

Optical Communications

Second Edition

Robert M. Gagliardi
Sherman Karp



A Wiley-Interscience Publication

JOHN WILEY & SONS, INC.

New York • Chichester • Brisbane • Toronto • Singapore

This text is printed on acid-free paper.

Copyright © 1995 by John Wiley & Sons, Inc.

All rights reserved. Published simultaneously in Canada.

Reproduction or translation of any part of this work beyond that permitted by Section 107 or 108 of the 1976 United States Copyright Act without the permission of the copyright owner is unlawful. Requests for permission or further information should be addressed to the Permissions Department, John Wiley & Sons, Inc., 605 Third Avenue, New York, NY 10158-0012.

Library of Congress Cataloging in Publication Data:

Gagliardi, Robert M., 1934–

Optical communications / R.M. Gagliardi, S. Karp. -- 2nd ed.

p. cm. -- (Wiley series in telecommunications and signal processing)

"A Wiley-Interscience publication."

Includes bibliographical references and index.

ISBN 0-471-54287-3

1. Optical communications. I. Karp, Sherman. II. Title.

III. Series.

TK5103.59.G33 1995

621.382'7--dc20

94-28420

Printed in the United States of America

10 9 8 7 6 5 4 3 2

THE OPTICAL COMMUNICATION SYSTEM

The objective of any communication system is the transfer of information from one point to another. This information transfer is accomplished most often by superimposing (modulating) the information onto an electromagnetic wave (carrier). The modulated carrier is then transmitted (propagated) to the destination, where the electromagnetic wave is received and the information recovered (demodulated). Such systems are often designated by the location of the carrier frequency in the electromagnetic spectrum (Fig. 1.1). In radio systems, the electromagnetic carrier wave is selected with a frequency from the radio frequency (RF) portion of the spectrum. Microwave or millimeter systems have carrier frequencies from those portions of the spectrum. In an optical communication system, the carrier is selected from the optical region, which includes the infrared, visible, and ultraviolet frequencies.

The principal advantages in communicating at optical frequencies are (1) the potential increase in modulation bandwidth, (2) the ability to concentrate power in extremely narrow beams, and (3) the significant reduction in component sizes. In any communication system, the amount of information transmitted is directly related to the bandwidth (frequency extent) of the modulated carrier, which is generally limited to a fixed portion of the carrier frequency itself. Thus, increasing the carrier frequency theoretically increases the available transmission bandwidth, and therefore the information capacity of the overall system. This means frequencies in the optical range will have a usable bandwidth approximately 10^5 times that of a carrier in the RF range. This available improvement is extremely inviting to a communication engineer vitally concerned with transmitting large amounts of information. In addition, the ability to concentrate available transmitter power within the transmitted electromagnetic wave also increases with carrier frequency. Thus, using higher carrier frequencies increases the capability of the system to achieve higher power densities, which generally leads to improved performance. Lastly, operation at the extremely small wavelengths of optics produces system devices

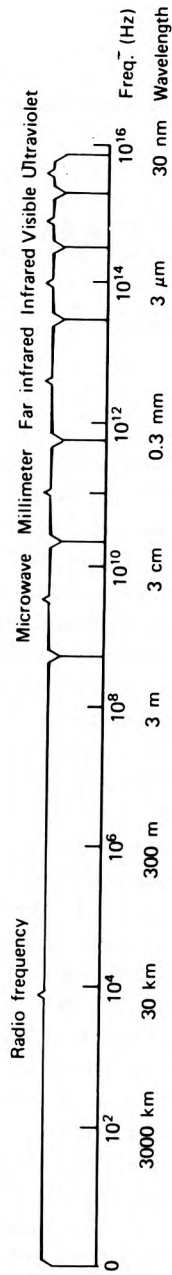


Figure 1.1. The electromagnetic spectrum.

and components that are much smaller than their equivalent electronic counterparts. For these reasons, optical communication has emerged as a field of special technological interest.

Communicating at optical frequencies has several major differences from RF communications. Because optical frequencies are accompanied by extremely small wavelengths, optical component design requires essentially its own technology, completely different from design techniques associated with RF, microwave, and millimeter devices. As a result, optical devices, although emulating equivalent electronic devices, may have performance characteristics significantly different from their electronic counterparts.

Another drawback to optical communications is the detrimental effect of the propagation path on the optical carrier wave. This is because optical wavelengths are commensurate with molecule and particle sizes, and propagation effects are generated that are uncommon to radio and microwave frequencies. Furthermore, these effects tend to be stochastic and time varying in nature, which hinders accurate propagation modeling. A vast amount of experimental data has been collected to aid in understanding this optical propagation phenomenon and, although certain models have been established, continued exploration is required for refinement and further justification.

The development of optical components and the derivation of propagation models, however, are only part of the overall system design. A communication engineer must also be concerned with the choice of components, the selection of system operations, and finally the interfacing or interconnecting of these operations in the best possible manner. These interfacing decisions require reasonably accurate mathematical models, which indicate component performance, anomalies, and degradations, knowledge of which can be used to advantage in system design. It is this aspect of optical communications that this book attempts to elucidate. Our objective is to understand system capability and to formulate system-design procedures and performance characteristics for the implementation of an overall optical communication system.

1.1 OPTICAL SYSTEMS

The block diagram of a generic optical communication system is shown in Figure 1.2. The diagram is composed of standard communication blocks, which are endemic to any communication system. A source producing some type of information (waveforms in time, digital systems, etc.) is to be transmitted to some remote destination. This source has its output modulated onto an optical carrier (a carrier frequency in the optical portion of the electromagnetic spectrum). This carrier is then transmitted as an optical light field, or beam, through the optical channel (free space, turbulent atmosphere, fiberoptic waveguide, etc.). At the receiver, the field is optically collected and processed (photodetected), generally in the presence of noise interference, signal distortion, and inherent background radiation (undesired light fields or other

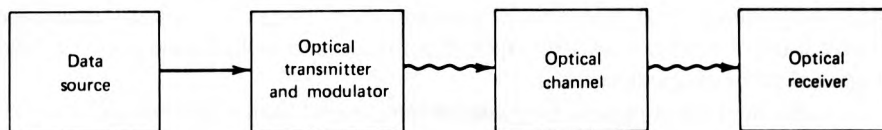


Figure 1.2. Optical communication system block diagram.

electromagnetic radiation). Of course, except for the fact that the transmission is accomplished in the optical range of carrier frequencies, the operations just mentioned describe any communication system using modulated carriers. Nevertheless, the optical system employs devices somewhat uncommon to the standard components of the RF system. These devices have significant differences in their operation and associated characteristics, often requiring variations in design procedures.

The modulation of the source information onto the optical carrier can be in the form of frequency modulation (FM), phase modulation (PM), or possibly amplitude modulation (AM), each of which can be theoretically implemented at any carrier frequency in the electromagnetic range [1]. In addition, however, several other less conventional modulation schemes are also often utilized with optical sources. These include intensity modulation (IM), in which information is used to modulate the intensity (to be defined subsequently) of the optical carrier, and polarization modulation (PLM), in which spatial characteristics of the optical field are modulated.

The optical receiver in Figure 1.2 collects the incident optical field and processes it to recover the transmitted information. A typical optical receiver can be represented by the three basic blocks shown in Figure 1.3, consisting of an optical receiving front end (usually containing some form of lens or focusing hardware), an optical photodetector, and a postdetection processor. The lens system filters and focuses the received field onto the photodetector, where the optical signal is converted to an electronic signal. The processor performs the necessary amplification, signal processing, and filtering operations to recover the desired information from the detector output.

Optical receivers can be divided into two basic types: power detecting receivers and heterodyning receivers. Power detecting receivers (often called direct detection, or noncoherent, receivers) have the front end system shown in Figure 1.4a. The lens system and photodetector operate to detect the instantaneous power in the collected field as it arrives at the receiver. Such receivers represent the simplest type for implementation and can be used whenever the transmitted information occurs in the power variation of the received field.

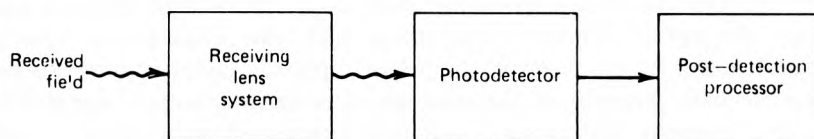


Figure 1.3. The optical receiver.

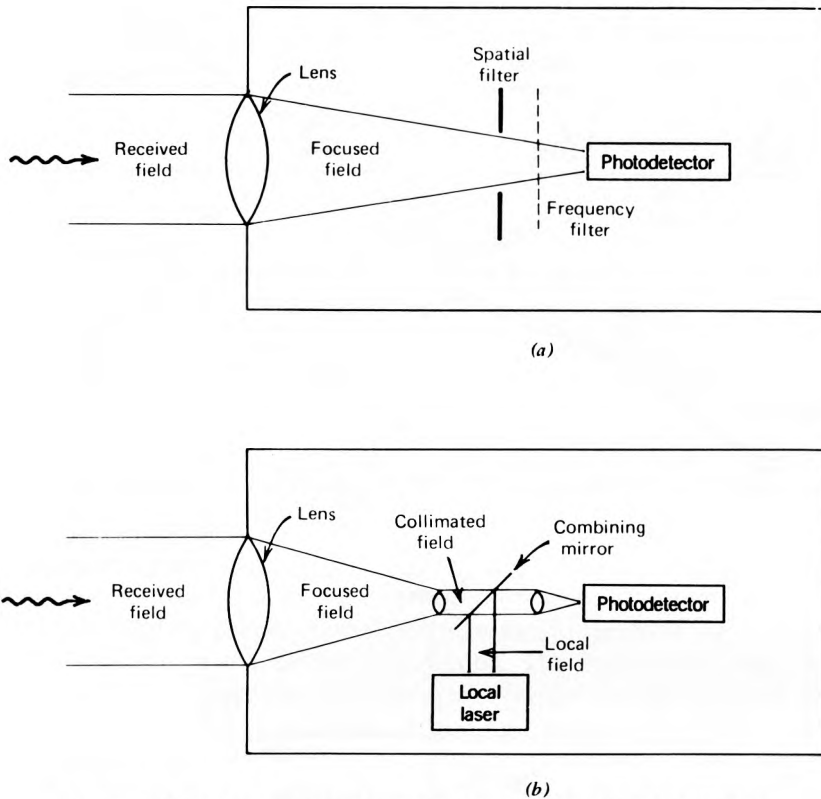


Figure 1.4. (a) Direct detection receiver. (b) Heterodyne detection receiver.

Heterodyning receivers have the front end system shown in Figure 1.4b. A locally generated lightwave field is optically mixed with the received field through a front end mirror, and the combined wave is photodetected. Such receivers are used whenever information is amplitude modulated, frequency modulated, or phase modulated onto the optical carrier. Heterodyning receivers are more difficult to implement and require close tolerances on the spatial coherence of the two optical fields being mixed. For this reason, heterodyned receivers are often called (spatially) coherent receivers. For either type of receiver, the front end lens system has the role of focusing the received or mixed field onto the photodetector surface. This focusing allows the photodetector area to be much smaller than that of the receiving lens.

The receiver front end, in addition to focusing the optical field onto the photodetector, also provides some degree of filtering, as shown in Figure 1.4. These filters are employed prior to photodetection to reduce the amount of undesired background radiation. Optical filters may operate on the spatial properties of the focused fields (polarization filters, field stops, etc.) or may filter in the frequency domain; that is, they pass certain bands of frequencies

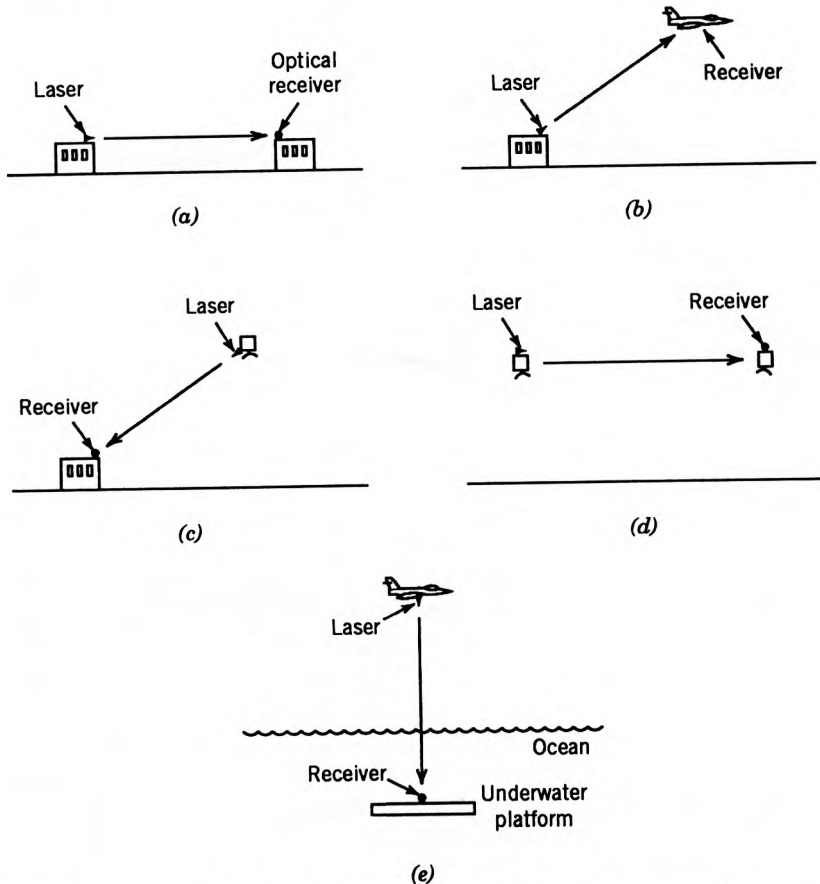


Figure 1.5. Optical space links. (a) Terrestrial, (b) ground based to air, (c) spaced based to ground, (d) intersatellite, (e) air to underwater.

and reject others. The latter filters determine the bandwidth of the resulting optical field subsequently photodetected.

Photodetectors convert the focused optical field into an electrical signal for processing. Although there are several types of detectors available, all behave according to quantum mechanical principles, utilizing photosensitive materials to produce current or voltage responses to changes in impinging optical field power. The basic model defining this interaction for all photodetectors is well accepted, although detectors may differ in their output response characteristics. This basic model, which is examined in detail in Chapter 3, is extremely important to the communication engineer because it generates the inherent statistics that must be utilized in design of the postdetection processing. The most common type of photodetectors are the phototubes, photodiodes, and photomultipliers.

The detection of optical fields is hampered by the various noise sources

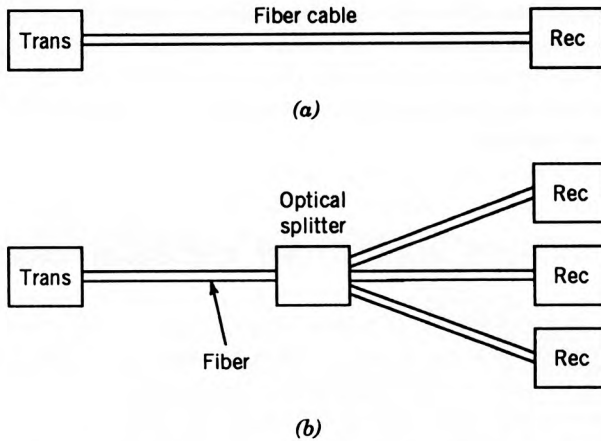


Figure 1.6. Fiber optic links. (a) Single fiber link, (b) fiber optic distribution.

present throughout the receiver. The most predominant in long-distance space communication is the background light or stray radiation that is collected at the receiver lens along with the desired optical field. Although this radiation may be reduced by proper spatial filtering, it still represents the most significant interference in the detection operation. The background effect can be eliminated when direct-coupled fiberoptic waveguides can be used for the transmission path. A second noise source is the photodetector itself, which, not being a purely ideal device, produces internal interference during the photo-detection operation. This induced noise is referred to as detector noise. The last noise source is the circuit and electronic thermal noise generated in the processing operations following photodetection. The thermal noise is accurately modeled as additive white Gaussian noise, whose spectral level is directly related to the receiver temperature, just as in any RF or microwave communication system. Each of these noise sources must be properly accounted for in any receiver analysis.

The models in Figures 1.3 and 1.4 are common to any optical communication system. In a space system, the transmitted optical field is focused into a beam of light and transmitted as a propagating electromagnetic field through a medium. Examples of these are shown in Figure 1.5. The system can be a terrestrial (ground-based) link, a ground-to-space (atmospheric) link, a space-to-space crosslink, or even a space-to-underwater link. All such systems use optical beams transmitted as unguided fields and are susceptible to the effects of the medium (atmosphere, clouds, water, etc.) over the communication path. The effects of atmospheric propagation in space links are examined in Chapter 9.

A fiberoptic system (Fig. 1.6a) confines the transmitted field to an optical waveguide (fiber) during its propagation. The system, therefore, is operated as is any cable-connected link. Because the field is guided, only the properties of the fiber itself affect the field transmission. In particular, atmospheric and

background noise effects are no longer important to system performance. The enormous improvement in fiber quality has permitted long communication links and fiberoptic distribution systems (Fig 1.6b) to be readily established. Today, fiberoptic systems are rapidly replacing the more traditional cable and wireline systems of the past.

1.2 OPTICAL SOURCES, MODULATORS, AND BEAM FORMERS

The key element in any optical communication system is the availability of a light source that can be easily modulated. Such a source should produce energy concentrated in a narrow wavelength band. The primary sources of light in modern optical systems are the light-emitting diodes (LED), the laser, and the laser diode (LD) [2-4]. Although the physical descriptions of these devices are beyond our scope here, their output properties and characteristics will be important in assessing the performance when used in an optical communication system.

An LED is formed from semiconductor junctions that interact when subjected to external current so as to radiate light energy, as indicated by the diagram in Figure 1.7a. A detailed theory of band energy is needed to describe this interaction and is not pursued here. The choice of the junction materials determines the emitted wavelength. Light-emitting diodes are typically formed from compounds of gallium arsenide, and produce light in the 0.8–0.9- μm wavelength bands. An LED is small in size (centimeters), relatively inexpensive, and can produce radiation with low-current drive levels. However, they are limited in output power (1 to 10mW), and the emitted light tends to be unfocused. Table 1.1 summarizes these basic characteristics.

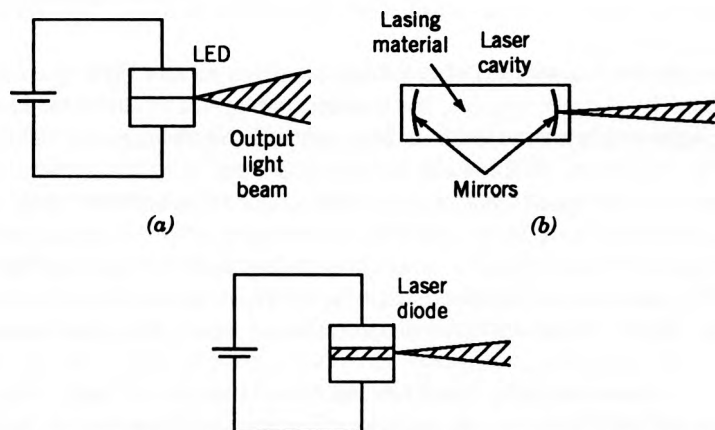


Figure 1.7. Light sources. (a) Light emitting diode (LED), (b) laser cavity, (c) laser diode (LD).

103 smaller

A laser tube is constructed as an optical cavity filled with light amplification material (gas or solid) and mirrored facets at each end as shown in Figure 1.7b. If the propagation gain of the material overcomes the reflection losses of the mirrors, then an initiated optical field reflected back and forth by the mirrors, will be self-sustaining. We say the cavity “lases,” and optical energy is produced within the cavity. By placing a small aperture in one mirror, some internal energy will escape as radiated light. As a result the laser can produce high power levels (0.1 to 1 W) with output radiation that can be more focused than an LED.

Many different materials can be used in the cavity to produce lasing (Table 1.1), each having specific atomic structures that form particular wavelength bands. In addition, the cavity length must be such that a propagating internal field will exactly reinforce (be phase aligned) after two-way mirror reflections. This means the cavity length must be precisely an integer multiple of a half wavelength of the internal field. That is, the lasing wavelength λ must be related to the cavity length L by $\lambda = 2L/n$, for some integer n . The lasing material sustains those wavelengths that are commensurate with its propagation gain profile. Laser tubes are therefore high-power devices, but are much bulkier than diode sources.

Laser diodes are semiconductor junction devices that contain substrates that are etched, or cleaved, to act as reflecting facets for field reinforcements over the junctions. They therefore combine the properties of an LED and the

TABLE 1.1 Optical Sources

Laser material	Wavelength (μm)	
Solid state		
GaAs	0.87	
InGaAs	1.0–1.3	
InGaAsP	0.9–1.7	
AlGaAs	0.8–0.89	
Ruby	0.694	
Nd-Yag	1.06	
Gas		
CO ₂	10.6	
HeNe	0.63	
Laser Types	Output Power	Linewidth (nm)
Diodes		
LED	0.1–10 mW	20–100
laser diode	1–40 mW	1–5
Distributed feedback laser	1–40 mW	0.1
Tubes		
CO ₂	1–5 W	0.01–1.0
HeNe	50–100 mW	0.01–1.0

cavity reflector, producing an external light radiation that is higher in power (10 to 50 mW) and better focused than a simple LED.

The important communication characteristics of any optical source are its modulation bandwidth (the rate at which the source can be modulated), its input-output power curve, and its frequency spectrum. Light sources, generated from extremely narrow cavities, can be modulated at bandwidths up to 1 to 4 GHz. This provides an enormous potential advantage over RF communications, where modulation bandwidths of only hundreds of megahertz are available.

Figure 1.8a sketches typical diode and laser power characteristics, plotting output light power versus external bias current. The threshold current is that needed to produce output light. The linear range defines the modulation range, where input current converts proportionally to output power. The saturation level limits the maximum available output power. Light-emitting diodes have low thresholds and can operate at low-current values, but they have limited peak powers. Laser diodes require more drive current, but have higher peak power. Laser tubes generally have to be pumped above threshold and are difficult to stabilize in the linear range. Hence, high-power lasers are usually operated as continuous-wave devices at their peak power capability.

The frequency spectrum of an optical source, as sketched in Figure 1.8b, indicates the spectral extent, or purity, of the light source. The spreading of spectrum around the desired wavelength indicates the presence of unwanted frequencies, or undesired noise modulations, superimposed on the output wavelength. This spectral spreading can hinder the ability to recognize desired information modulated on the source. Light-emitting diodes have relatively wide spectral extent (hundreds of Angstroms), whereas lasers significantly improve the light purity. The effect of source spreading on modulation performance is examined further later in this volume.

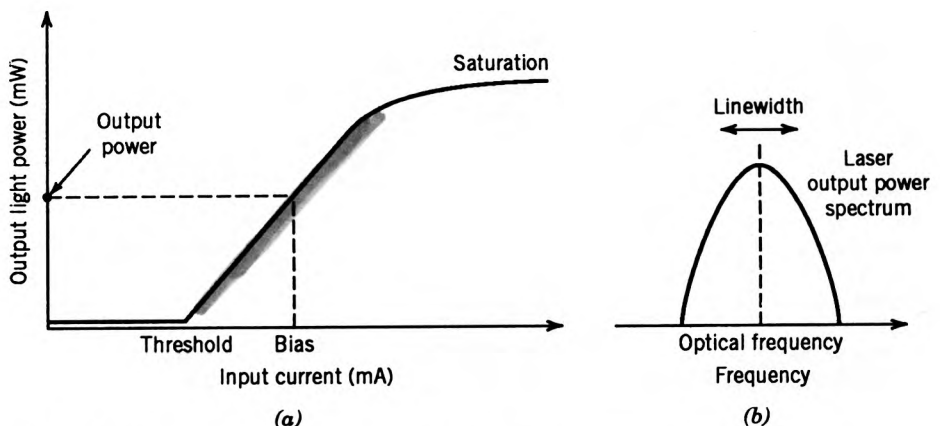


Figure 1.8. Characteristics of light sources. (a) Output light power vs. input drive current; (b) power spectrum of the output light.

Modulators superimpose the information signals (analog or digital) on to the source. Optical modulators [4] are of two basic types: internal or external (Fig. 1.9). An internal modulator is one in which the source itself is directly modified by the information signal to produce a modulated optical field. Amplitude or intensity modulation can be imposed by varying the bias current, as is evident from Figure 1.9a. Frequency or phase modulation can be inserted on a laser tube by varying its cavity length. Pulse modulation is easily applied to a diode by driving it above and below threshold. Such modulations are generally limited to the linear range of the power characteristic.

In external modulation (Fig. 1.9b), the source light is focused through an external device, whose propagation characteristics are altered by the modulating signal. Such systems have the advantage of utilizing the full power capability of the source. Modulation is achieved via the electrooptic or acoustooptic effect of the material, in which external currents can modify the transmission properties (index of refraction, polarization, direction of flow, etc.) of the inserted light. These effects produce delay variations (phase modulation) or polarization changes (intensity modulation) on the excited beam. Pulsed outputs can be achieved by blocking or deflecting the light path. Unfortunately, external modulators insert significant coupling losses, limit the modulation range, and generally require relatively higher modulation drive power.

Light fields from the radiating surfaces of optical sources and modulators are emitted with varying degrees of focusing, usually described by its emission angle, as shown in Figure 1.10. The light emission is further characterized by the source brightness function, $\mathcal{B}(\theta)$, which is in units of watts/steradian-area, that describes the normalized light power emitted in a given direction angle θ out from the source. The power is usually normalized to a unit solid angle per unit of source area. Hence the source brightness indicates the distribution of power radiated out from the source. A uniformly radiating source will have the same brightness at all angles within its emission solid angle Ω_s , as shown in Figure 1.10. The total power in watts emitted from a uniform source with area A_s and emission angle Ω_s is then

$$P_s = \mathcal{B} A_s \Omega_s \quad (1.2.1)$$

For symmetrically radiating sources, the solid angle Ω_s can be related to the

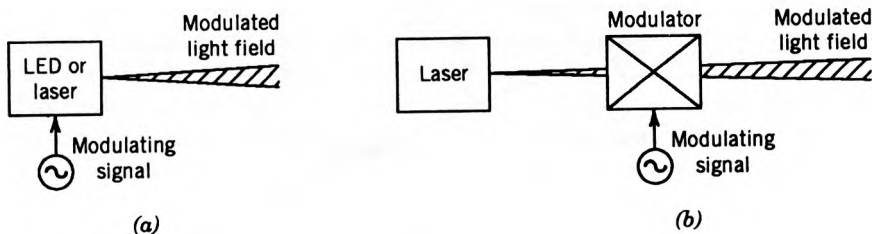


Figure 1.9. Optical modulators. (a) Internal, (b) external.

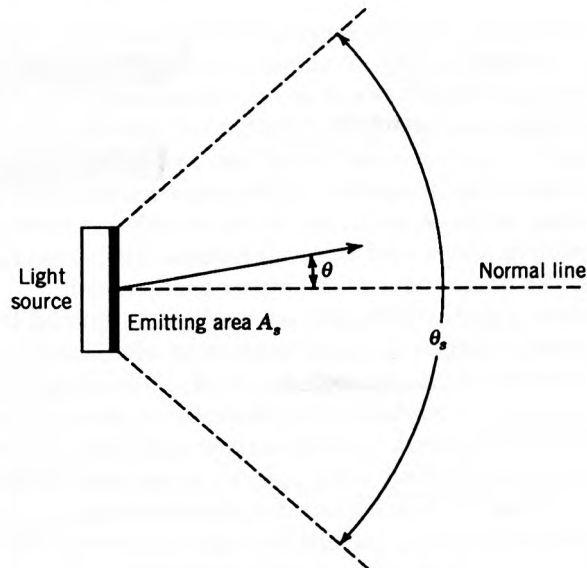


Figure 1.10. Light emission angles.

planar emission angle θ_s in Figure 1.10 by

$$\Omega_s = 2\pi[1 - \cos^2(\theta_s/2)] \quad (1.2.2)$$

A Lambertian source is a uniformly radiating source that emits in all forward directions $|\theta| < \pi/2$. For the Lambertian source, $\Omega_s = 2\pi$ steradians, and $P_s = \mathcal{B}2\pi A_s$. The assumption of uniformly radiating surfaces is generally used in describing most optically emitting sources.

Light fields from radiating sources can also be collected and refocused by means of beam-forming optics. The latter are usually combinations of various

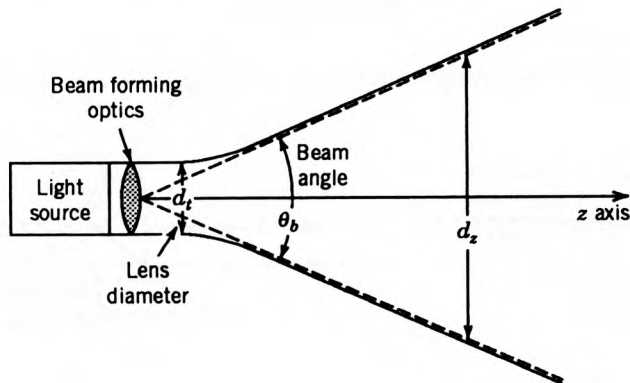


Figure 1.11. Beam forming and focused light.

types of lenses placed at the source or modulator output that orient the light into particular directions. Although light focusing will be considered in Chapter 2, Figure 1.11 shows a simple type of beam formation commonly used in long-range space links. A combination of a converging and diverging lens placed at the source is used to produce a collimated beam. Ideally, the converging lens focuses the source field light to a point, and the diverging lens expands it to a perfect beam. In practice, the source field is instead focused to a spot, and the expanded beam spreads during propagation with a planar beam diameter of approximately [5]

$$d_z = d_t \left[1 + \left(\frac{\lambda z}{d_t^2} \right)^2 \right]^{1/2} \quad (1.2.3)$$

where λ is the wavelength, d_t is the output lens diameter, and z is the distance from the lens. At points in the near field ($\lambda z/d_t^2 < 1$), the emerging light is collimated with a diameter equal to the lens diameter. That is, the light appears to uniformly exit over the entire lens. In the far field ($\lambda z/d_t^2 > 1$) the beam diameter expands with distance, and appears as if the light is emerging from a single point with a planar beam angle of approximately

$$\theta_b \cong \frac{\lambda}{d_t} \text{ rad} \quad (1.2.4)$$

The angle θ_b is called the diffraction-limited transmitter beam angle. The expanding field far from the source is therefore confined to a two-dimensional solid angle of approximately

$$\begin{aligned} \Omega_b &= 2\pi[1 - \cos(\theta_b/2)] \\ &\simeq \left(\frac{\pi}{4} \right) \theta_b^2 \end{aligned} \quad (1.2.5)$$

Figure 1.12 shows a plot of Eq. (1.2.4) as a function of diameter d_t for several optical wavelengths. For example, a 6-in. lens at an optical wavelength of $0.5 \mu\text{m}$ has a beamwidth on the order of 3×10^{-6} rad or approximately 0.16 mdeg. This is a spectacular advantage compared to RF transmitters where antenna beams are usually on the order of degrees. This ability to concentrate field flow to small beam angles with relatively small size optics is a significant advantage in long-range space communications.

The optical advantage of this source focusing can be further emphasized by converting to an effective antenna gain. From antenna theory [6], a transmitter with the beamwidth in Eq. (1.2.4) will have an effective antenna gain of

$$G_t = \frac{4\pi}{\Omega_b} \approx \left(\frac{4d_t}{\lambda} \right)^2 \quad (1.2.6)$$

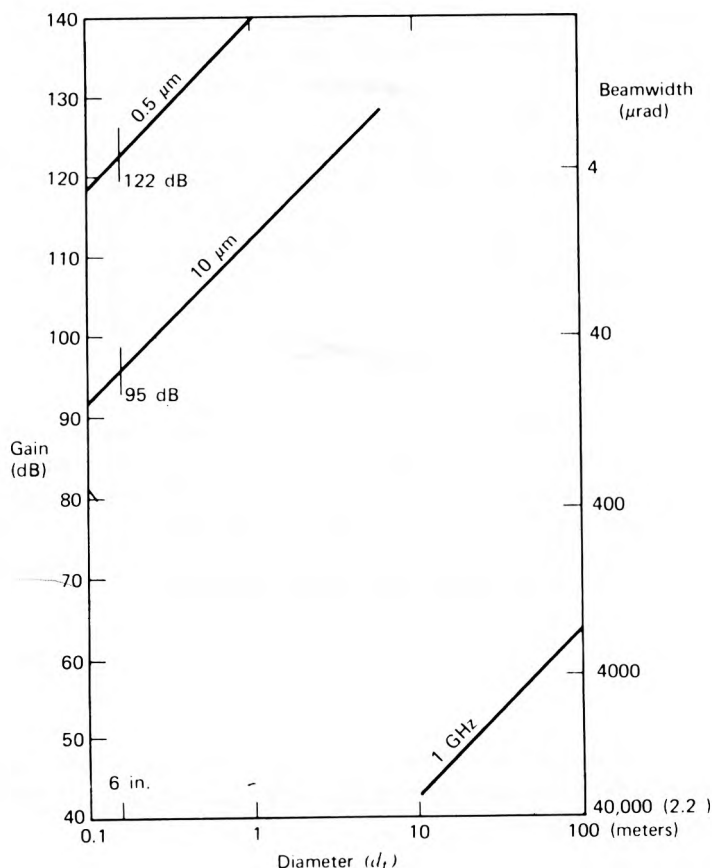


Figure 1.12. Antenna gain and beam vs. diameter.

This corresponding gain is also shown in Figure 1.12. The same 6-in. lens would have an effective gain of 122 dB, a sizable improvement over an RF antenna. (A 210-ft antenna generates only 60 dB of gain at RF.) It should be pointed out, of course, that with these thin beams, difficult pointing problems (aiming the transmitter at a particular spot) are now generated. This latter problem, which immediately becomes an integral part of the overall system design, is considered in Chapter 10.

1.3 TRANSMITTED OPTICAL FIELDS

The field from an optical source is transmitted as a propagating optical field. It is necessary to properly describe this emanating field to determine power flow. Because the transmitted field is an electromagnetic wave, it is described at any spatial point by solutions to Maxwell's equations. Let ξ represent a point of a selected coordinate system in which the field source is located at the origin, as

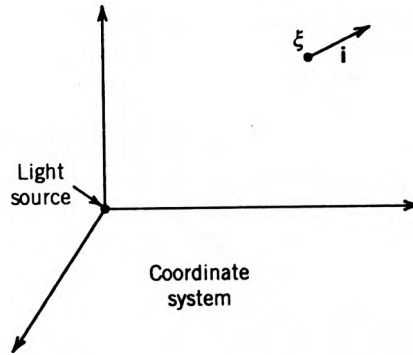


Figure 1.13. Three-dimensional coordinate space.

shown in Figure 1.13. At any time t , the radiation field is described at ξ by a unit vector $\mathbf{i}$ in the direction of field flow and its electrical field function

$$\text{Electric field} = \text{Real}\{f(t, \xi)\} \quad \text{V/length} \quad (1.3.1)$$

where $\text{Real}\{\cdot\}$ means "real part of" and $f(t, \xi)$ is referred to as the complex field. At each point ξ , the field has an intensity given by

$$I(t, \xi) = \frac{|f(t, \xi)|^2}{Z_w} \quad \text{W/area} \quad (1.3.2)$$

where Z_w is the wave impedance (i.e., the impedance of the medium; in free space, $Z_w = 377$ ohms). The *instantaneous power* of the complex field over a planar area A at time t is given by the surface integral of the intensity over the area,

$$P_A(t) = (\cos \theta) \int_A I(t, \xi) d\xi \quad \text{W} \quad (1.3.3)$$

where the integration is over all ξ in A , and θ is the angle between the normal line to the area and the unit vector $\mathbf{i}$ in the direction of power flow.* We point out that without loss of generality, the medium impedance Z_w can be directly incorporated into the field definition $f(t, \xi)$, so that the field intensity can be simply written as $I(t, \xi) = |f(t, \xi)|^2$. This normalized field $f(t, \xi)$ now has units of $\text{watts}^{1/2}/\text{length}$.

Equations (1.3.1) to (1.3.3) are key equations relating optical field, intensity, and power and are important to subsequent communication analysis. This is because optical detectors inherently respond to the intensity of the impinging fields. The most important optical field is the plane wave field that propagates in direction $\mathbf{i}$ and has its complex field defined over a normal planar area A

*The vector $I(t, \xi)\mathbf{i}$ through ξ , is called the Poynting vector. Equation (1.3.3) is the integration of the Poynting vector over the surface area.

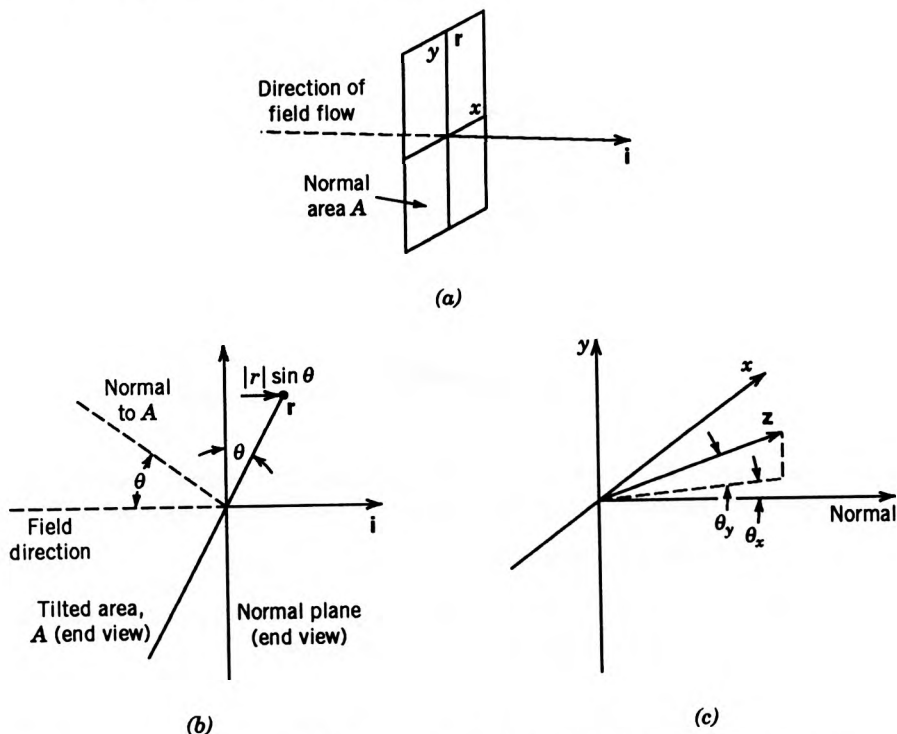


Figure 1.14. (a) Normal area to field propagation. (b) Tilted area to field propagation. (c) Subangles of the vector $\mathbf{z}$.

(Fig. 1.14a) containing points $\mathbf{r}$ as

$$f(t, \mathbf{r}) = a(t)e^{j\omega_o t} \quad \mathbf{r} \in A \quad (1.3.4)$$

where ω_o is the optical radian frequency of the transmitter and $a(t)$ is a complex amplitude function (normalized by Z_w) describing the field polarization in the plane (x, y) normal to the direction of propagation. This can be expanded as

$$a(t) = a_x(t)\mathbf{1}_x + a_y(t)\mathbf{1}_y \quad (1.3.5)$$

where $a_x(t)$ and $a_y(t)$ are the complex polarization components and $(\mathbf{1}_x, \mathbf{1}_y)$ are unit coordinate vectors in the (x, y) plane. The polarization components determine the polarization state of the plane wave. If $|a_x(t)| = |a_y(t)|$ and are 90 degrees out of phase with the other, the field is circularly polarized. If both are in phase, or one is zero, the field is linearly polarized. Unless otherwise stated, we assume linearly polarized fields throughout subsequent discussions, and the polarization notation can be dropped. In this case $a(t)$ represents the complex modulation applied to the optical carrier. If we write

$$a(t) = |a(t)|e^{j\phi(t)} \quad (1.3.6)$$

then $|a(t)|$ describes the amplitude modulation, and $\phi(t)$ the phase modulation ($d\phi/dt$ would describe the frequency modulation) occurring over the normal plane. The intensity modulation would then be given by $|a(t)|^2$ over the plane. Note that all points $\mathbf{r}$ in A vary in phase synchronism, and we say the plane wave is *phase coherent* over A . As the field propagates in the direction $\mathbf{i}$ through A , the coherent wavefront varies in time at each point in the plane.

Suppose we wish to describe a plane wave over an area A tilted from the normal, as shown in Figure 1.14b. Since the points in a normal plane are in phase, the points $\mathbf{r}$ on the tilted plane are phase shifted due to the field propagation. Let $f(t, 0) = a(t)e^{j\omega_0 t}$ be the complex field at the point $\mathbf{r} = 0$ on A . A point $\mathbf{r}$ will either lag or lead this point as the plane wave propagates through it. Thus we can write for any point $\mathbf{r}$ in A ,

$$f(t, \mathbf{r}) = a(t)e^{j\omega_0 t} \exp \left[\left(-j \left(\frac{2\pi}{\lambda} \right) |\mathbf{r}| \sin \theta \right) \right] \quad (1.3.7)$$

where again θ is the angle between $\mathbf{i}$ and the normal area. The exponential phase term accounts for the additional phase shift at point $\mathbf{r}$. This phase shift can be written in other ways. Define the propagation vector

$$\mathbf{z} = (2\pi/\lambda)\mathbf{i} \quad (1.3.8)$$

which is a vector passing through the origin of A with the direction of the field vector $\mathbf{i}$ and with amplitude $2\pi/\lambda$. We first note that we can write $(2\pi/\lambda)|\mathbf{r}| \sin \theta = |\mathbf{z}||\mathbf{r}| \cos(90^\circ - \theta) = \mathbf{z} \cdot \mathbf{r}$, denoting the vector dot product between the vectors $\mathbf{r}$ and $\mathbf{z}$. Thus we have the equivalent of Eq. (1.3.7),

$$f(t, \mathbf{r}) = a(t)e^{j\omega_0 t} e^{-j\mathbf{z} \cdot \mathbf{r}} \quad (1.3.9)$$

The above defines the field at any point $\mathbf{r}$ in A due to a plane wave field with complex envelope $a(t)$ arriving from angle θ relative to the normal. Note that the optical carrier is phase shifted from one point in the plane to another by an amount depending on the dot product at the point location.

We can also write the phase shift directly in terms of the coordinates of the point $\mathbf{r} = (x, y)$ by expanding the dot product, using the x and y coordinates of the vector $\mathbf{z}$ in Eq. (1.3.8). Denote $\mathbf{z} = (z_x, z_y)$ and let (θ_x, θ_y) be the subangles of the arrival angle θ , as shown in Figure 1.14c. Specifically, θ_y is the angle between $\mathbf{z}$ and the plane formed by the normal axis and the x axis of A (elevation angle). The subangle θ_x is the angle between the normal axis and the vector formed by projecting $\mathbf{z}$ on to this plane (azimuth angle). Then, in terms of these subangles,

$$\begin{aligned} z_y &= (2\pi/\lambda) \sin \theta_y \\ z_x &= (2\pi/\lambda) \cos(\theta_y) \sin(\theta_x) \end{aligned} \quad (1.3.10)$$

This expansion allows us to substitute $\mathbf{z} \cdot \mathbf{r} = xz_x + yz_y$, so that

$$f(t, \mathbf{r}) = a(t)e^{j\omega_0 t} e^{-j(xz_x + yz_y)} \quad (1.3.11)$$

The phase factor contains only the arrival subangles and the coordinates of $\mathbf{r}$. This further simplifies if the arrival subangles are small ($\mathbf{z}$ arrives close to the normal) so that $\cos \theta_y \cong 1$, $\sin \theta_y \cong \theta_y$ and $\sin \theta_x \cong \theta_x$. In this case, $z_x \cong (2\pi/\lambda)\theta_x$, $z_y = (2\pi/\lambda)\theta_y$, and

$$f(t, \mathbf{r}) \cong a(t)e^{j\omega_0 t} \exp \left[-j \left(\frac{2\pi}{\lambda} \right) (x\theta_x + y\theta_y) \right] \quad (1.3.12)$$

These phase relations between arrival angles of a plane wave and points $\mathbf{r}$ in the plane will be important in later focusing and detection analysis.

1.4 OPTICAL SPACE CHANNELS

In the optical space systems in Figure 1.5 the laser source produced an optical field that is intended for a receiving area that is located a significant distance away. This distance is usually far enough so that the laser source appears as a point source when viewed from the receiver. Systems in which this point source assumption can be made can be analyzed similar to RF communications.

If the source producing the optical field is a point source, operating in a free space medium and modulated with the field function in Eq. (1.3.6), then the field at a remote point ξ due to this source is given by

$$f(t, \xi) = \frac{\sqrt{G(\xi)} a \left(t - \frac{|\xi|}{c} \right) e^{j(\omega_0 t - 2\pi|\xi|/\lambda)}}{\sqrt{4\pi|\xi|}} \quad (1.4.1)$$

where $|\xi|$ is the distance to the point ξ , λ is the source wavelength corresponding to the frequency ω_0 , c , is the speed of light, and $G(\xi)$ is the transmitter power gain in the direction ξ . Equation (1.4.1) describes an optical field that propagates out from the source as a spherical wave with diminishing amplitude. However, to an area A located a distance z away, where $z \gg A^{1/2}$, all points on A are approximately the same distance from the source. Thus $|\xi|$ is the same over A and the field arriving at A from the point source is basically a plane wave field. That is, the field emanating from a point source appears as a plane wave when observed by a small area from a remote distance. The field intensity at a point ξ in A is then

$$I(t, \xi) = G_t \frac{\left| a\left(t - \frac{z}{c}\right) \right|^2}{4\pi z^2} \quad (1.4.2)$$

where G_t is the transmitting optics gain in the direction of A . Defining

$$P_s(t) = |a(t)|^2 \quad (1.4.3)$$

as the power variation of the point source, Eq. (1.4.2) reduces to

$$I(t, \xi) = G_t \frac{P_s(t - t_d)}{4\pi z^2} \quad (1.4.4)$$

where $t_d = z/c$ is the propagation delay from the source to the point ξ . Thus the field intensity produced by an optical point source with envelope function $a(t)$ propagates out from the transmitter according to the gain functions of the transmitting optics and decreases in strength as the square of the distance.

As an example, consider a point source that transmits a constant power P_s through beam-forming optics over a solid angle Ω_s , as shown in Figure 1.15. At a distance $|\xi| = z$ away, the field intensity within the beam will be

$$I(t, \xi) = \frac{G_t P_s}{4\pi z^2} \quad (1.4.5)$$

where G_t is the transmitter gain in Eq. (1.2.6). A normal receiving area A within the beam collects the field power, using Eq. (1.3.3), as

$$P_r = \left[\frac{G_t P_s}{4\pi z^2} \right] A \quad (1.4.6)$$

Note this corresponds to the standard power flow equation associated with any electromagnetic field produced from a source with the gain G_t . It is common practice to define a receiving gain related to the receiving area as

$$G_r = \left(\frac{4\pi}{\lambda^2} \right) A \quad (1.4.7)$$

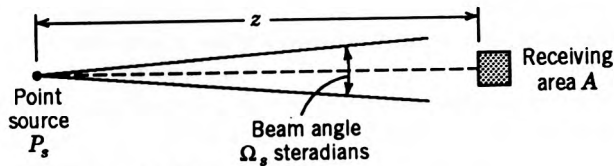


Figure 1.15. Space link model.

Equation (1.4.6) now becomes

$$P_r = P_s G_t L_p G_r \quad (1.4.8)$$

where

$$L_p = \left(\frac{\lambda}{4\pi z} \right)^2 \quad (1.4.9)$$

is the propagation loss associated with the transmission of an electromagnetic field of wavelength λ over a distance z . Thus the power flow equation (1.4.6) is identical to the link budget equations used in all standard communication link analysis. Note that the use of Eq. (1.4.7) in Eq. (1.4.6) inserts the optical wavelength λ into the power flow equation, giving the somewhat erroneous conclusion that collected power in a free space channel depends on wavelength.

If the receiving area A is tilted relative to the normal (i.e., the point source field arrives from an offset angle), Eqs. (1.3.3) must be used with the angle to determine received power. This inserts an additional power loss from Eq. (1.4.8) due to the cosine of the angle offset. This loss is negligible if the arrival angle is small (i.e., the field arrives close to the area's normal).

When the propagating medium is not free space, additional effects must be included on the propagating field. This is because when propagating through a nonfree space link, such as the atmosphere, the wave undergoes effects that tend to further alter the power levels and the field phase across the wavefront. These alterations depend on the wavelength of the wave and are due primarily to interactions with inhomogeneities and foreign particles comprising the medium. These effects become most predominant as the wavelength approaches the physical size of these particulates. For this reason, atmospheric distortion tends to be quite severe in the optical range, where wavelengths are commensurate with particulates as small as molecules. Thus the atmospheric space channel presents a major communication hurdle that must be of concern to optical system engineers.

The atmosphere primarily causes field absorption and scattering, the degree of which depends on the type of atmospheric conditions (clear air, clouds, rain, etc.) [7]. This absorption manifests itself in the propagating wavefront as both amplitude and spatial phase effects. Amplitude effects are exhibited in the time variation of the field and involve power loss, power fluctuations, and frequency filtering. Spatial effects appear as variations in the beam direction, beam spreading, or phase variations across the wavefront. In particular, it alters the phase coherency that would normally exist across the wavefront in standard plane waves.

Since the particulate structure of the atmosphere tends to be random in nature, these amplitude and phase effects evolve as random disturbances that can only be stochastically described. A detailed analysis of the atmospheric effects in optical space channels is considered in Chapter 9.

1.5 THE FIBER OPTICAL CHANNEL

An optical fiber is a thin pipe of glass, through which one can insert light and have it shine out at the output end. Hence a fiber can transmit a modulated light field from its input to its output. The basic construction of a fiber is shown in Figure 1.16, and corresponds to a central glass core encased in a plastic shielding. The core is made of high-quality glass through which the light can propagate. Glass, or a similar silicon compound, is used, because it offers the least attenuation for optical transmission. The core is supported by the shielding, or cladding, which provides mechanical protection, isolates the core from external radiation, and aids in confining the light to only the glass core. The cladding is opaque, and basically reflects any escaping light back into the core. Typically the core diameter is on the order of micrometers, whereas the fiber cross-section (core plus cladding) is on the order of several millimeters. Thus the fiber is literally a thread of transmission path over which the modulated light field can be propagated. This ability to pack large amounts of modulated data over an extremely small spatial area is an overwhelming advantage for optical fiber communications [8, 9].

The key parameters describing the propagation properties of the fiber light field are the refractive index of the core and cladding. The cladding is usually designed to have an index a few percent smaller than that of the core. (Glass typically has an index of approximately 1.5). These indices determine the propagation angles of the field in the fiber.

Consider the ray diagram in Figure 1.17, which shows an optical ray (optical field propagating in the ray direction) being inserted into the fiber core at angle θ . If θ is too large, the ray will be absorbed into the cladding, and will not propagate. Propagation will occur down the fiber only if the incident ray angle at the cladding boundary is shallow enough to be reflected back into the core. The ray will then continue to propagate down the core as it continually reflects off the core walls. The application of Snells law at the boundary requires that $\theta \leq \theta_p$, where

$$\theta_p = \cos^{-1} \left(\frac{n_2}{n_1} \right) \quad (1.5.1)$$

for the reflections to occur, n_1 is the core index and n_2 is the cladding index.

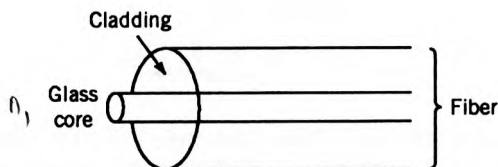


Figure 1.16. Fiber construction.

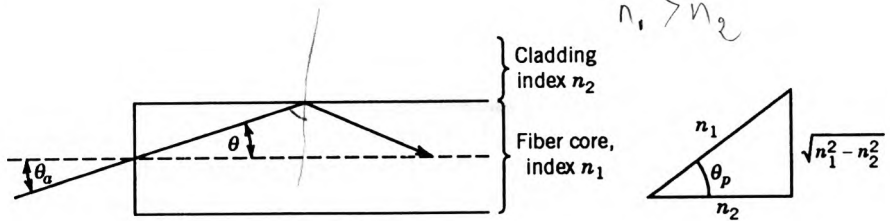


Figure 1.17. Fiber ray geometry.

The right-hand side of Eq. (1.5.1) is called the *critical angle* of the fiber and represents the maximum angle at which light will propagate down the fiber. Note that the angle depends only on the index ratio. By adjusting these indices one can control these light flow angles. In particular, by having $n_2 \gtrsim n_1$ the propagation angles can be made quite small, so the light that propagates does so at small angles, and the light will flow only down the center of the core. This means fiber boundary losses will be reduced, and less leakage will occur from the core into the cladding. Since n_2 will necessarily be close to the value of n_1 , we often deal with the fractional difference,

$$\Delta = \frac{n_1 - n_2}{n_1} \quad (1.5.2)$$

For small propagation angles, we generally desire a Δ of a few percent.

To propagate light into the core at angle θ_p , it is necessary that the light be properly inserted (because the light ray may bend at the fiber input). Let θ_a be the fiber insertion angle (i.e., the angle at which light is fed into the fiber from an external source), as shown in Figure 1.17. Again, by Snell's law applied to the index change from the external medium (assumed to be free space with index of unity) to the core glass, we have

$$n_1 \sin \theta_p = (1) \sin \theta_a = \sin \theta_a \quad (1.5.3)$$

The numerical aperture of the fiber is defined as

$$\begin{aligned} \text{NA} &\triangleq \sin \theta_a \\ &= n_1 \sin \theta_p \end{aligned} \quad (1.5.4)$$

We see that NA is an indication of the allowable insertion angle. For a given core index, we see that a small NA corresponds to a small propagation angle θ_p , and therefore a highly collimated field. Equations (1.5.3) and (1.5.4) suggest



the triangle relationship in Figure 1.17 and, from simple geometry, we obtain

$$\begin{aligned}
 n_1 \sin \theta_p &= n_2 \sin \theta_c \\
 \text{NA} &= n_1 \sin \theta_p \\
 &= \frac{n_1(n_1^2 - n_2^2)^{1/2}}{n_1} \\
 &= (n_1^2 - n_2^2)^{1/2} \\
 &= n_1 \left[\left(1 + \frac{n_2}{n_1}\right) \left(1 - \frac{n_2}{n_1}\right) \right]^{1/2}
 \end{aligned} \tag{1.5.5}$$

Thus, the numerical aperture NA is obtained directly from the indices. Since $n_2 \approx n_1$, we can substitute from Eq. (1.5.2) and use the approximation

$$\begin{aligned}
 \text{NA} &\cong n_1 \left[2 \left(1 - \frac{n_2}{n_1}\right) \right]^{1/2} \\
 &\cong n_1 \sqrt{2\Delta}
 \end{aligned} \tag{1.5.6}$$

When the core index is fixed, NA is actually a measure of the fractional difference of the indices. Because Δ is a few percent, NA typically takes on values between 0.05 and 0.2.

When light is inserted into the fiber, it will propagate at many of the ray angles θ satisfying Eq. (1.5.1). Each individual propagating ray line in the fiber is called a field mode, and each has its own spatial E -field satisfying the necessary boundary conditions. Each mode propagates as an individual fiber field with orthogonally polarized E -field vectors (corresponding to the various transverse electric fields similar to those found in microwave waveguides) and at slightly different angles. From Maxwells equations for cylindrical guides, the number of such modes [8, 9] is given by

$$\text{number of fiber modes} = \left(\frac{d^2}{2}\right) \left[\left(\frac{2\pi n_1}{\lambda}\right)^2 - \left(\frac{2\pi n_2}{\lambda}\right)^2 \right] \tag{1.5.7}$$

where d is the core diameter. By direct substitution from Eq. (1.5.5), we see that

$$\begin{aligned}
 \text{number of fiber modes} &= 2 \left(\frac{\pi d}{\lambda}\right)^2 [n_1^2 - n_2^2] \\
 &= 2 \left(\frac{\pi d}{\lambda}\right)^2 (\text{NA})^2
 \end{aligned} \tag{1.5.8}$$

Thus the numerical aperture NA also determines the number of propagating fiber modes. A fiber can therefore have hundreds of modes if the core diameter is many times the light wavelength.

A fiber will propagate only a single mode if

$$2 \left(\frac{\pi d}{\lambda} \right)^2 (\text{NA})^2 \leq 1 \quad (1.5.9)$$

or if

$$d \leq \frac{\lambda}{\sqrt{2} \pi \text{NA}} \approx (0.22) \left(\frac{\lambda}{\text{NA}} \right) \quad (1.5.10)$$

Single-mode fibers, therefore, require core diameters to be only several times the wavelength. A single-mode fiber will have the least loss during propagation (since the field flows primarily down the center of the fiber), but the extremely narrow core diameter will lead to obvious problems in fiber coupling and interconnections.

For large-core diameters the fiber becomes a multimode fiber with modes propagating at the various angles within θ_p . From a communication point of view having many modes is generally undesirable. This is due to the fact that power is usually distributed somewhat uniformly over the modes. (Even if the power is inserted at a single ray line direction, the internal reflections due to fiber bends will redistribute the power over all the modes.) The modes at the larger angles θ_p will be more lossy, and the overall attenuation will be greater than for a single-mode fiber. In addition, since modes propagate at different angles, their group velocities are different, and they are therefore delayed with respect to each other at the fiber output. This delay dispersion can cause an inherent filtering on the modulated light field on each mode. This effect will be considered in more detail in Chapter 7.

As in the space system, the key element is the source power that can be delivered to the fiber output. The important parameters are the efficiency of inserting source power into the fiber, and the power loss that occurs during the fiber propagation. Power insertion from an optical source into a fiber can be described by the diagram in Figure 1.18. The fiber can be butted against the source-emitting surface, or can be interconnected via lensing elements. Only the source power coupled into the acceptance angle of the fiber will produce propagating fields. Let the source have an emitting area of A_s , a brightness of $\mathcal{B}$ W/steradian-area and divergence cone angle Ω_s . If the fiber is connected directly to the source (Fig. 1.18a) the power into the fiber is

$$\begin{aligned} P_f &= \mathcal{B} \Omega_p A_f & \text{if } A_f < A_s \\ &= \mathcal{B} \Omega_p A_s & \text{if } A_s < A_f \end{aligned} \quad (1.5.11)$$

where the solid angle Ω_p is related to the planar angle θ_p by Eq. (1.2.5). From Eq. (1.2.1), the total source power is $P_s = \mathcal{B} \Omega_s A_s$. The fractional power loss

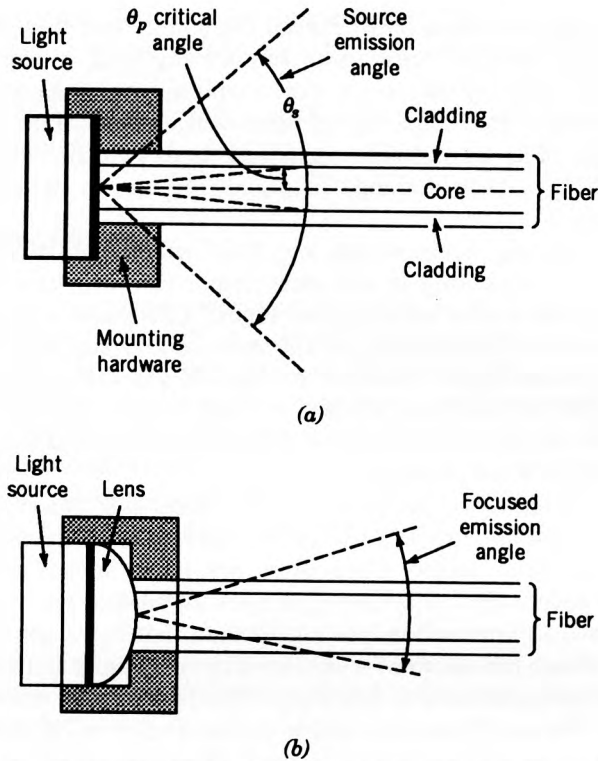


Figure 1.18. Power coupling with fibers. (a) Direct butting of the fiber to source. (b) Lens focusing into fiber.

while coupling into the fiber is then

$$\begin{aligned} \frac{P_f}{P_s} &= \left(\frac{\Omega_p}{\Omega_s} \right) \left(\frac{A_f}{A_s} \right) & \text{if } A_f < A_s \\ &= \left(\frac{\Omega_p}{\Omega_s} \right) & \text{if } A_s < A_f \end{aligned} \quad (1.5.12)$$

Thus the coupling loss may have an area mismatch term and an angle mismatch term. The area mismatch term will depend on the relative areas of the source and fiber core, with the loss contribution being most severe for the narrow-core fibers. If the light source produces unfocused Lambertian light in Eq. (1.5.12), then the angle coupling loss becomes

$$\frac{\Omega_p}{\Omega_s} = \frac{2\pi(1 - \cos \theta_p)}{2\pi} \approx \frac{\theta_p^2}{2} \quad (1.5.13)$$

Thus the angle coupling loss is related to the square of the numerical aperture, and will be greater for low NA fibers.

By inserting a focusing lens between the source and the fiber (Fig. 1.18b), the source light can be focused into smaller angles Ω_s in Eq. (1.5.12), which improves the coupling efficiency. For coupling into single-mode fibers it is almost mandatory that some type of field focusing be used for power insertion. Various types of lenses and index matching fluids (which provide some degree of focusing) have been designed in pigtail formats to permit reduced fiber coupling losses.

The propagating fields within the fiber will be attenuated due to the absorption and scattering of the glass impurities, with the scattering effect decreasing at the higher wavelengths. Figure 1.19 shows a typical plot of the overall attenuation in decibels per kilometer of a typical single-mode fiber as a function of wavelength. Minimal attenuation windows occur in the vicinity of wavelengths of approximately 0.8, 1.3, and 1.5 μm , where most of the fiber systems tend to operate. The latter two bands indicate that propagation losses less than 1 dB/km are possible.

As more modes are produced in the fiber, the power in the fiber is distributed over these modes with the outermost modes having higher attenuation factors. As a result, when there are many modes propagating, the attenuation effects tend to be averaged over all modes. The multimode fibers, therefore have an attenuation loss higher than that for single-mode fibers and tend to be somewhat constant with wavelength, as indicated in Figure 1.19.

The total propagation loss in a fiber is therefore directly related to its length, and can be obtained from appropriate curves similar to Figure 1.19. The fiber attenuation α_f in decibels/length is read off at the proper wavelength, and, knowing the fiber length Z , the total attenuation in decibels is $\alpha_f Z$. If P_f is the source power coupled into the fiber in Eq. (1.5.11), the fiber output power then follows as

$$P_r = P_f 10^{-\alpha_f Z/10} = P_f e^{-0.23\alpha_f Z} \quad (1.5.14)$$

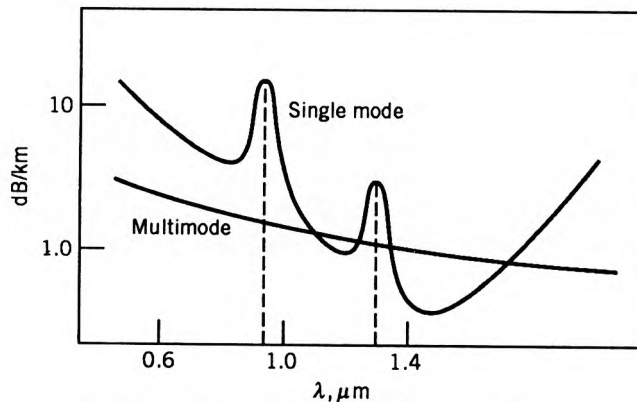


Figure 1.19. Fiber attenuation vs. wavelength.

Thus, fiberoptic power flow can be determined in this manner. Additional losses may have to be included for fiber splicings, cable interconnections, and fiber bending.

Other effects in fiber propagation, such as dispersion and filtering, will be considered in Chapter 7.

1.6 FIELD EXPANSIONS

In the analysis of communication systems, it is often convenient to use orthogonal expansions of signals to formulate a more general procedure, and sometimes to gain further insight into results [10]. Orthogonal expansions produce mathematically valid waveform representations that permit signals to be treated in pieces (multiple dimensions) rather than as a single entity. Likewise, in optical communications it is convenient to represent optical fields in terms of similar orthogonal expansions as an alternative to field description.

Consider again a general complex optical field defined over an area A as

$$f(t, \mathbf{r}) = a(t, \mathbf{r})e^{j\omega_o t} \quad \mathbf{r} \in A \quad (1.6.1)$$

where we have denoted $a(t, \mathbf{r})$ as the complex envelope and ω_o as the optical source frequency. Let us now consider a spatial orthogonal expansion of this complex envelope as

$$a(t, \mathbf{r}) = \sum_{i=1}^{\infty} a_i(t)\Phi_i(\mathbf{r}) \quad (1.6.2)$$

where $\{\Phi_i(\mathbf{r})\}$ are a set of complex two-dimensional orthogonal space functions defined over A . That is,

$$\int_A \Phi_i(\mathbf{r})\Phi_j^*(\mathbf{r}) \, d\mathbf{r} = 0 \quad \text{if } i \neq j \quad (1.6.3)$$

with the asterisk denoting complex conjugates. The corresponding expansions functions are given by

$$a_i(t) = \frac{\int_A a(t, \mathbf{r})\Phi_i^*(\mathbf{r}) \, d\mathbf{r}}{\int_A |\Phi_i(\mathbf{r})|^2 \, d\mathbf{r}} \quad (1.6.4)$$

That is, each $a_i(t)$ is the complex time function corresponding to the spatial function $\Phi_i(\mathbf{r})$. Equation (1.6.2) corresponds to the orthogonal expansion of the field envelope into the spatial orthogonal set $\{\Phi_i(\mathbf{r})\}$. Combining Eqs. (1.6.1)

and (1.6.2) produces an overall field description defined over A as

$$f(t, \mathbf{r}) = \sum_{i=1}^{\infty} a_i(t) e^{j\omega_i t} \Phi_i(\mathbf{r}) \quad (1.6.5)$$

The equality in Eq. (1.6.5) is in a squared integrable sense, and the convergence of the sum on the right to the field function on the left requires only a bounded energy constraint on the radiation field [10]. The advantage of an expansion of this form is primarily for mathematical convenience, although such expansions often yield physical insight into the related optical processing. It is natural to consider the functions $\{\Phi_i(\mathbf{r})\}$ as defining the spatial "modes" of the field over the area A . The time functions $\{a_i(t)\}$ then become the modal functions describing the temporal process in each mode. Because different orthogonal sets may be available, mode descriptions of a given field are not necessarily unique, each corresponding to a different expansion set.

For the fiber link, where the area A can represent the cross-sectional area of the fiber, the orthogonal expansion modes $\Phi_i(\mathbf{r})$ appear naturally as the orthogonal field modes that propagate in the fiber as the transverse E -fields, each satisfying the necessary boundary conditions. Hence the fiber modes discussed in Section 1.5 become the mathematical modes in Eq. (1.6.5), and the mode functions $a_i(t)$ are the time variations of these modes described over the area A . Thus the orthogonal expansion in Eq. (1.6.5) can represent the internal fiber field at any point along the fiber. In particular, it can be used to represent the field at the fiber output.

For an area collecting an arriving optical field in a space system, the concept of a mode is no longer as obvious. As an example, consider the complex function

$$\Phi_i(\mathbf{r}) = e^{-j(\mathbf{z}_i \cdot \mathbf{r})} \quad \mathbf{r} \in A \quad (1.6.6)$$

corresponding in Eq. (1.3.9) to a constant amplitude plane wave arriving from direction $\mathbf{z}_i$. Now consider two such functions with arrival vectors $\mathbf{z}_i$ and $\mathbf{z}_q$, and assume a square area A of dimension $l \times l$, as shown in Figure 1.20. The spatial integral of these functions become

$$\begin{aligned} \int_A \Phi_i(\mathbf{r}) \Phi_q^*(\mathbf{r}) d\mathbf{r} &= \int_{-l/2}^{l/2} \int_{-l/2}^{l/2} e^{-j[\mathbf{r} \cdot (\mathbf{z}_i - \mathbf{z}_q)]} dx dy \\ &= \int_{-l/2}^{l/2} \int_{-l/2}^{l/2} e^{-j(x\Delta z_x + y\Delta z_y)} dx dy \end{aligned} \quad (1.6.7)$$

where we have used $(\mathbf{z}_i - \mathbf{z}_q) = (\Delta z_x, \Delta z_y)$, with Δz_x and Δz_y the x and y components of the difference vector. Letting $(\theta_{ix}, \theta_{iy})$ be the subangles (azimuth

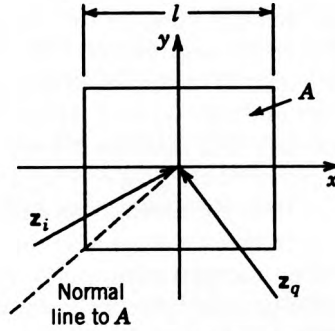


Figure 1.20. Receiver area geometry with impinging ray lines.

and elevation) of each vector $\mathbf{z}_i$, we can use Eq. (1.3.10) to write

$$\begin{aligned}\Delta z_y &= \left(\frac{2\pi}{\lambda}\right) (\sin \theta_{iy} - \sin \theta_{qy}) \\ \Delta z_x &= \left(\frac{2\pi}{\lambda}\right) (\cos \theta_{iy} \sin \theta_{ix} - \cos \theta_{qy} \sin \theta_{qx})\end{aligned}\quad (1.6.8)$$

Because

$$\int_{-l/2}^{l/2} e^{-jyx} dx = [l \sin(\gamma l/2)]/(\gamma l/2) \quad (1.6.9)$$

the integral in Eq. (1.6.7) is then

$$\begin{aligned}l^2 \left[\frac{\sin(l\Delta z_x/2)}{(l\Delta z_x/2)} \right] \left[\frac{\sin(l\Delta z_y/2)}{(l\Delta z_y/2)} \right] \\ = \begin{cases} A & \text{if } \Delta z_x = 0, \Delta z_y = 0 \\ 0 & \text{if } \Delta z_x = k_1(2\pi/l) \text{ or } \Delta z_y = k_2(2\pi/l) \end{cases}\end{aligned}\quad (1.6.10)$$

where k_1 and k_2 are any nonzero integers. Thus the plane wave fields corresponding to arrival vectors $\mathbf{z}_i$ and $\mathbf{z}_q$ are orthogonal fields over the area A if their angles produce coordinates satisfying $\Delta z_x = 2k\pi/l$ or $\Delta z_y = 2k\pi/l$ for any nonzero integer k . That is, the vectors will be orthogonal if their arrival angles satisfy the proper conditions. To expand on this, consider the small angle assumption, as in Eqs. (1.3.12), producing the requirement $\Delta z_y \simeq (2\pi/\lambda)(\theta_{iy} - \theta_{qy})$ and $\Delta z_x = (2\pi/\lambda)(\theta_{ix} - \theta_{qx})$. Under this assumption, the vectors $\mathbf{z}_i$ and $\mathbf{z}_q$ are orthogonal if

$$\begin{aligned}(\theta_{ix} - \theta_{qx}) &\approx k(\lambda/\sqrt{A}) \quad \text{or} \\ (\theta_{iy} - \theta_{qy}) &\approx k(\lambda/\sqrt{A})\end{aligned}\quad (1.6.11)$$

Thus complex plane waves arriving close to the normal at a squared area A having arrival angles separated by multiples of $\lambda/A^{1/2}$ are spatially orthogonal functions. The set of all such arrival vectors satisfying Eq. (1.6.11) will therefore constitute an orthogonal set and can be used as the modes of the expansion. We conclude, therefore, that any field function defined over A can be expanded into sets of these arriving plane waves properly separated in angle of arrival. The expansion in Eq. (1.6.5) then corresponds to describing the field $f(t, \mathbf{r})$ as sums of plane waves, each with its own amplitude time functions in Eq. (1.6.4). We find this rather simple interpretation quite useful in later analyses. Although the result was here derived specifically for a square area, the plane wave orthogonality concept can in fact be extended to more general area shapes [11], with approximately the same conditions as in Eq. (1.6.11). We use these generalizations in later studies.

1.7 RANDOM FIELDS

In an optical system, we are often forced to deal with stochastic or random fields. Such fields arise when dealing with atmospheric or weather effects (which cause random variations to occur in the transmitted field), or in describing background or stray light that may appear in the system. In addition, the laser itself may produce random contributions (self-emission noise) to its own emitted fields. These random variations lead to statistical fluctuations in the optical field, which can only be analyzed after associating proper statistics with the field itself. We wish to develop some basic properties for describing these general random fields.

The complex envelope of a stochastic field must be considered random at each point t and $\mathbf{r}$ describing the field over a designated area. As such, random fields are completely described by their probability densities, that is, the probability densities associated with each point, or set of points, of the field. These densities are often difficult to model exactly, and often assumptions must be imposed upon these statistics. For example, a Gaussian random field is one whose probability density at any point $(t, \mathbf{r})$ is taken as a complex Gaussian variable.

Stochastic field analysis is often confined to second-order statistics associated with the field; in particular, its coherence function [12]. In this regard, the time-space (mutual) coherence function of a stochastic field $f(t, \mathbf{r})$ at points $(t_1, \mathbf{r}_1)$ and $(t_2, \mathbf{r}_2)$ is formally defined as

$$R_f(t_1, t_2, \mathbf{r}_1, \mathbf{r}_2) = \overline{f(t_1, \mathbf{r}_1) f^*(t_2, \mathbf{r}_2)} \quad (1.7.1)$$

where the overbar represents the expectation operator over the joint field densities at the points involved, and the asterisk denotes the complex conjugate. The field coherence therefore describes the degree of correlation from one point to another.

Stochastic fields are often described by their inherent coherence properties. A stochastic field is said to be *temporally stationary* if the time dependence in its coherence function depends only on the time difference $t_1 - t_2$. The field is *spatially homogeneous* if the spatial dependence in the coherence function depends only on the spatial distance $(\mathbf{r}_1 - \mathbf{r}_2)$. A field is *completely homogeneous* if it is both temporally stationary and spatially homogeneous. A stochastic field is said to be *coherence-separable* if its coherence function factors as

$$R_f(t_1, t_2, \mathbf{r}_1, \mathbf{r}_2) = R_t(t_1, t_2)R_s(\mathbf{r}_1, \mathbf{r}_2) \quad (1.7.2)$$

The factor $R_s(\mathbf{r}_1, \mathbf{r}_2)$ describes the spatial coherence variation, whereas $R_t(t_1, t_2)$ defines the temporal correlation of the field.

The average intensity of a field at a point $(t, \mathbf{r})$ is defined as the mean squared value of the field at time t and point $\mathbf{r}$. Thus

$$\overline{|f(t, \mathbf{r})|^2} = R_f(t, t, \mathbf{r}, \mathbf{r}) \quad (1.7.3)$$

and therefore can be obtained directly from the mutual coherence. When applied to stationary coherence-separable fields, the average intensity at $\mathbf{r}$ becomes

$$\begin{aligned} \bar{I}(\mathbf{r}) &= R(t, t, \mathbf{r}, \mathbf{r}) \\ &= R_t(0)R_s(\mathbf{r}, \mathbf{r}) \end{aligned} \quad (1.7.4)$$

and depends only on the spatial coherence function. The value of $\bar{I}(\mathbf{r})$ is often called the field irradiance at the point $\mathbf{r}$.

A stochastic field is space coherent over an area A if $R_s(\mathbf{r}_1, \mathbf{r}_2) = R_s(\mathbf{r}_1, \mathbf{r}_1)$ for all $\mathbf{r}_1, \mathbf{r}_2$ in A . Note that in this case only the temporal randomness is exhibited over A . In this sense, the field is identical to a plane wave, except it has its field amplitude $a(t)$ as a random process in t . Hence, random fields coherent over an area behave as random plane waves over that area. A stochastic field is space incoherent if $R_s(\mathbf{r}_1, \mathbf{r}_2) = 0$, $\mathbf{r}_1 \neq \mathbf{r}_2$. Otherwise, it is partially space coherent.

Stochastic field envelopes also have infinite series expansions into orthogonal functions as in Eq. (1.6.5). That is, we can represent a random field as

$$f(t, \mathbf{r}) = \sum_{i=1}^{\infty} a_i(t) e^{j\omega_i t} \Phi_i(\mathbf{r}) \quad (1.7.5)$$

where the component functions $\{a_i(t)\}$ are now random processes. Equation (1.7.5) describes a field whose spatial mode functions have a random time envelope variation imposed. For example, the expansion can represent randomly varying arriving offset plane waves, or random spatial modes in a fiber,

at the frequency ω_0 . The convergence of the sum is now a mean square [13], and requires bounded average energy in the field over the expansion area. Again, any set of complete spatially orthogonal functions can be used for the expansion.

It is common practice to assume that the individual modal time processes $\{a_i(t)\}$ are pairwise uncorrelated random processes. This assumption has an implicit condition that the spatial orthogonal functions satisfy a set of required conditions (called the *Karhunen-Loeve*, or KL, conditions; see Problem 1.18). If the random field is assumed to be a Gaussian field with uncorrelated modal functions, then the $a_i(t)$ processes are in fact independent random processes, and the field is said to have independent random modes, as defined by Eq. (1.7.5). The random analysis presented here is useful later to better understand stochastic field effects in our optical receivers.

1.8 OPTICAL AMPLIFIERS

One of the more important advances in fiber technology is in the area of optical amplifiers. As with their electronic counterparts, optical amplifiers can increase the power level of inserted optical fields, thereby overcoming fiber propagation power losses and directly improving the detection performance in a fiberoptic communication system.

Optical amplifiers amplify light levels through various types of forced stimulated emission [14, 15]. Although a detailed discussion of the various amplification mechanisms are beyond the scope of our study here, the communication characteristics of such devices are important in assessing system performance and understanding potential applications. The most basic optical amplifier is the semiconductor laser amplifier (SLA) in which optical power gain is achieved by biasing a laser cavity below threshold, and inserting the light field at one end. The internal reflections from the mirrored facets build field intensity, similar to the lasing operation described in Section 1.2. Achievable power gain values in SLA devices will depend on the lasing material and the reflection coefficients of the cavity [14].

A fiberoptic amplifier (Fig. 1.21a) uses Raman or Brillouin scattering in the fiber core to transfer energy from an external optical pump field to the inserted field. The two fields are coupled together in the fiber amplifier, and energy transfer occurs during the joint propagation. The pump field is used to pump the core ions into an excited state, and the energy is transferred to the input signal field as it propagates as a traveling wave. At the amplifier output, the pump field is removed and a higher intensity signal field appears. Gain values depend on the pumping power and on the fiber amplifier length.

Improvements in fiber amplification have occurred with the insertion of core doping to improve the conversion from pump to field. Various dopant materials can be used, but erbium has the advantage of producing maximum gain at the center wavelength of $1.55 \mu\text{m}$, which matches the fiber loss window,

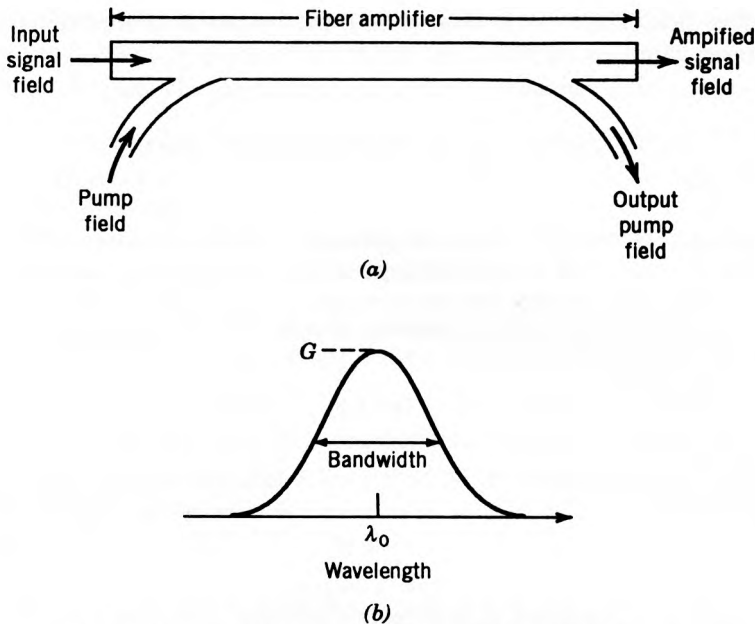


Figure 1.21. (a) Fiber amplifier diagram. (b) Gain function.

as shown in Figure 1.19. Hence the erbium-doped fiber amplifier (EDFA) has emerged as a critical fiberoptic element.

Any optical amplifier produces a wavelength-dependent power-gain transfer function, similar to electronic amplifiers. A typical gain function is sketched in Figure 1.21b and is characterized by its peak gain G , its optical bandwidth, and its center wavelength, all of which depend on the various design parameters of the device itself. Reported results [14–16] of present day amplifiers have gain values in the 20 to 40-dB range, with bandwidths of approximately 10 to 20 nm (approximately 10^{12} Hz). Hence optical amplifiers are extremely wideband devices with potentially high gain. However, as with any amplifier, gain levels may become saturated if the amplifier input signal power is increased. Thus peak gains may be higher when amplifying weaker optical signals than when amplifying a stronger signal.

In the process of generating the energy transfer that provides the power gain, the pump-signal coupling also generates a random spontaneous emission noise that is superimposed on the output signal field [17, 18]. This spontaneous noise field acts as a random independent field in the fiber, similar to the random optical fields described in Section 1.7. This noise is added to all the signal field modes that are being amplified [19].

From a communication point of view, the optical amplifier can be represented by a direct multiplication of the input optical field over all its modes and the insertion of the output spontaneous noise field. If we represent the general

input fiber field $f_1(t, \mathbf{r})$ as in Eqs. (1.6.5), the output field after amplification with power gain G is then

$$f_o(t, \mathbf{r}) = \sum_{i=1}^{\infty} [G^{1/2}a_i(t) + b_i(t)]e^{j\omega_o t}\Phi_i(\mathbf{r}) \quad (1.8.1)$$

Here G is the power gain at the wavelength λ corresponding to the optical frequency ω_o , and $b_i(t)$ is the envelope of the spontaneous noise in the i th mode. Studies have shown that the complex process $b_i(t)e^{j\omega_o t}$ will have a flat power distribution over the amplifier bandwidth [18, 19], with spectral level

$$N_{sp} = (G - 1)(h f_o)n_{sp} \quad \text{W/Hz} \quad (1.8.2)$$

where h is Planks constant, f_o is the optical frequency in hertz, and n_{sp} is the spontaneous emission coefficient, related to the population states of the gain material [14]. Values of n_{sp} typically range from slightly above 1 to approximately 5.

It is evident from Eq. (1.8.1) that amplification has produced an output signal field whose intensity is

$$\left| \sum_{i=1}^{\infty} G^{1/2}a_i(t)e^{j\omega_o t}\Phi_i(\mathbf{r}) \right|^2 = G|f_1(t, \mathbf{r})|^2 \quad (1.8.3)$$

and the amplifier has directly multiplied the input field intensity by the gain factor G . In later analyses, in Chapter 4 we examine the combined effects of the amplified field in Eq. (1.8.3) and the added field noise on the overall communication performance.

PROBLEMS

- 1.1 Communication engineers deal with bandwidth in terms of frequency. Physicists deal with bandwidth in terms of wavelength.
 - (a) Determine the conversion between these two bandwidths. That is, find the frequency band Δf corresponding to a given wavelength band $\Delta \lambda$.
 - (b) Use this result to sketch a plot of frequency band Δf in hertz versus optical wavelength bandwidth $\Delta \lambda$ in angstroms, assuming an optical center wavelength of $\lambda = 1 \mu\text{m}$.
- 1.2 A laser cavity produces a lightwave by sustaining an optical field between two reflecting surfaces located a distance L apart, such that the fields reinforce after each reflection.

- (a) Derive the relation between the field wavelength produced (there may be more than one) and the cavity length.
 - (b) Relate the change in field frequency to a change in cavity length.
- 1.3 A laser cavity 1-m long is to provide light at wavelength $\lambda = 0.5 \mu\text{m}$.
- (a) What is the nearest wavelength to $0.5 \mu\text{m}$ at which the laser can also produce light?
 - (b) What is the required wavelength bandwidth of the lasing material propagation gain to guarantee light at only $0.5 \mu\text{m}$?
- 1.4 An LED emits 10 mW of power when injected with 2.5 mA of current from a 5-V bias voltage. What is the power efficiency (power out/power in) of this LED?
- 1.5 A laser diode has the power characteristic shown in Figure 1.8a. Write an expression for the fraction of the output average power P_o that can be linearly modulated (equal positive and negative variations) in terms of P_o , the threshold power, and the saturation power level of the device.
- 1.6 A laser with power P_s transmits with a circular symmetric Gaussian beam pattern such that the source power transmitted at angle θ off boresight is given by

$$P(\theta) = \frac{P_s \pi}{\theta_b^2} e^{-(\theta/\theta_b)^2}$$

- (a) What is the effective antenna gain on boresight ($\theta = 0$)?
 - (b) How much is the field intensity reduced from boresight intensity at transmission angle $\theta = \theta_b$? Express in decibels.
 - (c) Repeat for $\theta = n\theta_b$.
- 1.7 Assume all optical power is confined to the diffraction-limited beam-width of the transmitter. Approximately how large must a receiver lens diameter be to collect all the source power from a transmitting lens of diameter d meters located L meters away?
- 1.8 A 10-mW laser transmits a red light beam through a 3-in. lens to a circular receiving lens having a 6-in. diameter located 10^7 m away.
- (a) In free space how much power will fall on the receiver lens?
 - (b) Recompute the receiver lens power if the laser is located only 10^5 m away.
- 1.9 A stationary satellite located 22,000 miles above Earth is to transmit a green light beam to cover a 1-mile radial distance on Earth. What is the approximate transmitting lens size needed on the satellite, and what is its associated antenna gain?

- 1.10** An Earth–Moon link is to communicate at $10.6\ \mu\text{m}$. The Earth beam is to have a solid angle just encompassing the Moon. The Moon beam must illuminate the United States (3000 miles diameter).
- (a) Design the transmitting and receiving lens that satisfy these specifications.
 - (b) What is the overall antenna gain (transmit + receive) of the system.
- 1.11** Determine the propagation loss L_p in Eq. (1.4.9) of an optical field at $1\ \mu\text{m}$ when transmitted over a 22,000-mile path. Express in decibels.
- 1.12** A fiber uses a core diameter of 1 mm, an index of 1.5, and a cladding with index 1.485. For an optical wavelength of $0.6\ \mu\text{m}$, find
- (a) the maximum propagation angle,
 - (b) the number of propagating orthogonal modes, and
 - (c) how small the core diameter should be for a single propagating mode.
 - (d) Repeat (a), (b), and (c) with a cladding index of 1.0.
- 1.13** A 1-mW Lambertian laser source is to be coupled to an equal diameter core fiber. Two fibers are available: (a) fiber with $NA = 0.1$ and a loss factor of 3 dB/km, or (b) a fiber with $NA = 0.4$ and a loss of 50 dB/km. For each fiber, estimate the power at the output for a 10-m fiber length and a 10-km fiber length.
- 1.14** A fiber has a glass core ($n_1 = 1.5$), an NA of 0.2, a diameter of $100\ \mu\text{m}$, and a loss coefficient of 2 dB/km. It is butted up against a uniformly emitting source of diameter 1 cm, power output 1 mW, and a total emitting angle of 30 degrees.
- (a) Determine the area mismatch loss.
 - (b) Determine the angle mismatch loss.
 - (c) If the fiber is 2 km long, what is the power at the fiber output?
- 1.15** Assume arriving orthogonal plane waves from offset angles satisfying Eq. (1.6.11) define field modes for a collecting area A . How many signal modes are contained in an arriving field from an optical source that extends over an area A_s located Z units away? *Hint*: assume the source area can be modeled as an array of point sources distributed over the area A_s .
- 1.16** Consider the set of orthogonal plane waves at wavelength λ defined over a square area $l \times l$. Now consider a general signal field from a source z -units away that produces a field expansion containing Q plane wave components. Estimate the effective transmitting area of the source as seen by the receiver.

- 1.17 Show that if the real and imaginary parts of a zero mean complex random field $f(t, \mathbf{r})$ are each uncorrelated fields, then its mutual coherence function is always real.
- 1.18 Show that the complex mode processes $a_i(t)$ and $a_j(t)$ generated from an orthogonal spatial expansion of a zero mean random field are uncorrelated time processes if the orthogonal function set satisfies the KL condition

$$\int_A R_f(t_1, t_2, \mathbf{r}_1, \mathbf{r}_2) \Phi_i(\mathbf{r}_2) d\mathbf{r}_2 = \gamma(t_1, t_2) \Phi_i(\mathbf{r}_1)$$

for some deterministic function set $\gamma(t_1, t_2)$.

- 1.19 (a) Show that a single plane wave field defined over an area A with a random envelope function $a(t)$ is always a coherence-separable field over this area.
- (b) Find the condition required for an arbitrary random field having the general orthogonal expansion in Eqs (1.7.5), with $\{a_i(t)\}$ as uncorrelated random processes, to be coherence separable.

REFERENCES

1. R. Gagliardi, *Introduction to Communication Engineering*, 2nd ed. Wiley, New York, 1988.
2. B. Lengyel, *Lasers*, 2nd ed., Wiley, New York, 1971.
3. B. Wherrett, *Laser Advances and Applications*, Wiley, New York, 1980.
4. A. Yariv, *Quantum Electronics*, 2nd ed., Wiley, New York, 1975.
5. M. Born and E. Wolf, *Principles of Optics*, 4th ed., Pergamon Press, London, 1970.
6. J. Kraus, *Antennas*, McGraw Hill, New York, 1950.
7. S. Karp, R. Gagliardi, S. Moran, and L. Stotts. *Optical Channels*, Plenum Press, New York, 1988.
8. J. Palais, *Fiber Optic Communications*, 2nd ed., Prentice Hall, Englewood Cliffs, NJ, 1988.
9. H. Killen, *Fiber Optic Communications*, Prentice Hall, Englewood Cliffs, NJ, 1991.
10. R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. 1, Wiley, New York, 1953.
11. C. Helstrom, Mode decomposition in aperture fields, *J. Opt. Soc. Am.*, 60(4), April, 1970.
12. E. O'Niell, *Introduction to Statistical Optics*, Addison-Wesley, Reading, MA, 1967.
13. R. Gagliardi and S. Karp, *Optical Communications*, 1st ed., Wiley, New York, 1978. Appendix A.
14. G. Agrawal, *Fiberoptic Communication Systems*, McGraw Hill, New York, 1992. Chapter 8.

15. Y. Yamamoto and T. Mukai, Fundamentals of optical amplifiers, *Opt. Quant. Electron.*, 27, pp. S1-S14, May 1989).
16. N. Olsson, Lightwave systems with optical amplifiers, *J. Lightwave Technol.*, 9, 505-513, March 1989.
17. W. Louisell, *Radiation and Noise in Quantum Electronics*, McGraw Hill, New York, 1964.
18. R. Loudon, Theory of noise in linear optical amplifier chains, *IEEE J. Quant. Electron.* QE-21, 766-773, April 1985.
19. B. Saleh and M. Teich, *Fundamentals of Photonics*, Wiley, New York, 1991.