RS	Reed-Solomon
RSOA	Reflective semiconductor optical amplifier
RWA	Routing and wavelength assignment
RZ	Return-to-zero
SAM	Separate-absorption-and-multiplication (APD)
SBS	Stimulated Brillouin scattering
SCM	Subcarrier modulation
SDH	Synchronous digital hierarchy
SFDR	Spur-free dynamic range
SFF	Small-form-factor
SFP	Small-form-factor (SFF) pluggable
SHF	Super-high frequency (3-to-30 GHz)
SLED	Superluminescent light emitting diode
SLM	Single longitudinal mode
SNMP	Simple network management protocol
SNR	Signal-to-noise ratio
SOA	Semiconductor optical amplifier
SONET	Synchronous optical network
SOP	State of polarization
SPE	Synchronous payload envelope
SPM	Self-phase modulation
SRS	Stimulated Raman scattering
SSMF	Standard single mode fiber

STM	Synchronous transport module
STS	Synchronous transport signal
SWP	Spatial walk-off polarizer
TCP	Transmission control protocol
TDFA	Thulium-doped fiber amplifier
TDM	Time-division multiplexing
TDMA	Time-division multiple access
TE	Transverse electric
TEC	Thermoelectric cooler
TFF	Thin-film filter
TIA	Telecommunications Industry Association
TM	Transverse magnetic
UHF	Ultra-high frequency (0.3-to-3 GHz)
UI	Unit interval
UPSR	Unidirectional path-switched ring
VAD	Vapor-phase axial deposition
VCSEL	Vertical-cavity surface-emitting laser
VECP	Vertical eye-closure penalty
VFL	Visual fault locator
VOA	Variable optical attenuator
VSB	Vestigial-sideband
WAN	Wide area network
WDM	Wavelength-division multiplexing
WRN	Wavelength routed network
WSS	Wavelength-selective switch
XPM	Cross-phase modulation
YIG	Yttrium iron garnet

APPENDIX F

Roman Symbols

Symbol	Definition
a	Fiber radius
A_{eff}	Effective area
B	Bit rate
B_e	Receiver electrical bandwidth
B_o	Optical bandwidth
c	Speed of light $= 2.99793 \times 10^8$ m/s
C_j	Detector junction capacitance
d	Hole diameter in a PCF
D	Dispersion
D_{mat}	Material dispersion
D_n	Electron diffusion coefficient
D_p	Hole diffusion coefficient
D_{wg}	Waveguide dispersion
DR	Dynamic range
E	Energy ($E = h\nu$)
E	Electric field
E_g	Bandgap energy
E_{LO}	Local oscillator field
f	Frequency of a wave
F	Finesse of a filter
$F(M)$	Noise figure for APD with gain M
$f(s)$	Probability density function
F_{EDFA}	EDFA noise figure
g	Gain coefficient (Fabry-Perot cavity)
G	Amplifier gain
g_B	Brillouin gain coefficient
h	Planck's constant $= 6.6256 \times 10^{-34}$ J-s $= 4.14 \times 10^{-15}$ eV-s
H	Magnetic field

Symbol	*Definition*
I	Electrical current
I	Optical field intensity
I_B	Bias current
I_D	Photodetector bulk dark current
I_{DD}	Directly detected optical intensity
I_M	Multiplied photocurrent
I_p	Primary photocurrent
$i_p(t)$	Signal photocurrent
I_{th}	Threshold current
$\langle i_s^2 \rangle$	Mean-square signal current
$\langle i_{shot}^2 \rangle$	Mean-square shot-noise current
$\langle i_{DB}^2 \rangle$	Mean-square detector bulk dark noise current
$\langle i_T^2 \rangle$	Mean-square thermal noise current
J	Current density
J_{th}	Threshold current density
k	Wave propagation constant ($k = 2\pi/\lambda$)
K	Stress intensity factor
k_B	Boltzmann's constant = 1.38054×10^{-23} J/K
L	Fiber length
L_c	Connection loss
L_{disp}	Dispersion length
L_{eff}	Effective length
L_F	Fiber coupling loss
L_i	Intrinsic loss
L_n	Electron diffusion length
L_p	Hole diffusion length
L_{period}	Soliton period
L_{split}	Splitting loss
L_{tap}	Tap loss
m	Modulation index or modulation depth
m	Order of a grating
M	Avalanche photodiode gain
M	Number of modes
m_e	Effective electron mass
m_h	Effective hole mass
n	Index of refraction
$\bar{N}$	Average number of electron-hole pairs
NA	Numerical aperture
n_i	Intrinsic n-type carrier concentration

N_{ph}	Photon density
n_{sp}	Population inversion factor
P	Optical power
$P_0(x)$	Probability distribution for a 0 pulse
$P_1(x)$	Probability distribution for a 1 pulse
$P_{amp,sat}$	Amplifier saturation power
P_{ASE}	ASE noise power
P_e	Probability of error
p_i	Intrinsic p-type carrier concentration
P_{in}	Incident optical power
P_{LO}	Local oscillator optical power
P_{peak}	Soliton peak power
PP_x	Power penalty for impairment x
P_{ref}	Reflected power
$P_{sensitivity}$	Receiver sensitivity
P_{th}	SBS threshold power
q	Electron charge = 1.60218×10^{-19} C
Q	BER parameter
Q	Q factor of a grating
R	Reflectivity or Fresnel reflection
r	Reflection coefficient
$\mathscr{R}$	Responsivity
$\mathscr{R}_{APD}$	APD responsivity
R_{nr}	Nonradiative recombination rate
R_r	Radiative recombination rate
R_{sp}	Spontaneous emission rate
$S(\lambda)$	Dispersion slope
T	Period of a wave
T	Absolute temperature
T_{10-90}	10-to-90 percent rise time
T_b	Bit interval, bit period, or bit time
t_{GVD}	Rise time from GVD
t_{rx}	Receiver rise time
t_{sys}	System rise time
V	Mode V number

APPENDIX G

Greek Symbols

Symbol	*Definition*
α	Refractive index profile shape
α	Optical fiber attenuation
α	Laser linewidth enhancement factor
$\alpha_s(\lambda)$	Photon absorption coefficient at a wavelength λ
β	Mode propagation factor
β_2	GVD parameter
β_3	Third-order dispersion
Γ	Optical field confinement factor
Δ	Core-cladding index difference
ΔL	Array waveguide path difference
Δv_B	Brillouin linewidth
Δv_{opt}	Optical bandwidth
η	Light coupling efficiency
η	Quantum efficiency
η_{ext}	External quantum efficiency
η_{int}	Internal quantum efficiency
θ_A	Acceptance angle
θ_c	Critical angle
λ	Wavelength
Λ	Period of a grating
Λ	Hole spacing or pitch of a PCF
λ_B	Bragg wavelength
λ_c	Cutoff wavelength
v	Frequency
σ_{DB}	Detector dark noise current variance
σ_λ	Optical source spectral width
σ_{mat}	Material-induced rms pulse broadening
σ_s	Signal current variance

σ_{shot}	Shot noise current variance
σ_{T}	Thermal noise current variance
σ_{wg}	Waveguide-induced pulse spreading
τ	Carrier lifetime
τ_{ph}	Photon lifetime
φ	Phase of a wave
Φ	Photon flux
χ	Fiber crack depth

Index