NEAR EAST UNIVRSITY

Faculty of Engineering

Department of Electrical and Electronics Engineering

FIBER OPTIC COMMUNICATION

Graduation Project
EE- 400

Student: OMER SHAHEEN ANSARI (981224)

Supervisor: Prof. Dr Fakhreddin Mamedov

NICOSIA 2003

ACKNOWLEDGMENT

First of all I like to thanks Allah (God) for the courage, He gave me for the completion of my project and engineering.

I wish to thanks to my patents who supported and encouraged me at every stage of my education and who still being generous for me as they are ever.

I would like to thanks my honorable supervisor Prof. Dr. Fakhreddin Mamedov also who was very generous with his help, valuable advises to accomplish this project and who will be always my respectful teacher.

All my thanks go to N.E.U. educational staff especially to electrical & electronic engineering teaching staff for their generosity and especial concern of me and all E.E. students.

Final acknowledgments go to my classmates and friends Mohammad Imran, Mohammad Ali, Atif Munir, Dawood, Ayoub and Raja Saqib who provided me with their valuable suggestions throughout the completion of my project.

ABSTRACT

Fiber optic system consists of transmitter, channel and receiver. Transmitter is consist of modular and circuitry that produce carrier and in fiber optic communication light is the carrier. Light sources are from two elements i.e. light emitting diode (LED) and injected light diode (ILD). LED is used for short distances while ILD is used for long ones.

An optical fiber consists of a core and a cladding layer. According to core diameters there are single mode and multimode fibers. Fiber attenuation and dispersion limit the transmission capacity and distance. Attenuation is caused by photon absorption, scattering, fiber bending and coupling.

There are two types of semiconductor receivers i.e. *p-i-n* (*p*-intrinsic-*n*) diode and APD (Avalanche Photo-Diode). Light-detectors, Photo-detectors, or optical-to-electrical converters are the form of receivers. Photo-detectors convert incident light into current from photon absorption and electron hole pair (EHP) generation. Noise and distortion are important performance limiting factors in signal detection. Thermal noise is white Gaussian noise due to random thermal radiations. Shot noise is caused by random EHP generations in photodiodes. APD noise is due to both random primary an secondary EHP generations. Phase noise is the phase fluctuation. Mode partition noise is caused by random power distribution among longitudinal modes.

INTRODUCTION

The aim of this project is to provide the knowledge of the components and subsystems, which make up fiber systems and of a wide variety of implemented and proposed applications for fiber technology. Optical Fiber Transmission system is a new technology which will have a large impact on near and far telecommunication feature, telecommunication networks, as well as video transmissions and computer interconnections. It provides several major advantages over conventional electronic transmission system. This includes immunity to electromagnetic interference, thinner and lighter cables, lower transmission losses and wider bandwidths.

The first chapter of this project is an introduction of fiber optic systems. It overviews fiber optic cables and cable specifications.

Chapter two is an overview of components and subsystems including sources and transmitters components. Describing LEDs and Lasers, their input-output characteristics, and coupling sources to fibers. Transmitter modules and circuitry diagram are explained.

Chapter three represents optical detectors and receivers. It explains *p-i-n* and APD detectors, receiver modules and APD control circuitry.

Chapter four surveys a wide range of existing and hypothetical applications, illustrating what is being done with currently available components and subsystems, and what might be done as a result of future developments. Applications addressed include the public network, military, industrial and computer systems.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION TO FIBRE OPTIC COMMUNICATION SYSTEMS

The fiber-optic is defined as branch of optics that deals with the transmission of light through ultra pure glass, plastic or some other form of transparent media. One of first noted experiment that demonstrated the transmission of light through a dielectric medium has been created to John Tyndall. In 1854 John Tyndall demonstrated that light could be guided through stream of water based on the principle of total internal reflection.

In 1880 Alexander Graham Bell invented the photo phone, a device that transmits voice signals over a beam of light.

Charles H. Townes created great interest in communication at optical frequencies in 1958 with the invention of the laser.

In 1966 Charles K. Kao and George Hockham of Standard Telecommunications Laboratories of England performed several experiments to prove that, if glass could be made more transparent by reducing its impurities, light loss could be minimized. Their research led to a publication in which they predicted that optical fiber could be made pure enough to transmit light several kilometers. In the next two decades researchers worked intensively to reduce the attenuation to 0.16 dB/km.

In 1988 the American National Standards Institute (ANSI) published the Synchronous Optical Network (SONET).

1995 Multimedia applications for business have become the major impetus for increased use of optical fiber within the LAN, MAN, and WAN environment.

1.1 Advantages and Disadvantages of the FOS

a) Advantages

The major advantages are:

Bandwidth One of the most significant advantages that fiber has over copper or other transmission media is a bandwidth. Bandwidth is directly related to the amount of information that can be transmitted per unit time. Today's advanced fiber optic systems are capable of transmitting several gigabits per second over hundreds of kilometers. Ten thousands of voice channels can now be multiplexed together and sent over a single fiber strand.

Less Lose. Currently, fiber is being manufactured to exhibit less than a few tenths of a decibel of loss per kilometer.

Less Weight and Volume. Fiber optic cables are substantially lighter in weight and occupy; much less volume than copper cables with the same information capacity. For example, a 3-inch diameter telephone cable consisting of 900 twisted-pair wires can be replaced with a single fiber strand 0.005 inch in diameter (approximately the diameter of a hair strand) and retain the same information-carrying capacity. Even with a rugged protective jacket surrounding the fiber, it occupies enormously less space and weights considerably less.

Security. Since light does not radiate from a fiber optic cable, it is nearly impossible to secretly tap into it without detection. For this reason, several applications requiring communication security employ fiber-optic systems. Military information, for example, can be transmitted over fiber to prevent eavesdropping. In addition, metal detectors cannot detect fiber-optic cables unless they are manufactured with steel reinforcement for strength.

Flexibility. The surface of glass fiber is much more refined than ordinary glass. This coupled with its small diameter, allows it to be flexible enough to wrap around a pencil. In terms strength, a 0.005-inch. strand of fiber is strong enough to cut one's finger before it breaks, if enough pressure is applied against it.

Economics. Presently, the cost of fiber is comparable to copper at approximately $0.20 to $0.50 per yard and is expected to drop as it becomes more widely used. Since transmission losses are considerably less than for coaxial cable, expensive repeaters can be spaced farther apart. Fewer repeaters mean a reduction in overall system cost and enhanced reliability.

Reliability. Once installed, a longer life span is expected with fiber over its metallic counterparts since it is more resistant to corrosion caused by environmental extremes such as temperature, corrosive gases, and liquids.

b) Disadvantages

In spite of the numerous advantages that fiber optic systems have **over** conventional methods of transmission, there are some disadvantages, particularly because of its newness. Many of these disadvantages are being overcome with new and competitive technology.

Interfacing costs. Electronic facilities must be converted to optics in order to interface to fiber. Often these costs are initially overlooked. Fiber-optic transmitter, receiver, couplers, and connectors, for example, must be employed as part of the communication system. Test and repair equipment is costly. If the fiber optic cable breaks, splicing can be a costly and tedious task

Strength. Fiber, by itself, has a tensile strength of approximately 1 lb, as compared the coaxial cable at 180 lb (RG59U) surrounding the fiber with stranded Kevlar and a protective PCV jacket can increase the pulling strength up to 500 lb. Installations requiring greater tensile strengths can be achieved with steel reinforcement.

Remote Powering of Devices. Occasionally it is necessary to provide electrical power to a remote device. Since this cannot be achieved through the fiber, metallic conductors are often included in the cable assembly. Several manufacturers now offer a complete line of cable types, including cables manufactured with both copper wire and fiber.

1.2 Theory of Light

In the seventeenth and eighteenth centuries, there were two schools of thought regarding the nature of light. Sir Isaac Newton and his followers believed that light consisted of rapidly moving particles (or corpuscles), whereas Dutch physicist Christian Huygens regarded light as being a series of waves.

The wave theory was strongly supported by an English doctor named Thomas Young. By 1905, quantum theory, introduced by dark Maxwell, showed that when light is emitted or absorbed it is not only as a wave, but also as an electromagnetic particle called a photon. Photon is said to possess energy that is proportional to its frequency. This is known as Planck's law, which states:

$$E = h \times v, \tag{1.1}$$

Where E = photon's energy (J),

h = plank's constant, 6.69×10^{-34} (J-s),

v = frequency of photon (Hz).

Using the particle theory, Einstein and Planck were able to explain photoelectric effect: when visible light or electromagnetic radiation of a hire frequency shines on a metallic surface, electrons are emitted, which is turning an electric current.

Electromagnetic Spectrum

Fundamentally, light has been accepted as a form of electromagnetic radiation that can be categorized into a portion of the entire electromagnetic spectrum, as shown in Table 1.1. In addition, each frequency can be specified in terms of its equivalent wavelength. Frequency or wavelength are directly related to the speed of light.

$$C = f \times \lambda \tag{1.2}$$

Where C - speed of light in a vacuum or free space, 3×10^8 (m/s)

f - frequency (Hz);

λ - wavelength (m).

Table 1.1 Electromagnetic spectrum

Range of wavelength, nm	Name of wavelength	
10^6 - 770	*Infrared*	
770 – 662	Red	*Visible*
662 – 597	Orange	*Visible*
597 – 577	Yellow	*Visible*
577 – 492	Green	*Visible*
492 – 455	Blue	*Visible*
455 – 390	Violet	*Visible*
390 –10	*Ultraviolet*	

The portion of the electromagnetic spectrum regarded as light has been expanded in Table 1.1 to illustrate three basic categories of light:

1, Infrared: that portion of the electromagnetic spectrum having a wavelength ranging from 770 to 10^6 nm. Fiber optic systems operate in this range.

2. ***Visible:*** that portion of the electromagnetic spectrum having a wavelength ranging from 390 to 770 nm. The human eye, responding to these wavelengths allows us to see the colors ranging from violet to red, respectively.

3. Ultraviolet: that portion of the electromagnetic spectrum ranging from 10 to 390 nm.

The light that we use for most fiber optic systems occupies a wavelength range from 800 to 1600 nm. This is slightly larger than visible red light and falls within the infrared portion of the spectrum.

Snell's Law : Total Interval Reflection

For light to propagate in any medium, the medium must be transparent to some degree. The degree of transparency determines how far light will propagate.

Transparent materials can be in the form of a liquid, gas, or a solid. Some examples are glass, plastic, air, and water.

One of the most fundamental principles of light is that when it strikes the interface between two transparent mediums, such as air and water, a portion of the light energy is reflected back into the first medium and a portion is transmitted into the second medium. The path in which light travels from one point to another is commonly referred to as the ray. Figure 1.1 illustrates the classic example of a ray of light incident upon the surface of water. Notice that part of the light is *reflected* off the surface of water and part of it penetrates the water. The ray penetrating to water is said to be refracted or bent toward the normal. The amount of refracted light is determined by the medium's index of refraction, generally denoted by the letter n. Index of refraction is the ratio of the speed of light in a vacuum - c, to the speed of light in the given medium - v. This relationship is given by the equation: $n = c / v$. Since the speed of light is lower in mediums other than a vacuum, the index of refraction in such mediums is always greater than 1.

Example for air $n = 1.003$, for water $n = 1.33$, for fiber-optic $n = 1.6$.

In 1621, the Dutch mathematician Willebrard Snell established that rays of light could be traced as they propagate from one medium to another based on their indices of refraction. Snell's low is stated by the equation:

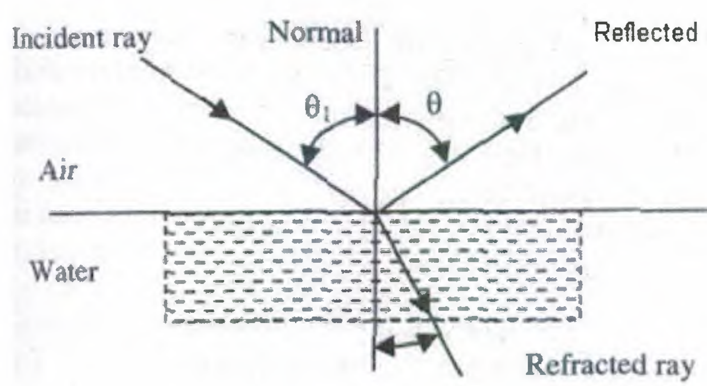

Figure 1.1 Theory of Light.

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \qquad (1.3)$$

$$\frac{n_1}{n_2} = \frac{\sin \theta_2}{\sin \theta_1}$$

Where n_1 - refractive index of material; θ_1 - angle of incidence; θ_2 - angle of refraction; n_2 - refractive index of material 2. When the angle of incidence, θ_1, becomes large enough to cause the sine of the refraction angle, θ_2, to exceed the value of 1, total internal reflection occurs. This angle is called the critical angle, θ_c. The critical angle, θ_c, can be derived from Snell's law as follows

$$n_i \sin \theta_1 = n_2 \sin \theta_2$$

$$\sin \theta_1 = n_2 \sin \theta_2 / n_1$$

When $\sin \theta_1 = \sin \theta_2$, then $\sin \theta_1 = n_2 / n_1$. Therefore, critical angle:

$$\theta_c = \sin^{-1} (n_2 / n_1)$$

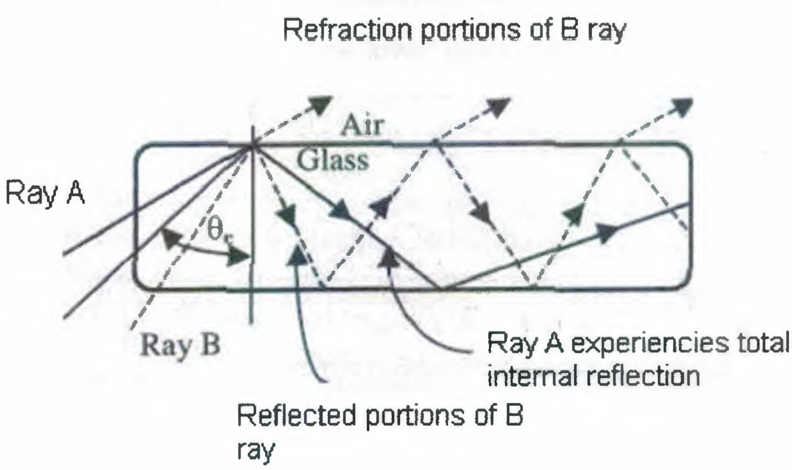

Figure 1.2 Refraction of light.

By surrounding glass with material whose refraction index is less than that of the glass, total internal reflection can be achieved. This is illustrated in Figure 1.2. Ray A penetrates the glass-air interface at an angle exceeding the critical angle, θ_c, and therefore experiences total internal Action. On the other hand, Ray B penetrates the glass air interface at an angle less than the critical angle. Total internal reflection does not occur. Instead, a portion of ray B escapes the glass and is refracted away from the normal as it enters the less dense medium of air. A portion is also reflected back into the glass. Ray B diminished in magnitude as it bounces back and forth between the glass-air interface. The foregoing principle is the basis for guiding light through optical fiber. Two key elements that permit light guiding through optical fibers are its core and its cladding fiber's core is manufactured of ultra pure glass (silicon dioxide) or plastic. Surrounding the core material called cladding. A fiber cladding is also made of glass or plastic. Its index of refraction, however, it is typically 1% less than that of its core. This permits internal reflection of rays entering the fiber and striking the core-cladding interface above critical angle of approximately 82-degree ($\sin^{-1}$ (1/1.01). The core of the fiber therefore guides the light and the cladding contains the light. The cladding material is much less transparent then the glass making up the core of the fiber. This causes light rays to be absorbed if they strike the core-cladding interface at an angle less than the critical angle.

In Figure 1.3, a light ray is transmitted into the core of an optical fiber. Total internal reflection occurs as it strikes the lower index cladding material.

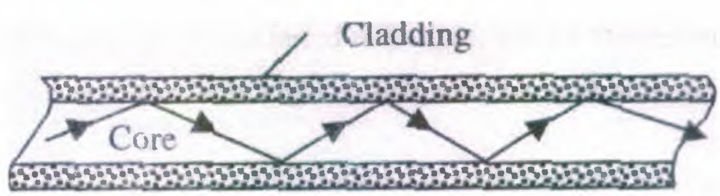

Figure 1.3 Light ray travels in core and refracts when strikes at cladding.

1.3 Block Diagram of the FOS

One of the main limitations of communication systems is their restricted information carrying capabilities. In more specific terms what this means is that the communications medium can only carry so many messages. And, as you have seen, this information-handling ability is directly proportional to the bandwidth of the communications channel. In telephone systems, bandwidth is limited by the characteristics of the cable used to carry the signals. As the demand for telephones has increased, better cables and wiring systems have been developed. Further, multiplexing techniques have been used to transmit multiple telephone conversations over a single cable.

In radio communication systems, the information modulates a high frequency carrier. The modulation produces sidebands, and therefore, the signal occupies a narrow portion of the RF spectrum. However, the RF spectrum is finite. There is only so much space for radio signals. To increase the information capacity of a channel, the bandwidth of the channel must be increased. This reduces available spectrum space. Multiplexing techniques are used to send more signals in a given channel bandwidth, and methods have been developed to transmit more information in less bandwidth.

The information-carrying capacity of the radio signal can be increased tremendously if higher carrier frequencies are used. As the demand for increased communications capacity has gone up over the years, higher and higher RF's are being used. Today, microwaves are the preferred radio channels for this reason, but it is more complex and expensive to use these higher frequencies because of the special equipment required.

One way to expand communications capability further is to use light as the transmission medium. Instead of using an electrical signal traveling over a cable or electromagnetic waves traveling through space, the information is put on a light beam and transmitted through space through a special cable. In the late nineteenth century, Alexander Graham Bell, the inventor of the telephone, demonstrated that information could be transmitted by light.

Light beam communication was made more practical with the invention of the laser. The laser is a special high-intensity, single frequency light source. It produces a very narrow beam of brilliant light of a specific wavelength (color). Because of its great intensity, the laser beam can penetrate atmospheric obstacles better than other types of light, thereby making light-beam communication more reliable over longer distances. The primary problem with such free-space light beam communication is that the transmitter and receiver must be perfectly aligned with one another.

Instead of using free space, some type of light carrying cable can also be used. For centuries it has been known that light is easily transmitted through various types of transparent media such as glass and water, but it wasn't until the early in 1900s that scientist were able to develop practical light carrying media. By the mid-1950s glass fibers were developed that permitted long light carrying cables to be constructed. Over the years, these glass fibers have been perfected. Further, low cost plastic fiber cable also developed. Developments in these cables permitted them to be made longer with less attenuation of the light.

Today the fiber optic cables have been highly refined. Cables many miles long can be constructed and interconnected for the purpose of transmitting information on a light beam over very long distances. Its great advantage is that light beams have an incredible information carrying capacity. Whereas hundreds of telephone conversations may be transmitted simultaneously at microwave frequencies, many thousands of signals can be carried on a light beam through a fiber optic cable. Using multiplexing techniques similar to those used in telephone and radio systems, fiber optic communications systems have an almost limitless capacity for information transfer.

The components of a typical fiber optic communications system are illustrated in Figure 1.4.

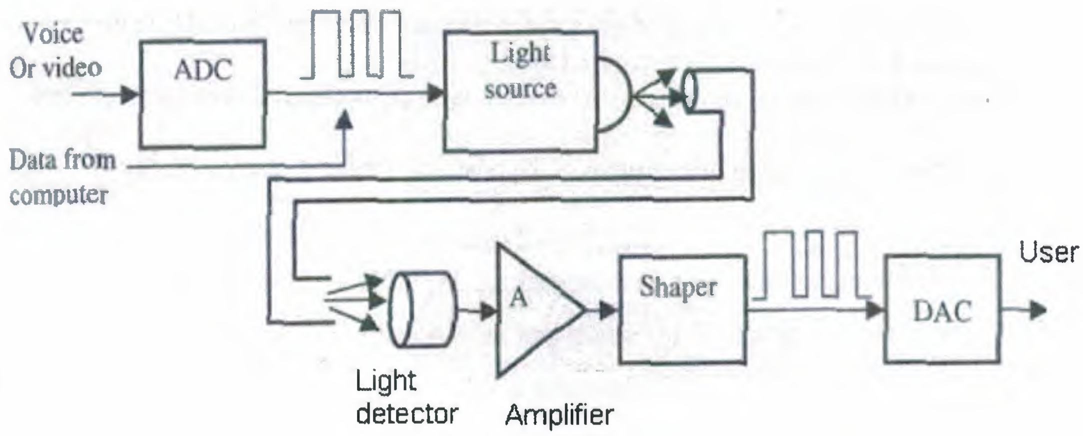

Figure 1.4 Fiber Optic Systems.

The information signal to be transmitted may be voice, video, or computer data. The first step is to convert the information into a form compatible with the communications medium. This is usually done by converting continuous analog signals such as voice and video (TV) signals into a series of digital pulses. An Analog-to-Digital Converter (ADC) is used for this purpose. Computer data is already in digital form. These digital pulses are then used to flash a powerful light source off and on very rapidly. In simple low cost systems that transmit over short distances, the light source is usually a light-emitting diode (LED). This is a semiconductor device that puts out a low intensity red light beam. Other colors are also used. Infrared beams like those used in TV remote controls are also used in transmission. Another commonly used light source is the laser emitting diode. This is also a semiconductor device that generates an extremely intense single frequency light beam.

The light beam pulses are then fed into a fiber optic cable where they are transmitted over long distances. At the receiving end, a light sensitive device known as a photocell or light detector is used to detect the light pulses. This photocell or photo detector converts the light pulses into an electrical signal. The electrical pulses are amplified and reshaped back into digital form. They are fed to a decoder, such as a

Digital-to-Analog Converter (DAC), where the original voice or video is recovered for user.

1.4 Fiber Optic Cables

Just as standard electric cables come in a variety of sizes, shapes, and types, fiber optic cables are available in different configurations. The simplest cable is just a single strand of fiber, whereas complex cables are made up of multiple fibers with different layers and other elements. The portion of a fiber optic cable (core) that carries the light is made from either glass or plastic. Another name for glass is silica. Special techniques have been developed to create nearly perfect optical glass or plastic, which is transparent to light. Such materials can carry light over a long distance. Glass has superior optical characteristics over plastic. However, glass is far more expensive and more fragile than plastic. Although the plastic is less expensive and more flexible, its attenuation of light is greater. For a given intensity, light will travel a greater distance in glass than in plastic. For very long distance transmission, glass is certainly preferred. For shorter distances, plastic is much more practical.

All fibers consist of a number of substructures including (see Figure 1.5):

• A core, which carries most of the light, surrounded by

• A cladding, which bends the light and confines it to the core, surrounded by

• A substrate layer (in some fibers) of glass which does not carry light, but adds to the diameter and strength of the fiber, covered by

• A primary buffer coating, which provides the first layer of mechanical protection, covered by

• A secondary buffer coating, which protects the relatively fragile primary coating

•

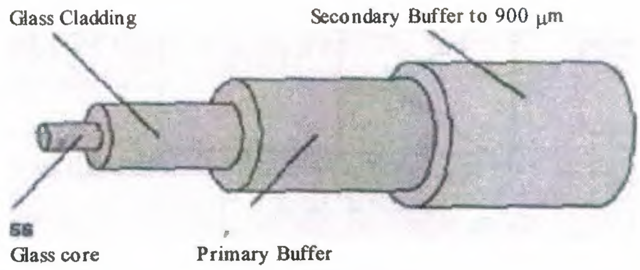

Figure 1.5 Fiber Optic Cable

The cladding is also made of glass or plastic but has a lower index of refraction. This ensures that the proper interface is achieved so that the light waves remain within the core. In addition to protecting the fiber core from nicks and scratches, the cladding adds strength. Some fiber optic cables have a glass core with a glass cladding. Others have a plastic core with a plastic cladding. Another common arrangement is a glass core with a plastic cladding. It is called plastic-clad silica (PCS) cable.

1.4.1 Basic Construction of the Fiber-Optic Cables

There are two basic ways of classifying fiber optic cables. The first way is an indication of how the index of refraction varies across the cross section of the cable. The second way of classification is by mode. Mode refers to the various paths that the light rays can take in passing through the fiber. Usually these two methods of classification are combined to define the types of cable. There are two basic ways of defining the index of refraction variation across a cable. These are step index and graded index. Step index refers to the fact that there is a sharply defined step in the index of refraction where the fiber core and the cladding interface. It means that the core has one constant index of refraction N1, while the cladding has another constant index of refraction N2.

The other type of cable has a graded index. In this type of cable, the index of refraction of the core is not constant. Instead, the index of refraction varies smoothly and continuously over the diameter of the core. As you get closer to the center of the core, the index of refraction gradually increases, reaching a peak at the center and then declining as the other outer edge of the core is reached. The index of refraction of the cladding is constant.

Mode refers to the number of paths for the light rays in the cable. There are two classifications: single mode and multimode. In single mode, light follows a single path through the core. In multimode, the light takes many paths through the core.

Each type of fiber optic cable is classified by one of these methods of rating the index or mode. In practice, there are three commonly used types of fiber optic cable: multimode step index, single mode step index and multimode graded index cables.

1. The multimode step-index fiber. This cable (see Figure 1.6 (a)) is the most common and widely used type. It is also the easiest to make and, therefore, the least expensive. It is widely used for short to medium distances at relatively low pulse frequencies.

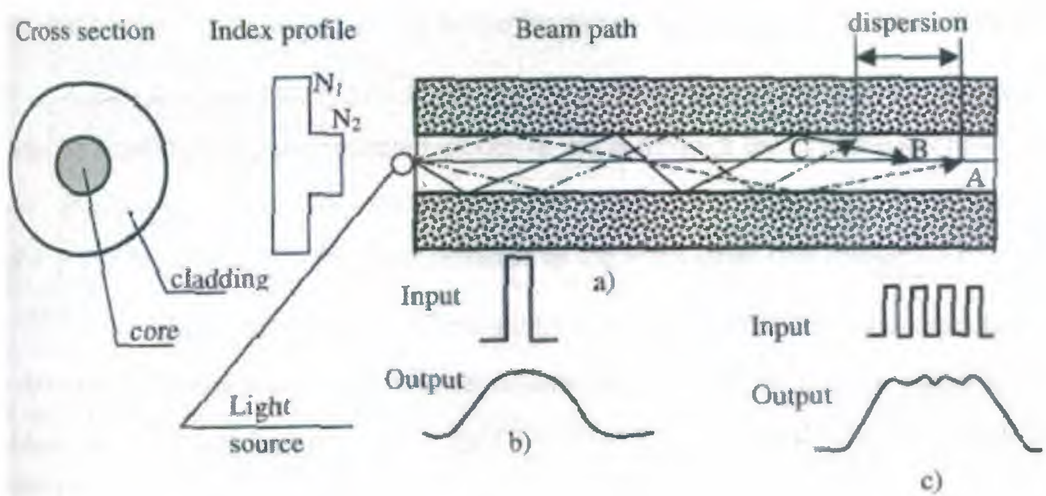

Figure 1.6 Multimode Step-Index Fiber Optic.

The main advantage of a multimode step index fiber is the large size. Typical core diameters are in the 50-to-1000 micrometers (μm) range. Such large diameter cores are excellent at gathering light and transmitting it efficiently. This means that an inexpensive light source such as LED can be used to produce the light pulses. The light takes many hundreds of even thousands of paths through the core before exiting. Because of the different lengths of these paths, some of the light rays take longer to reach the end of the cable than others. The problem with this is that it stretches the light pulses (Figure 1.6 (b). In Figure 1.6 ray A reaches the end first, then B, and C. The result is a pulse at the other end of the cable that is lower in amplitude due to the attenuation of the light in the cable and increased in duration due to the different arrival times of the various light rays. The stretching of the pulse is referred to as modal dispersion. Because the pulse has been stretched, input pulses cannot occur at a rate faster than the output pulse duration permits. Otherwise the pulses will essentially merge together as shown in Figure 1.6 (c). At the output, one long pulse will occur and will be indistinguishable from the three separate pulses originally transmitted. This means that incorrect information will be received. The only core for this problem is to reduce the pulse repetition rate. When this is done, proper operation occurs. But with pulses at a lower frequency, less information can be handled.

2. Single mode cable. In a single mode, or mono-mode, step-index fiber cable the core is so small that the total number modes or paths through the core are minimized and modal dispersion is essentially eliminated. The typical core sizes are 5 to 15 μm. The output pulse has essentially the same duration as the input pulse (see Figure 1.7).

The single mode step index fibers are by far the best since the pulse repetition rate can be high and the maximum amount of information can be carried. For very long distance transmission and maximum information content, single-mode step-index fiber cables should be used.

The main problem with this type of cable is that because of its extremely small size, it is difficult to make and is, therefore, very expensive. Handling, splicing, and making interconnections are also more difficult. Finally, for proper operation an expensive, super intense light source such as a laser must be used. For long distances, however, this is the type of cable preferred.

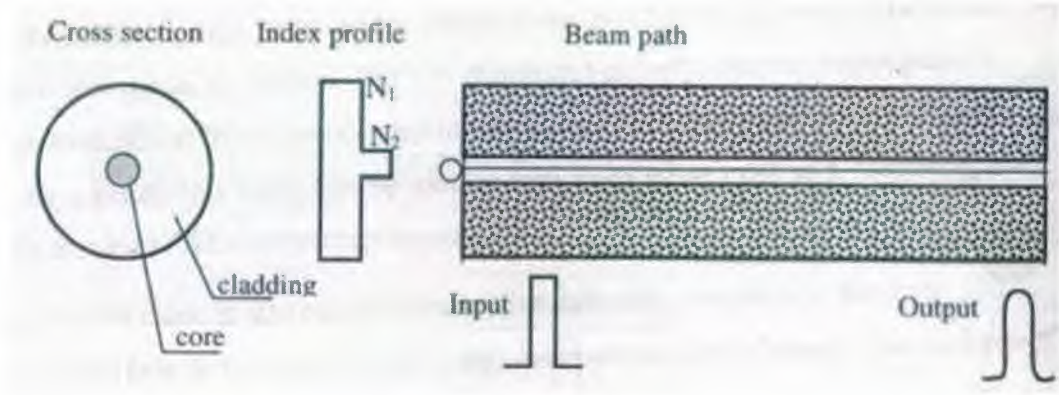

Figure 1.7 Single Mode Fiber Optic.

3. Multimode graded-index fiber cables. These cables have a several modes or paths of transmission through the cable, but they are much more orderly and predictable. Figure 1.8 shows the typical paths of the light beams. Because of the continuously varying index of refraction across the core, the light rays are bent smoothly and converge repeatedly at points along the cable. The light rays near the edge of the core take a longer path but travel faster since the index of refraction is lower.

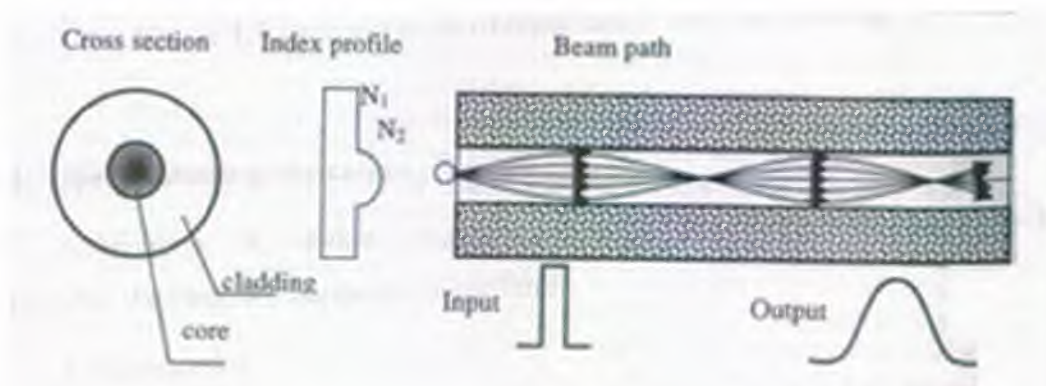

Figure 1.8 Multimode Graded-Index Fiber Optic.

All the modes or light paths tend to arrive at one point simultaneously. The result is that there is less modal dispersion. It is not eliminated entirely, but the output pulse is not nearly as stretched as in multimode step index cable. The output pulse is only slightly elongated. As a result, this cable can be used at very high pulse rates and, therefore, a considerable amount of information can be carried on it.

This type of cable is also much wider in diameter with core sizes in the 50 to 100 (μm) range. Therefore, it is easier to splice and interconnect, and cheaper, less-intense light sources may be used. The most popular fiber-optic cables that are used in LAN: Multimode-step index cable - 65.5/125; multimode-graded index cable - 50/125. The multimode-graded index cable - 100/140 or 200/300 are recommended for industrial control applications because its large size. In high data rate systems is used single mode fiber 9/125. Typical core and cladding diameters of these cables are shown in Figure 1.9.

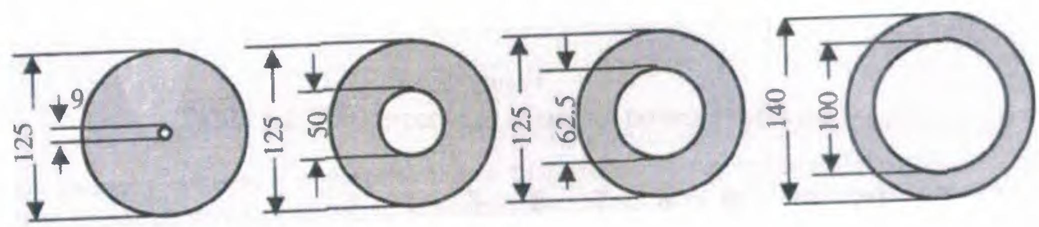

Figure 1.9 Different types of diameters of core and cladding.

1.4.2 Specifications of the cables

Attenuation, A, db/km; Numeric aperture, NA and Dispersion, ns/km, characterize the fiber as a transmission medium.

a) Attenuation

The main specification of a fiber optic cable is its attenuation. Light power, which does not reach the other end of the fiber, has either left the fiber or been absorbed (converted to heat) in it. The amount of attenuation varies with the type of cable and its size. Glass has less attenuation than plastic. Wider cores have less attenuation than narrower cores. But more importantly, the attenuation is directly proportional to the length of the cable. It is obvious that the longer the distance the light has to travel the greater the loss due to absorption, scattering, and dispersion. Doubling the length of a cable doubles the attenuation, and so on.

The attenuation of a fiber optic cable is expressed in decibels per unit of length. The standard specification for fiber-optic cable is the attenuation expressed in terms of decibels per kilometers. The standard decibel formula used is

$$\text{Loss, dB} = 10 \log (P_o/P_i) \tag{1.3}$$

where P_o is the output power and P_i is the input power.

The table 1.2 shows the percentage of output power for various decibel loss. The attenuation ratings of fiber-optic cables vary over a considerable range.

The table 1.2 shows the percentage of output power for various decibel loss.

Table 1.2 The percentage of output power expressed by dB

Loss, (dB)	1	2	3	4	5	6	7	8	9	10	20	30
Po (%)	79	63	50	40	31	25	20	14	12	10	1	0.1

The finest single mode step-index cables have an attenuation of only 1 dB/km. However, a very large core plastic fiber cables can have an attenuation of several thousands decibels per kilometer. Temperature dependence of the attenuation of fiber optic cable is shown in Figure 1.10

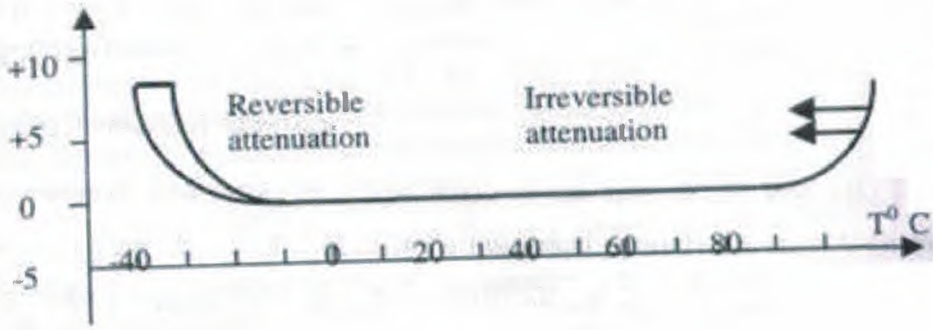

Figure 1.10 Reversible and Irreversible attenuation.

The following contribute to the attenuation:

Reyleigh-scattering. A mechanism called Rayleigh scattering prevents any further improvement in attenuation loss. Rayleigh scattering is caused by micro irregularities in the random molecular structure of glass. These irregularities are formed as the fiber cools from a molten state. Normally, electrons in glass molecules interact with transmitted light by absorbing and reradiating light at the same wavelength. A portion of the light, however, strikes these micro irregularities and becomes scattered in all directions of the fiber, some of which is lost in the cladding. Consequently the intensity of the beam is diminished.

Radiation Losses.

A phenomenon called micro bending can cause radiation losses in optical fibers in excess of its intrinsic losses. Micro bends are miniature bends and geometric imperfections along the axis of the fiber that occur during the manufacturing or installation of the fiber. Mechanical stress such as pressure, tension, and twist can cause micro bending. This geometric imperfection causes light to get coupled to various unguided electromagnetic modes that radiate and escape the fiber.

Absorption. The following contribute to the absorption: Intrinsic impurities, irregularities in core diameter, IR-absorption (infrared), OH- absorption (hydroxy, humidity) and molecular agitation.

b) Numerical Aperture

Numerical Aperture tells how much of the light can be pass into the fiber. An important characteristic of a fiber is its numerical aperture (NA). NA characterizes a fiber's light-gathering capability. Mathematically, it is defined as the sine of half the angle of a fiber's light *acceptance cone.* For multimode step index fiber

$$NA = \sqrt{N^2_1 - N^2_2} \qquad (1.4)$$

Typical values for NA are 0.25 to 0.4 for multimode step-index fiber and 0.2 to 0.3 for multimode graded-index fiber.

c) Dispersion:

Dispersions classified: material dispersion, Ψ_{mat} (ns/km) and modal dispersion, Ψ_{mod} (ns/km)

Material dispersion.

A light pulse is composed of light of different wavelengths depending on the spectral width of the light source. The refractive index depends weakly on the wavelength. This causes the material dispersion.

Modal dispersion. As shown above the modal dispersion due to the different arrival times of the various light rays.

The Table 1.3 shows the characteristics of the dispersions.

Table 1.3 Characteristics of the dispersions

Fiber type	Dispersion (ns / km) Modal	Material
Step index	$\Psi_{mod} = t * (\Delta / 2)$	$\Psi_{mat} = 0.1\ \Delta\lambda$
Graded index	$\Psi_{mod} = t * (\Delta^2 / 2)$	
Single mode	$\Psi_{mod} = 0$	

Note: t- traveling time per km $t=N/C$, for $N=1.5$, $t= 5$ µs/km;

$\Delta\lambda, = \Delta N/N$, in practice $\Delta=0.01$

$\Delta\lambda$ - Bandwidth of the light;

The total dispersion equals $\Psi_{total} = (\Psi_{mat} + \Psi_{mod})^{1/2}$

CHAPTER 2

OPTICAL SOURCES AND TRANSMITTERS

2.1 LEDs and Lasers

To construct an optical communications system one requires a source of optical power and a means for modulating that source. A suitable source for an optical fiber communications system must have certain characteristics, which include the following: emission at a wavelength within a window of low fiber transmission loss, efficient conversion of prime power to light coupled into the fiber, high reliability, ease of modulation, adequate modulation speed capability, sufficient ruggedness, ease of coupling the source output into the fiber, adequately low noise, adequately high linearity of modulation, sufficiently narrow spectral width (range of wavelengths in the emitted light), and other more subtle requirements. No source can provide ideal characteristics, but the requirements above eliminate many candidate sources either because they are totally impractical for the fiber system application or they are clearly inferior to other alternatives. (Obviously all of these requirements must be met at an acceptable cost if the source and the system are to be practical.) In the early days of fiber optic system research two candidate sources emerged as adequately meeting most or all of the requirements listed above. These are the semiconductor light-emitting diode (LED) and the semiconductor injection laser diode (ILD). (There was a third source which was given some consideration as a possible candidate, but was dropped because it appeared to be more expensive to fabricate and more difficult to modulate. This was the miniature Nd-YAG laser pumped by light-emitting diodes. Because of its long intrinsic time constants, it requires an external modulator for most applications. It is possible that interest in this device will reemerge in the future.)

The LED and ILD are both solid-state semiconductor devices, which can be fabricated by batch processes. They can be fabricated from various semiconductor material systems, which allow the device designer to select the desired wavelength of emission. In particular, devices fabricated in the gallium—aluminum- arsenide material system can emit in the range of wavelengths between 0.8 and 0.9 μm. Devices fabricated in the indium—gallium—arsenide—phosphide material system can emit in the range of wavelengths between 1.0 and 1.6 μm. Both LEDs and ILDs can be modulated by varying the electrical current used to power the devices (direct modulation). The achievable direct, modulation rates range from 20 MHz to beyond 1 GHz for LEDs (depending upon the materials, the device design, and tradeoffs against other parameters), and up to 5—10 GHz for the fastest ILDs. Although the amount of light power coupled into the fiber is typically a small fraction of the drive power, this electrical-to-optical conversion efficiency is adequate for most applications (and is typically much better than what can be obtained with alternative devices). The spectral width of an LED is relatively large, which limits its range of applications. The spectral width of a laser can be very small (a single frequency of emission) depending upon the device design.

Both LEDs and lasers can have very high reliability, and arc compact, mechanically stable, devices. Lasers are susceptible to damage from electrical abuse and are somewhat sensitive to high temperatures. In summary, both LEDs and ILDs have found important applications in a wide range of systems.

Figure 2.1 shows a schematic drawing of a typical light-emitting diode. The device consists of a number of layers of semiconductor material, some of which are *p* -doped and some of which are *n* -doped as shown. Where the *p* -doped and *n* -doped materials come together one has a *p-n* junction. When holes and electrons are injected into the junction (by applying a current in the forward biased direction) they combine and give up an amount of energy equal to the charge of a electron multiplied by the semiconductor band gap (in volts) between the valence and conduction bands. This energy can be given up as a photon of light or in the form of mechanical lattice vibration (heat).

The simplest type of light-emitting diode would just have two layers— one *p* -doped and one n-doped. Unfortunately such a simple structure would not be capable of efficient conversion of electrical drive power into light power captured by a fiber. There are several reasons for this. The light emitted by the LED can be reabsorbed before it leaves the device.

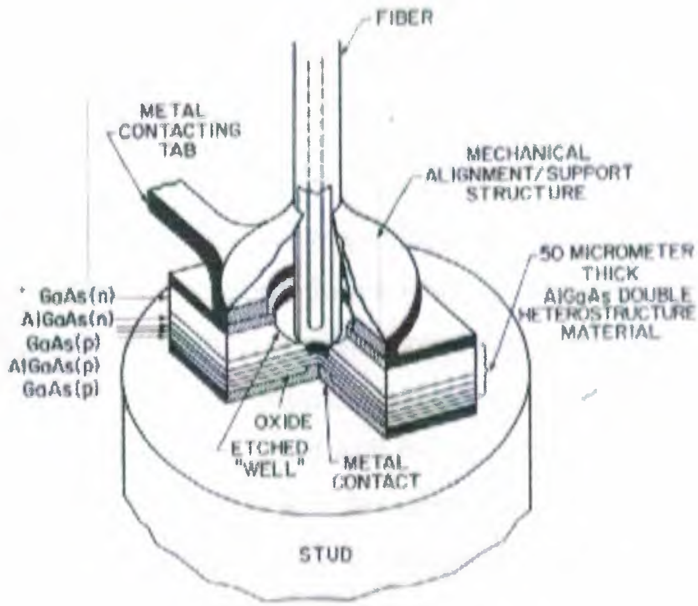

Figure 2.1 Schematic Drawing of a Burrus Type LED.

To minimize this problem a well is etched to allow the fiber to be brought close to the junction. However, in a simple two-layer device the holes and electrons combine in a relatively thick layer on either side of the junction. Furthermore the process of etching the well too close to the junction can introduce mechanical damage to the material, which results in nonradiative recombination (producing heat instead of light). To alleviate these problems a multi layer structure is used, where the layers are made of semiconductor compounds of varying composition. For example, the active layer, shown might be made of pure GaAs material while the layers on either side might be made of GaAlAs where the ratio of gallium to aluminum is 90/10. Note that the aluminum is not a small quantity dopant but is present in substantial percentage. The use of this layered "heterojunction" approach leads to some interesting results. The energy band structures of the layers are different (both the absolute levels and the band gap). As a result potential barriers are formed on either side of the active layer, which confine the holes and electrons to a thin volume within the active layer. Thus all of the photons are created within this thin volume. In addition, the layers containing aluminum are relatively transparent to the light emitted by the active layer. Thus it is not necessary to etch the well all the way down to the active layer (thereby avoiding the nonradiative recombination due to damage). Thus the layered structure allows much more of the light generated within the device to reach the fiber.

The fiber can only capture that portion of the light, which illuminates its core. The use of lenses cannot provide coupling from a large light-emitting area to a small fiber core. Thus it is wasteful of electrical drive power to inject current into the junction over a large area. For this reason a dot contact is defined by an insulating layer to create a column of current of limited area aligned with the well. Finally, to provide efficient heat sinking, the device is often mounted with the substrate up to bring the junction close to the heat sink. A device with this structure is called a Burrus type LED after its inventor **C.A.** Burrus.

Figure 2.2 shows the structure of the typical ILD. In this case we again have a series of heterogeneous layers (heterostructure), but the light is emitted from the side of the device. A laser is an oscillator (as opposed to an LED which is a broadband optical noise source). To obtain oscillation one needs gain, feedback, and saturation. Nature will always provide saturation. We must provide the other two items. To obtain gain one injects sufficient current into the device so that a condition called a population inversion exists in the active layer. When holes and electrons combine there is a brief period of time when they are about to emit a photon of light, but have not done so yet.

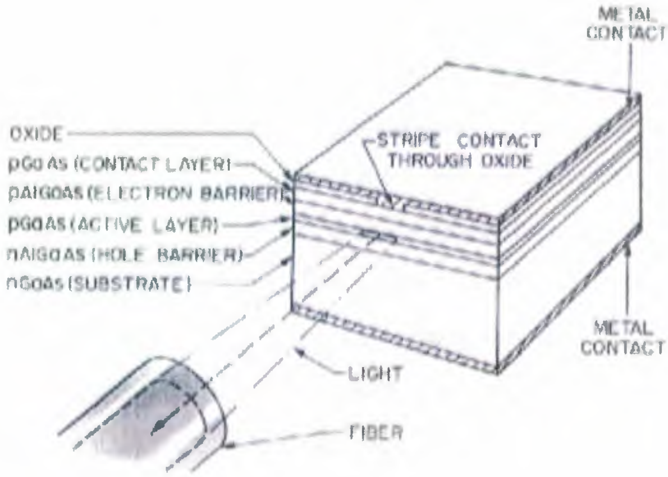

Figure 2.2 Schematic Drawing of a Stripe-Contact Gain-Guided Injection Laser Diode.

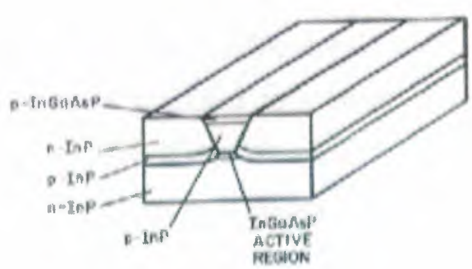

Figure 2.3 Schematic Drawing of an Index-Guided Injection Laser Diode.

If in that time period an existing photon of light passes by, it can stimulate the hole and electron to add their light energy synchronously to the existing field. Thus the existing field grows in amplitude as it travels through the medium. The reverse process can also occur. That is, an existing photon can be absorbed to produce a hole—electron pair. In order to have gain dominate over loss one must inject a sufficient density of hole—electron pairs into the junction to have the stimulated emission process dominate the absorption process.

To provide feedback an optical cavity is formed as follows. For the direction front-to-back, perpendicular surfaces are cleaved (fabricated by scoring and breaking) onto each end of the device. The interface between the semiconductor material and air produces an approximately 30% reflection. This reflection, when combined with the gain in the active layer, is adequate to result in a unity round-trip gain from any point within the active layer to the front face, across the active layer to the back face, and back. The layers on either side of the active layer form feedback in the top-to-bottom direction. By good fortune, they have a lower index of refraction and thus the light is guided within the active layer by total internal reflection. Feedback in the left to right direction is provided by one of two mechanisms. If the cross-section of the layers is uniform left-to-right as shown in Figure 2.2, feedback can be obtained by the higher effective refractive index of the region in which current is flowing (defined by a stripe contact on the top of the device). This is called gain guiding. If the layers have a more complex structure (as shown in Figure 2.3), then the layers themselves form a three-dimensional wave-guide for guiding the light. This is called index guiding.

In order for the device to operate (achieve gain and lasing) at room temperature, the population inversion must be confined to a small volume of space (to limit the injected current per unit area). The layers on either side of the active layer not only provide wave guiding for the light, but form potential barriers which confine carriers to a small cross-sectional area.

Thus with this design we obtain carrier confinement, gain at acceptable current densities, and field confinement (feedback) — all of which are needed to produce a practical laser (oscillator).

2.1.1 Difference Between the LED and the ILD

a) Light Emitting Diode

The major difference between the LED and the ILD is the manner in which light is emitted from each source. The LED is an incoherent light source that emits light in a disorderly way. A laser is a light source that emits coherent monochromatic light. Monochromatic light has a pure single frequency. Coherent refers to the fact that all the light waves emitted are in phase with one another. Coherent light waves are focused into a narrow beam, which, as a result, is extremely intense. The effect is somewhat similar to that of using highly directional antenna to focus radio waves into a narrow beam, which also increases the intensity of the signal. Figure 2.4 illustrates the differences in radiation patterns. Both devices are extremely rugged, reliable, and small in size.

In terms of spectral purity, the LED's half power spectral width is approximately 50 nm, whereas the ILD's spectral width is only a few nanometers. This is shown in Figure 2.4.

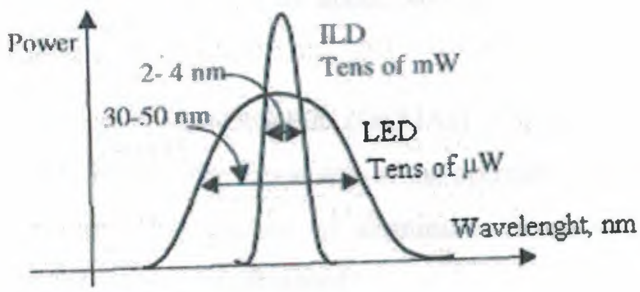

Figure 2.4 Power Spectral Widths of LED and ILD.

Ideally, a single spectral line is desirable. As the spectral width of the emitter increases, attenuation and pulse dispersion increase. The spectral purity for the ILD and its ability to couple much more power into a fiber make it better suited for long-distances telecommunications links. In addition, injection laser can be turned on and off at much higher rates than an LED. The drawback, however, is its cost, which may approach several hundreds of dollars as compared to a few dollars for LED's in large quantities.

Table 2.1 lists the differences in operating characteristics between the LED and the ILD.

Table 2.1 Typical source characteristics for LED and ILD

	Output Power, μ.W	Peak wavelength, nm	Spectral width, nm	Rise time, ns
LED	250	820	35	12
	700	820	35	6
	1500	820	35	6
Laser	4000	820	4	1
	6000	1300	2	1

Various semiconductor materials are used to achieve this. Pure *gallium arsenide* (GaAs) emits light at a wavelength of about 900 nm. By adding a mixture of 10% aluminium (Al) to 90%

GaAs, *gallium-aluminium-arsenide* (GaAlAs) is formed, which emits light at a wavelength of 820 nm. Recall that this is one of the optimum wavelengths for fiber optic transmission. By tailoring the amount of aluminum mixed with GaAs, wavelengths ranging from 800 to 900 nm can be obtained.

To take advantage of the reduced attenuation losses at longer wavelengths, it is necessary to include even more exotic materials. For wavelengths in the range 1000 to 1550 nm, a combination of four elements is typically used: *indium, gallium, arsenic* and *phosphorus.* These devices are commonly referred to as *quaternary* devices. Combining these four elements produces the compound *indium- gallium-arsenide-phosphide* (InGaAsP). Transfer characteristic of LED and ILD are shown Figure 2.5 (a) and (b) respectively.

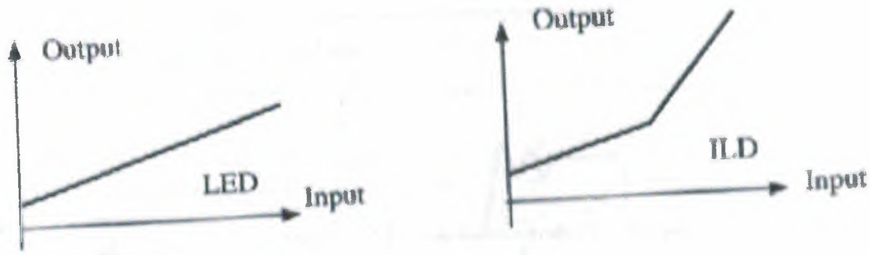

Figure 2.5 Transfer characteristic of LED and ILD.

b) Injection Laser Diode

The term *laser* is an acronym for *light amplification by stimulated emissions of radiation.* There are many types of lasers on the market. They are constructed of gases, liquids, and solids. Laser diodes are also called injection laser diodes (ILD), because when current is injected across the pn- junction, light is emitted relatively large and sophisticated device that outputs a highly intense beam of visible light. Although this is in part true, the laser industry is currently devoting a great deal of effort toward the manufacture of miniature semiconductor laser diodes. Figure 2.4 illustrates the spectrum ILD. ILDs are ideally suited for use within the fiber-optic industry due to their small size, reliability, and ruggedness. Step response of ILD is shown Figure 2.6.

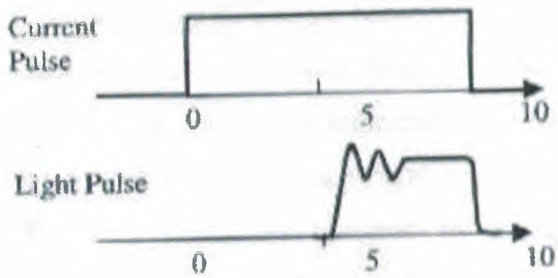

Figure 2.6 Step Response of ILD.

The most widely used light source in fiber optic systems is ILD. Like the LED, it is a pn junction diode usually made of GaAs. Injection laser diodes are capable of developing light power up to several watts. They are far more powerful than LEDs and, therefore, are capable of transmitting over much longer distances. Another advantage ILDs have over LEDs is theirs speed. High-speed laser diodes are capable of gigabit per second digital data rates.

2.1.2 Input —Output Characteristics

When we apply sufficient current to an LED or ILD light will be emitted. In order to understand the capabilities and limitations of these devices in system applications it is necessary to understand some details regarding their output power vs applied current characteristics, the spectral characteristics of their emitted light, their modulation speed capabilities and sources of limitations, the spatial properties of their emitted light fields, and the effects of temperature and aging on relevant characteristics.

a) Input -Output Characteristics of LEDs

An LED is a noise source, which emits its noise in a band of wavelengths centered about its nominal optical wavelength. The nominal wavelength of the device is determined by the nominal energy gap between the semiconductor valence band and conduction band. For gallium—aluminum—arsenide material this band gap is approximately 2×10^{-19} joules (J), corresponding to a photon at a wavelength of approximately 0.85 μm (depending upon the composition of the light-emitting active layer). However, the actual energy difference between a hole in the valence band and an electron in the conduction band can deviate from the band gap energy by about the Boltzmann energy kT= 4×10^{-21} (at room temperature). Thus the actual energy of the emitted photon when a hole in the valence band combines with an electron in the conduction band can vary about its nominal by about $4 \times 10^{-21} / 2 \times 10^{-19} = 2\%$. In other words the light emitted by the LED at room temperature occupies a spectral region centered at about 0.85 μm and having a spectral width of ± 2% of 0.85 μm or ± 17 nm. The nominal optical frequency is about 3×10^{14} Hz. This spectral width of ± 2% corresponds to ± 6×10^{12} Hz (i.e.. a bandwidth of around 12 THz). Thus it is clear that the band of frequencies emitted by the LED is much larger than the bandwidth of an information-bearing signal, which might modulate the power emitted by the LED.

When we apply a forward bias current to the LED holes and electrons are injected into the junction and combine to produce photons of light. Figure 2.7 illustrates the characteristic of total output power vs applied forward bias current for a typical high radiance (efficient light emitter) GaAlAs LED. We see that for moderate applied currents (below a few hundred milliamperes in this case) the output light power increases approximately linearly with the applied current.

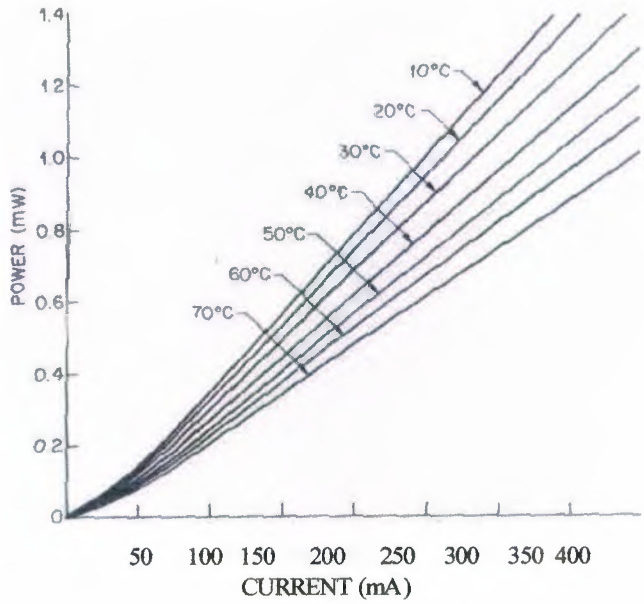

Figure 2.7 Total Output Power vs Applied Current for a Typical GaAlAs LED.

Above this level of drive, heating effects cause saturation of the light output vs current drive characteristic. We also observe that the slope of the curve decreases when the device ambient temperature increases. Higher device temperatures reduce the efficiency of conversion of hole—electron pairs to photons of light (at higher temperatures more hole—electron pairs combine nonradiatively to produce heat). Figure 2.8 shows a similar characteristic for a long wavelength (1.3 μm) In GaAsP LED. Note that the temperature sensitivity is stronger for this material.

In addition to the dc input—output characteristic of the device, we are also interested in the modulation capabilities. Modulation is accomplished by varying the drive current. The varying drive current causes the emitted output power to vary in response. If we vary the drive current slowly the output light power will follow reasonably faithfully (there is some non-linearity in the input—output characteristic, but we can neglect it in this discussion). However, if we vary the drive current rapidly the output light power may not be able to track these variations.

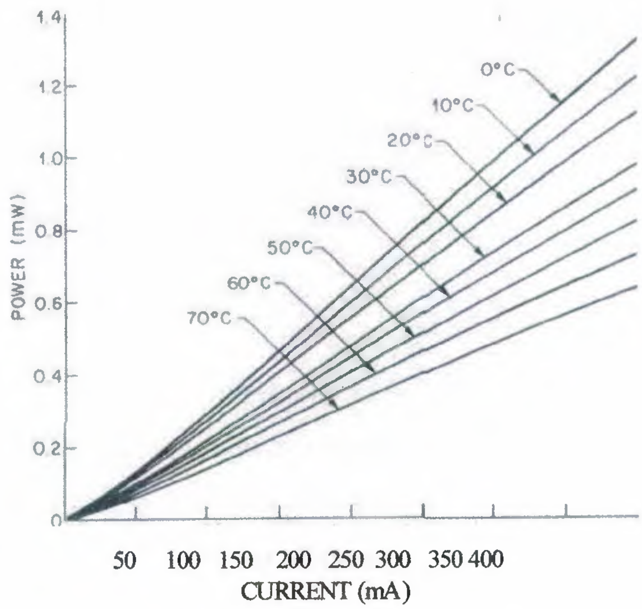

Figure 2.8 Total Output Power vs Applied Current for a Typical InGaAsP LED.

This frequency response limitation can be caused by ordinary circuit limitations (capacitances and inductances which prevent the high frequencies in the modulated current applied to the device terminal conductors from reaching the junction as injected current variations) and by the intrinsic time constants of the device itself. Circuit limitations can be minimized by careful transmitter design, minimization of lead lengths, device capacitance, and series resistance, etc. The intrinsic device modulation speed limitation is associated with the recombination lifetime of a hole—electron pair in the junction.

If a given number of hole—electron pairs are injected into the junction at some time To then this number of pairs will decay exponentially with a decay time constant known as the recombination lifetime T_r. The recombination lifetime of a material can be decreased by adding dopants, which usually (but not always) increase the rate of nonradiative recombination. This reduces the conversion efficiency (ratio of light output to electrical power input) of the device. Recently results have been reported of doping techniques which can reduce the recombination lifetime while maintaining a high ratio of radiative to nonradiative recombination. The recombination lifetime relates to the modulation bandwidth of the device through the following simple equation:

$$p(t) = C \int_{-\infty} i\ (\tau)\exp[(\tau-t) / T_r]\, d_\tau$$

$$B_{modulation} = 1 / (2\pi\ T_r) \qquad (2.1)$$

Where $i(t)$ is the applied current, $p(t)$ is the modulated emitted light, and C is a constant. We see that the emitted light is a low-pass filtered replica of the drive current with a 3-dB (0.707 amplitude) roll of at frequency $1 / (2\pi\ T_r)$Hz. Note that the current, $i(t)$, in (2.1) is assumed to be positive for all times t.

The light emitted by the LED can be most simply characterized in terms of its total power. We have already mentioned the fact that the device emits light over a band of wavelengths. The device also emits light simultaneously and independently in a large number (typically thousands) of field patterns. In a sense, the LED acts like a large number of independent light emitters all operating in the same light-emitting cross-sectional area. Consider a light-emitting diode with a circular light-emitting area of diameter D operating at nominal wavelength X. Within this area one can fit N complex field patterns $E(x,y)$, where x and y correspond to the cross-sectional coordinates, which are orthogonal. That is they satisfy the following equations:

$$\int E_i\ (x,y)\ E_j^*\ (x,y)\ d \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \qquad (2.2)$$

Each of these field patterns corresponds to light emitted in a different direction relative to the direction perpendicular to the emitting area. In a sense, these field patterns are very much like the modes guided in a multimode fiber. The light emitted in each of these LED spatial modes is in a random phase relationship to the light emitted in the other modes. Further more the light in any mode has a random amplitude and phase, which changes every T_s seconds, where T_s is the reciprocal of the LED spectral width. The fact that the emitted light in each mode randomizes its amplitude and phase every T_s seconds, and the emitted spatial modes have statistically independent amplitudes and phases is the reason we have referred to the light emitted by this device as noise.

When an LED is modulated, there is a small delay between the modulation waveform applied to the drive current and the corresponding response in the output power vs time waveform. It has been observed that if one examines narrow bands of wavelengths within the LED output spectrum with an optical spectrum analyzer (monochromatic) one finds that this delay is not uniform across the LED output spectrum. This phenomenon is called chirping.

The curve of optical output power vs electrical drive current of the LED is slightly nonlinear (below saturation). It has been observed that if one uses an optical system to select portions of the total output field emitted by the device then there is a lack of tracking of this nonlinearity between the selected portions. That is, the nonlinearity in the input drive current vs optical power produced characteristic may not be the same for different spatial modes or combinations of spatial modes emitted by the LED.

b) Input—Output Characteristics of Lasers

When current is applied to a semiconductor injection laser, it behaves at first like a light-emitting diode. However, when the current reaches the threshold value, the process of stimulated emission begins to dominate the LED process of spontaneous emission, and the device begins to oscillate (laser action begins). Figure 2.9 shows a typical set of curves of light output vs applied current for a GaAlAs injection laser. We can see the temperature dependence of the threshold current and some slight non-linearity on the light output vs applied current characteristic above threshold.

For most modern injection lasers used for optical fiber communication systems the light emitted by the device is in a well-defined spatial field pattern. In the early days of injection lasers (early 1970s) the device designs were such that this was not necessarily the case. A laser which can emit light in more than one spatial field pattern will typically show a kink (sharp bend) in its curve of output power vs drive current, corresponding to the onset of oscillation of a second spatial mode (field pattern). If a device can oscillate in several spatial field patterns (transverse modes) it will tend to divide its total output power in an unpredictable manner (randomly) amongst the allowed modes, and this subdivision will tend to change randomly in time.

The laser will oscillate at a frequency, which corresponds to a resonance of the optical cavity. Refer to Figure 2.10. The wavelength of the light in the cavity is the free space wavelength divided by the index of refraction of the cavity material. The condition for resonance is that the round trip length of the cavity, $2L$, be a multiple of the wavelength of the light in the cavity at the frequency of oscillation.

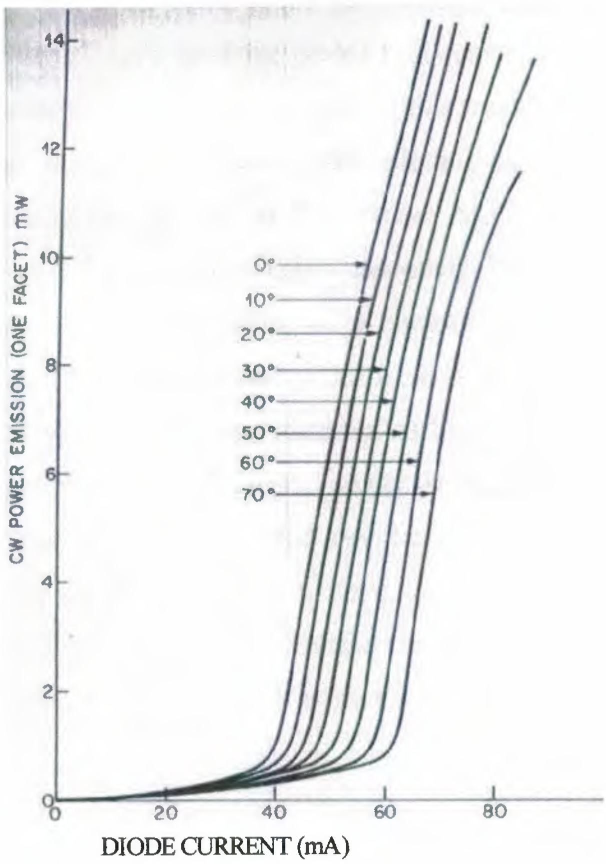

Figure 2.9 Total Output Power vs Applied Current for a Typical GaAlAs Injection Laser.

Consider a cavity of length 100 μm made of GaAs and oscillating at a nominal free space wavelength of 0.85 μm. The wavelength in the cavity is 0.28 μm. Thus the round trip length of the cavity is 200/0.28 = 706 wavelengths. The cavity will also resonate at a free space wavelength where the round trip length of the cavity is 706 ± 1 wavelengths, corresponding to 0.85 μm ± 1.2 nm. Indeed, the cavity will have a series of resonant wavelengths spaced 1.2 nm apart. If the cavity is 200 μm long, then the spacing between cavity resonances will be 0.6 nm. What mechanism will determine the particular wavelength of oscillation? In today's lasers there is a weak selection mechanism corresponding to the curve of gain in the cavity vs wavelength associated with the active layer. However, this gain vs wavelength curve has a full width to half maximum of about 25—50 nm (corresponding to the width of the spontaneous emission). Thus the round trip gain afforded to two wavelengths near the peak of this curve, but spaced by only a few nanometers, is very nearly the same. As a result of this, typical lasers can oscillate at a number of cavity resonances. The frequency of oscillation can jump from one resonance to another in an unpredictable fashion, and the device can oscillate in several resonant longitudinal modes (several wavelengths) simultaneously, with its output power divided unpredictably between these modes. There are a number of methods, which have been proposed to stabilize the frequency of operation of an injection laser. These approaches all provide for a frequency selection mechanism of sufficiently high selectivity to resolve the subnanometer spacing between normal cavity resonances. The methods include cleaving the laser cavity into two coupled cavities having slightly different sets of resonant modes, adding an external mirror to form a second cavity outside of the laser, and using corrugations formed on the laser surface to form a frequency selective mirror. In the multicavity approaches the laser will oscillate at a frequency, which is common to all (both) cavities.

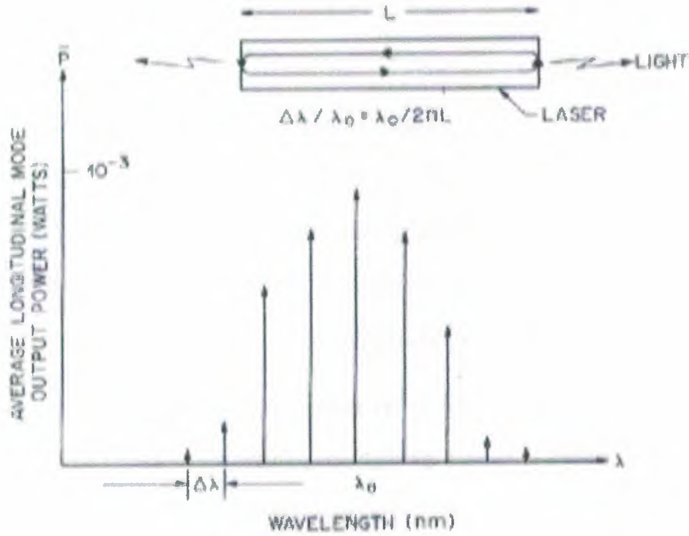

Figure 2.10 Laser Longitudinal Mode Spectrum.

The multicavity approaches tend to force the laser into a single frequency of oscillation, but stabilizing that frequency of oscillation (preventing occasional jumps), in the presence of modulation, temperature changes, and bias complex driver circuitry. Fabrication of the distributed mirrors on the surface of the device presents difficult challenges, but much progress has been reported recently.

In addition to multifrequency oscillation, lasers are also observed to suffer noise like fluctuations in their total output power and in the power in any particular longitudinal mode (frequency). Recently there has been an attempt to understand this behavior more clearly by studying the equations, which describe the relationship between carrier (holes and electrons) populations and photon populations within the laser cavity. Attempts have been made to model the laser as a near-unity-gain amplifier of spontaneous emission, again using the above-mentioned equations. Simulations based on these models have predicted some of the observed noise like behavior of laser emission. These noiselike fluctuations limit the signal-to-noise ratios, which can be achieved at the outputs of communication links, which employ lasers.

Modulation of the laser output power is accomplished in most present applications by varying the drive current. Typically lasers are operated with a fixed amount of drive current to bring the device just below the threshold of lasing. Incremental current is then added to modulate the device from low output (off) to high output (on) for digital applications. For analog modulation, the device is operated with a bias current which places it at an intermediate level of output power. The device is then modulated above and below this level of output by a superimposed modulating current.

In digital applications the injection of a fixed bias current serves two purposes. First, there is a delay of several nanoseconds to bring a typical laser from zero applied current to the threshold of lasing. By applying a fixed bias current, slightly below the lasing threshold current, this delay can be avoided. The response speed of the device to the incremental currents that are added to the fixed bias can be a small fraction of a nanosecond. Thus modulation rates of several gigahertz are possible, depending upon the details of the device. Second, it is easier to produce high-speed digitally modulated currents at a level equal to the incremental values, rather than to modulate the current from zero.

As in the case of the LED, the modulation limits are imposed not only by the intrinsic time constants of the interactions between carriers and photons but by circuit limitations as well. High-speed modulation of lasers requires circuit design techniques, which account for parasitic inductances, device capacitance, and series resistance, current changes, requires

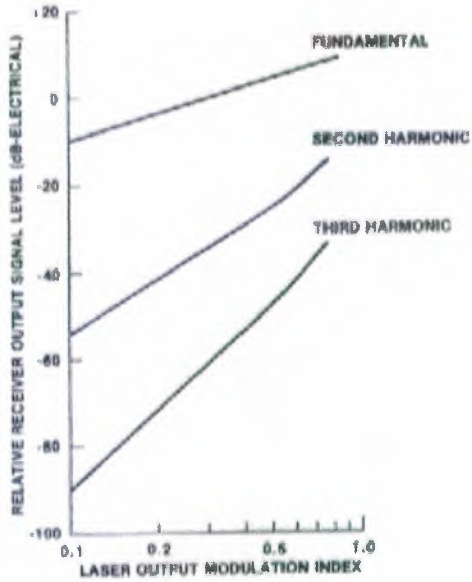

Figure 2.11 Laser Response Harmonics vs Modulation Index for a Typical GaAs Laser.

Typical lasers have a current input—light output characteristic, which appears to be linear above threshold. However, if one modulates the current sinusoidally about an operating point above threshold one observes harmonics in the detected output power. Figure 2.11 shows some characteristics of a typical injection laser, demonstrating how the amplitudes of the second and third harmonics increase with the modulation index (amplitude of modulation relative to the level above threshold). These curves can be very sensitive to the operating point. For example if the device is operated near a point of inflection, then the second harmonic may be very small, for small modulation indices. The limited linearity of lasers makes the design of analog modulation methods difficult.

2.1.3 Fabrication of LEDs and Lasers

Both LEDs and lasers are fabricated by batch processes where a number of device chips are formed on a wafer using photolithographic techniques. Layers of material are deposited to form the heterojunctions and *p-n* junctions starting with a wafer of GaAs or InP. The growth processes used are liquid phase epitaxy, vapor phase epitaxy, and molecular beam epitaxy. In liquid phase epitaxy the growing wafer is brought in contact with a series of supersaturated liquids which cause new crystalline material to grow on the surface.

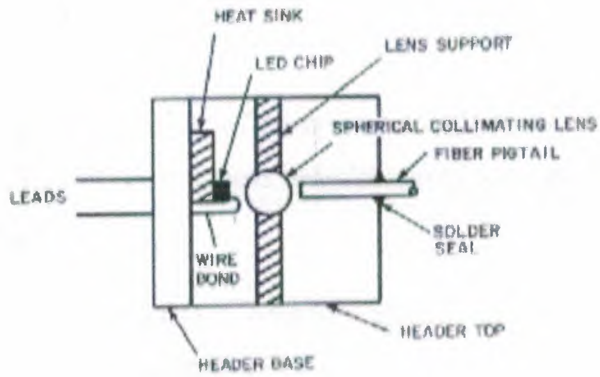

Figure 2.12 Schematic of a Packaged LED.

Precise control of the composition and temperature of the melts is required. In vapor phase epitaxy the wafer is exposed to gas mixtures of varying composition, which form new crystalline layers on its surface. In molecular beam epitaxy, the wafer is exposed to beams of molecules, evaporated from heated molecular sources, which can be turned on and off.

In the process of fabrication of the wafer containing individual devices various etching procedures may be used, either after deposition or as an intermediate step, to implement the details of the device geometry. After the wafer is fabricated, individual chips are diced from it. These must be bonded on heat sinks, bonded to conducting wires, and aligned with their output fibers. At present, this packaging operation is a substantial portion of the cost of fabricating useful sources. Figure 2.12 shows a schematic of a packaged LED.

The fabrication of a reliable device requires not only careful control of the wafer growth process, but attention to many details in the bonding (to minimize strain) and fiber alignment procedures, as well as proper sealing of the finished package against contamination. The difficulty in implementing all of these details in a production environment may explain some of the discrepancies in the reliability of field devices compared with laboratory results. These discrepancies tend to diminish with production/field experience, which leads to improvements in manufacturing controls and methods.

2.2 Coupling Sources to Fibers

In order for the output power of an optical source to be useful in a fiber optic system application some fraction of it must be successfully and stably coupled to a fiber.

2.2.1 Coupling LEDs to Fibers

The output of a typical LED is generally emitted from a surface whose area is equal to or larger than that of a multimode fiber (diameter equal to 50 μm or larger), and generally fills a solid angle much larger than the acceptance solid angle of a multimode fiber (the range of angles captured in the core via total internal reflection). Figure 2.13 illustrates this typical situation in two dimensions. Note that with the fiber butted up against the emitting area a substantial fraction of the emitted rays are not captured because their angles relative to the fiber axis are too large, or because they do not strike the core (using the geometric optics model as shown).

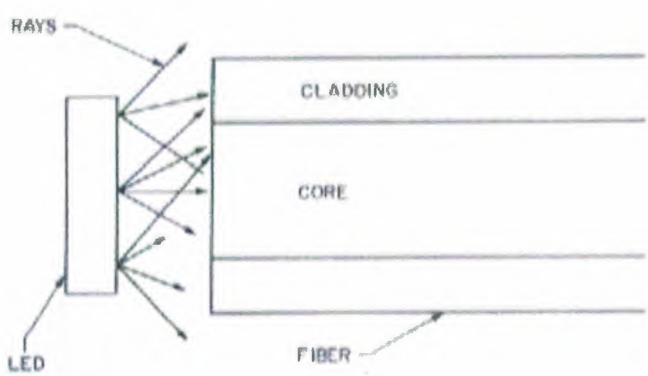

Figure 2.13 Coupling an LED to a Fiber Without a Lens.

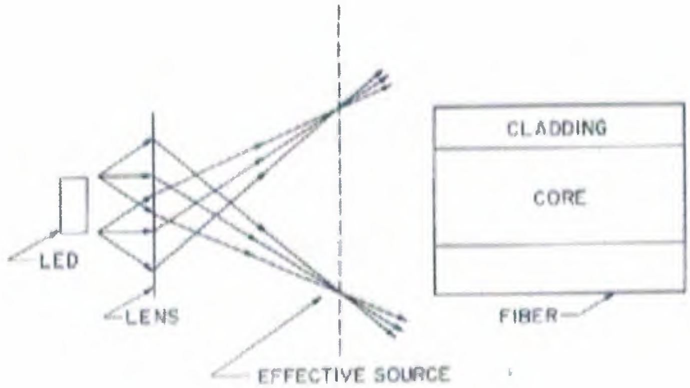

Figure 2.14 Coupling an LED to a Fiber Using a Lens.

At first one might assume that the fraction of coupled light could be improved by the use of a lens. However, this is generally not the case. The LED is emitting independently in thousands of spatial modes (orthogonal field patterns), while the multimode fiber can guide only a few hundred spatial modes. One can show that the coupling efficiency from the LED to the fiber can, at most, equal the ratio of the number of guided fiber modes to the number of modes emitted by the LED. It is possible to have a much lower coupling efficiency, and sometimes lenses are necessary to approach this theoretical maximum. The situation illustrated in Figure 2.14 gives an example of how a lens can be used to improve the coupling efficiency. The light-emitting area, in this example, is smaller than the fiber core. The lens images the emitting area into an "effective source" whose area is larger but whose range of angles is smaller. The lens trades off unfilled area against noncaptured high angle emissions.

A lens can also be useful when it is not practical to bring the fiber close to the light-emitting area. In this case the lens images the light to a position away from the emitting surface. However, the constraint on coupling efficiency (known as the law of brightness) cannot be violated. In effect, the source emits a certain amount of power per spatial mode. The total power that can be coupled into the fiber is the power per spatial mode multiplied by the number of fiber modes.

The constraint on coupling efficiency between LEDs and fibers does have a positive side. If, on the one hand, the coupling efficiency is only 2%, there are, on the other hand, roughly 50 independent positions, which achieve this coupling efficiency. Thus alignment is not as critical as it is in the laser-fiber case to be described below.

2.2.2 Coupling Lasers to Fibers

Most modern injection lasers emit light in a well-defined spatial field pattern, which remains stable in time, with modulation, aging, etc. This implies, in principle, that the laser can be coupled with 100% efficiency to any fiber— even a single mode fiber. However, the field pattern emitted by .the laser is not matched to the field pattern that is accepted by a single mode fiber, or any linear combination of field patterns accepted by a multimode fiber. Thus, if the fiber end is simply butted up against the laser, the coupling will be relatively low (a few percent for a single mode fiber, and a few tens of a percent for a multimode fiber). To correct this mismatch problem a field-matching device must be placed between the fiber and the laser. For multimode fibers this can take the form of a crude lens melted on the end of the fiber. For single mode fibers a popular technique for implementing the field-matching device is to heat and taper the fiber end. For properly aligned interfaces coupling efficiencies from lasers to multimode fibers are typically over 50%. Coupling efficiencies from lasers to single mode fibers are typically 15%—20%, with somewhat higher results reported in the laboratory.

There are noise problems caused by reflections from the coupled fiber back into the laser. These problems can be particularly troublesome in single mode fiber systems, which typically operate at higher modulation speeds (and are, therefore, more susceptible to rapid fluctuations).

In designing the interface between the fiber and the laser some compromise may be made to minimize reflections at the expense of coupling efficiency. The alignment of the laser and the fiber is critical and must be stable with temperature (of the laser package) and time. It is possible to design the interface to sacrifice some coupling efficiency in order to reduce sensitivity to mechanical alignment errors.

2.3 Transmitter Modules

In order for an LED or a laser to be useful in a system application it must be successfully interfaced to the electronic terminal, which is generating the modulating signal. Important considerations in the design of a transmitter module, which includes the electronic interface circuitry, are power consumption, modulation bandwidth, available power supply voltages, transient and over voltage protection, impedance matching, and stabilization against temperature and aging effects. For analog modulation applications one is also concerned with flatness of frequency response and linearity.

The light transmitter consists of the LED and its associated driving circuitry. An optical transmitter circuit using the LED is shown in Figure 2.15. The binary pulses are applied to a logic gate, which, in turn, operates a transistor switch T that turns the LED off and on. A positive pulse at the NAND gate input causes the NAND output to go to zero. This turns off T, so the LED is forward-biased through R_1 and turns on. With zero input, the NAND output is 1, so T turns on and shunts current away from the LED.

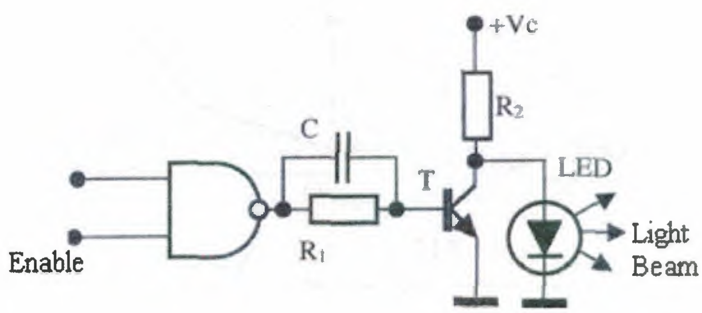

Figure 2.15 Basic Circuit of LED Transmitter

Very high current pulses are used to ensure brilliant high-transmission rates are limited. Most LED like transmitters are used for short-distance, low speed digital fiber-optic systems. With zero input, the NAND output is 1, so T turns on and shunts current away from the LED. Most LEDs are capable of generating power levels up to approximately several hundred μW. With such low intensity, LED transmitters are good for only short distances. Further the speed of the LED is limited. Turn-off and turn-on times are no faster than several nanoseconds.

A typical injection laser transmitter circuit is shown in Figure 2.16. When the input is zero, the AND gate output is zero, so T is off and so is the laser. Capacitor C_2 charges through R_3 to the high voltage. When a binary 1 input occurs, T conducts connecting C_2 to the ILD. Then C_2 discharges a very high current pulse into the laser, turning it on briefly and creating the light pulse.

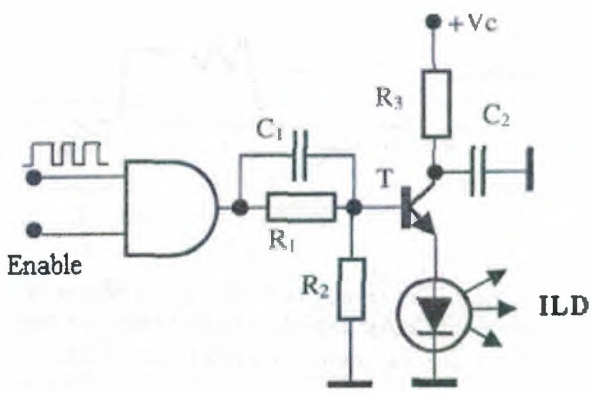

Figure 2.16 Basic Circuit of ILD Transmitter

2.3.1 Driver Circuitry for LEDs

Figure 2.7 shows a typical dc characteristic of light output vs. drive current for a light-emitting diode. Figure 2.17 shows a typical small signal frequency response of a light-emitting diode (the modulation depth of the detected output power normalized by the modulation depth of the applied current). A very simple driver interface circuit for an LED, suitable for modest modulation rates, is shown in Figure 2.18.

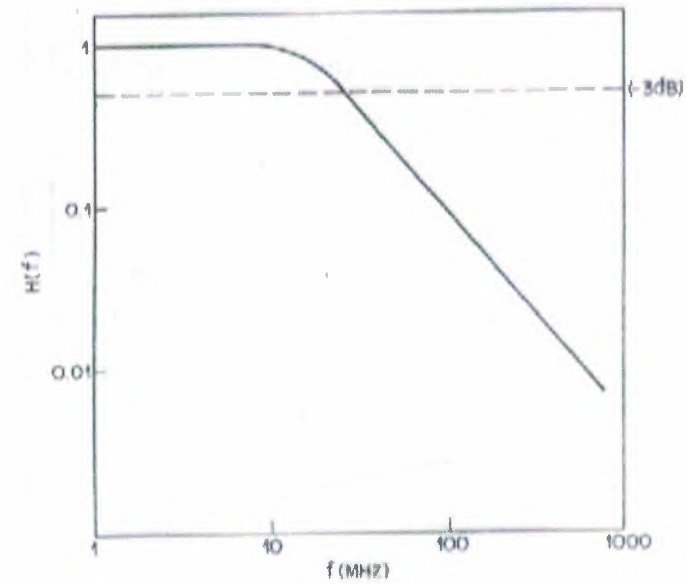

Figure 2.17 Typical Frequency Response for a Light-Emitting Diode

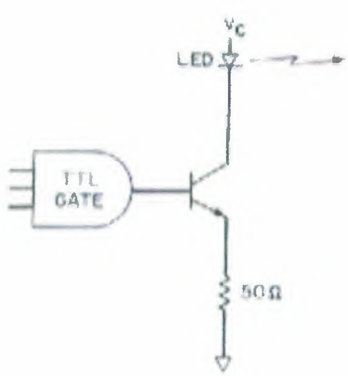

Figure 2.18 Simple LED Driver Circuit.

Here the single transistor acts as a current amplifier allowing a relatively high impedance terminal to produce adequate current to modulate the LED. This circuit is suitable for analog or digital modulation. If we assume that the output of the TTL gate is

5 V in the on state, and if we assume a 0.7 V drop from the base to the emitter of the driver transistor, then we obtain 4.3 V across the 50 Ω resistor. This implies 85 mA of current through the LED for a transistor current gain of 100

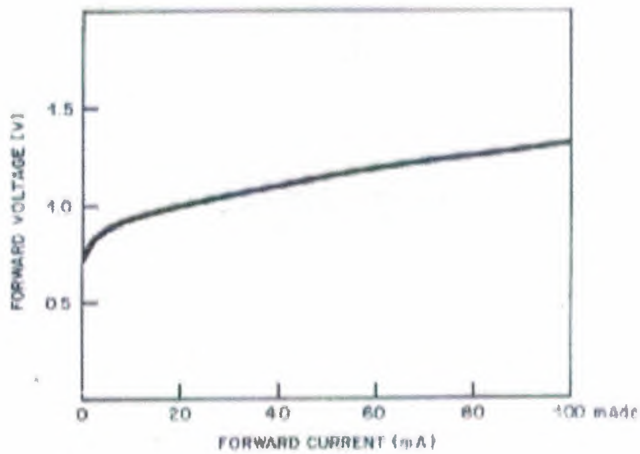

Figure 2.19 Voltage vs Applied Current for a Typical InGaAsP LED.

Figure 2.19 shows the voltage vs applied current characteristic for a typical InGaAsP LED. We see that the 85 mA forward current results in a voltage drop across the LED of about 1.3 V. Adding up the voltages across the emitter resistor and the LED we see that we require about 7 V for V+ in order for the circuit to operate as described (no transistor saturation).

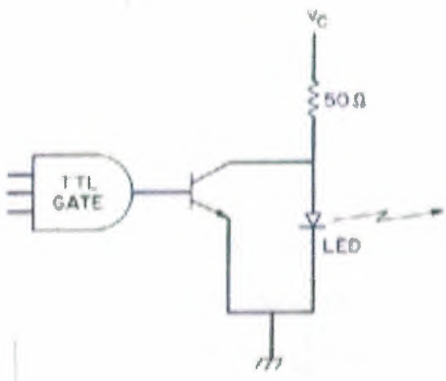

Figure 2.20 Shunt-Type LED Driver Circuit.

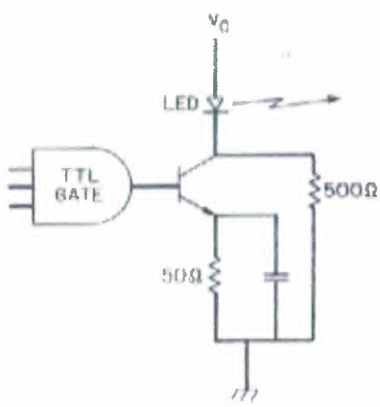

Figure 2.21 LED Driver with Speed-up Circuitry.

Figure 2.20 shows a configuration in which the driver interface transistor shunts current away from the LED. The shunt design also allows the LED "space charge" capacitance to discharge relatively rapidly through the 50-Ω resistor, reducing pattern-dependent modulation effects associated with this space charge in digital modulation applications. (In Figure 2.18 it is difficult for this capacitance to discharge rapidly when the transistor turns off, unless additional circuit components are added for this purpose.)

In addition, during turn-on the LED is driven by a low impedance source, which can charge its space charge capacitance more rapidly than the current source of Figure 2.18.

Figure 2.21 shows a popular approach for speeding up the turn on time of the LED by applying a larger initial current. In addition, an unmodulated component of current is drawn through the LED to keep its space charge capacitance charged near the turn-on voltage, thus reducing the effects associated with a long discharge time. In general driver designs, which attempt to squeeze out more modulation bandwidth from an LED, consume increasingly large amounts of power.

2.3.2 Driver Circuitry for Lasers

Driver circuitry for laser sources is typically more complex than that for LED sources for two reasons. First the temperature dependence of the laser threshold must be accommodated. Second, lasers are typically used at higher modulation speeds. (Lasers are also more prone to damage from transients.) A typical set of input drive current—light output characteristics for an injection laser is given in Figure 2.9. In practical operation it is necessary to apply an adjustable bias current to keep the laser emitting at a fixed level of short-term-average output in the presence of threshold variations, caused by temperature changes and aging.

This is often accomplished by monitoring the short-term average output power with a detector on the back face of the laser (the fiber is attached to the front face) and using this short-term average output indicator in a feedback loop to adjust the bias. Figure 2.22 shows a practical circuit for accomplishing this. Note that the short-term average output indication is compared to a reference derived from the modulating signal. In this way, if the modulating signal itself contains a changing short-term average value, the control circuit will not track this out. Further more, if the modulating signal calls for the laser to be off for an extended time, the control circuit will not interpret this as a result of inadequate bias. The adjustment of such a circuit for proper tracking with the reference derived from the modulating signal is a nontrivial matter.

In some laser driver designs a thermoelectric cooler is used to adjust the threshold current value of the laser, rather than adjusting the bias, to implement a controlled average output power. Cooling the laser reduces its threshold as shown in Figure 2.9.

Figure 2.23 shows a balanced driver circuit, which can supply the incremental current to modulate the laser. In this circuit, designed for digital modulation, a current source is switched between the collectors of two transistors, one of which contains the laser. The bias is controlled to have a value just below the threshold current. The incremental modulation current drives the laser from this nearly off level to the "on" state. An on—off ratio of around 10:1 is typically obtained for the light coupled into the fiber, using this type of configuration.

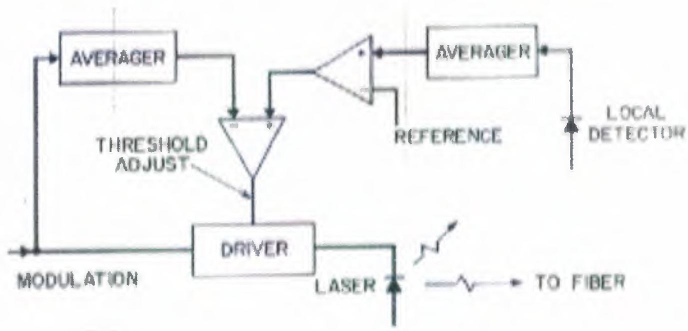

Figure 2.22 Practical Laser Driver Circuit with Feedback Bias Control.

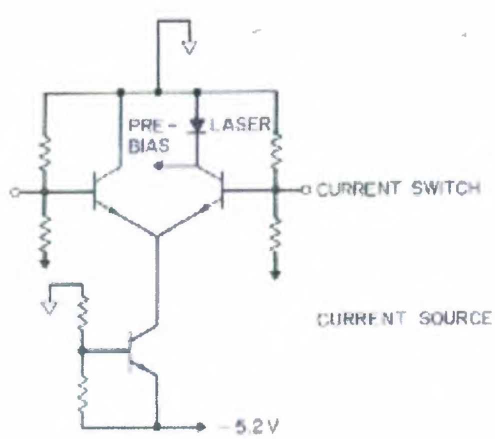

Figure 2.23 Balanced Laser Modulating Circuitry.

Recently, as interest has increased in the implementation of stabilized single frequency lasers and (more exotically) in lasers that are stabilized in frequency to an extent suitable for coherent systems laser driver circuitry has become potentially more complex. Essentially, provisions must be added to monitor the spectral content or the frequency of the source and to adjust the bias current, temperature, or some other control mechanism to provide the required stabilization.

Interest in very high frequency modulation of lasers (beyond 1 GHz) has led to the use of microwave circuit design techniques to properly match the laser impedance characteristics and to compensate for parasitic components in the laser and its package.

CHAPTER 3

OPTICAL DETECTORS AND RECEIVERS

To implement an optical communication or sensing system one requires a means of extracting information contained in the modulation of an optical signal. It is possible to perform some all-optical functions on a received optical signal, such as optical amplification. It is also possible for a received optical signal to turn on another optical signal. However, in the optical communication and sensing systems used today (and foreseen for some time to come) information is extracted from the optical signal by converting it to an electrical signal, and by then performing electrical signal-processing functions. The device that converts the optical signal into a voltage or current is called the optical detector. The electronic circuit which interfaces the detector (with its particular characteristics and (requirements) to conventional electronic terminal equipment is called a preamplifier. The combination of the detector and the preamplifier is called a receiver.

3.1 *p-i-n* and APD Detectors

3.1.1 Definition and Input—Output Characteristics

An optical detector provides a mechanism for converting optical power into an electrical voltage or current. One can imagine a variety of methods of doing this. For example, one could absorb optical energy in a device, causing it to heat up. This change in temperature could modulate a physical property (e.g., resistance), which could, in turn, be sensed by an electronic circuit. Detectors which work on absorption of optical power to produce heat (thermal detectors) are too slow in their response speeds for most optical fiber system applications (except power meters, for example).

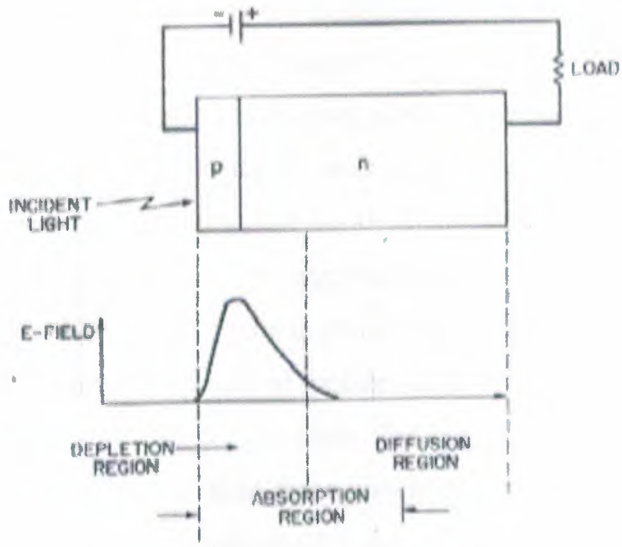

Figure 3.1 Principle of Operation of a *p-n* Diode Detector.

Another approach to implementing an optical detector is to allow the incident optical power to illuminate a semiconductor device, resulting in the generation of hole—electron pairs by absorbed photons. These pairs can, in turn, flow in the presence of an electric field to produce an observable current. This process can produce detectors with response speeds of tens of picoseconds. This is the principle of operation of *p-i-n* and avalanche photodiodes to be described below.

Figure 3.1 shows an example of a simple semiconductor device that can serve as a high-speed optical detector. The device is a p-n diode made of a material that strongly absorbs light at the wavelength of interest. For example, for the wavelength range 0.8—0.9 μm silicon is an excellent material choice; the device is operated in a back-biased manner. In the absence of illumination only a very small current flows due to thermally generated hole—electron pairs. This small leakage current is often referred to as dark current. When the device is illuminated as shown optical energy (photons) is absorbed in the material, causing hole—electron pairs to be generated at a rate of one pair per absorbed photon. These pairs separate within the device in the presence of the local fields and produce a displacement current observable in the external electrical circuit. The total current produced in the external circuit due to the generation of a single hole—electron pair has an area (charge) equal to the charge of an electron, e. One can define three regions within the device as shown. The absorption region is the volume of Space in which the light is absorbed, and extends for some depth into the device. The absorption depth depends upon the wavelength of the incident light, the material from which the device is made, and the definition of how much attenuation of the incoming light signal corresponds to absorption (e.g., 99%). The depletion region corresponds to the volume of space within the device that is depleted of mobile carriers due to the application of the reverse bias voltage (leaving behind immobile donor and acceptor ions). Within the depletion region there is an electric field that accelerates carriers. Furthermore, it can be shown that the displacement current observed in the external circuit is present only when carriers are moving in the depletion region. The volume of space, which does not contain a field, is called the diffusion region. Carriers within this region tend to move randomly at relatively low speeds and produce no displacement current. Thus when a hole—electron pair is generated in the diffusion region, the displacement current response will be delayed until the hole randomly finds its way to the depletion region, and is swept across the junction.

In designing the detector we want the absorption region to be long enough to absorb nearly all of the light (say 99%), but we want the resulting carriers to produce their displacement current quickly, in order to obtain a high-speed device. Thus we want the depletion region to extend the full length of the absorption region. Basic semiconductor diode theory reveals that the depth of the depletion region to the right of the junction depends upon the applied voltage and upon the doping level in the *n*-type material. The depth increases with the square root of the applied voltage, and it increases inversely with the square root of the doping level. Practical constraints limit the voltage that can be applied to a few hundred volts. Therefore, to obtain a deep depletion region, the n-type material to 'the right of the junction is fabricated so as to be very lightly doped. In fact, it is so lightly doped that it is nearly an intrinsic or *i*-type material. It is difficult to make a good nonrectifying metallic contact to an *i*-type material. Therefore a heavily doped n-type material is appended to the end of the structure as shown in Figure 3.2. Thus we end up with a *p-i-n (p* -doped, intrinsic, *n* -doped) design.

The performance of a detector is measured by the efficiency with which it converts optical power to electrical current, and by its speed of response. (When one begins to design a system in detail, parameters such as dark current, junction capacitance, and linearity are also important.) The conversion efficiency is characterized by one of two equivalent measures. The first measure is the fraction of incident photons that produce electron—hole pairs. This is a number less than or equal to unity and is called the quantum efficiency of the detector η.

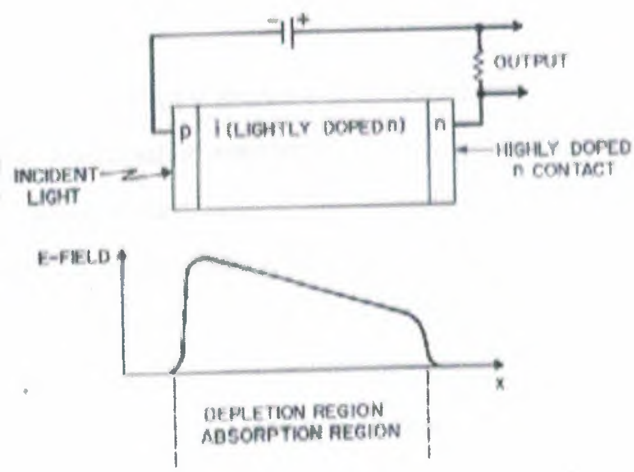

Figure 3.2 Principle of Operation of a Silicon *p-i-n* Diode Detector.

For a *p-i-n* photodiode the quantum efficiency can be less than unity for any of the following reasons: some incident light power (photons per second) is reflected at the interface between the surrounding medium and the detector surface; some of the light passes through the *i*-type region without being absorbed (carriers created by photon absorption in the heavily n-doped region recombine before producing a useful current); carriers generated in the *i*-typed region can recombine before producing a useful displacement current. For a typical detector the *i*-type region is designed to be thick enough so that light that is not reflected from the front face is absorbed in that region. Recombination in the *i*-type region is usually negligible. The second measure of conversion efficiency is the ratio of displacement current produced to incident optical power (amperes/watt). The displacement current is equal to the number of hole—electron pairs per second produced in the detector, multiplied by the electron charge, *e*. The incident power is the number of photons per second incident upon the detector, multiplied by the energy in a photon, *hf*. When units were being defined in the MKS system the electron charge turned out to be 1.6 x 10^{-19} coulombs (C). The energy in a photon at 1 μm wavelength turned out to be about 2 x 10^{-19}. The responsivity of the detector in amperes per watt is given by η x *e/hf* where η is the quantum efficiency. The ratio *e/hf* has the dimensions of amperes/watt but numerically it is a number near unity (0.8 for 1 μm wavelength).

The speed of response of a detector is characterized by the rise and 'all times of the current waveform produced in response to an optical pulse with a fast rise and fall time when the detector is biased with a low impedance (typically 50 Ω) circuit. The response speed is limited by the time it takes for carriers to cross the depletion region (assuming the depletion region encompasses the absorption region). Typical silicon detectors designed to operate over the wavelength range 0.8—0.9 μm have a response speed of about 0.5 ns (10%—90%). If the device is optimized to work at only the shorter wavelengths in this band, then it can have a thinner /-type region and a faster response (silicon absorbs more strongly at the short wavelengths). However, the thinner *i*-region results in more capacitance for a given light sensitive area. It is possible to trade off quantum efficiency (responsivity) against response speed by making a thin *i*-type region and allowing the absorption region to penetrate into the heavily doped *n* material (where generated carriers recombine before producing current). Long wavelength detectors have been made from both germanium and InGaAsP materials. Properly selected compounds of InGaAsP can be very highly absorbing to light in the 1.0—1.55 μm range of wavelengths, allowing for a thin *i*-region with high quantum efficiency and response speeds below 100 ps. These thin devices can also operate with conveniently low voltages.

When the detector is integrated with a preamplifier to make a receiver one finds that the performance of the receiver is limited by the noises introduced by active and passive components in the preamplifier. To combat the limitations in receiver performance caused by preamplifier noise one can use a device called an avalanche photodiode (APD).

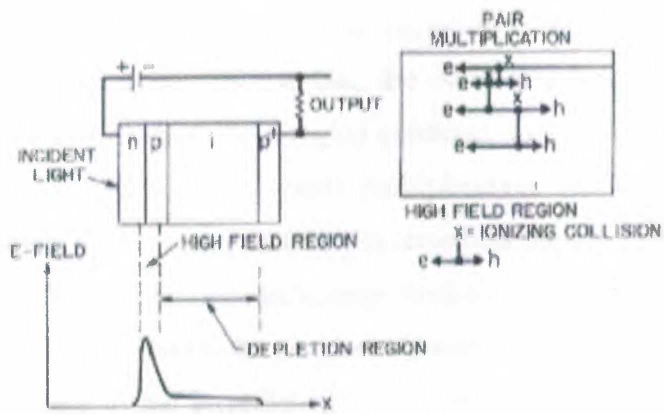

Figure 3.3 Principle of Operation of an Avalanche Photodiode.

The principle of the APD is shown in Figure 3.3. Doping levels are adjusted to create a region around the junction where the electric field is very high under backbiased conditions. Light that is absorbed in the *i*-type region to the right of the junction produces electron—hole pairs, the electrons of which (for silicon devices) drift into the high field region. The field levels in the high field region are sufficient to cause moving carriers to occasionally suffer ionizing collisions, where an additional hole—electron pair is produced. Carriers produced by ionizing collisions plus the original photogenerated carriers can in turn produce further pairs by the same mechanism. Thus we have a carrier multiplication process which results in a larger displacement current than we obtain with the nonmultiplying *p-i-n* photodiode. This carrier multiplication process is statistical. That is, one cannot predict exactly how many secondary pairs will result from an initial primary pair. Careful design and control of fabrication of the APD is required to produce a statistical distribution that gives the best tradeoff between average multiplication and noise associated with the statistical nature of the multiplication process. APDs have been fabricated in silicon, germanium, and InGaAsP materials for both short and long wavelength fiber optic systems. Figure 3.4 shows a curve of multiplication factor (and equivalently responsivity) vs applied reverse bias voltage for typical silicon APD. As reverse bias is applied, holes and electrons (mobile carriers) move away from the junction to form the depletion region. At low reverse bias, only the high field region is depleted (partially) and there are no fields in the *i*-type region to accelerate carriers produced by absorbed photons. Thus at low reverse bias the APD is a low-speed device. As the reverse bias is increased the high field region continues to deplete and the fields within that region begin to be sufficient to cause multiplication. However, at that point the device is still a low-speed detector. At larger reverse bias the high field region, becomes fully depleted; and further increases in voltage deplete the *i*-type region. The device is now a high-speed detector. When the *i*-region depletes, the width of the field increases, resulting in an increase in the incremental voltage required to increase the field level. Thus the curve of gain vs voltage has a kink at the point of *i*-region depletion as shown in Figure 3.4. If the reverse bias voltage is sufficiently high one obtains a condition of infinite multiplication, corresponding to sustained avalanche breakdown.

The multiplication factor of the APD changes with temperature as shown in Figure 3.4. Higher temperatures reduce the mean free path of moving carriers between nonionizing collisions, and reduce the chances of obtaining an ionizing collision.

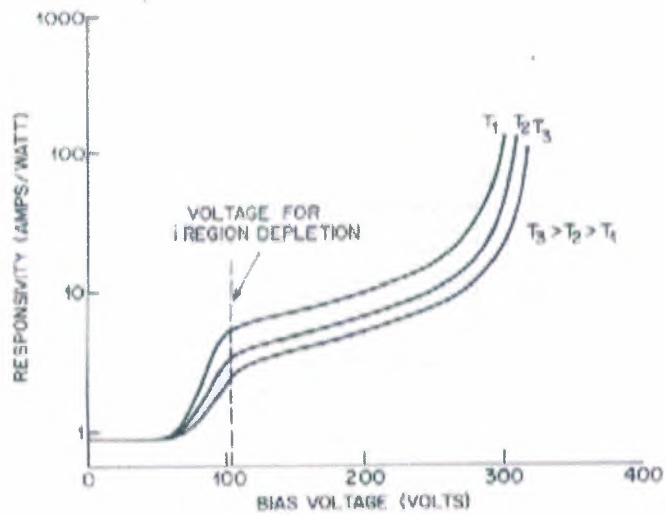

Figure 3.4 Responsivity vs Reverse Bias for a Typical Silicon APD.

At high values of multiplication this temperature dependence makes it difficult to bias the device for a desired level of multiplication. Receivers, which incorporate APDs generally, must implement some method of compensation of this temperature dependence.

Figure 3.5 shows a simplified schematic of a typical receiver incorporating a *p-i-n* or APD detector. Figure 3.6 shows the equivalent electronic circuit. A current represents the photodetection process source with responsivity R amperes/watt.

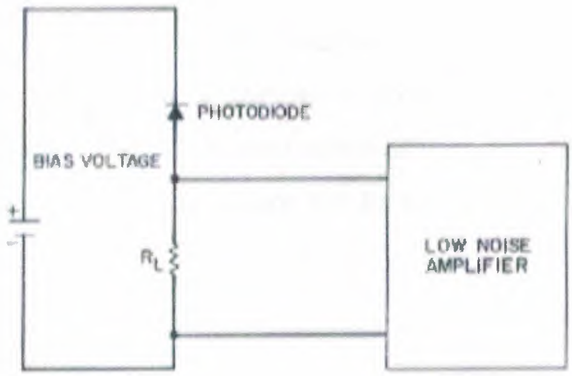

Figure 3.5 Simple Receiver Schematic.

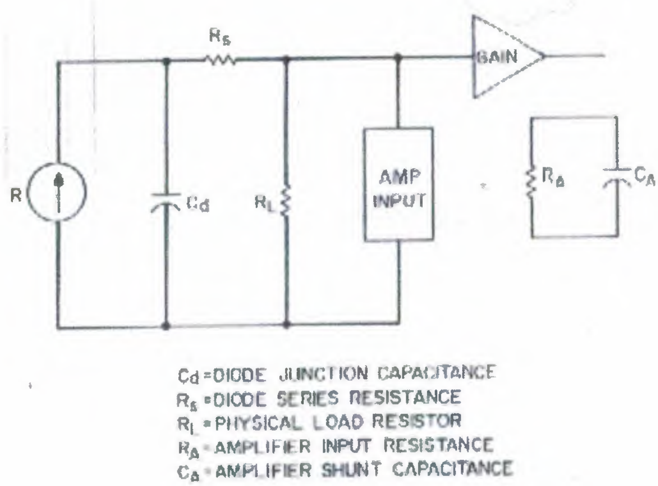

Figure 3.6 Equivalent Circuit of Simple Receiver.

The detector junction capacitance appears in parallel with this current source. A series resistance (typically assumed to be a few ohms or a few tens of ohms) is also shown. A parallel capacitance and resistance represent the preamplifier input impedance.

Note that *p-i-n* detectors are linear (*R* is a constant) over many orders of magnitude of incident optical power level. (They are essentially limited by noise levels at low illumination, and by power supply limitations or excessive dissipation at high optical illumination.) On the other hand APDs are somewhat nonlinear because the presence of carriers in the high field region can reduce the local field levels and thus reduce the multiplication.

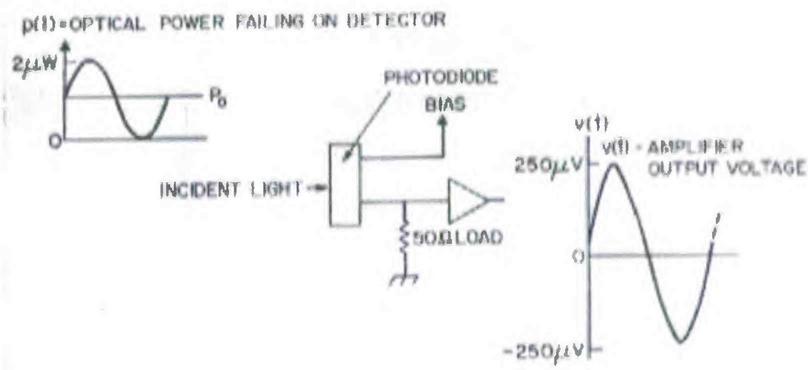

Figure 3.7 Illustration of a Photodetector in Operation.

As an example of how a photodetector might work in a circuit. Figure 3.7 provides a simple illustration. An optical signal of average power level 1 μW is shown incident upon a detector. The optical signal is modulated sinusoidally at a moderate frequency from 0 to 2 μW. The detector is in series with an amplifier having a 50-Ω input impedance and a gain of 10 (20 dB). The detector responsivity is assumed to be 0.5 A/W. The 1 μW peak sinusoidal optical modulation produces a sinusoidal detector output current of 0.5 *μA* (0.5 A/W). This current flows into the 50-Ω amplifier input impedance to produce an amplifier input of 25 *μV* peak. This results in an amplifier output of 250 *μV*. Further gain of a factor of 1000 (60 dB) is required to bring this up to a level of about 0.25 V.

When designing high speed and low noise receivers the capacitance of the detector is of importance. The capacitance depends upon the area of the *i*- region, the thickness that is depleted, and the dielectric constant of the material (plus parasitic capacitances). Silicon detectors are typically fabricated with a relatively thick depletion region (100 μm or more) and a relatively large area (250—500 μm diameter). The thick depletion region is required for high quantum efficiency (to absorb the light). This thick *i*- region results in a low capacitance since the capacitance is proportional to the region area divided by its thickness. InGaAs devices are typically fabricated with a relatively thin *i*-region. This is possible because of the high absorption constant of this material attainable in the wavelength range 1.0—1.55 μm. This results in very fast response speed but requires a small area of the *i*-region (typically 100 μm diameter) to maintain a low capacitance. The capacitance of a typical unpackaged detector is below 2 pf and can be as low as a small fraction of a pf for small area devices. A small area device is naturally more difficult to couple light into.

The volume in which hole—electron pairs are generated is relatively large compared to the volumes of typical components in integrated circuits. As a result, x-rays, gamma rays, and heavy particles can produce objectional amounts of dark current in environments where such radiation is contemplated. This can be particularly problematic for APDs that operate with very low photocurrents. Dark current can manifest itself in the form of shot noise or can overload electronics in the preamplifier interfacing with the detector.

3.1.2 Fabrication *of p-i-n* and APD Detectors

p-i-n and APD detectors are typically fabricated in a batch process where many devices are produced at once on a common wafer. The layers of *p, i,* and *n* type material can be formed by epitaxial growth or by the diffusion of dopants into a starting substrate.

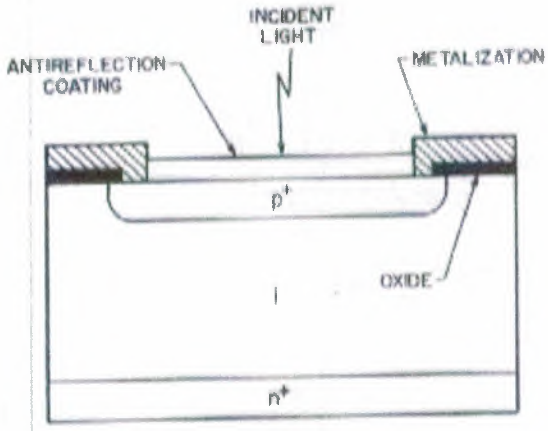

Figure 3.8 Schematic of a Typical Silicon *p-i-n* Detector.

Figure 3.8 shows a silicon *p-i-n* detector fabricated by diffusing dopants though a series of masks to form the *p* and *n* regions in an *i*-type substrate. An oxide layer provides insulation. Also shown is a thin-layered antireflection coating applied to reduce the reflection from the surface of the device (thus increasing its quantum efficiency). Figure 3.9 shows a silicon avalanche photodiode also fabricated by diffusion processes. The high field region is formed by consecutive diffusion of *p* and *n* dopants. The artifact labeled "guard ring" is incorporated to prevent premature break down due to fields at the edge of the device.

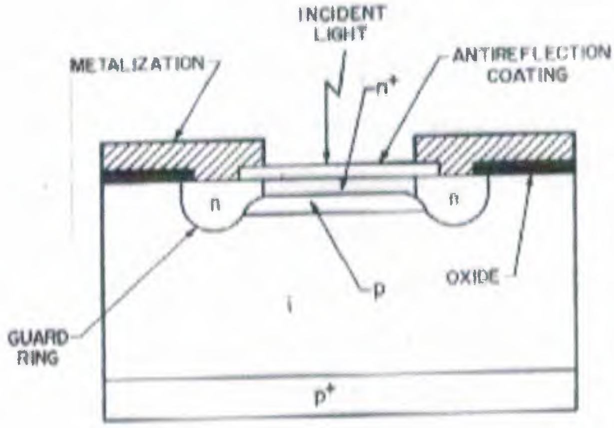

Figure 3.9 Schematic of a Typical Silicon Avalanche Photodiode.

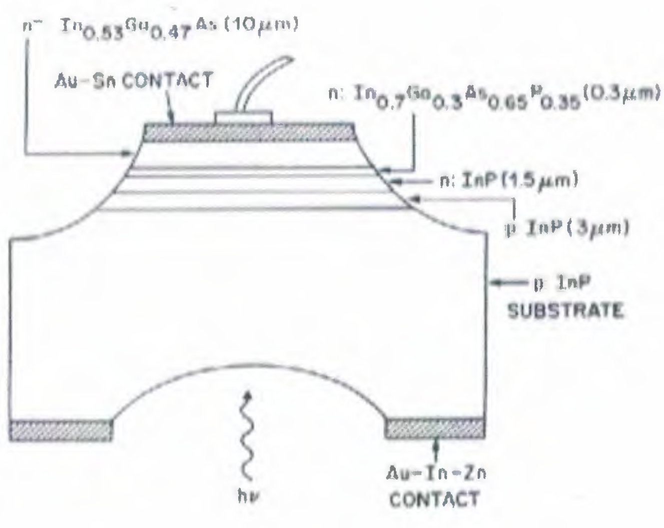

Figure 3.10 Schematic of an InGaAsP Avalanche Photodiode. Courtesy of J. C. Campbell, AT&T Bell Laboratories.

Uniformity and parameter control are critical in the fabrication of useful APDs. Variations in the high field region of the local field required for breakdown (due to material imperfections) can cause excessive noise due to random occurrences of very high multiplication (microplasma breakdown) of carriers passing through those regions. Non-uniformity of the high field region multiplication characteristics adds noise by enhancing the unpredictability of the multiplication experienced by individual carriers. Poor control of the width or doping levels in the high field region can result in devices which have a slow response speed (i-region not depleted) at the desired multiplication levels, have too large a required voltage for multiplication, or which have poor multiplication statistics. Figure 3.10 shows an avalanche photodiode fabricated in the InGaAsP material system using epitaxial growth and etching processes. Here a complex layered structure is grown in order to balance a number of important factors. The *i*-region (layer at the top labled n^-) is fabricated from a material with a narrow band-gap (energy difference between valence band and conduction band) in order to absorb the long wavelength light strongly. However the high field region (two layers below) is fabricated from a higher band gap material to prevent excessive leakage (dark) current from being generated in this volume. A transition layer between the high field region and the *i*-region prevents carriers from being blocked by abrupt energy level discontinuities (which form barriers).

The problem of coupling the fiber to the detector is not as difficult as in the case of the optical source. The detector has a relatively large area (compared to the fiber core) and a large acceptance angle. The fiber must be placed in reasonable alignment with the sensitive area and fixed to prevent motion from temperature-induced expansion of the assembly or mechanical stress.

To obtain high sensitivity, particularly with *p-i-n* detectors, it is important to minimize the total capacitance of the detector, the input of the preamplifier, and wiring. In addition, the low level and high impedance interface between the detector and the preamplifier is susceptible to induced noise. To accommodate these constraints the detector is often mounted as a chip component on a hybrid circuit board in close proximity to the preamplifier. The entire assembly, including the attached fiber, can then be enclosed within an environmentally sealed and shielded metallic package.

3.1.3 Detection Statistics

In this section we shall review the statistics of the detection process, both for *p-i-n* and APD detectors, to understand how these statistics affect the performance of receivers.

Classical radio systems are limited by thermal background noise. Since an optical system is essentially a very high frequency radio system (fiber systems operate at around 3×10^{14} Hz) one might ask why the limiting noises in optical systems should differ from those of radio systems. The answer is a matter of approximations. Classical radio systems span the frequency range from below 100 Hz to tens of gigahertz. Over this nearly nine order of magnitude range the models used to predict the performance of radio receivers have held up pretty well. However, when we make the additional four orders of magnitude leap to fiber system frequencies certain approximations begin to break down. Noise sources that were dominant at classical radio frequencies become negligible at optical frequencies (as we shall see shortly). Effects that were negligible at radio frequencies become important or dominating at optical frequencies. Consider the classical formula for thermal noise. It is given by kT watts per mode per hertz of bandwidth in a classical multimode transmission line, where k is Boltzmann's constant and T is the temperature of the medium producing the excitation of the transmission line. For example, for a low loss transmission line fed by an antenna, T would be the temperature of the medium in the field of view of the antenna. However, we can recall a result from physics, where it has been pointed out that such a flat noise spectrum implies infinite noise power if one considers the infinite possible range of frequencies. Quantum theory leads to a correction of the expression for thermal noise, and the corrected expression is given as follows:

$$N(f) = hf\,[\exp(-hf/kT) - 1]^{-1} \qquad (3.1)$$

Where f is the frequency, $N(f)$ is the noise spectral density per transmission line mode, and h is Planck's constant. Note that at room temperature kT is approximately 4×10^{-21} J. At 10^9 Hz hf is approximately 10^{-24} J. Thus, we see that for classical radio frequencies kT is much greater than hf. In this case equation (3.1) reduces to $N(f) - kT$; the classical expression. However, at 3×10^{14} Hz (1 μm wavelength) hf is approximately 2×10^{-19} J. In that case hf/kT is approximately 50. We see then that $N(f)$ is much less than kT

at these frequencies. In fact, unless we are looking into the sun or at some other very hot medium, thermal background noise is typically negligible in optical systems at wavelengths shorter than 2 μm, and in particular fiber systems.

If background noise is negligible, it might seem that there is no lower limit on the theoretical sensitivity of an optical receiver. However, this is not the case. A phenomenon known as quantum noise sets a lower limit on the detectable optical power level. What this lower limit is depends upon the detection problem at hand.

Suppose we have an ideal photodetector with no dark current. Suppose we illuminate the detector with a pulse of optical energy E joules. It can be shown that the number of hole—electron pairs produced within this (nonmultiplying) detector is not predictable; but follows a statistical distribution. The probability that exactly n hole—electron pairs will be produced is given by the Poisson distribution

$$p(n) = \Lambda^{n} e^{-\Lambda} / n! \qquad (3.2)$$

Where Λ is the average number of photons in the received pulse, Λ— E/hf, and e *is* the natural log base. From equation (3.2) we see that even though E is known in advance, we may get anywhere from zero to an infinite number of pairs produced. The average of this statistical distribution is E/hf pairs. Thus the average number of pairs produced (if we performed this experiment many times over) is equal to the average number of photons in the received pulse. The fact that the number of pairs produced is not predictable is not a defect in the detector; it is a manifestation of the phenomenon of quantum noise.

Consider the following situation. We produce, at a transmitter, an Optical pulse having the shape shown in Figure 3.11 with a peak power level of 2 mW and a duration of 1 ns. The total energy in the pulse is 1 pJ (10^{-12}.J). This pulse propagates through a medium (e.g., a fiber), which introduces a loss of 60 dB.

Figure 3.11 Transmitted Pulse Shape.

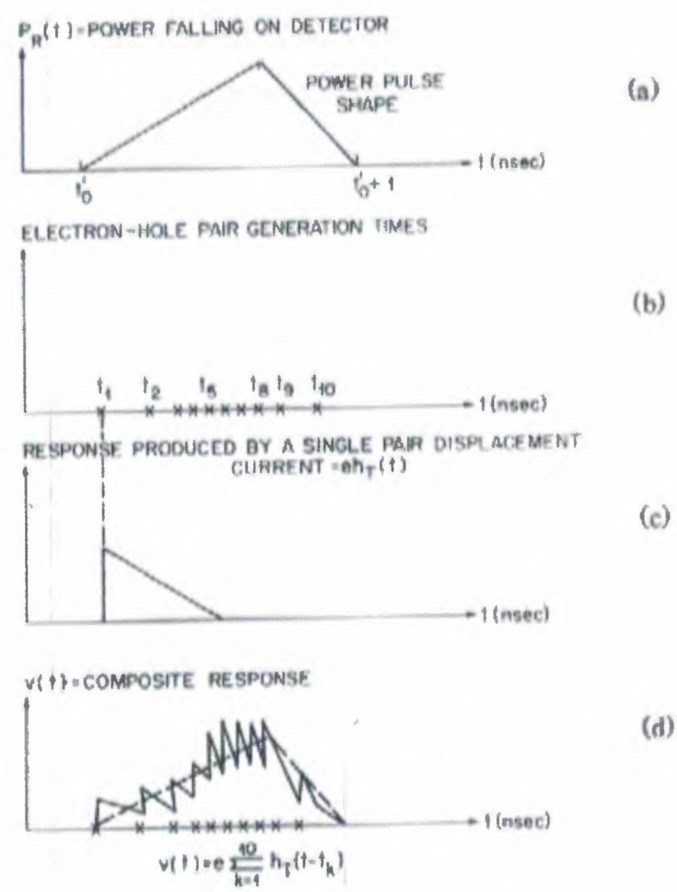

Figure 3.12 Realization of Detector Hole—Electron Pair Production Times, (a) Received Pulse Shape, (b) Times at Which Pairs are Produced, (c) Individual Hole-Electron Pair Displacement Current Response Waveform, (d) Composite Response Resulting from all Hole—Electron Pairs.

Thus the pulse received at the other end of the medium has energy 10^{-18} J. The energy in a photon is assumed to be 10^{-19} J at the wavelength being used. Thus the received pulse contains 10 photons (in a sort of average sense). This received pulse, shown in Figure 3.12a, will illuminate a detector that will produce hole—electron pairs in response. We cannot predict exactly how many hole—electron pairs will be produced, or what times (statistical uncertainty due to quantum effects). Figure 3.12b shows one possible realization of this experiment where 10 hole—electron pairs are produced at the times shown by the crosses. We assume that the detector produces a displacement current in response to each pair, having the shape shown in Figure 3.12c. Figure 3.12d shows the composite response from all of the displacement currents (from all ten pairs). We see that the composite response is reminiscent of the incident power waveform, but only a noisy approximation. Further, there is no arrangement of ten of the responses shown in Figure 3.12c, which superimpose to produce the shape shown in Figure 3.12a.

Suppose on the other hand that the photon energy was five orders of magnitude smaller (as it would be for a microwave frequency). In that case we would have 10^6 photons in the received pulse. This would result in about 10^6 hole—electron pairs produced in the detector. With 10^6 individual responses to superimpose, we can construct a composite displacement current waveform that is an excellent approximation to the shape shown in Figure 3.12a. Thus we see that quantum effects can be important at optical frequencies and negligible at microwave frequencies and below.

Consider next the following simple communication system. A transmitter is turned on or off to produce a pulse of either E_T J or 0 J. Thus the transmitter has a perfect extinction (off—on) ratio. This pulse arrives, attenuated, at a receiver containing an ideal *p-i-n* detector and an ideal electronic preamplifier. The received pulse energy is either E_r J or 0 J. The ideal *p-i-n* detector produces no dark current. Further there is no background light. Thus no hole—electron pairs are produced in the absence of a signal from the distant transmitter. The preamplifier electronics are so noise free that the displacement current of a single hole—electron pair is observable. At the receiver we use the following simple rule to determine if a pulse of light was sent from the distant end (or not). If no hole—electron pairs are produced in the detector we decide that no pulse was sent. If even a single hole—electron pair is observed (via its displacement current) we decide that a pulse was sent.

We note that if no pulse is sent by the distant transmitter, hole—electron pairs can be produced. Therefore the only way we can make a decision error is if a pulse of optical energy is sent, but still no pairs are produced. The probability that n pairs are produced if the received energy is E_r is given by

$$p(n) = (E_r/hf)^n \; e^{-(Er/hf)} \; / \, n\,! \qquad (3.3)$$

Thus the probability that no pairs are produced is given by

$$p(0) = exp[-(E_r/hf)] \qquad (3.4)$$

Suppose we wish to have the probability of such an error be equal to 10^{-9} then by algebra we have the requirement that E_r/hf be equal to 21. That is, in order to guarantee that only one time in a billion no pairs will be produced in response to a received pulse, we require that the received pulse contain 21 photons (on the average). This requirement of 21 photons per received pulse is called the quantum limit for binary direct detection (detection with a *p-i-n* detector). If we wish to communicate a sequence of pulses at rate B pulses per second, and if on the average half the pulses are on and half are off, then the power required at the receiver is

$$P_r = 1/2 \; x \; 21 \; x \; hf \, x \; B \qquad (3.5)$$

An actual receiver will not perform as well as the ideal receiver described above. To understand why, we must examine the assumptions we made. We assumed that the transmitter was perfectly off in the "off" state. We also assumed that there was no dark current in the detector, and no background light. These assumptions guaranteed that there were no pairs produced in the detector when the transmitter was not sending a pulse. In reality, the transmitter extinction ratio will not be perfect, and there will be some dark current in the detector. However, the effects of these departures from ideal on the performance of the receiver are not large. Typically we might find that the sensitivity of the receiver would be reduced from 21 photons required per pulse to about four times that value (a 6-dB sensitivity penalty) as a result of these deviations from ideal assumptions. The assumption that is most unrealistic is that the displacement currents produced by individual hole—electron pairs produced in the detector are observable in the background of preamplifier noise.

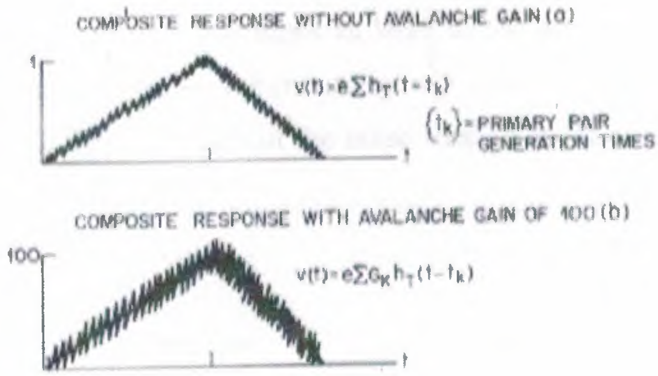

Figure 3.13 Composite Pulse Produced by a Photodiodc, (a) *p-i-n* detector, (b) APD detector.

Even with heroic design effort, preamplifier noise is hundreds of times as large as the response produced by a single hole— electron pair generated in the detector. Thus only the cumulative response of hundreds of hole—electron pairs is observable. For this reason e sensitivity of a typical receiver with a *p-i-n* detector is about 30 dB worse than the ideal of 21 photons required per received pulse. To obtain a performance closer to ideal, one requires a detector that produces many hole— electron pairs per detected photon.

As described above, an avalanche photodiode can produce a "bunch" secondary pairs in response to a primary photon- generated pair. This in mechanism can help to overcome the limitations of noises added by the preamplifier electronics. However, the number of secondary pairs produced by a given primary pair is only statistically predictable. Figure 3.13*a.* shows the composite current response of a *p-i-n* detector illuminated the optical pulse shown in Figure 3.11. The quantum noise effects due the finite number of hole—electron pairs contributing to the response are visible. Figure 3.13b shows the composite response of an APD, where each primary hole—electron pair produces an average of 100 secondary pairs via collision ionization. Note that although each primary produces, on the average, 100 secondary pairs, the actual number of secondary pairs produced by any primary pair is not predictable. The waveform shown in Figure 3.13b is, on the average, 100 times as large as the waveform shown in Figure 3.13a. However/the noise on the waveform of Figure 3.13b is more than 100 times larger than that of Figure 3.13a. (The width of the noise band is larger even though the scale factor has been increased by 100.)

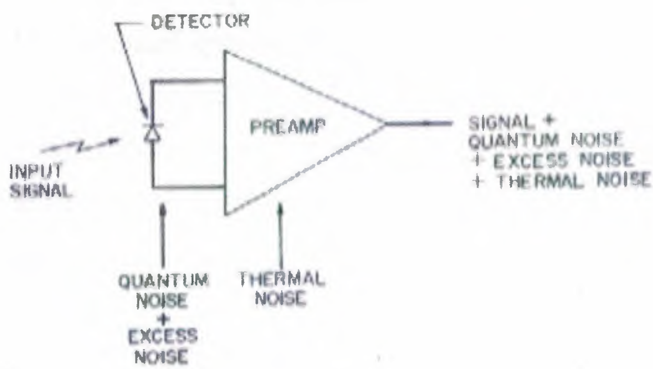

Figure 3.14 Noise Sources at the Preamplifier Output.

Thus the signal-to-noise ratio is smaller in Figure 3.13b. Why then would we use avalanche gain? The reason is that in both Figure 3.13a and Figure 3.13b we have not shown any additive amplifier noise. Suppose the amplifier noise had an rms value of 5 units height on the scales used in these figures. In Figure 3.13a with the *p-i-n* detector in use, this noise would completely mask the signal (which is only 1 unit in height). In Figure 3.13b the signal is 100 units in height, and would be clearly visible in the preamplifier noise of 5 units height. Thus, the net signal-to-noise ratio, taking into account the quantum noise, the enhancement of quantum noise caused by the randomness of the avalanche gain, and the preamplifier noise, is improved by the use of avalanche gain.

As the avalanche gain is increased from unity (no gain) the signal-to-noise ratio improves as described above. However, as the avalanche gain is increased to very large values one can reach a point of diminishing returns where the increased noise from the randomness of the gain mechanism may be worse than the preamplifier noise. This can be seen mathematically as follows. If we look at the output of the preamplifier shown in Figure 3.14 we see three components: the desired signal from the detector, quantum noise which is enhanced by the random avalanche mechanism, and preamplifier noise. The desired signal has an amplitude proportional to the average avalanche gain, G; and also proportional to the received optical power level P_r (watts). The preamplifier noise has some variance N_a. The enhanced quantum noise can be shown to have a variance given by the product of the received optical power level P_r, the square of the average avalanche gain, G^2, and an excess noise factor $F(G)$ which is larger than unity and grows with increasing average avalanche gain. Thus the signal-to-noise ratio, SNR, (ratio of desired signal squared-to-noise variance) at the preamplifier output is given by

$$SNR = aP_r^2\, G^2\, [N_a + bP_r\, G^2\, F(G)]^1 \qquad (3.6)$$

where a and b are constants.

We see that as G is increased from unity, the signal-to-noise ratio increases until the first term in the denominator is smaller than the second term. Beyond this, further increases in G can decrease the signal-to-noise ratio because $F\{G)$ increases with increasing G.

The statistics of avalanche multiplication are complex. Details are available in the references. However, expressions have been derived for the excess noise factor $F(G)$. In an APD, both holes and electrons can produce ionizing collisions. For a given distance of travel in the high field region one carrier is usually more likely to suffer an ionizing collision (is more highly ionizing). It can be shown that the best multiplication statistics are obtained if the more highly ionizing carrier initiates the multiplication process. That is, the device should be designed such that when photons are absorbed in the i-region, the more ionizing carrier drifts into the high field region. It can also be shown that the best avalanche statistics are obtained (minimum gain variance for a given average gain) when the ratio of ionizing ability of the two carrier types is large. For example in an ideal APD only one type of carrier (electrons or holes) would cause ionizations. The ratio of the probability of ionization per unit distance of motion in the high field region for holes and electrons is called the ionization ratio. If electrons are the more ionizing carrier, then we want this ratio to be as small as possible. The APD designer has some control over this ratio because the ionization probabilities are dependent upon the field levels in the high field region, and they scale differently for the two carriers. The ratio is called k (not to be confused with Boltzmann's constant) in the literature. Assuming that the multiplication process is initiated by the more ionizing carrier, we obtain the following formula for $F(G)$:

$$F(G) = \left[2 - \frac{1}{G}\right](1-k) + kG \tag{3.7}$$

Figure 3.15 shows a plot of $F(G)$ vs G for various values of k. The performance of a receiver using an avalanche detector depends upon the details of the avalanche gain distribution. However, certain approximate expressions for the performance depend only on the variance of the avalanche gain, which is given by G^2 $F(G)$. We will not reproduce the complicated derivations of the avalanche gain distribution here.

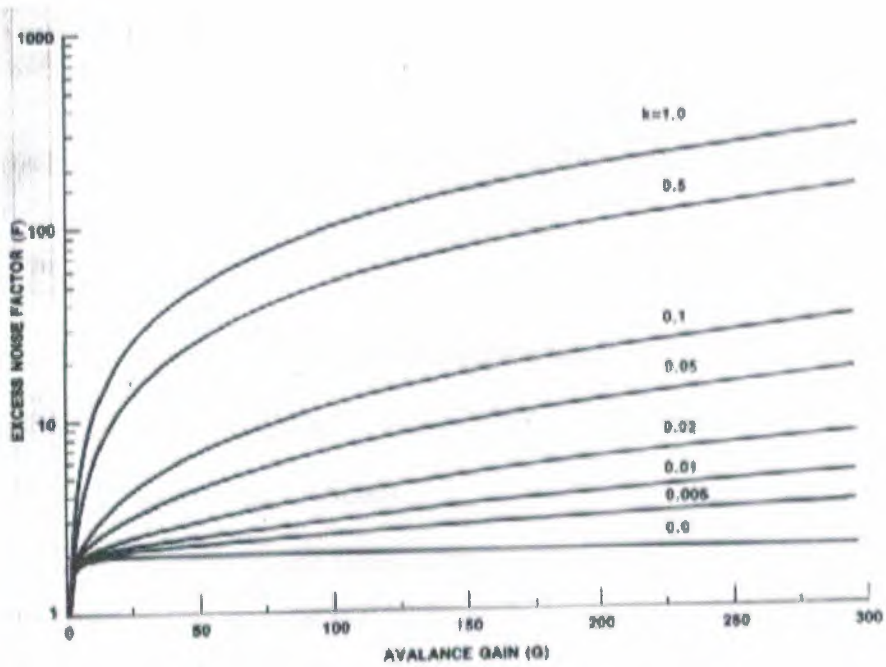

Figure 3.15 Excess Noise Factors vs Gain of APD Detectors.

However, we can state a few interesting results. For an ideal unilateral gain detector, where only one carrier suffers ionizing collisions ($k = 0$ or k = infinity), the gain distribution is geometric. That is, the probability of exactly g secondaries produced by a primary pair (including the primary pair itself) is

$$P_G(g) = \frac{1}{G-1}\left[1-\frac{1}{G}\right]^g \quad \text{for } g \geq 1 \tag{3.8}$$

Where G is the average gain (multiplication)

This is approximately an exponential distribution when G is larger than 10. When k is not 0 or infinity the distribution of the gain g is more complex, but always has an exponentially shaped tail. An approximate expression for this distribution, valid for G greater than 10 is

$$P_G(g) = \frac{(2\pi)^{-1/2}\, G^{1/2}\, F(G)(1/g)^{3/2}}{(F(G)-1)^{3/4}} \exp[-g/2G[F(G) - 1]^{1/2}] \tag{3.9}$$

Where G is the average value of the gain g and $F(G)$ is the excess noise factor, given in equation (3.7).

3.2 Receiver Modules

Two kinds of semiconductor in receivers are used:

• *p-i-n* (*p*-intrinsic-*n*) diode;

• APD (Avalanche Photo-Diode).

Instead of "receiver" some times is used light detector, photo-detector or optical-to-electric converter. The receiver part of the optical communications system is relatively simple. It consists of a detector that will sense the light pulses and convert them into an electrical signal. This signal is then amplified and shaped into the original digital serial data. The most critical component, of course, is the light sensor.

The most widely used light detector is a photo diode. This is silicon *pn* junction diode that is sensitive to light. This diode is normally reverse-biased. Whenever light strikes the diode, this leakage current will increase significantly. It will flow through a resistor and develop a voltage drop across it. The result is an output voltage pulse.

The resulting voltage pulse is very small, so it must be amplified. This can be done by using a phototransistor. Thus the transistor amplifies the small leakage current into a larger.

The sensitivity and response time of a photo diode can be increased by adding an undraped or intrinsic (*i*) layer between the *p* and *n* semiconductors to form a *p-i-n* diode.

Although the *p-i-n* photodiode is extremely well suited for most fiber optic applications, its sensitively to light (responsively) is not as great as the avalanche photodiode (APD). Due to their inherent gain, typical values of responsively for APDs may range from 5 A/W to as high as 100A/W. This is considerably higher than the *p-i-n* photodiode, which makes it extremely attractive for fiber-optic communications receivers.

Figure 3.16 shows the basic circuit used in most receivers. The current through the photodiode (PD) generated when light is sensed produced a current, which is then amplified in an amplifier (A). The pulses to ensure fast rise and fall times. The output is passed through a logic gate so that the correct binary voltage levels are produced. Most systems have a data rate of 10 to 500 Mbits/s.

The product of the bit rate and the distance usually indicates a system performance. This rating tells the fastest bit rate that can be produced over a 1-km cable. Assume a system with a 100 Mbits-km/s rating. If the distance increases, the bit rate decreases in proportion.

Another important consideration is the maximum distance between repeaters. Obviously the fewer the repeaters are better. The average distance between repeaters is now up to 100 km range.

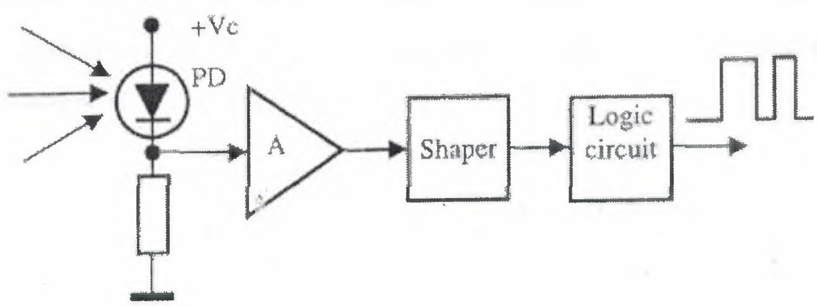

Figure 3.16 Basic Receiver Circuit.

A receiver consists of the combination of a detector and a preamplifier that accommodates the characteristics and requirements of the detector to optimally interface it to conventional electronics. The objective is to build a receiver subsystem, which provides a combination of high sensitivity (low required incident optical power to achieve a desired performance), adequate dynamic range (ability to accommodate a range of incident optical power levels), adequate bandwidth, and a proper interface to the remaining electronics beyond the receiver.

3.2.1 Preamplifier Design

The preamplifier in an optical receiver must be designed to act as an interface between the detector and conventional electronics. In particular it must amplify the weak current produced by a photodetector; provide a low impedance level interface to following electronic stages; and it must do so with as little additive noise as possible. Preamplifiers are required in conventional microwave radio and metallic cable systems as well. However, the characteristics of the optical detector are different from those of an antenna feed or a coaxial cable. Figure 3.17 shows a simplified schematic of a resistive signal source, characteristic of many (but not all) classical systems. The signal is represented by a current source in parallel with a resistance. Also included is a noise current source associated with the thermal noise of the resistance. Figure 3.18 shows a schematic of the equivalent circuit of a photodiode. The signal is represented by a current source in parallel with a capacitance. (The series resistance of the photodiode is neglected here). If we neglect the shot noise of leakage current, there is no intrinsic noise source apparent in this model.

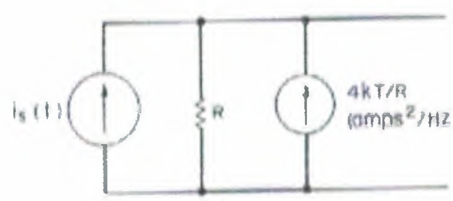

Figure 3.17 Resistive Signal Source.

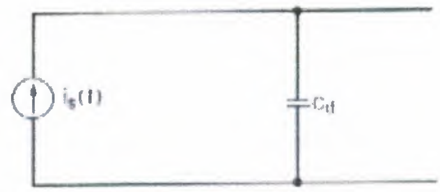

Figure 3.18 Capacitive Signal Source.

A photodiode is often referred to as a capacitive source. Other examples of capacitive sources are vidicon tubes and biological signal sources. When we design a preamplifier to work with a resistive source Such as the one shown in Figure 3.17 we characterize the noise added by the preamplifier electronics by a measure called the noise figure. The noise figure is the ratio of total noise at the preamplifier output to the portion of the output noise due to the intrinsic noise of the signal source itself. If the preamplifier adds a noise, which is small compared to the intrinsic noise of the signal source, then the noise figure is close to unity. If the noise figure is close to unity, there is little benefit in trying to further reduce the preamplifier noise contribution.

When we have a capacitive source as shown in Figure 3.18 there is no meaningful intrinsic noise against which to compare the preamplifier noise. The noise figure in this case is always very large, since the preamplifier noise dominates the total. For a capacitive source, a more meaningful measure of preamplifier noise can be obtained by asking whether or not individual hole—electron pairs generated in the detector produce an observable signal at the preamplifier output. If the individual hole—electron pair responses are observable, than the receiver fundamental quantum effects will limit performance. We can define a parameter, *Z*, as follows:

$$Z = \frac{\text{output rms noise}}{\text{output response to an electron–hole pair}} = \sigma_o / v_e \qquad (3.10)$$

Where σ_0 is the rms noise at the receiver output in the band of frequencies occupied by the modulation of the input signal, and where v_e is the peak response produced at the receiver output by an individual hole—electron pair generated in the detector.

If Z is a number less than unity, then individual pair responses are observable. If Z is larger than unity, then only the cumulative output due to many generated pairs is observable. Consider the following example. Figure 3.19 shows a simple optical receiver where a photodetector is coupled to a conventional amplifier modeled as shown. In this example we assume that the conventional amplifier has a physical 50-Ω resistor at its input, and that the resistor produces Johnson (thermal) noise of spectral density $4kT/R$ (A^2/Hz). The receiver has a bandlimiting filter of bandwidth B. Presumably B is the bandwidth required to accommodate the modulation of the incoming light signal. To calculate Z we must look at the output of the receiver. We must calculate the rms noise due to the input resistance of the amplifier; and we must calculate the response at the same point produced by a pair generated in the detector. Using standard noise theory we obtain the variance of the noise at the receiver output as follows:

$$\sigma_o^2 = 4kT50A^2B \tag{3.11}$$

Where A is the amplifier transconductance (G_m) x the filter gain and $kT = 4\text{x}10^{-21}$ J (at room temperature).

We can derive the amplitude of the output pulse produced by a pair generated in the detector as follows. The generated pair produces a displacement current whose area is e, the charge of an electron. This current flows through the 50-Ω resistor (we neglect capacitance in this example) to produce an input voltage waveform whose area is 50 x e (V-s). This results in an output waveform whose area is 50 x e x A (V-s). The bandlimiting filter has bandwidth B. Thus the pulse produced at the filter output has duration of roughly $1/B$. Thus the peak of the response produced at the filter output is roughly 50 x e x A x B (V).

We then obtain the following expression for Z:

$$Z = \frac{(4kT50A^2B)^{1/2}}{50eAB} = \left[\frac{4kT}{50e^2B}\right]^{1/2} = 10^8\left[\frac{1}{B}\right]^{1/2} \tag{3.12}$$

We see that Z decreases as the bandwidth B increases. This is because the peak response grows linearly with B, while the rms noise grows only as the square root of B.

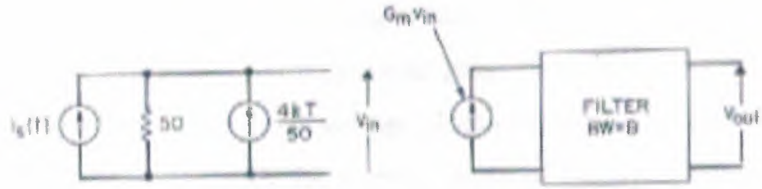

Figure 3.19 50-Ω Amplifier Example.

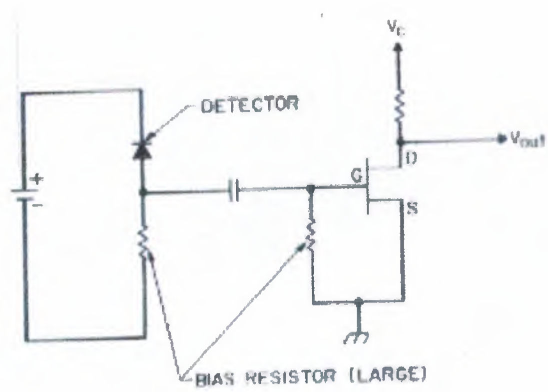

Figure 3.20 FET Preamplifier Example.

However, at very high bandwidths other noise sources that we have neglected (due to the preamplifier active components) become important and cause Z to increase with bandwidth. As an example, if the bandwidth B is 10^8 Hz, then Z is approximately 10,000. This implies that 10,000 pairs must be generated in the detector to produce a cumulative response, which is just equal to the noise at the preamplifier output.

With careful electrical and physical design, we can implement preamplifiers with values of Z of between several hundred and several thousand, depending upon the bandwidth and other tradeoffs. Figure 3.20 shows a preamplifier based on a field-effect transistor. Figure 3.21 shows an equivalent circuit for this simple preamplifier. The input of the field effect transistor (FET) is a capacitance. The transistor produces an output current proportional to the input voltage, with proportionality constant G_m. There is a noise source in series with the input voltage produced across the parallel detector and FET capacitances, associated with the thermal noise of the source-drain channel. This noise source has single-sided spectral density $2.8kT/G_m$ (V^2/Hz).

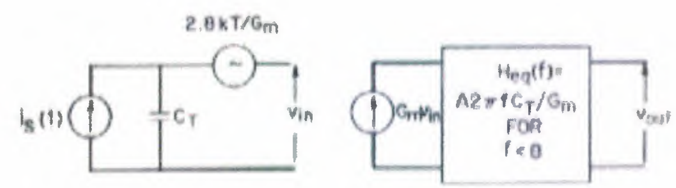

Figure 3.21 FET Preamplifier Equivalent Circuit.

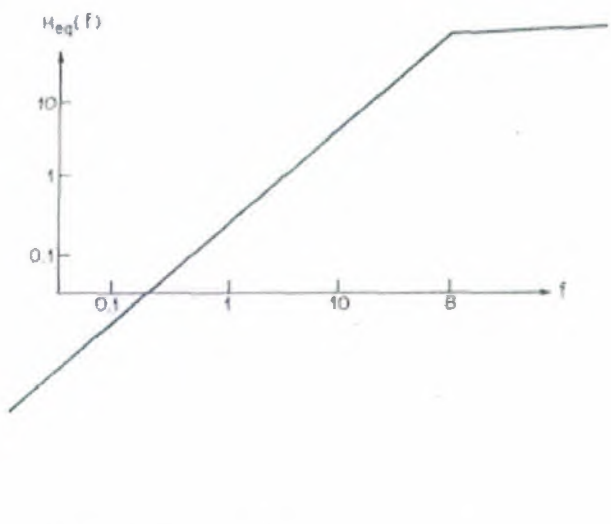

Figure 3.22 Equalizer Frequency Response.

The parallel capacitance at the input of the FET causes the signal current produced in the detector to be integrated. This is compensated by a differentiating network at the output of the FET as shown in Figures 3.22 and 3.23, resulting in a net frequency response which is flat out to some desired bandwidth *B*. Note that the preamplifier input noise source is not integrated at the input. Thus it produces a noise at the receiver output, which does not have a flat spectrum. We can next calculate *Z* for a typical FET preamplifier using the following parameter values. We shall assume that the detector capacitance is 0.5 pf and that the preamplifier input capacitance plus any stray capacitance at the input totals 0.5 pf. Thus the net input capacitance, C_T, totals 1.0 pf. We shall assume that the FET is a GaAs device with a transconductance G_m, of 10 millisiemens (mS). The rms noise at the receiver output can be obtained from standard noise theory by integrating the frequency dependent noise produced by the input noise source over the band of frequencies *B*.

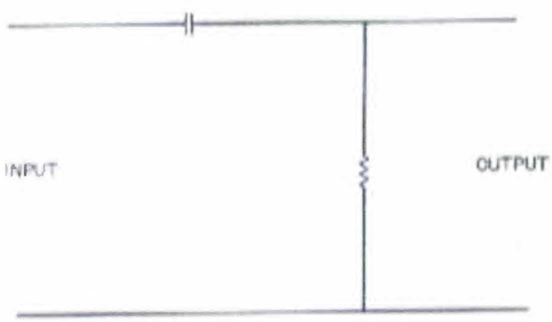

Figure 3.23 Equalizer Schematic.

We obtain the following result:

$$\sigma_n^2 = \frac{2.8kT}{3G_m}(2\pi C_T)^2 A^2 B^3 \qquad (3.13)$$

The peak of the response produced at the receiver output by a pair generated in the detector can be calculated (as we did in the 50-Ω preamplifier case above) to be

$$V_e = eAB \tag{3.14}$$

Thus we obtain the following expression for Z:

$$Z = \frac{\sigma_o}{v_e} = \left[\frac{2.8kTB}{3G_m e^2}(2\pi C_T)^2\right]^{1/2} = 2.4 \times 10^{-2} B^{1/2} \tag{3.15}$$

Note that Z increases as B increases due to the input noise source of the transistor. Note also that Z is proportional to $C_T/G_m^{1/2}$. Thus to minimize Z we wish to minimize the total input capacitance while at the same time using a device with a large value of G_m. If we set B = 10^8 Hz and if we use the parameter values given above then we obtain a value of Z equal to 240.

We can also make low noise preamplifiers for capacitive sources using bipolar transistors. One can show that with bipolar transistor preamplifiers the value of Z for a given selection of transistors is independent of B over a wide range of frequencies, provided the preamplifier is optimally biased for the value of B selected. A typical value of Z for a bipolar transistor preamplifier in the range of bandwidths from 100 MHz to 2 GHz is around 1000, provided appropriate transistors are used.

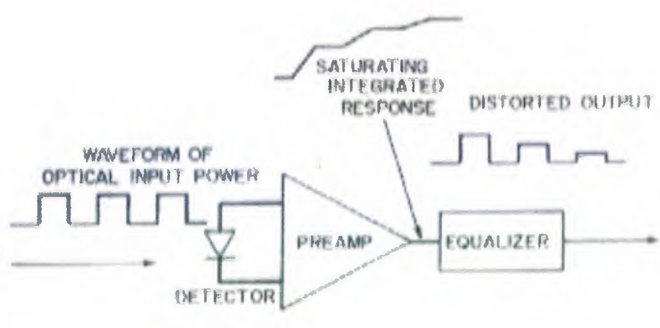

Figure 3.24 Nonlinear Effects in High Impedance (Integrating) Preamplifiers.

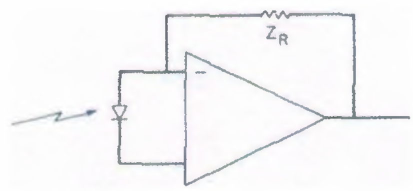

Figure 3.25 Transimpcdancc Preamplifier.

Preamplifiers that are designed to minimize Z often have very limited dynamic range. If we reexamine the FET preamplifier described above we observe that the compensation network at the output of the preamplifier can only remove the integration caused at the preamplifier input if the preamplifier operates linearly. Low frequencies in the input light signal modulation result in low frequency currents in the detector response that can be integrated by the total capacitance at the preamplifier input- to produce large voltages. Thus if the optical signal is above its minimal detectable level, and if the modulation contains low frequency components, the preamplifier may overload (become nonlinear). Figure 4-23 illustrates how this process might distort a sequence of pulses at the input to the preamplifier might be distorted by this process.

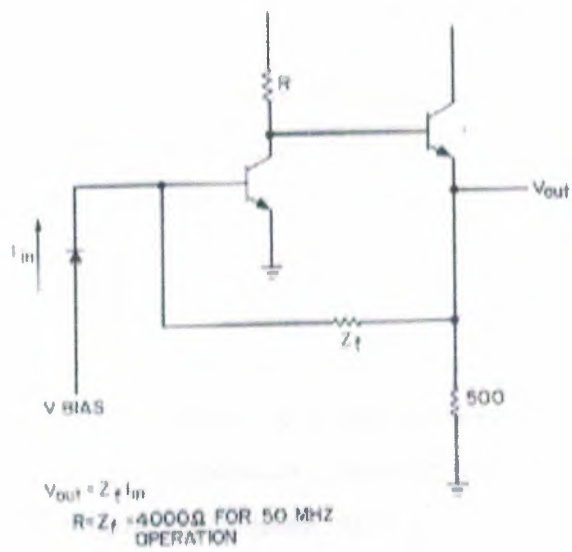

Figure 3.26 Example of a Transimpedance Preamplifier.

To increase the dynamic range of the preamplifier, several approaches have been used. Figure 3.25 shows a transimpedance preamplifier design, where the preamplifier is used as an operational amplifier. In order for the transimpedance design to work properly, the gain around the feedback path (loop) must be large Compared to unity. Since the feedback resistance and the amplifier input impedance form a voltage divider, the allowable value of the feedback resistance is limited. The feedback resistance adds noise at the amplifier input that typically dominates the noises from the active components (transistors) in the preamplifier. Thus a transimpedance preamplifier typically has more noise than an integrating design made from the same components. A useful rule of thumb is that Z is 2—3 times larger for the transimpedance design (vs the integrating design). Figure 3.26 shows a transimpedance preamplifier using bipolar transistors and having a bandwidth of about 100 MHz. The Z value for this preamplifier is about 3000.

Other approaches to improving the dynamic range of a preamplifier (beyond the use of a transimpedance design) include the use of nonlinear elements at the input of the preamplifier (to prevent overload of the preamplifier itself) and the use of an adjustable transimpedance (to reduce the transimpedance value at high signal levels).

3.2.2 APD Control Circuitry

The average multiplication (gain) of an avalanche photodiode is dependent upon the applied reverse bias and the temperature of the device. Figure 4-4 shows curves of responsivity vs voltage and temperature for a typical APD. In. a practical system application it is usually necessary to provide some mechanism for assuring that the APD gain is correctly set. When relatively low values of gain are used (i.e., where the temperature sensitivity is modest) open loop compensation of the applied reverse bias can be sufficient. Thus for those situations one might implement a controller for the high voltage APD supply whose temperature characteristic approximately tracks the requirements for maintaining nearly constant APD gain. When higher values of APD gain are used, closed-loop control mechanisms may be necessary (unless the APD is maintained in a controlled ambient). Figure 3.27 shows a receiver design often implemented in digital systems where the APD is included in the automatic gain control loop of the receiver.

The output level of the receiver can be adjusted via either the gain of the APD or the gain of the variable gain amplifier (VGA). The output peak-to-peak level is measured by a peak-to-peak detector (a nontrivial circuit at high bandwidths), averaged, and compared to a reference. The low frequency control circuitry of the AGC system produces an error voltage in response to the difference of the reference and the averaged peak-to-peak measurement.

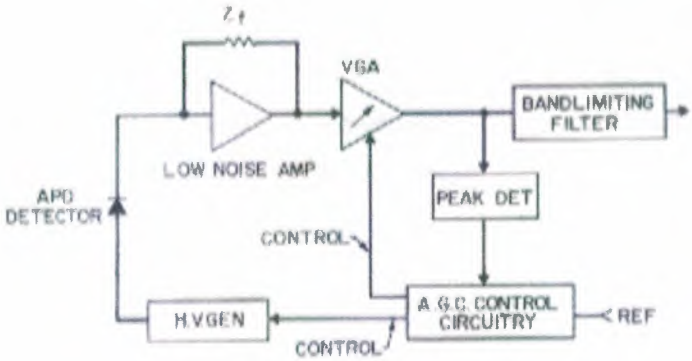

Figure 3.27 Feedback Control of APD Gain.

This error voltage is used to adjust either of the available gain mechanisms (or both). One possible strategy is to have the APD high voltage supply controller respond over a range of AGC error voltages which is separate from the range of error voltages which control the VGA. For example, the VGA could be adjusted to be at full gain for error voltages between 0 and 6 V, and to be at reduced gains for error voltages between 6 and 12V. Meanwhile the high voltage controller could be adjusted to produce some minimum level of output voltage (required for proper APD operation) for error voltages above 5 V, but to produce higher voltages for error voltages between 0 and 5 V. In this scheme, if the output of the receiver is too high, the AGC control loop will first reduce the APD gain down to some minimum permissible value (set by the requirement that the *i*-region in the device be swept out), and will then reduce the gain of the VGA. This scheme will automatically adjust the gain of the APD to compensate for temperature changes when the APD is operating in the high range of gains (where the AGC is controlling its voltage). When the APD voltage is fixed at its lower limit, the APD will be operating at low gains where the precise gain value is not critical and where the APD gain is less sensitive to temperature changes. Figure 3.28 shows an example of a high voltage controller circuit that uses a shunt current path to drop the voltage applied to the APD.

The Zener diodes limit the maximum and minimum APD voltages. There are some cautions that must be observed in using this design. The APD will not operate as a high-speed device unless it is biased with sufficient voltage to deplete its *i*-region. The required voltage may vary amongst APDs, which are nominally the same. Thus some specification needs to be made regarding the voltage required, for *i*-region depletion, and the lower limit of the bias voltage must be set accordingly.

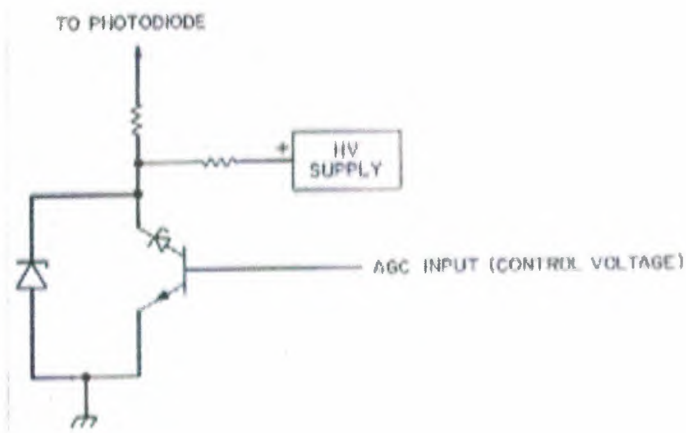

Figure 3.28 High Voltage Control Circuit.

The voltage required to produce the desired gain in one APD may produce breakdown in another nominally similar device. In some applications caution must be exercised to limit the reverse bias to a value below breakdown (although this is not always a problem). In designing the feedback loop one must be careful that the differing time constants of the multiple gain control mechanisms do not result in oscillatory behavior. It can also be shown that if the control loop is set up initially to adjust the APD gain to the optimal value at some minimum anticipated optical signal level, then the gain will not be optimal (too low) at higher signal levels. (The loop is keeping the product of received power level and avalanche gain constant.) It can be argued that when the optical signal level increases, the receiver performance increases, even with the suboptimal APD gain setting. However, in some applications this may not be sufficient.

In analog systems, the required avalanche gain is usually low enough that the temperature sensitivity of the APD gain can be ignored or tracked by open loop methods. However, this would not be the case if the absolute end-to-end gain of the link is specified to tight tolerances. In such a situation some sort of pilot tone or other measure of APD gain must be available at the receiver output, in order to set the gain by closed loop control.

CHAPTER 4

APPLICATIONS AND FURURE DEVELOPMENTS

4.1 Introduction

In order to appreciate the many areas in which the application of light wave transmission via optical fibers may be beneficial, it is useful to review the advantages and special features provided by this method of communication. The primary advantages obtained using optical fibers for line transmission may be summarized as follows:

(a) enormous potential bandwidth;

(b) small size and weight;

(c) electrical isolation;

(d) immunity to interference and crosswalk;

(e) signal security;

(f) low transmission loss;

(g) raggedness and flexibility;

(h) system reliability and ease of maintenance;

(I) potential low cost.

The use of fibers for optical communication does have some drawbacks in practice. Hence to provide a balanced picture these disadvantages must be considered. They are:

(a) the fragility of the bare fibers;

(b) the small size of the fibers and cables which creates some difficulties with splicing and forming connectors;

(c) some problems involved with forming low loss T-couplers;

(d) some doubt in relation to the long-term reliability of optical fibers in the presence of moisture;

(e) an independent electrical power feed is required for any repeaters;

(f) new equipment and field practices are required;

(g) testing procedures tend to be more complex.

A number of these disadvantages are not just inherent to optical fiber systems but are always present at the introduction of a new technology. Furthermore, both continuing developments and experience with optical fiber systems are generally reducing the other problems.

In this chapter we consider current and potential applications of optical fiber communication systems together with some likely future developments in the general area of optical transmission and associated components. The discussion is primarily centered around application areas including the public network, military, industrial and computer systems.

4.2 Public Network Applications

The public telecommunications network provides a variety of applications for optical fiber communication systems. It was in this general area that the suitability of optical fibers for line transmission first made an impact. The current plans of the major PTT administrations around the world feature the installation of increasing numbers of optical fiber links as an alternative to coaxial and high frequency pair cable systems.

4.2.1 Trunk Network

The trunk or tool network is used for carrying telephone traffic between major conurbations. Hence there is generally a requirement for the use of transmission systems, which have a high capacity in order to minimize costs per circuit. The transmission distance for trunk systems can vary enormously from under 20 dm to over 300 dm. and occasionally to as much as 1000 km. Therefore transmission systems which exhibit low attenuation and hence give a maximum distance of unrepeated operation are the most economically viable.

4.2.2 Junction Network

The junction or interoffice network usually consists of routes with major conurbation's over distances of typically 5-20km. However, the distribution of distances between switching centers (telephone exchanges) or offices in the junction network of large urban areas varies considerably for various countries as indicated in Figure 4.1.

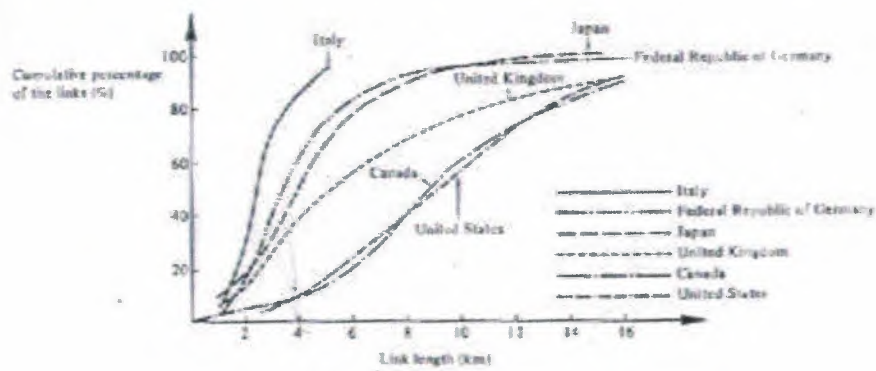

Figure 4.1 Distribution of distances between switching centers in metropolitan areas.

It may be observed from Figure 4.1 that the benefits of long unreported transmission distances offered by optical fiber systems are not as apparent in the junction network due to be generally shorter link lengths. Nevertheless optical fiber junction systems are often able to operate using no intermediate repeaters whilst alleviating duct congestion in urban areas.

4.2.3 Local and Rural Networks

The local and rural network or subscriber loop connects telephone subscribers to the local switching center or office. Possible network configuration are shown in Figure 4.2 and include a ring, tree and star topology from the local switching center. In a ring network (Figure 4.2(a)) any information fed into the network by a subscriber passes through all the network nodes and hence a number of transmission channels must be provided between all nodes. The tree network, which consists of several branches as indicated in Figure 4.2(b), must also provide a number of transmission channels on its common links.

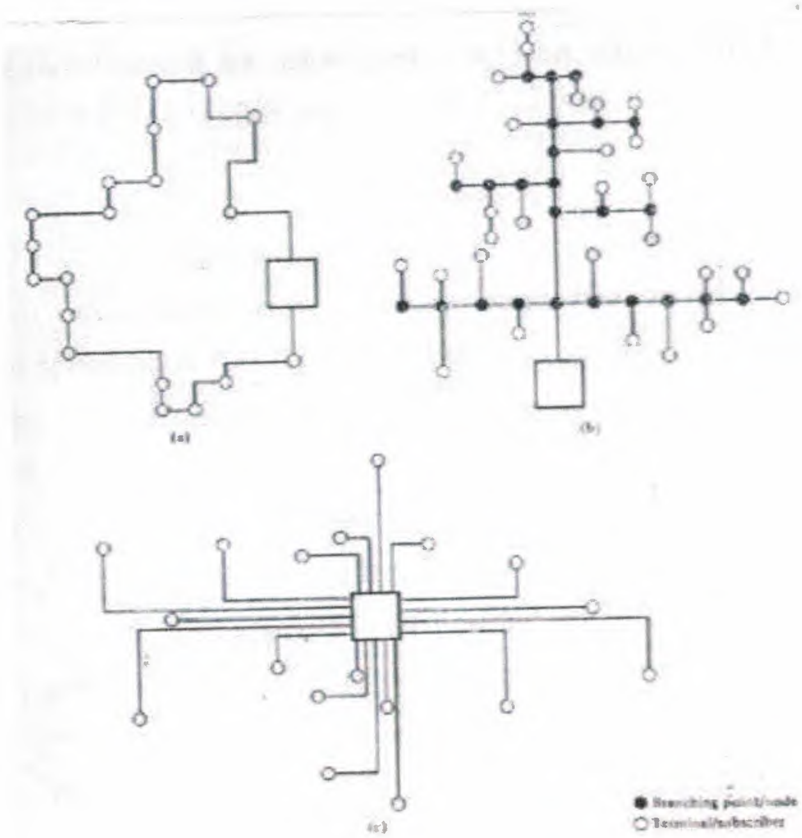

Figure 4.2 Local and rural network configurations: (a) ring network; (b) tree network; (c) Star network.

In contrast, the star network (Fig. 4.2(c)) provides a separate link for very subscriber to the local switching center. Hence the amount of cable required is considerably increased over the ring or tree network, but is offset by enhanced reliability and availability for the subscribers. In addition simple subscriber equipment is adequate (i.e. no TDM) and network expansion is straightforward. Thus virtually all local and rural telephone networks utilize a star configuration based on copper conductors (twisted pair) for full duplex (both way) speech transmission.

4.3 Military Applications

In these applications, although economics are important, there are usually other, possibly overriding, considerations such as size, weight, deployability, survivability (in both conventional and nuclear attack and security. The special attributes of optical fiber communication systems therefore often lend themselves to military use.

4.3.1 Mobiles

One of the most promising areas of military application for optical fiber communications is within military mobiles such as aircraft, ships and tanks. The small size and weight of optical fibers provide an attractive solution to space problems in these mobiles that are increasingly equipped with sophisticated electronics. Also the wideband nature of optical fiber transmission will allow the multiplexing of a number of signals onto a common bus. Furthermore, the immunity of optical transmission to electromagnetic interference (EMI) in the often noisy environment of military mobiles is a tremendous advantage. This also applies to the immunity of optical fibers to lightning and electromagnetic pulses (EMP) especially within avionics. The electrical isolation, and therefore safety, aspect of optical fiber communications also proves invaluable in these applications, allowing routing through both fuel tanks and magazines.

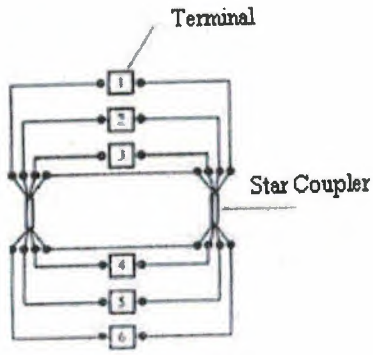

Figure 4.3 Multiple access bus showing the interconnection of six terminals using two four-way transmission star couplers.

The above advantages were demonstrated with preliminary investigations involving fiber bundles and design approaches now include multi terminal data systems using single fibers and use of an optical data bus. In the former case, the time division multiplex system allows ring or star configurations to be realized or mixtures of both to create bus networks. The multiple access data highway allows an optical signal injected at any access point to appear at all other access points. An example is shown in Figure 4.3, which illustrates the interconnection of six terminals using two four-way transitive star couplers. These devices give typically 10 dB attenuation between any pair of ports.

4.3.2 Communication Links

The other major area for the application of optical fiber communications in the military sphere includes both short and long distance communication links. Short distance optical fiber systems may be utilized to connect closely spaced items of electronic equipment in such areas as operations rooms and computer installations. A large number of these systems have already been installed, in military installations in the UK.

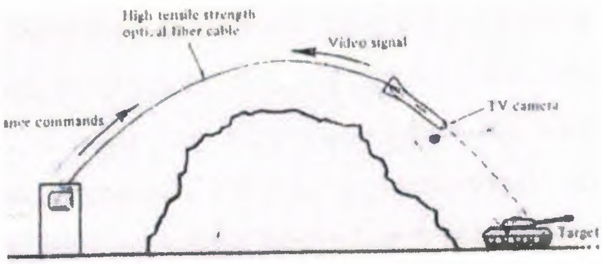

Figure 4.4 A fiber guided weapons system.

Other long distance applications include torpedo and missile guidance, information links between military vessels and maritime, towed sensor arrays. In these areas the available bandwidth and long unrepeated transmission distances of optical fiber systems provide a solution that is not generally available with conventional technology. A fiber-guided weapons system is illustrated in Figure 4.4 whereby an operator uses a low loss, high tensile strength fiber to relay a video signal back to control station to facilitate targeting.

In summary, it appears that confidence is being established in this new technology such that its wide scale use in military applications in the future is ensured.

4.4 Computer Applications

Modern computer systems consist of a large number of interconnections. These range from lengths of a few micrometers (when considering on chip very large scale integration (VLSI) connections) to perhaps thousands kilometers for terrestrial links in computer networks.

Nevertheless it is likely that optical transmission techniques and optical fibers themselves will find application with in data processing equipment. In addition, investigations have already taken place into the use of optical fibers for mains isolators and digital data buses within both digital telephone exchanges and computers. Their small size, low loss, low radiation properties and freedom from ground loops provide obvious advantages in these applications.

At present, however, a primary potential application for optical fiber communications occurs in interequipment connections. These provide noise immunity, security and removal of earthy loop problems, together with increased bandwidth and reduced cable size in comparison with conventional coaxial cable computer system interconnections. The inter equipment connection topology for a typical mainframe computer system (host computer) is illustrated in Figure 4.5.

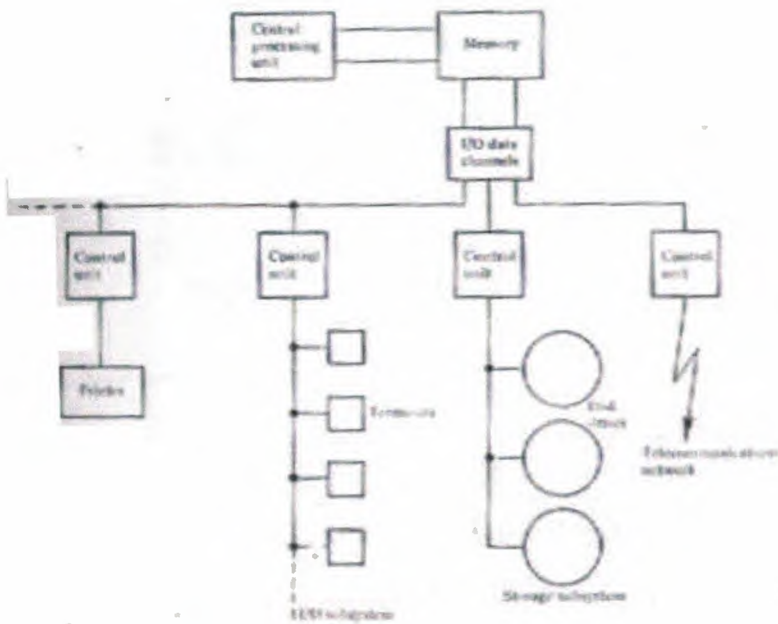

Figure 4.5 Block schematic of a typical mainframe computer system.

4.4.1 Local Area Networks

A local area network (LAN) is generally defined as an interconnection topology entirely confined within a geographical area of a few square kilometers. It is therefore usually confined to either a single building or a group of buildings contained within a site or establishment (industrial, military, educational, etc). Hence, the data processing and peripheral equipment together with any communication links are usually under the control of the owning body rather than a common carrier.

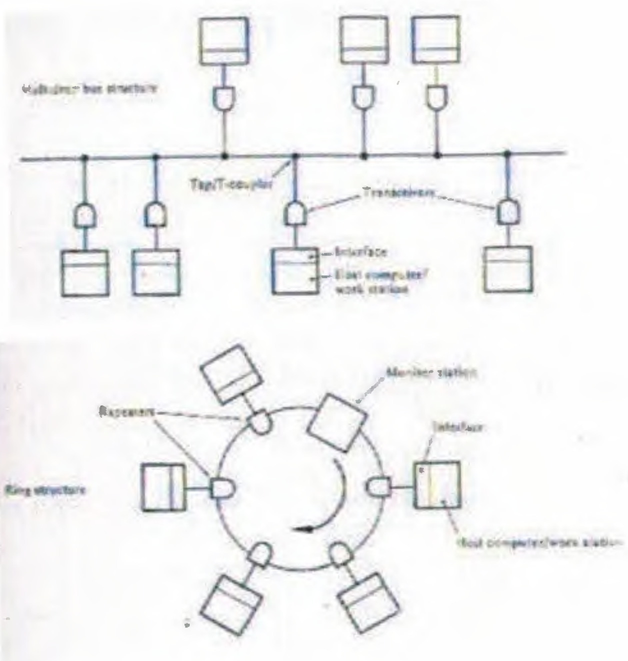

Figure 4.6 Local area networks: (a) the Ethernet network topology and packet format;(b) the Cambridge ring topology and packet format.

4.5 Beam Splitters and Switches

Beam splitters are a basic element of many optical fiber communication systems often providing a Y-junction by which signals from separate sources can be combined, or the received power divided between two or more channels. A passive Y-junction beam splitter fabricated from $LiNbO_3$ is shown in Figure 4.7. Unfortunately the power transmission through such a splitter decreases sharply with increasing half angle γ, the power being radiated into the substrate. Hence the total power transmission depends critically upon γ, which for the example chosen must not exceed $0.5°$ if an acceptable insertion loss is to be achieved. In order to provide effective separation of the output arms so that access to each is possible, the junction must be many times the width of the guide. For example around 3000 wavelengths are required to give a separation of about 30 μm between the output arms. Therefore, for practical reasons, the device is relatively long.

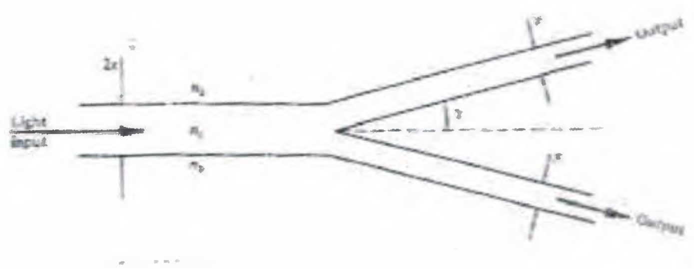

Figure 4.7 A passive Y-junction beam splitter.

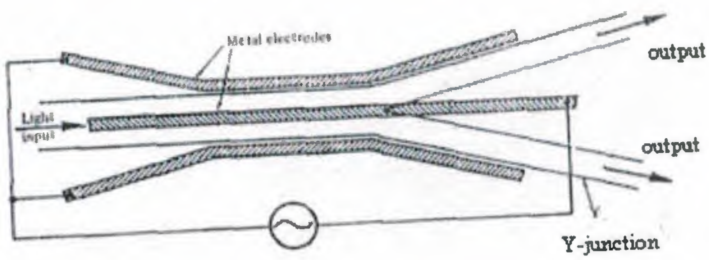

Figure 4.8 An electro-optic Y-junction switch.

The passive Y-junction beam splitter finds application where equal power division of the incident beam is required. However, the Y-junction is of wider interest when it is fabricated from an electro-optic material, in which case it may be used as a switch. Such materials exhibit a change in refractive index δn, which is directly proportional to an applied electric field E following,

$$\delta n = \pm \tfrac{1}{2} n_1^3 r E \qquad (4.1)$$

Where n_1 is the original refractive index, and r is the electro-optic coefficient.

CONCLUSION

Fiber optic communication technology is developed several generations. Larger transmission capacity and longer transmission distance are the two primary objectives. Higher output power and smaller fiber attenuation are essential to longer transmission distance, and good spectral coherence is key to higher transmission speeds. Due to this reason most efforts have been made to improve the output power and spectral coherence of light sources and to reduce fiber attenuation and dispersion.

In past, signals were detected incoherently when optical communications was not well known. But as the time passes coherent detection was used to enhance the receiver's sensitivity. With coherent detection, received signals are amplified by the local carrier, which makes the system performance limited by shot noise.

Optical amplifiers have been developed to eliminate the attenuation and dispersion limits. Optical amplifiers amplify optical signals directly in the optical domain. Researches in optical amplifiers have become sufficiently mature now a days, and repeaterless systems have been demonstrated over hundreds of kilometers.

Fiber optic communication has been essentially used in point-to-point long distance transmission such as telephone networks. As the technology improves, it also becomes attractive to use optical fiber transmission for networking applications. For example, optical communication has been used for local area networks such as fiber distributed data interface (FDDI). At the turn of the century, literally thousands of intersatellite links – radio-frequency (RF) and optical - are expected to be in operation in commercial multi-satellite constellations providing mobile communications, video conferencing and multimedia services. Optical technology offers too many advantages in terms of mass, power, system flexibility and cost, to leave the field entirely to RF. In the near future, we will see optical fibers wired near or even in our homes. These are called fiber-in-the-loop and fiber-to-the-home.

REFERENCES

[1] Mamedov F.S., *Telecommunication*, Lecture notes, Near East University Press, Nicosia, 2000.

[2] Josephc. Palais, *Fiber Optic Communication Design Hand Book*, Prentice Hall, 1994.

[3] John M. Senior, *Optical Fiber Communication Principles and Practice*, 2nd ed., Prentice-Hall Inc., 1992.

[4] Wim van Etten, Jan van der Plaats, *Fundamentals of Optical Fiber Communication*, Prentice-Hall Inc., 1991.

[5] Stewart D. Personick, *Fiber Optics*, John Wiley & sons Inc., 1998.