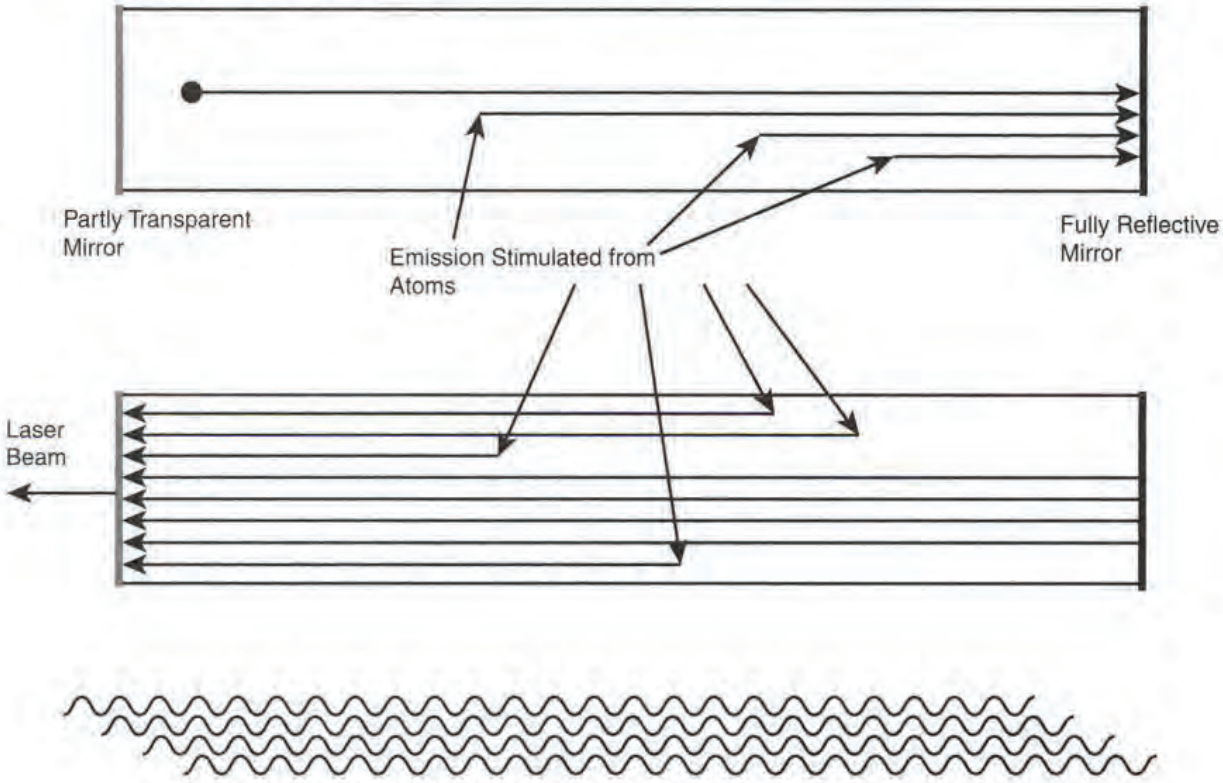
A laser is a light oscillator, which generates its own signal.

The word *laser* was coined as an acronym for *light amplification by the stimulated emission of radiation*, but that phrase glosses over the critical distinction between amplification and *oscillation*. A laser is a *light oscillator*, not a light amplifier.

An amplifier boosts the strength of an external signal, but doesn't generate a signal on its own. An oscillator generates a signal internally at a wavelength or frequency determined by its structure. The spark that starts laser oscillation is a spontaneously emitted photon, which stimulates emission of another photon, starting a cascade of other photons. A pair of mirrors on opposite ends of the device keeps the oscillation going. Light bounces back and forth between the mirrors, as shown in Figure 9.8, stimulating the emission of more photons on each pass. The pair of mirrors form a *resonant cavity*. One mirror reflects all light that strikes it, but the other mirror transmits a fraction of the light, which becomes the laser beam.

Reflection of light back and forth between the mirrors makes it pass multiple times through more excited laser material, amplifying the light more than is possible on a single



An initial spontaneous emission (black dot) stimulates emission of more photons. Mirrors at the ends of the laser cavity reflect light back and forth, building up stimulated emission. A fraction of the light leaks through a partly transparent mirror to form the laser beam.

Stimulated emission generates coherent light, with all the waves lined up in phase.

FIGURE 9.8

Laser emission from a resonant cavity.

pass. The mirrors select light that bounces back and forth in a line between them, so only light aimed in that direction is amplified, which concentrates the light emission into a narrow beam, as shown in Figure 9.8.

Each laser material has its own characteristic gain, which varies with the wavelength and conditions in the laser medium. Stimulated emission produces the strongest amplification at wavelengths where the gain is strongest, and laser oscillation further narrows the range of wavelengths. The laser structure determines how much the spectral width is narrowed.

Optical Amplifiers in Fiber Optics

Optical amplifiers boost the strength of weak signals.

Optical amplifiers deserve special attention here because they are very important in fiber-optic communications. By amplifying a weak optical signal, an optical amplifier increases the distance the signal can be transmitted. An optical amplifier is essentially a laser without mirrors at the ends. The light makes a single pass, instead of bouncing back and forth between mirrors, and is amplified by the gain within the amplifier, as shown in Figure 9.7. Two types of optical amplifiers are particularly important in fiber optics.

The *semiconductor optical amplifier* is a diode laser with its ends coated or integrated with semiconductor waveguides so they don't reflect light. (Semiconductors have a high refractive index, so they reflect some of the light trying to leave the crystal.) You'll learn more about this amplifier in Chapter 12.

The *erbium-doped fiber amplifier* is an optical fiber with erbium added to its core, as described in Chapter 7. It can amplify weak light signals that pass through it under the proper conditions. You'll learn more about its operation in Chapter 12.

We'll turn now to lasers because this chapter is about signal sources. We focus first on semiconductor lasers, starting with the simplest common type, because they are the usual type used in fiber optics. Later in this chapter, you'll learn about fiber lasers.

Simple Semiconductor Lasers

Semiconductor lasers resemble LEDs in important ways.

Like LEDs, semiconductor lasers are two-terminal devices called *diodes* in which holes flow through a *p region* and electrons flow through an *n region* to cause recombination in a junction layer separating the regions. Generally, diode lasers and LEDs use the same materials. The key differences are in their manner of operation and in the internal structures that control their operation.

LEDs produce spontaneous emission from electrons that release their surplus energy as they fall from the conduction band into the valence band. Diode lasers produce stimulated emission, which involves extracting light energy from the recombining electrons before they can spontaneously emit light. This extraction process requires concentrating the excitation energy to produce the population inversion required for laser action in the junction layer. This in turn requires a laser resonator, higher drive currents than those used in LEDs, and confinement of both the excitation and the generated light. We'll start with the simplest type of laser, called a *Fabry-Perot laser,* and describe each of these factors, then move on to other types of lasers.

Fabry-Perot Laser Cavities

Fabry-Perot diode lasers emit from the edge of the chip.

As you learned earlier, laser action occurs when light bounces back and forth within a resonant cavity, stimulating emission from excited atoms. The simplest type of resonant cavity is a pair of parallel mirrors, as shown in Figure 9.8. This is called a *Fabry-Perot cavity* after the two men who first used this arrangement to observe the interference of light waves.

In a gas laser, the two mirrors are on opposite ends of a tube containing the laser gas. In a semiconductor laser, the two mirrors are opposite edges of the semiconductor chip. One edge has a coating that reflects most of the light back into the semiconductor. (Typically a small amount of the light is transmitted so laser power can be monitored.) The opposite edge transmits more light, which emerges as the laser beam. (Semiconductors have high refractive indexes, which means they naturally reflect much of the light back into the solid and so may not require special reflective coatings.)

Some semiconductor lasers are called *edge emitters* because the light emerges from the edge of the chip, not from the surface. As with LEDs, the light is generated in the junction layer. Recall that edge-emitting LEDs are used in communications, as shown in Figure 9.5.

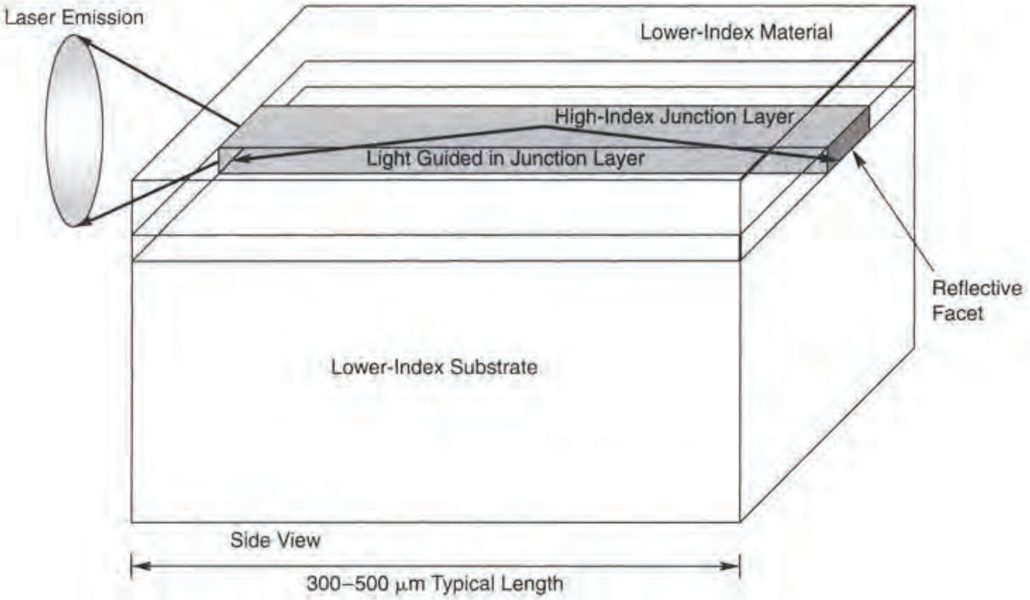
Stripe-Geometry Lasers

A stripe-geometry laser confines light in the junction layer.

A stripe-geometry laser, shown in Figure 9.9, confines light both vertically and horizontally within the junction layer. The junction layer itself is thin, typically a fraction of one micrometer.

The vertical confinement is created by making the *p-n* junction—called the *active layer* of the laser—of a semiconductor compound that has a refractive index slightly higher than

FIGURE 9.9
A double-heterojunction stripe-geometry laser



that of the *p* and *n* layers above and below it. This is done by changing its composition slightly, such as by adding a small amount of aluminum to gallium arsenide, which creates a boundary with different refractive index called a *heterojunction* between the two layers. The refractive-index difference between the junction layer and the layers above and below it creates a waveguide effect, which confines light in the junction plane, just as the lower-index cladding confines light in the core of an optical fiber. This layered structure is called a *double heterojunction* or *double heterostructure* and was a crucial step in the development of semiconductor devices, for which Herbert Kroemer and Zhores Alferov shared the 2000 Nobel Prize in Physics.

A double heterojunction confines light in the junction plane.

A stripe-geometry laser also confines laser action within a narrow stripe, typically only a few micrometers wide, in the junction layer. In telecommunications lasers, this area is usually a narrow high-index stripe in the junction layer so the difference in refractive index guides light in the horizontal plane just as the heterojunction guides light vertically. This index-guiding limits the laser to oscillating in a single mode and matches the core size of single-mode fibers.

Another approach is to limit current flow to a narrow stripe in the junction layer by depositing insulating layers that block current flow in other regions. This limits the population inversion to the narrow stripe where current flows, thus confining laser gain to the stripe. Lasers that only use *gain-guiding* do not confine light as well as index-guided lasers, but they are adequate for some purposes, and the two types of guiding can be combined in the same device.

There are many variations in the internal design of stripe-geometry diode lasers. Several layers of various compositions may be used to control the flow of current and light. In general, the more tightly the layers confine light, the more efficiently they produce a population inversion in the junction layer, and the more efficient the laser. Quantum well structures fabricated in the junction layer constrain where electrons and holes can recombine, improving light confinement and enhancing the performance of diode lasers.

We won't go into depth on these designs because this book focuses on fiber optics. Instead, we will concentrate on differences in design that affect the function and performance of diode lasers.

Laser and LED Performance

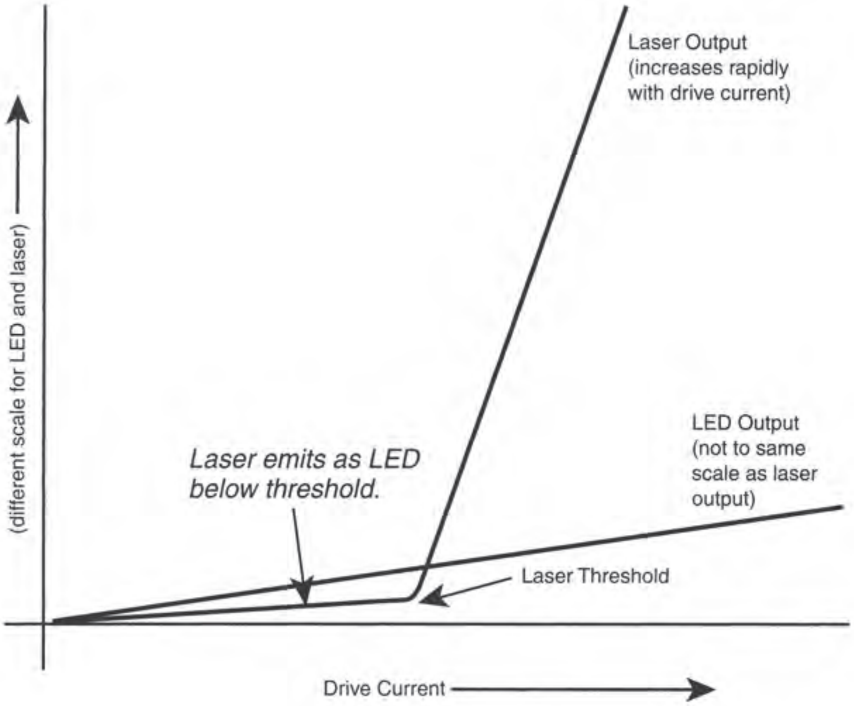
A first step in understanding diode laser performance is to compare the operation of lasers and LEDs. Like an LED, a laser requires a drive voltage greater than the bandgap voltage in order to generate light. Both diode lasers and LEDs must be forward-biased, with positive bias applied to the *p*-type material and negative bias to the *n* material.

Laser operation begins above a threshold current.

A profound difference between lasers and LEDs is their behavior as the drive current increases from zero. At low currents, both devices generate some light by spontaneous emission from recombining carriers, although lasers in general are inefficient. However, once the drive current exceeds a threshold value, the laser begins to generate stimulated emission, which increases much faster with drive current, as shown in Figure 9.10. Above this *threshold current,* a diode laser generates light much more efficiently than an LED. No such threshold exists for an LED.

The threshold is the point where the optical gain in the laser cavity exceeds the loss. As drive current increases, more carriers recombine, and are available for stimulated emission.

FIGURE 9.10
LED and laser power/current curves.



This increases the gain within the laser resonator. Below threshold, the gain that light makes in a round trip of the laser cavity is lower than the losses it suffers from absorption and light escaping through the end mirrors. At threshold, the gain exceeds the loss, and above threshold stimulated emission increases very rapidly with drive current, as shown in Figure 9.10. Although the curve looks steep, the increment in output power does not exceed the increment in input power. Laser efficiency can be measured in two ways, either overall efficiency comparing the output power to input power, or the *slope efficiency,* which measures the extra power generated per increment in drive current. Diode lasers are much more efficient than LEDs, with slope efficiencies that can reach tens of percent.

Diode lasers are much more efficient than LEDs.

LEDs don't have resonant cavities, because their light-emitting surfaces are made to suppress reflection, so they don't produce stimulated emission or have a threshold. That means their output increases steadily as drive current increases from zero, but the rate of increase is much less.

The threshold current is an important figure of merit for diode lasers. Below the threshold, most of the input energy must be dissipated as heat; above the threshold, much of the input energy emerges as light. In general, the lower the threshold, the better the laser's efficiency and performance. Reducing the laser threshold also tends to increase laser lifetime because it reduces the heat dissipation and operating temperature.

Another important difference between lasers and LEDs is that the laser emits a much narrower range of wavelengths. Figure 9.10 shows that spontaneous emission from an LED varies across a range of wavelengths. Stimulated emission varies in the same way, but the amplification process builds up the difference because photons at the peak wavelength are more likely to stimulate emission than those away from the peak. You'll learn more about how this works when we describe laser wavelengths later in this chapter.

Vertical Cavity Diode Lasers

VCSELs emit from their surface, perpendicular to the junction layer.

The edge-emitting laser is a tried-and-true design, but its emission from the edge of the chip can be a significant practical disadvantage. Many edge-emitting lasers can be fabricated at one time on a single wafer, but the entire wafer must be cleaved before each individual laser can be mounted and tested. Lasers that emit from the surface of the chip can be tested while still on the wafer, making them easier to package and produce, and thus lower in cost. This has led to the development of the *vertical-cavity surface-emitting laser* or VCSEL.

The resonant cavity in a VCSEL is perpendicular to the junction layer and vertical in the wafer, so the laser output emerges from the surface, as shown in Figure 9.11. This is a device rather different from an edge-emitting laser. Amplification in a VCSEL occurs only in the thin junction layer, so the light can be amplified by only a small amount in each pass through the laser cavity. In contrast, light in an edge-emitting laser passes through the entire length of the junction layer (a few hundred micrometers) on each pass, so it is amplified much more. To compensate for the lower gain within the VCSEL cavity, the mirrors must reflect more light than those used in edge-emitting lasers.

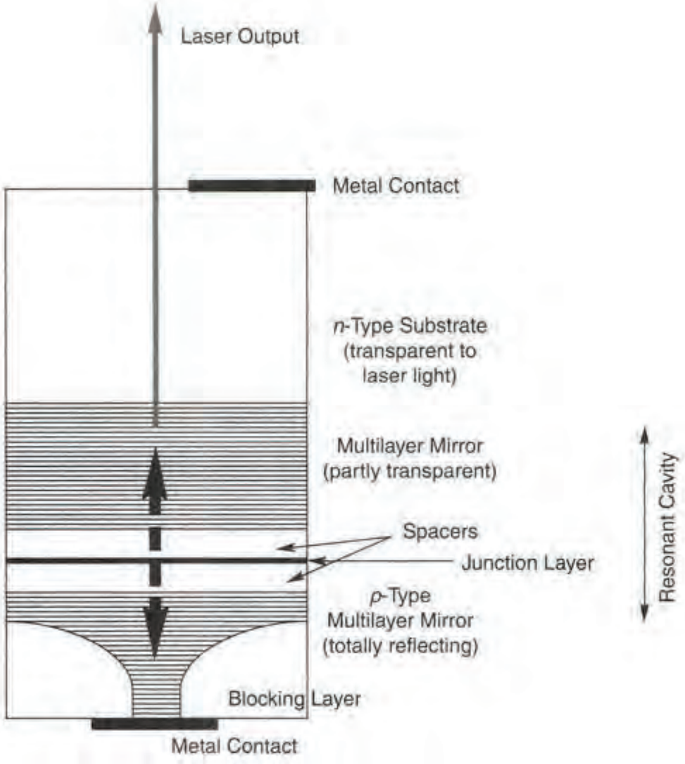


FIGURE 9.11
A vertical-cavity surface-emitting laser.

VCSEL mirrors normally are made with standard semiconductor processing techniques. They are composed of a series of layers with alternating compositions, so they selectively reflect a narrow range of wavelengths. Which wavelengths the mirrors reflect depends on the thicknesses and refractive indexes of the layers. (You may remember the concept from the discussion of fiber Bragg gratings in Chapter 7. It's a basic principle of optics that was developed decades ago.)

A VCSEL emits from a round spot typically 5 to 30 μm across on the surface of the wafer. This spot is larger than the core of a single-mode fiber, but smaller than the core of a multimode fiber. It is also larger than the output spot of an edge-emitting laser, so diffractive effects do not make a VCSEL beam spread out as rapidly. The beam is also circular, unlike the oval beam from an edge emitter.

In addition to being easy to manufacture and package, VCSELs have low threshold currents and are quite efficient in converting input electrical power into light. This means they consume less power and dissipate less heat than edge-emitters. They also have a longer lifetime and can be directly modulated at data rates well above 1 Gbit/s.

VCSELs can be made in two-dimensional arrays. They are easy to test and mount.

Unlike other diode lasers, VCSELs can be made in two-dimensional arrays covering the surface of a wafer, and these individual lasers can be modulated separately. Such arrays are attractive for optical switching and signal processing, to produce beams for transmission through optical fibers or free space.

Currently, VCSELs are the favorite fiber-optic lasers for wavelengths of 750 to 1000 nm. Longer wavelengths have proved more difficult because materials that emit light at those wavelengths don't work well for the multilayer mirrors used in VCSELs.

Laser Reliability

LEDs are more reliable than edge-emitting lasers.

Early GaAs lasers were unreliable, but great improvements have been made. Nonetheless, LEDs are more reliable than edge-emitting diode lasers because the lasers have higher current densities and optical power outputs. Threshold current can be an index of laser reliability; the lower the threshold, the longer-lived the laser. VCSELs, which have very low thresholds, are the most reliable lasers.

Operating temperature is a major factor in laser reliability; increasing temperature shortens lifetimes, and elevated temperatures are used in accelerated-aging tests. Threshold currents increase with operating temperature, increasing the waste heat generated within the laser, which further increases temperature and degrades efficiency. Gallium arsenide is more vulnerable than InGaAsP to this problem, which can lead to thermal runaway. To control heat buildup and efficiency decreases, many laser transmitters are built with active temperature stabilization, such as thermoelectric coolers. Most lasers are packaged with heat sinks, even if active cooling is not required.

Laser output declines with age.

Output power of diode lasers tends to decline slowly with age. To compensate for this decline, the transmitter can be designed to slowly increase drive current so the output power remains constant. A laser operated in this way is said to fail when it no longer delivers the required output power.

Diode lasers are particularly vulnerable to damage from electrostatic discharges. Careful handling and proper packaging can overcome this problem, but you should be aware of its potential and always ground yourself when handling lasers.

Laser Wavelength

The output wavelength of diode lasers is central to their use in fiber-optic systems. Both the peak wavelength emitted and the range of wavelengths are important for system performance. The composition of the semiconductor in the junction layer determines the wavelengths where a laser (or LED) can emit light. The device structure determines which wavelengths in that range the laser can emit. We'll start by looking at the materials, then turn to the structures used to produce particular effects.

Semiconductor Laser Materials

Earlier in this chapter, you learned that the LEDs and diode lasers used in fiber-optic systems are made of *III-V semiconductor* compounds. The laser structure consists of a series of layers on a substrate wafer. The composition of the substrate is chosen to be compatible with the composition of the other layers, which in turn are chosen to work with the composition of the junction layer. In practice, substrates are made of two-element compounds, gallium arsenide or indium phosphide, which are easier to produce in bulk than the three- or four-element compounds used for the active layers.

Diode lasers are made on GaAs or InP substrates.

The principal compounds used are:

- $Ga_{(1-x)}Al_xAs$ on GaAs for 780 to 850 nm
- $In_{(1-x)}Ga_xAs$ on GaAs for 980 nm
- $In_{(1-x)}Ga_xAs_{(1-y)}P_y$ on InP for 1100 to 1700 nm

The subscripts indicate the relative fractions of each element. Indium, gallium, and aluminum are all Group III elements and can be interchanged with each other. Arsenic and phosphorus are Group V elements and can be interchanged with each other. Compounds containing three elements are called *ternary*, and those containing four elements are called *quaternary*. As you will learn later, 980- and 1480-nm lasers are used to pump optical amplifiers, while others are used as signal sources.

Emission wavelength depends on the composition of the junction layer.

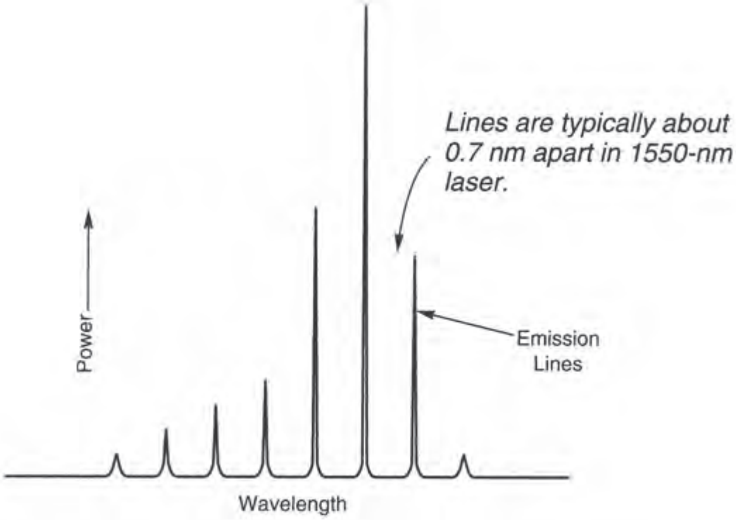
The composition of the active layer determines the peak gain wavelength. For example, the gain of InGaAsP peaks at 1310 nm for a composition of $In_{0.73}Ga_{0.27}As_{0.58}P_{0.42}$. As you learned earlier, the process of stimulated emission amplifies the peak wavelength more than other wavelengths, narrowing the range of output wavelengths from the broad spectrum seen in an LED to the narrower range of a diode laser, shown in Figure 9.1. Other layers may have slightly different composition, and the whole structure is deposited on a substrate of either GaAs or InP, which are much easier to make in the large volumes needed for substrates.

Laser Spectral Range

The spectral range of diode lasers depends on their structure as well as their composition. The simple edge-emitting Fabry-Perot laser described earlier, with one mirror at each end, has a bandwidth of 1 to 3 nm concentrated on multiple narrow lines, as shown in Figure 9.12. These multiple lines arise from the nature of the resonant cavity.

Fabry-Perot lasers have spectral widths of 1 to 3 nm.

FIGURE 9.12
Wavelengths in multiple longitudinal modes.



For light to resonate within the laser cavity, the round-trip distance between the mirrors must equal an integral number of wavelengths. Thus, a laser cavity with length L in a laser material with refractive index n can resonate at wavelengths λ defined by

$$2nL = N\lambda$$

where N is an integer and where the laser material has large enough gain. Each wavelength spike in Figure 9.12 corresponds to a different value of N. The spikes span the range of wavelengths where the gain is highest.

Each spike in Figure 9.12 is a separate *longitudinal mode* of the laser, which means a resonance along the length of the laser cavity. (The modes across the width of a laser or an optical fiber are *transverse modes,* defined by the width of the laser or fiber; narrow-stripe diode lasers operate in a single transverse mode.) Each of these longitudinal modes has much narrower spectral width than the entire envelope of modes emitted by the laser. The spacing between longitudinal modes depends on the cavity length and wavelength. The longer the cavity length (measured in wavelengths), the closer the modes are spaced. Edge-emitting Fabry-Perot diode lasers have short cavities, only about 500 μm long, and their modes are about 0.6 nm apart at 1300 nm or about 0.7 nm apart at 1550 nm.

Minor fluctuations during operation can make edge-emitting Fabry-Perot lasers "hop" between modes, shifting the emission wavelength suddenly. The emission peak in Figure 9.12 moves from one longitudinal mode to another. This and other problems typically limit Fabry-Perot edge-emitting lasers to transmission rates less than 1 Gbit/s and to coarse versions of wavelength-division multiplexing with widely separated optical channels.

The same principles apply to VCSELs, but their cavities are much shorter than those of edge-emitters, so their longitudinal modes are tens of nanometers apart instead of a fraction of a nanometer. This means that in practice VCSELs emit a single longitudinal mode and can transmit at much higher speeds than edge-emitting Fabry-Perot lasers.

Single-Frequency Lasers

For high performance, low dispersion, and closer spacing of optical channels, laser emission must be limited to a single longitudinal mode or, equivalently, to a single frequency. This has led to development of more elaborate laser resonators. Figure 9.13 shows three leading approaches.

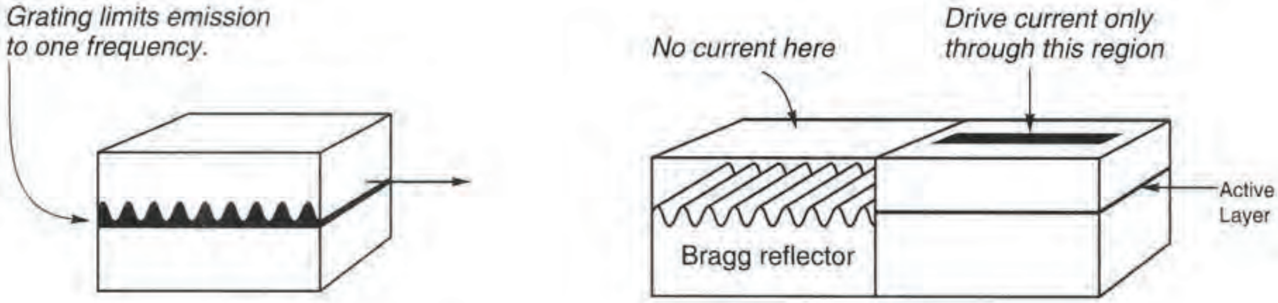
The *distributed-feedback (DFB) laser,* in Figure 9.13(a), has a series of corrugated ridges on the semiconductor substrate, which scatter light back into the active layer. This provides feedback like the cavity mirrors on a Fabry-Perot laser, although the details of the physics are different. The *distributed Bragg reflection (DBR) laser* shown in Figure 9.13(b) works in much the same way, but the grating is etched in a region outside the zone that is pumped by electric current. In both cases, the grating ridges are spaced evenly so they scatter only a narrow range of wavelengths back into the active layer of the laser. The active layer of the laser amplifies only this selected range of wavelengths, producing very narrow spectral bandwidths at a nominal "single frequency." The wavelength depends on the line spacing in the grating and the refractive index of the semiconductor. Recall that Bragg reflection is the same effect that selects the wavelengths reflected by a fiber Bragg grating.

●

Single-frequency lasers are needed for high-speed transmission.

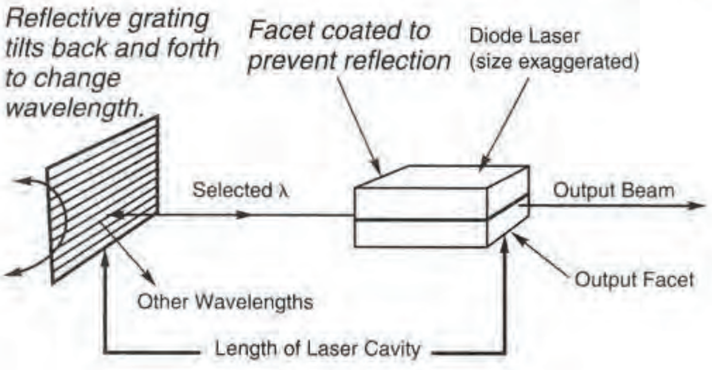
●

Distributed-feedback and distributed Bragg reflection lasers emit only a single frequency.



a. Distributed-Feedback Laser.

b. Distributed Bragg Reflection Laser.



c. External Cavity Tunable Laser.

FIGURE 9.13
Three single-frequency lasers.

A different way to stabilize laser wavelength is by placing an edge-emitting semiconductor laser within an external cavity, which selects the emission wavelength. This requires coating one or both facets to suppress reflection back into the semiconductor, and adding one or two external mirrors to extend the resonator cavity beyond the laser chip. A wavelength-selective element also is added to the laser cavity. In the simple design of Figure 9.13(c), the tuning element is a diffraction grating, which serves as one external mirror, reflecting light at an angle that depends on its wavelength. (You can get the same effect by inserting a prism or some other wavelength-selective component into the laser cavity, but diffraction gratings are easier to use.) The laser chip emits a range of wavelengths, but when they strike the grating, most wavelengths are reflected at angles that take them away from the laser chip. Only a very narrow range of wavelengths are at the right angle to be reflected back into the laser chip for further amplification. This limits output to a single frequency.

Distributed-feedback and distributed Bragg reflection lasers are the types most often used to generate a single, fixed wavelength with very narrow spectral width. However, there is growing interest in lasers with output that can be tuned to emit at precise wavelengths.

Tunable Lasers

Tunable lasers can simplify logistics.

An external cavity laser is a good starting point for discussing *tunable lasers,* which can be changed in wavelength. You have already seen how a diffraction grating can reflect a single wavelength back into the laser chip for amplification, building up emission at a single wavelength. Turning the grating changes what wavelengths are reflected back to the laser chip, which changes the laser's output wavelength.

As you will learn later, wavelength tunability is an attractive property for lasers used in WDM systems and measurement instruments. Standard lasers emit only a fixed wavelength, so a system with 80 different wavelengths requires 80 different models of laser. Moreover, the service department needs spares for every one of those 80 different laser models. If a telephone company wants to install 80-channel systems, its maintenance department would need to stock every site with spares for each of the 80 wavelengths. The logistics could become a nightmare.

Tunability also can enhance the flexibility of optical networking. For example, it sometimes may be necessary to move an optical channel from one wavelength to another, because the same wavelength isn't available along its entire route. Using a tunable laser to generate the new signal would allow the wavelength to be changed without switching lasers. The laser might be tunable continuously across the spectrum, but system design would be easier if lasers were preset to emit precisely at standard wavelengths. Users would then select an optical channel, just as viewers select a channel on a modern television set, without having to adjust the laser to match the desired frequency. The technology is still young, but several approaches have been developed.

Changing the length of a VCSEL cavity can tune its wavelength.

We saw earlier that the resonant wavelength depends on the length of the laser cavity. Changing the cavity length has little effect on edge-emitting lasers, because their longitudinal modes are less than a nanometer apart. However, it can tune the wavelength significantly because a VCSEL cavity can be made very short, micrometers long rather than hundreds of micrometers long for edge-emitters. With a short enough cavity, only a single

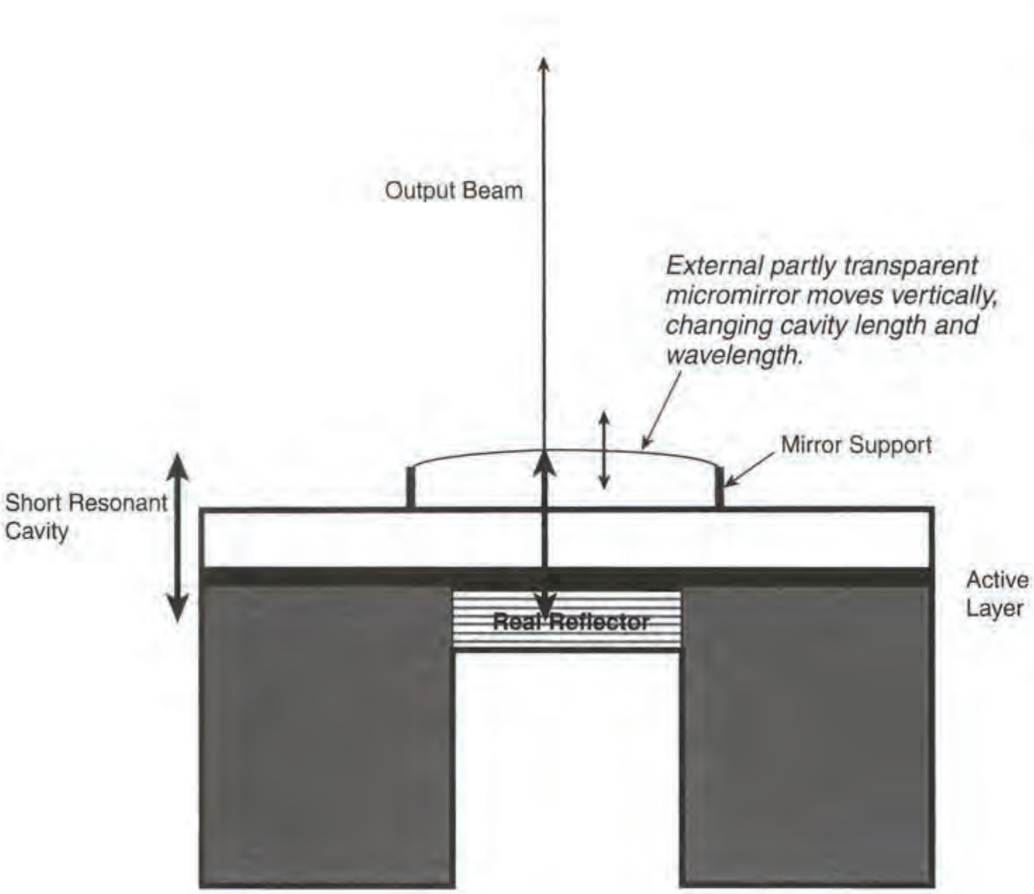
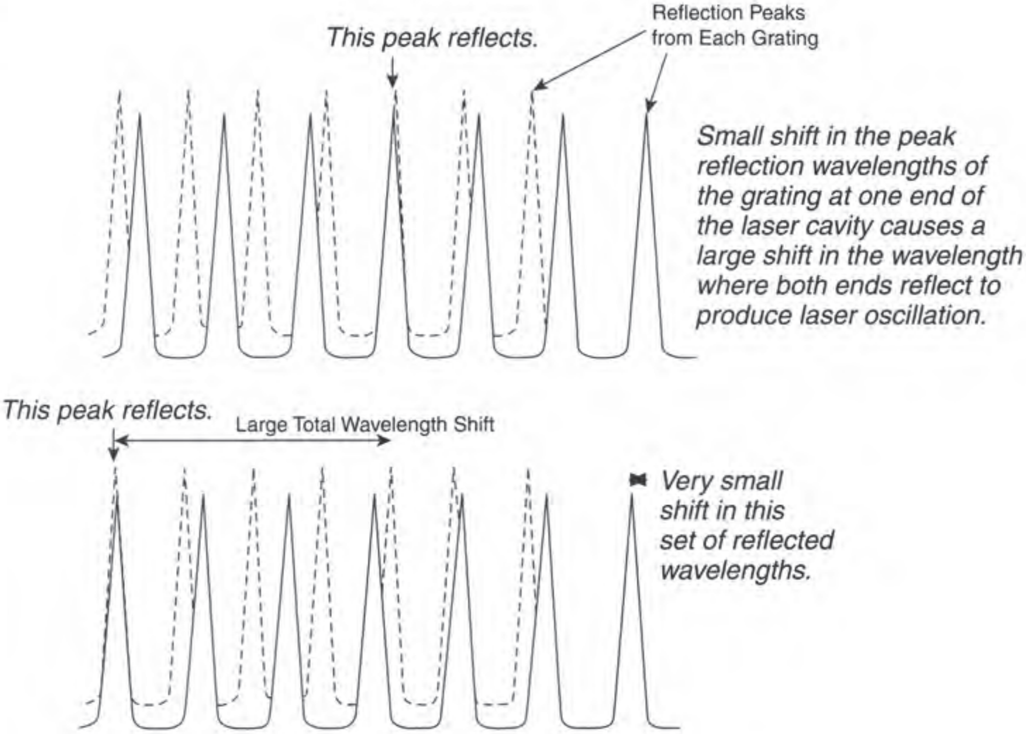


FIGURE 9.14
Tunable VCSEL relies on a moving external micromirror to change cavity length and thus wavelength.

longitudinal mode falls within the laser's gain band. This makes possible the sort of tunable VCSEL shown in Figure 9.14. A thin, partly transparent mirror is held above the VCSEL by a movable micro-electromechanical system (MEMS) device. Vertical motion of the MEMS mirror changes the resonant wavelength in the VCSEL cavity, tuning the wavelength by over 30 nm in the laboratory.

Distributed-feedback and distributed Bragg reflection lasers can be tuned in other ways. As you saw earlier, the wavelengths selected depend both on the grating period and on the refractive index of the material. Changing the temperature of the material can change both, by causing thermal expansion (or contraction) of the laser material as well as by directly affecting refractive index. Passing a current through a material also affects the refractive index. Generally these changes are relatively small, and allow turning over only several nanometers.

Tuning ranges can be extended to tens of nanometers by using more elaborate distributed Bragg reflectors. One example is the *sampled-grating distributed Bragg reflector (SG-DBR)*, which contains regions with different grating spaces that reflect a comb-like series of regularly spaced wavelengths. To make a tunable laser, slightly different separate sampled-grating reflectors are fabricated on each end of the active region of the laser. The laser can oscillate only at a wavelength reflected by the gratings on both ends, which can be tuned independently by changing current level or temperature. These changes shift the reflection peaks of the individual grating only slightly, but this small shift causes a much larger shift



Small shift in the peak reflection wavelengths of the grating at one end of the laser cavity causes a large shift in the wavelength where both ends reflect to produce laser oscillation.

Very small shift in this set of reflected wavelengths.

FIGURE 9.15
Tuning of sampled-grating distributed Bragg reflector laser.

in the wavelength at which both gratings reflect to allow laser oscillation, as shown in Figure 9.15. This is sometimes called a vernier effect, because a vernier scale works in the same way to amplify the size of a small change. A related approach is the *grating-assisted coupler and sampled reflector* (GCSR) laser, which combines a sampled-grating reflector with other elements to tune the output wavelength.

Another approach to tuning is selecting one laser stripe from an array of several on a single semiconductor substrate. For example, if each of 12 stripes had a tuning range of 3 nm, and their center wavelengths were 3 nm apart, they could combine to cover a 36-nm range. The device could tune across one laser's 3-nm range, then switch to the next laser and tune over its range.

Other tunable lasers also are in development. They face a number of practical challenges. Tunable lasers must be locked to the right wavelengths so they don't drift during operation. They also must be affordable, reliable, and compatible with standard telecommunications equipment. The spread of tunable lasers was stalled by the telecommunications downturn.

Modulation and Wavelength Chirp

External modulation avoids laser wavelength chirp, which causes chromatic dispersion.

Direct modulation via changing the drive current is the simplest and cheapest way to modulate the output of a diode laser. Unfortunately, it has the undesirable side effect of shifting the laser's wavelength during the emitted pulse. The electron density in the semiconductor changes as the current changes, and the semiconductor's refractive index varies with the electron density. This means that modulating the current effectively changes the optical path length in the semiconductor, which equals the refractive index *n* times the physical distance

THINGS TO THINK ABOUT

Lasers and the Bubble

The telecommunications bubble promoted the development of new types of lasers with very narrow spectral width and tunable output wavelength. System manufacturers wanted narrow-line lasers so they could pack as many wavelengths into as many channel slots as possible, which would enable a single optical fiber to carry huge volumes of telecommunications traffic. They also wanted the tunable lasers so they could change the wavelength, which would enable them to switch signals in their networks and to reduce their inventory of different-wavelength lasers. This led to the development of new types of lasers that could offer the desired performance, although at a high cost.

After the bubble collapsed, system operators discovered that they really didn't need all that bandwidth, or all those wavelengths going through a single fiber. Instead, they wanted lower-cost systems that could match their more modest transmission requirements. The bubble made new technology available, but some of it is still sitting on the shelf, waiting for demand to increase.

---

through the semiconductor L. From the earlier equation for the resonant wavelength in an optical cavity, you can see that this means wavelength λ changes by an amount $\Delta\lambda$:

$$\Delta\lambda = \frac{2(\Delta n \times L)}{N}$$

where Δn is the change in refractive index and N is an integer, the number of wavelengths needed to make a round trip in the cavity. Although the change, called *chirp,* is small, it occurs during every laser pulse, so every pulse contains a broader range of wavelengths than it otherwise would include. The resulting dispersion can impair long-distance transmission at speeds above about 1 Gbit/s.

The cure for chirp is to drive the laser with a steady current, then modulate the steady beam externally. This is done by applying the signal to a modulator, so the fraction of the laser beam it transmits varies in proportion to the signal. You'll learn more about modulators in Chapter 16; for now all you need to know is that they modulate beam intensity but do not affect its wavelength. External modulators also are very fast, working at speeds to 40 Gbit/s. They generally are used in long-distance systems transmitting at 2.5 Gbit/s or higher.

Driving the laser source with a steady current also improves its inherent wavelength stability, because any modulation induces fluctuations in laser properties.

As mentioned earlier, wavelength is temperature-sensitive, so stabilizing the laser's operating temperature controls variations in its wavelength. Stabilization is required for high-speed transmitters.

Fiber Lasers

Fiber lasers are excited by light from pump lasers, not by electric current.

Fiber lasers and amplifiers rely on the same laser principle as the semiconductor lasers we have covered so far. However, they use a different type of laser medium: a glass or crystalline material that is transparent but doesn't conduct electricity. Instead of getting their energy directly from an electric current passing through a semiconductor junction, fiber lasers are excited by light from an external source, usually a semiconductor laser that is called the *pump laser.* In this section you'll learn about fiber lasers; fiber amplifiers, which lack mirrors and a resonant cavity, are covered in Chapter 12.

A semiconductor laser is not considered a solid-state laser.

Fiber lasers are part of a larger group called *solid-state lasers,* in which light from an external source excites atoms in a glass or crystal rod to produce stimulated emission. The first laser was a solid-state laser, in which a photographic flashlamp excited chromium atoms in a ruby rod that had its ends coated with silver. From an optical standpoint, a fiber is just a very thin rod. And from a laser standpoint, a semiconductor laser is *not* a solid-state laser.

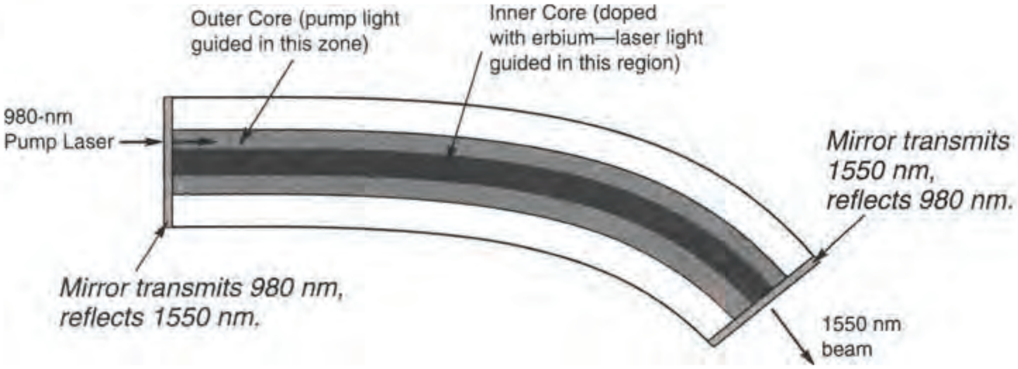
Solid-state lasers can be pumped by light from the sides or the ends. Fiber lasers often are excited by a pump laser on one end, as shown for an erbium-doped fiber laser in Figure 9.16. The light-emitting atoms (erbium in this case) are concentrated in an inner core, while the pump light is guided by an outer core, an arrangement described in Chapter 7. This concentrates the laser beam in the small inner core. In the design shown, one mirror transmits pump light and reflects the laser wavelength, while the other mirror transmits the laser beam and reflects the pump light. Generally one mirror reflects all light at the laser wavelength, while the output mirror transmits some of the light and reflects the rest to sustain laser oscillation. The thin fiber can be curved, and if it is long, it may be coiled.

Fiber lasers cannot be directly modulated.

Fiber lasers cannot be directly modulated like semiconductor lasers. They can be powered by a steady source to generate a continuous beam, which can be modulated externally with a signal. Alternatively, they can use optics to produce a series of pulses shorter than one picosecond (10^{-12} s) spaced at regular intervals; this process is called *modelocking.* Simple modelocked pulses don't carry information, but external modulation can add a signal by switching pulses on or off.

Like semiconductor lasers, fiber lasers can emit light at a range of wavelengths, depending on their composition. The specific wavelength emitted by a continuous laser is determined by the cavity optics: The output can be at a fixed wavelength, or tuned across a range of wavelengths. Modelocked pulses inherently contain a much wider range of wavelengths than a continuous beam, so they can be used as broad-spectrum light sources.

FIGURE 9.16
Erbium-doped fiber laser pumped at 980 nm emits 1550 nm.



Types of Fiber Lasers

Many types of fiber lasers have been developed, but only a few have found significant applications. The most important fiber lasers are:

- *Erbium-doped fiber lasers,* which typically emit at 1530 to 1620 nm and can be pulsed or continuous. Like erbium-doped fiber amplifiers, they are excited by light at wavelengths of 980 or 1480 nm. Their main applications are in communications, measurement instruments, and research.
- *Ytterbium-doped fiber lasers,* which operate at 1030 to 1120 nm and can generate powers above one kilowatt at their peak wavelength (near 1070 nm). They are being developed for industrial and military applications.
- *Thulium-doped fiber lasers,* emitting at 1750 to 2200 nm for research use.

Other Solid-State Laser Sources

Other solid-state lasers also can be signal sources in fiber-optic systems. As in fiber lasers, light from an external source excites atoms in a solid material between a pair of mirrors. The laser material is a glass or crystal host doped with atoms that emits light by stimulated emission. However, the laser material is usually shaped as a rod, is much shorter than a typical fiber laser, and lacks an internal light-guiding structure.

Diode-pumped neodymium lasers can generate more than a watt near 1300 nm.

The most important of these are crystalline neodymium lasers, in which the rare earth neodymium is doped into crystals called YAG (for yttrium aluminum garnet) or YLF (for yttrium lithium fluoride). These lasers can be excited by GaAs diode lasers emitting near 800 nm. The primary neodymium output line is near 1060 nm, but there are secondary lines at 1313 and 1321 nm in YLF and 1319 nm in YAG. Those fall right in the 1300-nm fiber window, but are not available from fiber lasers.

Like erbium-fiber lasers, solid-state neodymium lasers cannot be modulated directly; they require external modulators. Their big attraction is their ability to generate high power—more than a watt near 1300 nm. That's more than you want to send signals through a single length of fiber, but it can be split among many fibers to carry the same signals to many terminals, for network communications or cable television signal distribution.

What Have You Learned?

1. Wavelength, spectral width, output power, and modulation speed are key considerations in fiber-optic light sources.
2. Modulating the light source by changing the drive current is called direct modulation; it works for diode lasers and LEDs. External modulation uses a separate external device to modulate a steady beam from a light source.
3. LEDs and semiconductor lasers are both semiconductor diodes that emit light when current in the diode causes electrons and holes to recombine at the

junction between *p*- and *n*-type material. The electrons drop from the conduction band into the valance band.

4. The output wavelength of an LED or diode laser depends on the composition of its junction layer. GaAlAs junctions, fabricated on GaAs substrates, emit at 780 to 850 nm. InGaAsP junctions, made on InP substrates, emit at 1200 to 1700 nm.
5. LEDs produce spontaneous emission; lasers produce stimulated emission. The word *laser* comes from *light amplification by the stimulated emission of radiation.*
6. Red LEDs emitting at 650 nm are used with plastic fibers. GaAlAs LEDs emitting at 820 or 850 nm are used with short runs of glass fiber.
7. Laser light is produced by the amplification of stimulated emission as light is reflected back and forth between a pair of mirrors. A population inversion is needed to produce stimulated emission and laser action. A laser is a light oscillator and emits more power than an LED.
8. Stimulated emission amplifies light; gain measures the strength of the amplifier. Optical amplifiers boost the strength of weak signals, but lack mirrors and do not oscillate.
9. Fabry-Perot diode lasers emit from the edge of the junction layer and the side of the chip. They have spectral widths of 1 to 3 nm and emit multiple longitudinal modes.
10. A double-heterojunction laser sandwiches the junction layer between layers of lower refractive index to confine light vertically. A narrow stripe of high-index material in the junction plane confines light horizontally.
11. Laser operation begins when the drive current exceeds a threshold value. Laser power rises rapidly above threshold.
12. Vertical-cavity surface-emitting lasers (VCSELs) have mirror layers fabricated above and below the junction layer, so they emit their beam from their surface rather than their edge. Their output can be coupled easily into a multimode fiber.
13. Single-frequency lasers oscillate in a single longitudinal mode to generate the narrow-line output needed for high-speed transmission. The most common types are distributed-feedback (DFB), distributed Bragg reflection (DBR), and external-cavity lasers.
14. Laser wavelength can be tuned by adjusting the cavity mirrors or changing the refractive index in the cavity. Tunable lasers simplify logistics in WDM systems and optical networks and are used in measurement instruments.
15. Fibers that have cores doped with light-emitting elements are used in fiber lasers. Light from a laser or other external source excites the light-emitting elements, producing stimulated emission that oscillates between mirrors on the ends of the fiber. The most important fiber lasers in fiber optics are erbium-doped fiber lasers, which emit at 1530 to 1620 nm.

What's Next?

Now that I have described light sources, Chapter 10 will cover fiber-optic transmitters that use these light sources.

Further Reading

Govind P. Agrawal, *Semiconductor Lasers: Past, Present and Future* (AIP Press, 1995)

Jeff Hecht, *Understanding Lasers* (IEEE Press, 1994)

C. Breck Hitz, J. J. Ewing, and Jeff Hecht, *Introduction to Laser Technology,* 3rd ed. (IEEE Press, 2001)

Questions to Think About

1. Below threshold a diode laser emits some light by spontaneous emission. Why does it behave like an LED?
2. A diode laser has a threshold current of 12 milliamperes. That current passes through a stripe 5 μm wide and 500 μm long. What is the current density in amperes per square centimeter?
3. The diode laser in Problem 2 emits 5 mW of light. The junction layer is 0.5 μm thick. If the light is evenly distributed across the end of the active layer, what is the optical power density in W/cm^2?
4. InGaAsP has a refractive index of about 3.5. A Fabry-Perot laser emitting multiple longitudinal modes at a nominal wavelength of 1550 nm has a cavity 500 μm long. What is the wavelength difference between two adjacent longitudinal modes?
5. Suppose the laser in Problem 4 was a VCSEL with cavity length only 10 μm, and for the time being forget about the difficulty of making InGaAsP VCSELS. Calculate the separation between two modes using the same technique.

Chapter Quiz

1. Operating wavelengths of GaAlAs LEDs and lasers include
   a. 820 and 850 nm.
   b. 500 nm.
   c. 1300 nm.
   d. 1550 nm.
   e. none of the above

2. Light emission from an LED is modulated by

a. voltage applied across the diode.
b. current passing through the diode.
c. illumination of the diode.
d. all of the above

3. Which of the following statements about the difference between semiconductor lasers and LEDs are true?

a. Lasers emit higher power at the same drive current.
b. Lasers emit light only if drive current is above a threshold value.
c. Output from LEDs spreads out over a broader angle.
d. LEDs do not have reflective end facets.
e. All of the above

4. Laser light is produced by

a. stimulated emission.
b. spontaneous emission.
c. black magic.
d. electricity.

5. The spectral width of a Fabry-Perot semiconductor laser is about

a. 2 nm.
b. 30 nm.
c. 40 nm.
d. 850 nm.
e. 1300 nm.

6. A distributed-feedback laser is

a. a laser that emits multiple longitudinal modes from a narrow stripe.
b. a laser with a corrugated substrate that oscillates on a single longitudinal mode.
c. a laser made of two segments that are optically coupled but electrically separated.
d. a laser that requires liquid-nitrogen cooling to operate.

7. Which of the following is an important advantage of external modulation of lasers?

a. simpler operation
b. does not require electrical power
c. no extra devices needed
d. avoids wavelength chirp that could cause dispersion

8. What guides light in a narrow-stripe edge-emitting laser?

   a. reflective layers on the edges of the laser wafer

   b. The stripe has higher refractive index than surrounding material, so it functions as a waveguide.

   c. coatings applied above and below the junction

   d. light entering it from an external optical fiber

9. Which of the following is *not* true for VCSELs?

   a. VCSELs emit light from their surfaces.

   b. VCSEL beams are rounder than those from edge-emitting lasers.

   c. VCSELs can be made easily from GaAs or InGaAsP compounds.

   d. VCSELs have low-threshold currents.

   e. VCSELs have multilayer coatings as their resonator mirrors.

10. A Fabry-Perot diode laser operating at 1.3 μm has a cavity length of 500 μm and a refractive index of 3.2. How far apart are its longitudinal modes? (*Hint:* First estimate the number of waves that could fit into the cavity; then calculate the wavelengths of modes N and $N + 1$.)

   a. 0.013 nm

   b. 0.053 nm

   c. 0.53 nm

   d. 5.3 nm

   e. 0.13 μm

CHAPTER

10

Transmitters

About This Chapter

Optical transmitters convert an input electronic signal into the optical form sent through optical fibers. This task involves both electronics and optics, but this chapter, like the rest of the book, concentrates on the optics. It first introduces the basic operational concepts involved in transmitters, then covers how multiplexing generates signals and modulation converts them to optical form. Finally it looks inside the box to show the functional components of transmitters.

Transmitters contain the light sources described in Chapter 9 and encode signals for the receivers covered in Chapter 11. Chapter 12 covers signal amplifiers and regenerators.

Transmitter Terminology

Strictly speaking, a fiber-optic transmitter is a device that generates an optical signal from an electronic input. An optical transmitter always contains a light source. However, the terminology can get muddled, particularly once transmitters are packaged into commercial equipment, so let's go through a few definitions before getting started.

A *transmitter* generates optical signals; a *receiver* detects them and converts them back into electronic form for equipment at the other end of the system. A *link* is the combination of a transmitter, fiber-optic cable, and a receiver used to send a signal between points. A *data link* is specifically a digital link, usually for computer data transmission, but something merely called a "link" may be digital or analog.

A *system* can be pretty much anything you want it to be. In practice, it usually means the equipment needed to generate an optical signal, transmit it, and receive it. Often this includes the electronics that process input signals to convert them into the form required to drive the light source in the transmitter. Likewise, it often includes electronics, which process the received signals. *Networks* are systems that link many points and contain many transmitters and receivers as well as cables. Later chapters will teach you more

A transmitter generates optical signals from an electronic input. It contains a light source.

A transceiver includes both a transmitter and a receiver that serve one terminal.

about the many different kinds of systems and networks used for telecommunications. A *terminal* is a device attached to a system or network, such as a computer or a phone.

Two-way communications requires that each terminal be equipped with one transmitter and one receiver. It's common to package the transmitter and receiver together in a unit called a *transceiver*, which both transmits and receives signals. Transceivers are also used in many other communication systems. For example, your telephone is a transceiver because it transmits and receives speech; so is the modem that links your computer to the Internet.

If you're familiar with electronics, you'll recognize these terms because they come directly from electronic communications. A broadcast television station has a transmitter; your home television set is a receiver. The television transmitter puts a video signal into the right form for broadcasting; your television receiver converts that signal into a form you can watch. Remember that everything called a transmitter is not fiber optic.

Some of these definitions can be a bit hazy because they depend on packaging, and what the engineers decide to put in the boxes. Although the terms are widely accepted by engineers, sometimes marketing departments have their own ideas. Some are flat-out wrong, like short analog systems called "data links." Others are ambiguous, like "solutions," which can be anything from transmitters to systems.

Operational Considerations

A number of operational considerations shape the design and performance of a fiber-optic transmitter. These include the type of system, the type of modulation, the data rate or bandwidth, the number of optical channels, and optical power requirements.

The higher the transmission speed, the more complex the transmitter.

The higher the system performance, particularly measured as data rate, the more the transmitter has to do and, in general, the more complex it becomes. It's possible to directly modulate the output intensity of an LED by applying an electrical signal across the diode—if the signal is a simple one like analog speech or a slow stream of digital bits. The higher the performance, the more care is needed in modulation. A diode laser needs a bias current in addition to a modulation signal. A high-speed laser in a WDM system needs an external modulator and a thermoelectric cooler to stabilize its operating wavelength. Very high speed circuits also require special electronic circuits able to process the high frequencies involved.

Transmitter and System Packaging

All transmitters include a light source and optical and electronic connections.

Most fiber-optic transmitters are sold as small and fairly simple packages that contain only the essentials for converting electronic signals to optical form. Even the simplest transmitters include a light source and optical and electronic connections. More complex transmitters may include circuits that put the input signal into the proper form for modulation, circuits that drive the light source with the input signal (or with a stable input to generate a steady output), devices to modulate and stabilize temperature, laser output monitors, and external modulators. You will learn more about these functions later in this chapter.

Transmitter modules vary considerably in design. All require some electronic interface for the input signal, and some optical interface to transfer the output signal to an optical fiber. The simplest are little more than an LED built into a connector adapter. Figure 10.1

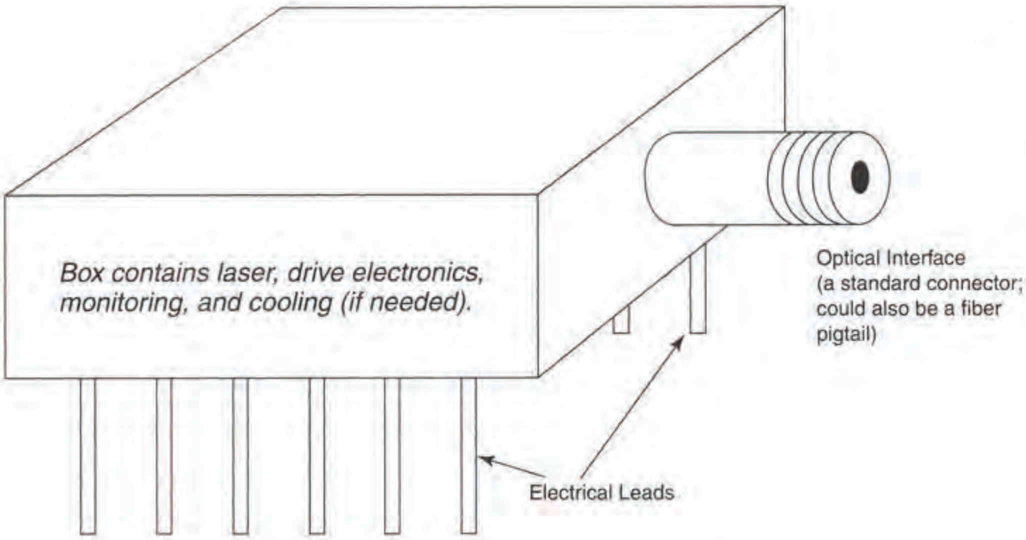


FIGURE 10.1
A simple transmitter module.

shows a more typical module, with a laser source and drive electronics packaged in a small multipin module. Generally a fiber is butted against the light source inside the package, and that fiber delivers the light to the outside world. In a package equipped with a connector adapter, like the module in Figure 10.1, a jumper fiber may deliver the signal to the connector interface; in very simple transmitters, the light source may be built into the connector housing. In either case, a fiber cable mates to the connector interface. Other transmitter modules deliver output via a fiber pigtail, which can be spliced to an output fiber in a patch panel or other splice enclosure.

Standardized transmitter modules are widely used.

The industry increasingly uses standardized modules available from a number of vendors in a common format. Figure 10.2 shows a modular Xenpak transceiver. The transmitter portion decodes four input channels encoded in one format, and combines them into a 10-Gbit/s signal encoded in the standard 10-Gigabit Ethernet format that is transmitted out a fiber. Optical, mechanical, and electronic interfaces for the 4.8-by-1.4-by-0.7-inch

FIGURE 10.2
A Xenpak transceiver for 10-Gigabit Ethernet. (Courtesy of Intel)

package are standard. The module also includes a clock, signal processing circuits, laser control circuits, and a laser driver, as well as a receiver.

These transmitter or transceiver modules typically are built into larger systems, buried inside boxes with only their optical interfaces visible. Larger systems often package transmitter and receiver modules with electronics that perform other functions, such as switching or combining low-speed electronic signals to produce a single higher-speed stream of output data. Although these electronics perform operations on the signals being transmitted, they are more properly considered part of the communication system rather than part of the transmitter.

A WDM transmitter includes a module for each optical channel.

A WDM transmitter is more complex because it contains many light sources operating at different wavelengths, each with its own associated electronics. You can think of a WDM transmitter as an array of transmitters for separate optical channels, each containing a single light source driven by a separate signal at a distinct wavelength. Because our concern is the optics, we'll look at them both as individual single-wavelength transmitters, and as collective multiwavelength transmitters.

Transmitter Performance

The light source or external modulator usually limits transmitter speed.

Electronic and optical components combine to determine the performance of an optical transmitter. The light source or external modulator limits the raw speed and power. No matter what drive electronics you use, you can't make the light signals change faster than the light source (or external modulator) is capable of changing. Likewise, the electronics can't extract more power than the light source is designed to deliver—except, perhaps, in the brief interval between the time the drive power overloads the light source and the moment the light source burns out.

The output wavelength depends mainly on the light source, but also may depend on control circuits that stabilize it at an assigned value. How much stabilization is needed depends on the application. Single-channel systems generally require no wavelength stabilization, but dense-WDM systems need active stabilization to lock the laser at a standard wavelength. Temperature also must be controlled because laser wavelength is temperature-sensitive. Cooling may also be required to limit operating temperature, because laser lifetime decreases as temperature increases.

Transmitter electronics and input signal set the modulation format and clock rate.

The transmitter electronics and the input signal set the optical signal modulation format and data rate (called *clock rate* in communications). As long as the laser or external modulator can handle the speed and power, it can transmit any format the electronics can support. Normally the transmitter electronics are designed to support standard formats.

Electronics also monitor the operation of the light source. An internal sensor may monitor output power from the rear facet of a laser, which usually transmits a small fraction of light for this purpose. Other circuits check drive-current levels, important because aging lasers need more drive current to generate the desired power level.

Analog and Digital Transmission

Optical transmitters can generate digital or analog signals.

Light sources are inherently analog devices, with output proportional to the modulating signal. An analog modulation circuit generates analog signals; a digital driver generates digital pulses.

Chapter 3 introduced analog and digital transmission. Recall that each has its virtues. Our eyes and ears are analog devices, so audio and video signals have to start and end in analog format. On the other hand, analog signals require much more precise reproduction and transmission than digital signals. This makes analog signals far more vulnerable to noise and distortion during transmission and processing. Digital signals can better tolerate distortion because they usually need only to detect whether a binary signal is off or on, not its shape or level. (Multilevel digital codes exist, but so far their only uses in fiber optics have been in the laboratory.) Digital electronics also are easier to design and cheaper to buy.

Digital signals are more robust than analog.

Digital transmission demands faster response than analog signals carrying the same information. Earlier you learned that a single telephone voice line requires only 3000 Hz of analog bandwidth, but needs 64,000 bits per second of digital transmission capacity. The two signals carry the same information. The digital version contains frequencies much higher than the 3000 Hz analog signal, but the digital signal does not have to be reproduced with the exacting precision needed for analog signals. The edges of digital pulses can be blurred as long as the receiver can tell the difference between on and off, but analog distortion is noise.

Although digital signals are far more common than analog in modern communication systems, some analog optical transmitters remain. As you will learn later, their most important applications are in cable television systems.

Bandwidth and Data Rate

Frequency bandwidth measures analog capacity. Bit rate measures digital capacity.

Transmission speeds are measured as bandwidth in analog systems and data rate or bit rate for digital systems. Analog bandwidth normally is defined as the frequency where the amplitude of the modulated signal drops 3 dB below the value at low frequencies, a 50% reduction in power. The digital data rate is the number of bits per second that can be transmitted with no more than a specified fraction of errors, often one incorrect bit in every trillion bits (10^{12}).

Rise time limits transmitter speed.

Both of these are quantities measured at the receiver, and don't directly measure the response or *rise time* that limits transmitters. Rise time is defined as the interval it takes light to rise from 10% to 90% of the steady-state high power level. *Fall time* is the inverse, the delay needed for the signal to drop from 90% to 10% of the maximum. If the rise and fall times are equal (they aren't always), this can be used to approximate bandwidth,

$$\text{Bandwidth (MHz)} = \frac{350}{\text{rise time (ns)}}$$

The precise relationship between bandwidth and rise time differs among light sources and transmitters.

Rise time is an important variable in selecting light sources and modulation techniques. LEDs have rise times ranging from a few nanoseconds to a few hundred nanoseconds. Directly modulated diode lasers have much faster rise times than LEDs (typically a fraction of a nanosecond), with VCSELs faster than edge-emitting diode lasers. In practice, edge-emitters can be directly modulated at rates to 2.5 Gbit/s, and VCSELs can reach 10 Gbit/s.

As you learned in Chapter 9, direct modulation induces a chirp in wavelength that can contribute to dispersion in long-distance transmission. Direct modulation normally is not used at speeds above 2.5 Gbit/s in long-distance systems, as external modulation usually is used at higher rates. Direct modulation can be used in short 10-Gbit/s systems, but external modulation is necessary at 40 Gbit/s.

Multiplexing

As you've seen earlier, multiplexing is the combining of multiple signals into a single entity that can be transmitted more economically. Multiplexing takes place at or before the transmitter. There are three important types, and their handling at the transmitter differs.

Time-Division Multiplexing

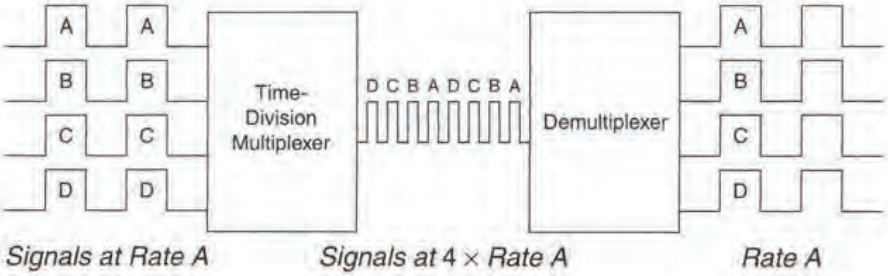
Time-division multiplexing interleaves bits from data streams to form a faster signal.

Traditional time-division multiplexing (TDM) combines two or more digital signals by interleaving bits or bytes from separate data streams to give one faster signal, as shown in Figure 10.3. For example, 24 voice phone lines, digitized at 64,000 bits per second, can be combined into one 1.55-Mbit/s digital signal. The combined signal carries all the bits from the 24 digitized phone signals, plus extra bits that help organize the combined signals. Appendix C lists the standardized hierarchy of successively higher data rates developed for telecommunication systems.

Alternatively, incoming streams of data can be broken into groups of bits called *packets*, which are transmitted together rather than interleaved with bits from other data streams. Incoming packets are held in an electronic buffer until they can be transmitted. This technique, used on the Internet, is sometimes called *statistical time-division multiplexing* to distinguish it from the traditional interleaving of bits shown in Figure 10.3. You'll learn more about the differences in Chapter 19.

In current systems, signals are time-division multiplexed by electronic circuits before they reach the transmitter. The electronic multiplexers that combine slower bit streams to make a higher-data-rate signal may be in the same box or rack as the transmitter. The higher the multiplexed data rate, the more important it is to perform electronic multiplexing close to the transmitter, because of problems with high-speed electronic transmission. Care must be taken in laying out electronic components inside transmitters operating at 10 Gbit/s or above. At the present time, the highest time-division multiplexing rate in regular use is 10 Gbit/s. Externally modulated transmitters have been developed for 40 Gbit/s, but there is little demand for that extra capacity.

FIGURE 10.3
Time-division multiplexing.



Electronic time-division multiplexing becomes extremely difficult at higher data rates. An alternative is *optical time-division multiplexing*. Interleaved optical TDM has been demonstrated in the laboratory, but applications are far off.

Frequency-Division Multiplexing

Frequency-division multiplexing combines signals at different frequencies.

Frequency-division multiplexing combines two or more signals modulated on carriers at different frequencies to produce one signal covering a range of frequencies. This system is used in broadcast radio and television and in cable television. Traditionally frequency-division multiplexing is used for analog signals, but it also can be used for digital signals.

Frequency-division multiplexing divides a range of frequencies into a set of distinct channels with fixed bandwidths that do not overlap. Each input signal modulates the carrier frequency for one channel, producing a signal that fits into one channel slot. Combining the modulated carriers produces a signal that spreads across a wider range of frequencies. You can see this on an AM or FM radio dial, where each station has a nominal frequency (the carrier) that actually spreads across a wider range, but should not overlap with the signals of other stations. The radio band contains dozens of stations; you tune your receiver to select a specific station. Television works the same way, but the channels are numbered rather than named by frequency, and some frequencies are reserved for services other than television.

Cable television also uses frequency-division multiplexing, but transmits the combined signals through optical fiber and coaxial cable rather than broadcasting them. Like broadcast television, cable systems assign channel numbers, which identify assigned frequencies. The different channels may carry different types of signals, including analog video, digital video, voice, and data.

Like time-division multiplexing, frequency-division multiplexing is done electronically before the signal reaches the transmitter. In practice, the term *frequency-division* is limited to multiplexing at radio frequencies.

Wavelength-Division Multiplexing

Wavelength-division multiplexing is the optical version of frequency-division multiplexing.

Think of wavelength-division multiplexing as being the optical version of frequency-division multiplexing. It transmits multiple optical signals through the same optical fiber at different wavelengths, just as frequency-division multiplexing transmits multiple radio signals at different frequencies. The basic principles are the same, but the terminology and details are different.

In practice, one optical transmitter is assigned per wavelength, as shown in Figure 10.4. The input is (generally) a digital signal from an electronic multiplexer. The optical transmitter modulates that signal onto the output of a light source at a wavelength λ. Other signals modulate the outputs of other transmitters with light sources at other wavelengths.

Separate fibers deliver the output of each transmitter to an optical multiplexer, which combines them for transmission through a single output fiber. Each wavelength in the signal is a separate optical channel, which if the system is designed properly does not affect other optical channels, as in frequency-division multiplexing of radio broadcast through the air.

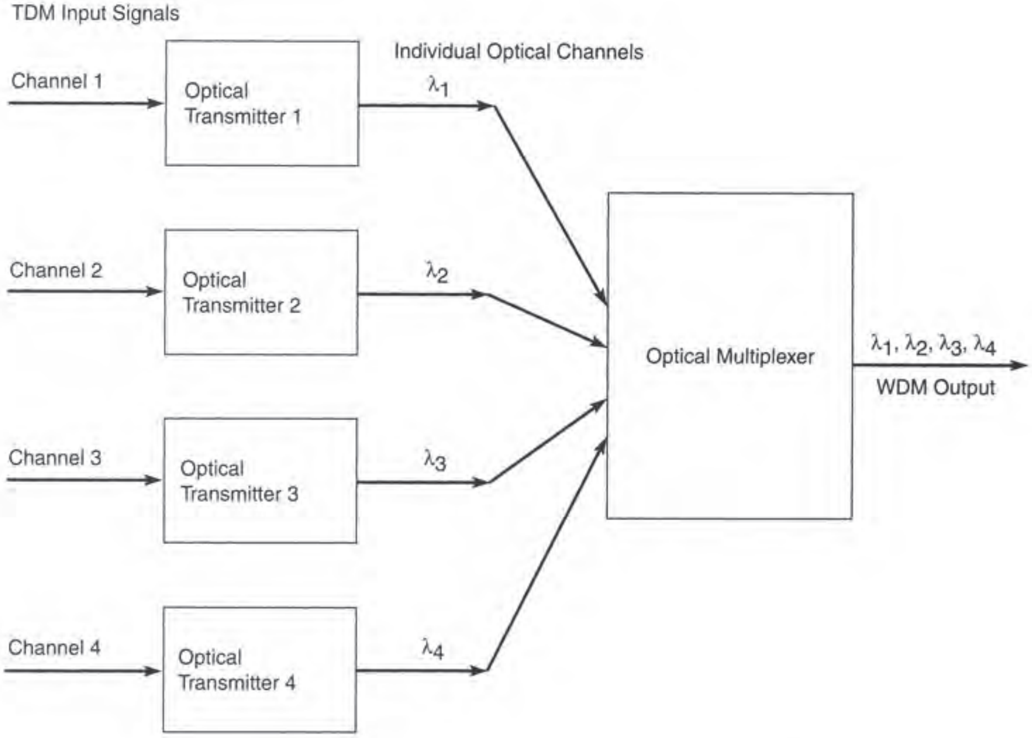


FIGURE 10.4 *Wavelength-division multiplexing at the transmitter.*

Wavelength-division multiplexing comes after the light source.

Note that wavelength-division multiplexing takes place *after* optical transmitters generate modulated optical signals at different wavelengths. Time-division and frequency-division multiplexing are both done electronically, *before* the signals modulate a transmitter.

Later chapters will cover other aspects of wavelength-division multiplexing in more detail, particularly its implications in overall system design and performance.

Modulation

So far, we have discussed modulation in very general terms. Transmitter performance depends strongly on how the light source is modulated, so let's take a look at the details.

Amplitude or Intensity Modulation

Most optical carrier signals are modulated in intensity.

Most optical transmitters are modulated by signals that change the amplitude or intensity of the light they generate. Figure 10.5 shows how this works for an ideal light source that normally generates a steady coherent light wave.

The ideal light wave at the top is the *carrier signal,* like the carrier frequency of a broadcast radio or television station. *Amplitude modulation* changes the amplitude or intensity of the output light as the amplitude of the drive signal changes. This is a straightforward approach because the output power of an LED or diode laser increases with drive current.

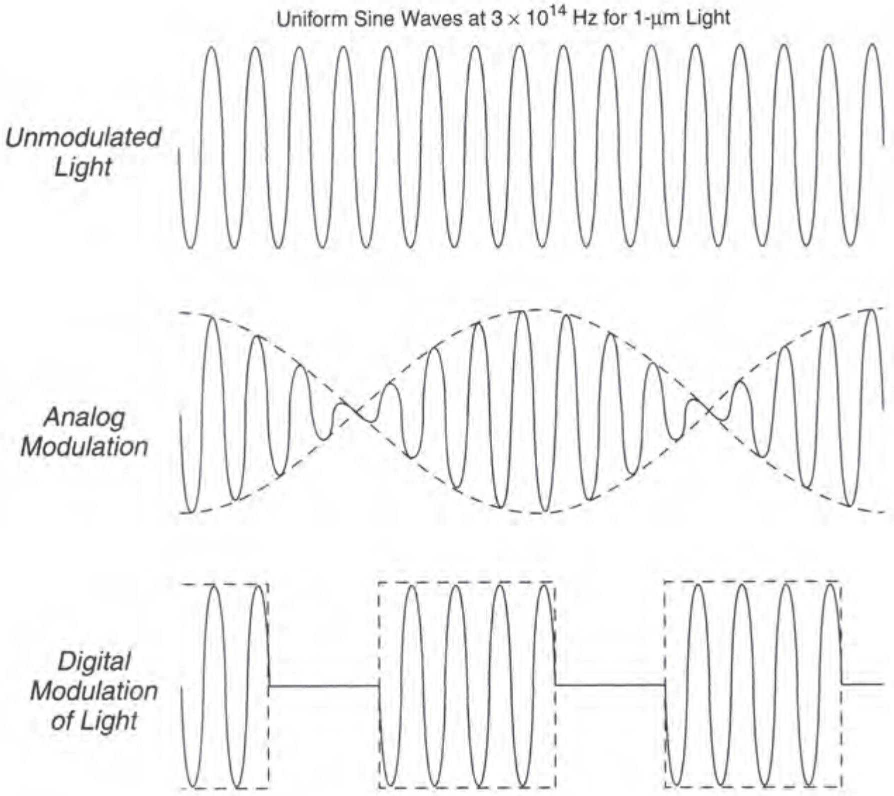


FIGURE 10.5 *Amplitude modulation by digital and analog signals.*

This can produce the analog signal at the middle of Figure 10.5 or the digital signal at the bottom.

The figure was distorted in one important way to show the scale of the modulation. Light at 1500 nm has a frequency of about 200 THz, or 10^{14} Hz. The signals modulating the light wave are at much lower frequencies, measured in gigabits per second for a digital signal. If the figure was drawn to scale, one pulse from a 10-Gbit/s data stream would contain around 20,000 waves of light. That's impossible to draw, so the figure pretends the optical and modulation frequencies are much closer.

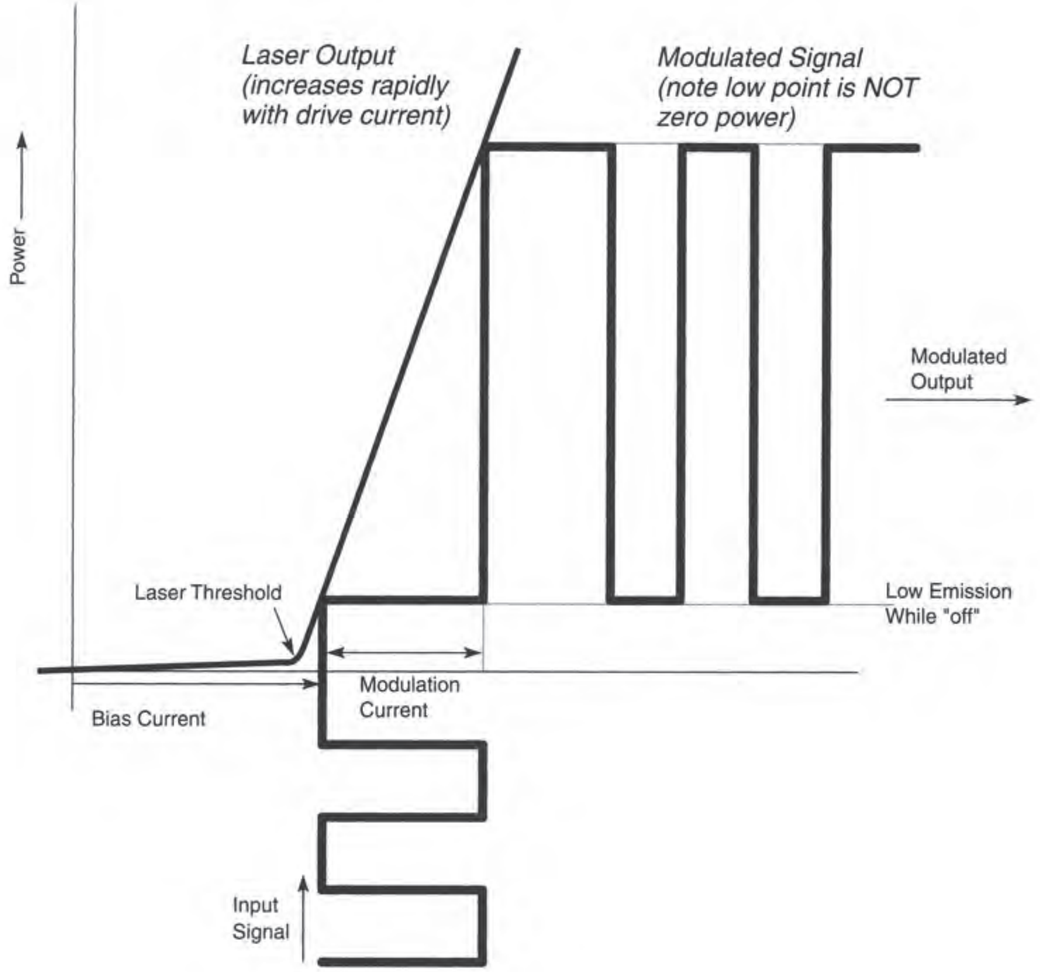
Amplitude modulation is standard in fiber-optic systems because it's simple and easy. Light source output naturally varies with drive current, and intensity modulators are relatively easy to build.

Direct and External Modulation

As you've learned, semiconductor light sources can be modulated either directly or externally. Increasing the drive current increases the power emitted by either LEDs or semiconductor lasers. Diode lasers respond very quickly to the increase, but LEDs have slower response, limiting their modulation bandwidths.

Direct modulation is ideal for inexpensive transmitters, but it causes an undesirable wavelength chirp, which causes excessive chromatic dispersion at high speeds. In addition, lasers develop undesirable relaxation oscillations at frequencies of a few gigahertz, which limits the

FIGURE 10.6
Direct modulation of a laser diode.



maximum frequency for direct modulation. External modulators are needed at higher speeds, or when the light source cannot be directly modulated. They are covered in Chapter 16.

Lasers are modulated by adding a signal current to a steady bias current.

LEDs normally are switched off and on because they turn on as soon as the current starts. The threshold effect complicates the operation of diode lasers. When a diode laser is switched on, emission does not begin until the drive current exceeds the threshold, delaying the output pulse. This can be a problem at high speeds, where delay can significantly shorten the pulse. To prevent this effect, diode lasers normally have two sources of drive current, which add together. One produces a steady bias current, which typically slightly exceeds laser threshold. The other is the time-varying signal, which actually modulates the laser output by increasing the current above the bias level, as shown in Figure 10.6.

This bias arrangement leaves the laser emitting some light when it is nominally "off," but much more when it is "on." Receivers can accommodate this difference by setting proper decision thresholds for the presence or absence of pulses, as you will learn in Chapter 11.

Don't forget that semiconductor diodes respond differently to current than to voltage. As you saw in Chapter 9, a minimum voltage has to be applied across an LED or laser to give the charges carrying current through the device enough energy to recombine at the junction level. Once that *voltage* is exceeded, current starts flowing. That current is enough to start an LED to emit some light, but the drive current must exceed a threshold to produce stimulated emission from a diode laser.

Coherent Transmission

If you're familiar with radio, you know that amplitude modulation is not particularly sophisticated. The amplitude-modulated AM-radio band is justly notorious for its static and background noise. Modulating the frequency of the carrier signal rather than the amplitude gives the FM (frequency-modulated) radio band much better signal quality. AM receivers pick up random spikes of amplitude noise from spark plugs and power lines, and the signal strength fades with distance from the receiver. In contrast, FM receivers do not pick up random noise spikes because they don't change the frequency of the radio signal. In addition, the strength of an FM signal depends not on its intensity but on how it changes the carrier frequency, so signals don't fade into the background; they stay strong until they start breaking up. (Other differences in AM and FM reception come from the different transmission frequencies.)

You might think it logical to try frequency modulation to improve optical transmission. In fact, it's been tried, but hasn't worked out very well.

Frequency modulation is one type of what optical engineers call *coherent transmission*, which works like a heterodyne (FM) radio system. The trick is to combine the incoming frequency-modulated signal with another signal kept at a constant frequency. Processing the difference between the input signal and the constant frequency (called a local oscillator) reproduces the radio program. As shown in Figure 10.7, the optical version requires a pair of lasers, one at the transmitter and a second (the local oscillator) at the receiver. The two lasers emit slightly different frequencies, v_1 and v_2. The transmitter modulates the frequency of the outgoing laser beam by passing it through a suitable external modulator. Alternatively, a modulator could delay the phase of the transmitter beam, causing a phase shift. At the receiver, the incoming signal at the transmitter frequency is mixed with the local oscillator beam, to give a microwave signal at the difference between the frequencies of the two light waves. That microwave signal carries the frequency- or phase-modulated signal, which can be extracted by further processing.

Coherent optical transmission works like an FM-band radio, but has not yet proved practical.

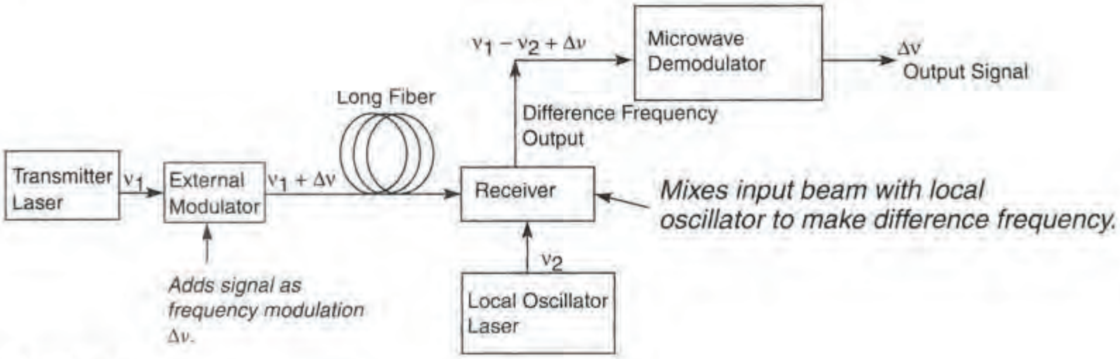


FIGURE 10.7
Coherent optical transmission.

Experiments have shown that coherent transmission can work, but so far it has not proven practical, and it's largely been forgotten in the explosive development of wavelength-division multiplexing and optical networking.

Single-Channel Transmitter Design

We've already covered several components of optical transmitters. To pull the whole picture together, we'll now consider a generalized single-channel transmitter as a whole. Our example will include the elements needed in high-performance laser transmitters (some won't be present in low-speed LED transmitters). To keep the picture clear, we assume the transmitter receives its input signal in final electronic form from an external source, and doesn't have to encode any signals. This type of transmitter is usually a module inside a larger system. For convenience, we assume the wavelength is fixed, not tunable.

You should remember that the real world is more complex. For example, a transmitter and a receiver are often packaged together as a transceiver; pairs of transceivers are attached to opposite ends of a fiber pair for two-way transmission; transmitters operating at separate wavelengths may be packaged together if their outputs are directed into the same fiber; sometimes some signal encoding is done inside a transmitter module; tunable lasers are starting to come into use; and sometimes a transmitter may be packaged with other equipment in a larger box that is also called a "transmitter," although it may also perform other electronic and optical functions.

Figure 10.8 is a functional diagram of a transmitter module. (You'll learn about receivers in Chapter 11.) Its main functional elements are:

- Housing
- Electronic interfaces

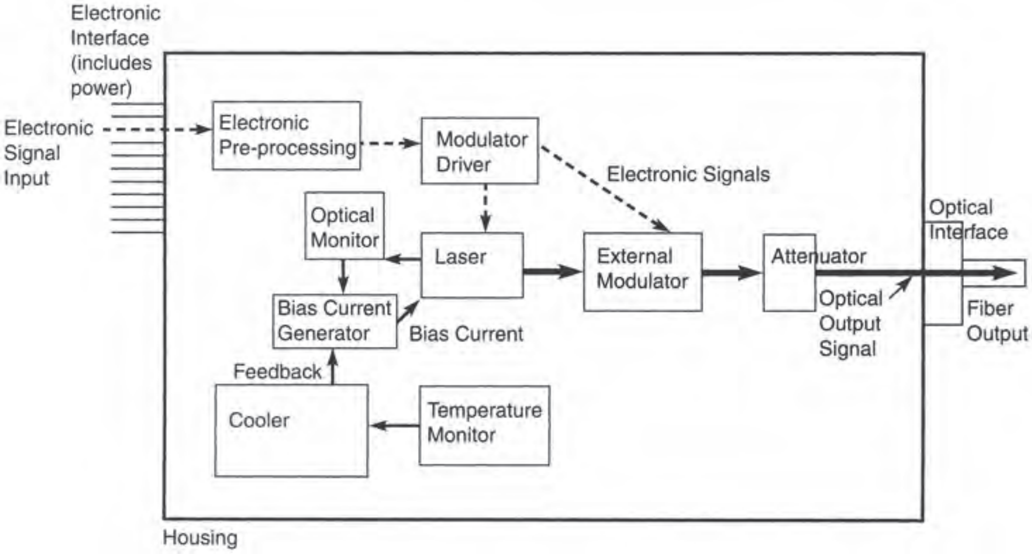


FIGURE 10.8
Generic optical transmitter.

- Internal signal-processing electronics
- Laser monitoring and bias circuits
- Temperature control
- Optics (including laser and external modulator)
- Optical interfaces

Some control functions can be performed by circuits outside the packaged transmitter. For example, output of the sensor that monitors laser output can be routed to external circuits that adjust bias current. Temperature data also can be routed to external controls. Here we will consider their functions rather than their locations.

Housing and Electronic Interfaces

The mechanical housing of an optical transmitter is a box designed to mount conveniently within other equipment. Screws, solder bonds, or other fasteners attach the housing mechanically to a printed circuit board or other substrate. More complex transmitters are packaged in cases that mount in standard equipment racks.

Typically, pins on the package provide electronic connections with other devices. They deliver power to drive the laser, the input signal, and other required signals. The transmitter also may return some information to other equipment, such as transmitter status.

The mechanical housing is designed to mount conveniently in other equipment.

Signal Processing Electronics

Internal electronics typically do some processing of the input signal to put it into a form suitable for modulating the light sources. One function is converting signals from the voltage variations used in electronic circuits to the current variations needed to directly modulate diode lasers and LEDs. Other processing may change signals to formats better suited for fiber transmission. Transmitters used in some networks may include buffers to hold input data until it can be sent, because the networks can deliver data faster than the transmitter can send the data.

Internal electronics put the input signal into a form suitable for modulating the light source.

Figure 10.8 shows the path of the signal from the outside connection through the electronic preprocessor, and modulator driver to either the diode laser or the external modulator. In direct-modulation systems, the modulation current is added to the bias current to drive changes in the optical output.

Laser Drive and Monitoring Circuits

The handling of laser drive, bias, and monitoring circuits depends on the type of modulation.

For direct modulation, a bias circuit provides the bias current that drives the laser to a point just above threshold so it generates a weak beam. A separate drive circuit adds the modulating current, which for digital transmission is a series of pulses. When the pulses are "on," the laser output is high.

For external modulation, a bias circuit drives the laser so it emits a steady output power high enough to compensate for any losses in the external modulator. A separate drive circuit delivers the signal to the external modulator, causing it to vary from opaque to transparent and switch the beam off and on.

In both types of laser transmitter, some light from the laser is directed to a sensor that monitors the laser output. Typically this is a small amount of light coupled out the back of the laser, but it could be a small part of the output beam. The sensor measures the intensity of this light and uses that information to drive a feedback circuit, which adjusts the bias current to keep the output power steady. Laser power declines as the laser ages, so this feedback circuit turns up the bias current to compensate for this decline in power. (Eventually, however, the laser will no longer be able to meet power specifications.)

Temperature Control

Temperature stabilization maintains a fixed wavelength.

The threshold current, output power, and wavelength of a diode laser vary with its operating temperature, making temperature stabilization important in high-speed or narrow-line transmitters.

Power and threshold current variations are important in determining laser lifetime. The threshold current, I_{thresh}, increases exponentially with temperature T:

$$I_{\text{thresh}}(T) = I_0 e^{(T/T_0)}$$

where I_0 is a constant and T_0 is the characteristic temperature of the laser material. Simple double-heterostructure InGaAsP lasers have T_0 of 50° to 70°K, but strained layer structures, quantum wells, and improved confinement can increase T_0 values to as high as 180°K. The characteristic temperature is higher for GaAs lasers of comparable structures; for simple lasers, T_0 is 120°K. Increasing the temperature reduces the slope efficiency of light generation above laser threshold, further decreasing overall efficiency and output power produced at a particular drive current. Reduced efficiency increases heat generation, which warms the laser and, without adequate heat dissipation, can lead to thermal runaway.

Wavelength variations are particularly important for lasers used in dense wavelength-division multiplexing, where the wavelength must be kept constant within a small fraction of a nanometer to keep the signal in the desired optical channel. The temperature affects both the refractive index of the semiconductor and the physical length of the cavity, two factors that determine the output wavelength. Details vary with the material and laser structure, but in general changing the temperature of a laser by 10°C changes the wavelength about 1 nm. Therefore, a change of just a few degrees can shift the laser's output to the adjacent channel in a dense-WDM system.

Wavelength-insensitive transmitters can operate uncooled at 10 Gbit/s.

Cooling requirements depend on the system. Heat sinks and passive cooling can suffice for data rates to 10 Gbit/s if the output power is not high and if the system design allows the laser wavelength to vary by 20 nm. In practice, that variation is allowable in single-channel systems transmitting through standard single-mode fiber at its zero-dispersion wavelength of 1310 nm. Heat sinks also might be used in WDM systems with widely spaced wavelengths. But in practice there is little interest in uncooled operation in the 1550-nm window, because most systems operating in that band use closely spaced WDM channels. In such cases, wavelength tolerances are tight, output powers are high, or the system is very sensitive to wavelength chirp and dispersion, so active cooling is required.

Active cooling normally uses a temperature monitor to control a thermoelectric cooler, which can maintain transmitter tolerance within the required range.

Optics and Optical Interfaces

The optical part of the transmitter starts with the laser, which generates either a steady beam or a modulated output depending on the input. Monitoring the laser's output through the rear facet controls the laser's output power through feedback to the bias-current generator.

Modulation either adds to or subtracts from the light generated by the laser source. If the laser is directly modulated, the modulation drive current increases the laser's output power above the biased level where it emits a little light. If the laser is modulated externally, the bias current drives it at a high level, and the modulator varies between fully transparent and essentially opaque. That means the external modulator *reduces* the power level in the laser beam. As with direct modulation, a small amount of power may be transmitted in the "off" state.

Direct modulation adds to the laser output from a low-bias current.

An attenuator is optional. It may be needed if the receiver is close to the transmitter, and might be overloaded by the transmitter's unattenuated power. Attenuators are most likely to be used in networks that use standard transmitters, but where the attenuation in the cable can vary over a wide range.

Wavelength Tunability

Including tunable-wavelength lasers will increase transmitter complexity. In addition to monitoring temperature and power, the transmitter will need to monitor and stabilize the operating wavelength within system margins. Many applications also will require the operator to adjust laser wavelength manually or remotely, either so tunable lasers can be used as replacements, or so the transmitter can be reconfigured. Few tunable lasers are now used in transmitters, so we won't worry about the details.

Sample Transmitters

The best way to get a feeling for the internal workings of transmitters is to look at examples, deliberately simplified to aid in understanding basic concepts.

Figure 10.9 shows a very simple drive circuit for an LED. A bias voltage $+V$ is applied across a transistor, LED, and current-limiting resistor, going to ground. This provides the

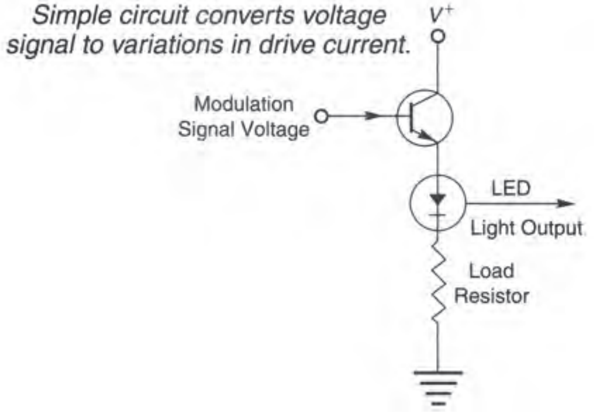
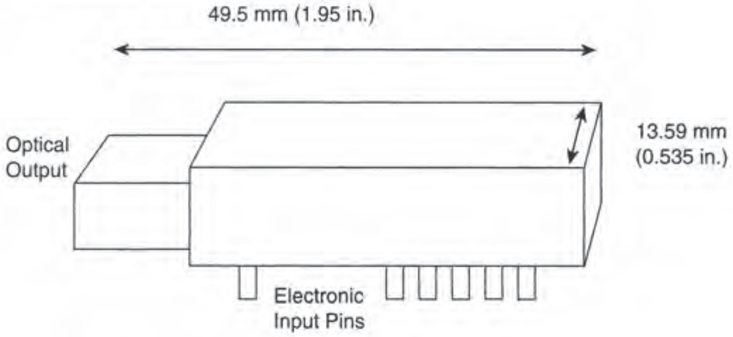


FIGURE 10.9
Simplified drive circuit for LED.

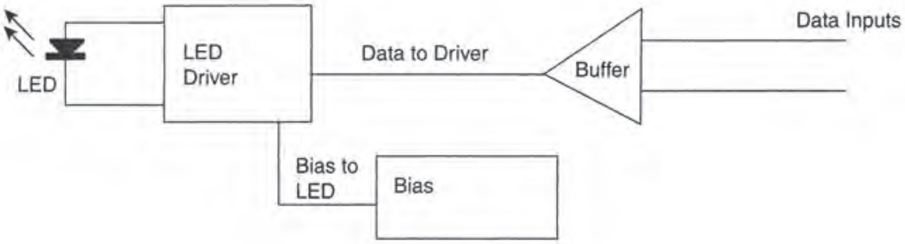
FIGURE 10.10

156-Mbit/s fiber-optic transceiver, packaged with integral small-form-factor connector.

Dimensional Sketch of Transceiver



Block Diagram of Transmitter Side (receiver not shown)



forward-voltage bias needed by the LED (about 1.5 V for 850-nm emitters and around 1 V for 1300-nm LEDs). The input signal is delivered as a modulated voltage applied to the base of the transistor, which causes variations in the current passing through the LED and load resistor. The LED could be on either side of the transistor, but the circuit requires a transistor or other circuit element to modulate the drive current. The resistor limits LED current. The circuits that generate the input signal are not shown. The higher the speed and performance level, the more complex the circuit.

These transmitters can be made compact. Figure 10.10 shows the dimensional outlines of a transceiver, which includes a 1300-nm LED transmitter able to operate at speeds to 156 Mbit/s. The figure also shows a block diagram of the transmitter portion. The whole integrated package (transmitter and receiver) is just under 50 mm (2 in.) long, including a built-in connector interface at the left. (One thing to watch out for is that some mechanical specification sheets still don't label the units of dimensions. They usually include both inches and millimeters, with the metric units the higher numbers, but they may not label them. Remember what happened to Mars Climate Orbiter when a NASA contractor forgot to label units!)

Diode lasers have electrical characteristics similar to those of LEDs, but their optical operation differs. Lasers emit little light until drive current passes a threshold value, but above that threshold tend to emit higher power with more efficiency. Lasers draw much higher drive currents than LEDs, so they normally are used with smaller current-limiting resistors. Laser transmitters also may require additional components to meet the more demanding requirements of laser operation, including a sensor to monitor laser output, a temperature-sensing

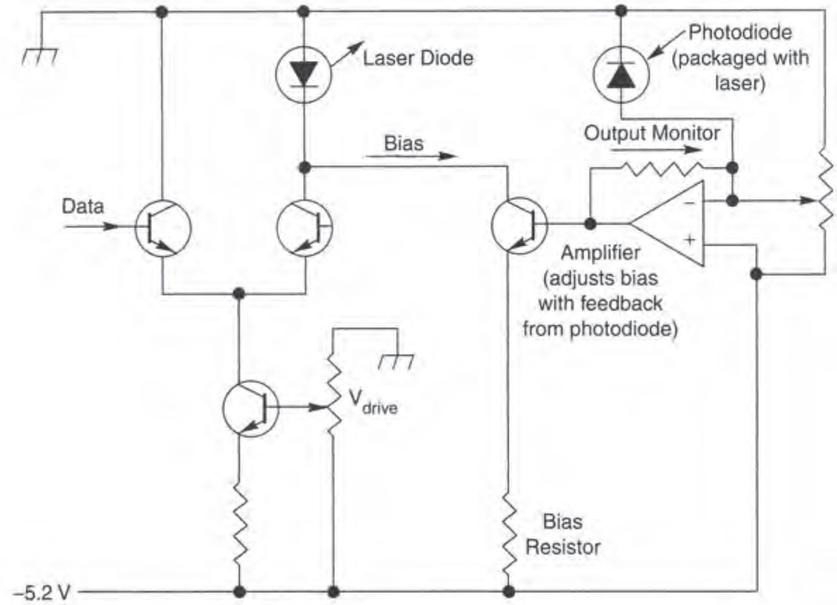


FIGURE 10.11

Laser diode drive circuitry. (Adapted with permission from a figure by Paul Shumate)

thermistor to monitor transmitter temperature, and a thermoelectric cooler. Typical laser transmitters are somewhat larger than the one shown in Figure 10.10, to accommodate these extra components.

The circuit in Figure 10.11 is an example of the type used in a directly modulated laser transmitter. The drive signal is applied as a voltage to the base of a transistor, where it adds to a bias voltage and modulates drive current through the laser diode. The detector packaged with the laser monitors its output, providing feedback to an amplifier (at right), which adjusts the bias applied to the laser diode so average power remains constant. Externally modulated lasers require simpler drive circuits because they need only provide constant output power, but the transmitter then requires drivers for the external modulator.

Externally modulated lasers have simpler drive circuits, but the transmitters require separate electronics to drive the external modulator.

Typically the drive circuit prebiases the laser with a current close to threshold. Adding the modulated drive signal raises it above the threshold, so it starts emitting light. This approach enhances speed by avoiding the delay needed to raise the drive current above threshold. At low speeds, the prebias may be below the threshold, to avoid generating low-level emission that might confuse the receiver. At higher speeds, the laser may be prebiased to emit a little light in the off state, with addition of the modulated drive signal increasing power greatly. Prebiasing requires careful adjustments so the receiver does not mistake the low emission from the laser in its off state for a faint pulse in the on state.

Steady advances in electronic packaging and integration have squeezed high-performance transmitters into quite compact packages. Figure 10.12 shows an XFP transceiver module along with a functional diagram. An entire 10-Gbit/s laser transmitter and receiver fit into a case that measures 78 mm long, 18.4 mm wide, and 8.5 mm high. It operates on 3.3 volts and draws 1 to 2 watts of electrical power. It operates uncooled, generating up to a milliwatt of light from a 1310-nm laser source. Optical interfaces are in the front—a pair of standard LC fiber connectors, one for the input and the other for the output. The electrical interface is a 30-pin electrical connector in the back.

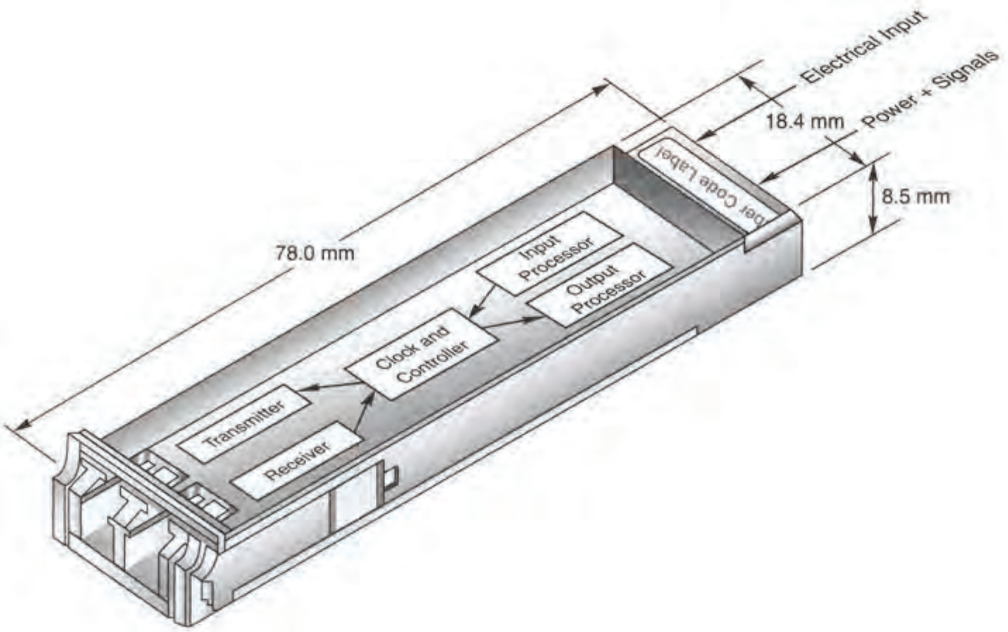


FIGURE 10.12
An XFP transceiver with a block diagram of its components.

The XFP is one of several standardized modules developed for 10-Gbit/s transmission and offered by multiple developers. The design reflects a growing trend in the communications industry of standardizing mass-produced, interchangeable modules. The modules differ somewhat in design philosophy. The XFP module is particularly small because it includes only the minimal electronics needed to drive the transmitter at 9.95 to 11.1 Gbit/s, as shown in the block diagram of Figure 10.8. Other standard modules such as Xenpak include electronics that format input signals for transmission standards such as 10-Gigabit Ethernet. That approach requires extra circuits, so those modules must be larger.

What Have You Learned?

1. A transmitter contains a light source and generates optical signals from an electronic input. It includes optical and electronic connections.
2. A transceiver is a package that includes both the transmitter and receiver on one end of a system. It connects to a pair of fibers.
3. Transmitters range from simple to complex. The complexity increases with transmission speed.
4. Wavelength-division multiplexing requires a transmitter module for each optical channel that generates a separate wavelength.
5. Transmitter electronics and the input signal define the modulation format and the clock rate for a transmitter.
6. Optical transmitters may generate digital or analog signals. Digital signals are more robust than analog, but require higher data rates to carry the same information.

7. Frequency bandwidth is the capacity of an analog transmitter. Data rate is the capacity of a digital transmitter. Both depend on the rise time of the transmitter.
8. Time-division multiplexing combines bits from data streams to form a faster signal. Frequency-division multiplexing assigns signals to different frequencies combined into one signal. Both are done electronically before the signal is fed to the optical transmitter.
9. Wavelength-division multiplexing transmits separate signals through one fiber at many different wavelengths. It requires one transmitter per optical channel. The signals are combined on the fiber after the optical transmitter.
10. Optical carrier signals are modulated in intensity.
11. Major elements of a typical transmitter are the housing, the electronic interfaces, the internal signal-processing electronics, the detector and circuits to monitor laser bias, temperature control, the laser and (optional) external modulator, and the optical interfaces.
12. Temperature stabilization is important in maintaining stable wavelength and output power from the laser.

What's Next?

Now that you have learned about fiber-optic transmitters, Chapter 11 will examine the other end of the system, the receiver.

Further Reading

HP Fiber Optic Technical Training Manual, available from *http://www.agilent.com/*

Paul W. Shumate, "Lightwave Transmitters," in Stewart E. Miller and Ivan P. Kaminow, eds., *Optical Fiber Telecommunications II* (Academic Press, 1988)

Questions to Think About

1. The difference between the "off" and "on" states of a laser transmitter is 20 dB. The "on" output is 1 mW. What is the "off" output at the output port of the transmitter? What is the "off" output after 30 dB of loss?
2. The diode laser used in a transmitter has a threshold current of 10 mA and normally operates at 15 mA in the "on" state. Suppose it is directly modulated with a signal without a bias current, which turns the laser on from zero current. The signal pulse has 30-ps rise time, 30-ps "on" time, and 30-ps fall time. How long is the output optical pulse, starting from the time the current crosses laser threshold?
3. Suppose you could find electronics fast enough to generate time-division multiplexed signals at 640 Gbit/s on a single optical channel. How many light waves would one bit correspond to at a wavelength of 1550 nm?

4. Time-division multiplexing requires reducing the duration of low-speed pulses and interleaving them into a combined signal. It now is done electronically. Suppose you had a way to reduce the duration of optical pulses. Can you think of a way to interleave these shortened optical pulses for time-division multiplexing? Consider a system that multiplies data rate by a factor of four. (*Hint:* Think about ways to delay pulses.)

Chapter Quiz

1. Digital transmission capacity is measured as
   a. bandwidth in megahertz.
   b. rise time in microseconds.
   c. frequency of 3-dB point.
   d. number of bits transmitted per second.

2. Analog transmission capacity is measured as
   a. bandwidth in megahertz.
   b. rise time in microseconds.
   c. frequency of 3-dB point.
   d. number of bits transmitted per second.

3. If the rise time of a transmitter is 1 ns, what is its theoretical bandwidth?
   a. 1 Gbit/s
   b. 100 MHz
   c. 350 MHz
   d. 350 kHz
   e. 350 Mbit/s

4. Standard optical interfaces for transmitters are
   a. integral optical connectors or fiber pigtails.
   b. integral electronic connectors or fiber pigtails.
   c. output windows or fiber pigtails.
   d. 14-pin DIP packages.
   e. only fiber pigtails.

5. How many transmitters are needed for wavelength-division multiplexing?
   a. none; the signals are combined in the optical fiber
   b. one
   c. one per optical channel
   d. depends on the data rate per channel

6. What provides feedback to stabilize laser intensity in a transmitter?
   a. a signal relayed from the receiver
   b. changes in input impedance

c. output from the rear facet of the laser monitored by a photodiode
d. light scattered from the optical interface with the input fiber
e. no feedback is required

7. What is the usual modulation method for fiber-optic transmitters?

a. intensity modulation
b. frequency modulation
c. wavelength modulation
d. voltage modulation

8. What does time-division multiplexing do?

a. transmits different signals at different wavelengths
b. shifts the frequencies of several analog signals to combine them into a single input
c. encrypts signals for secure transmission
d. interleaves several digital signals into a single data stream

9. What is the total data rate of a WDM system carrying 2.5-Gbit/s signals at 1550, 1552, 1554, 1556, 1558, 1560, 1562, and 1564 nm?

a. 2.5 Gbit/s
b. 10 Gbit/s
c. 12.5 Gbit/s
d. 20 Gbit/s
e. 25 Gbit/s

10. With wavelength-division multiplexing, how many fibers do you need for two-way transmission of 2.5-Gbit/s signals at 1550, 1552, 1554, 1556, 1558, 1560, 1562, and 1564 nm?

a. one
b. two
c. four
d. eight
e. sixteen

11. A 1300-nm LED-based transmitter has rise time and fall time of 1.2 nanoseconds. What is the best application?

a. transmitting 10 Gbit/s through single-mode fiber
b. transmitting Gigabit Ethernet (1 Gbit/s data rate) through graded-index multimode fiber
c. transmitting fast Ethernet (100 Mbit/s data rate) through graded-index fiber
d. transmitting ATM signals (155 Mbit/s data rate) through single-mode fiber
e. transmitting 1.5 Mbit/s through graded-index fiber

12. The LED-based transmitter in Question 11 has minimum output of −19 dBm into 62.5/125 graded-index multimode fiber. The receiver requires an input of at least −30 dBm. If fiber attenuation is 2 dB/km and other losses total 3 dB, how far away can you put the receiver?

a. 1 km
b. 3 km
c. 4 km
d. 8 km
e. 11 km

13. How is LED output modulated?

a. externally by varying current delivered to an external modulator
b. directly by varying drive current
c. directly by varying voltage applied to an external modulator
d. directly by varying LED operating temperature

14. Temperature stabilization of a laser transmitter also stabilizes what operating characteristics?

a. laser threshold
b. output power
c. wavelength
d. all of the above
e. none of the above

CHAPTER

11

Receivers

About This Chapter

A fiber-optic receiver converts the optical signal transmitted through a fiber into an electronic form that can serve as input to other devices or communication systems at the receiver end of the system. This device contains two essential components—a detector, which converts an optical input into an electronic signal, and receiver electronics, which convert the raw detector output into a form usable by other equipment. In practice, this conversion may require some processing of the signal.

After defining the concept of receiver, this chapter first describes optical detectors, then covers their electronic functions. The chapter closes with examples of receiver circuits. Optical amplifiers and wavelength-division multiplexing, which sometimes are used at the receiver end of the system, are covered in Chapters 12 and 15, respectively.

Defining Receivers

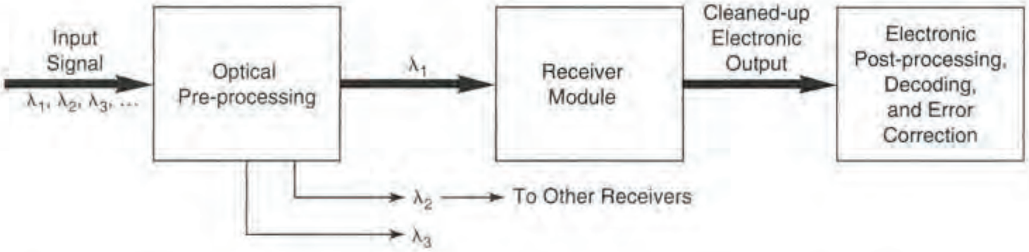
Fiber-optic receivers detect optical signals and convert them to electronic signals usable by other equipment. In general, receivers include both optical and electronic devices.

Figure 11.1 shows major functions that receivers may perform. They fall into three general categories:

Detectors are color-blind.

- *Optical pre-processing* of the input optical signal. This may include optical preamplification of the input signal to raise it to a level that improves receiver performance. If the input signal includes multiple optical channels, those must be separated by a *wavelength-division demultiplexer*, which directs each wavelength to a separate optical detector. This is vital because the detectors themselves are color-blind and cannot separate signals at different wavelengths. Like black-and-white film, detectors tell you the brightness of the input light, but not the color. Signals of different wavelengths must be separated before the detector, as shown

FIGURE 11.1
Processing functions at the receiver end of a fiber system.



in Figure 11.2. (This stage is not required if the input signal includes only one wavelength.) You'll learn about these functions in Chapters 12 and 15.

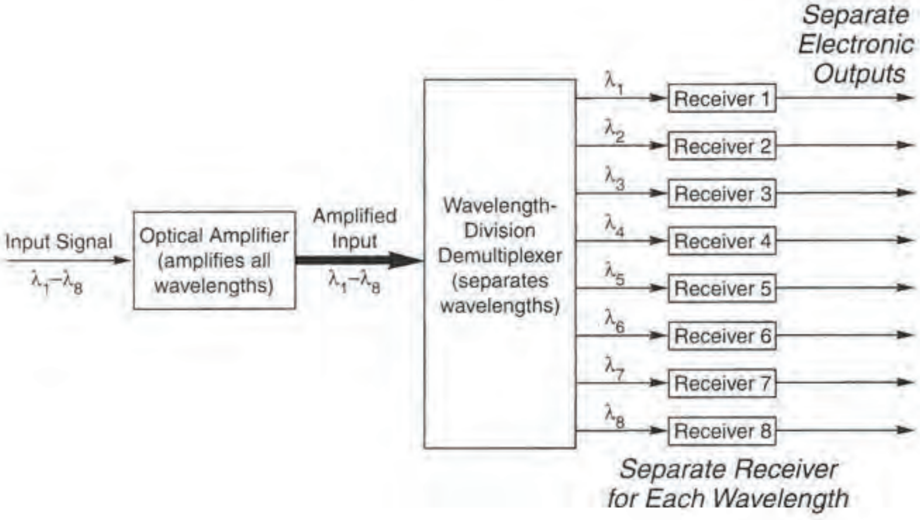
A digital receiver module produces a string of output bits.

- *Receiver module,* which converts the input light signal at one wavelength into an electronic form that represents either an analog signal or a series of input bits. The *detector* responds to input photons by producing a flow of electrons, so its raw output is an electrical current (with a few rare exceptions). All but the simplest receiver modules have electronic conversion and amplification stages that convert the current to a voltage, then amplify it to produce a stronger signal. Digital receivers include an additional stage that cleans up the output to produce a clear set of pulses. The output of the receiver module is a fresh signal, amplified and cleaned up. Figure 11.3 shows these stages. Note that the digital receiver includes an input analog section. You also can view the digital stage as an addition to an analog receiver. In either case, the digital version produces a string of output bits.

Electronic post-processing decodes bits from the receiver module.

- *Electronic post-processing* converts the raw output bits to the form needed by external systems. For digital systems, this can involve two stages: error correction and decoding. Error correction performs internal checks to verify that all bits have been received correctly, then tries to fix any mistakes. Decoding identifies how bits should be grouped and what they mean. These functions depend on the system, and usually are grouped with the system electronics rather than inside the receiver module. For

FIGURE 11.2
Optical pre-processing amplifies and splits a WDM signal among receiver modules.



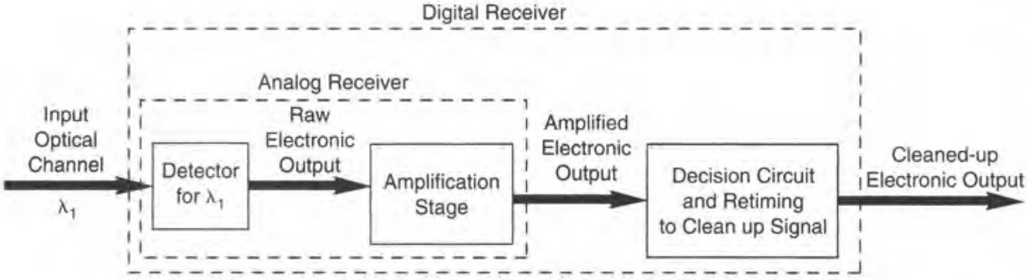


FIGURE 11.3

Receiver module includes analog and digital stages.

example, the receiver side of the XFP transceiver described in Chapter 10 produces a raw string of bits at a rate around 10 Gbit/s. Separate electronic post-processing circuits decode the signal to meet system requirements. The same receiver modules can be used for 10-Gigabit Ethernet and 10-Gbit/s telephone standards, but different post-processing is needed to decode the different signal formats. You'll learn more about those formats and standards in Chapters 19 and 20.

The receiver module, like the transmitter module, is often called an *opto-electronic* device because it converts optical signals into electronic form. As you learned in Chapter 10, the receiver module is often packaged with the transmitter module in a transceiver. We'll start by learning about detectors, then move on to the electronic side of the receiver module.

Detector Basics

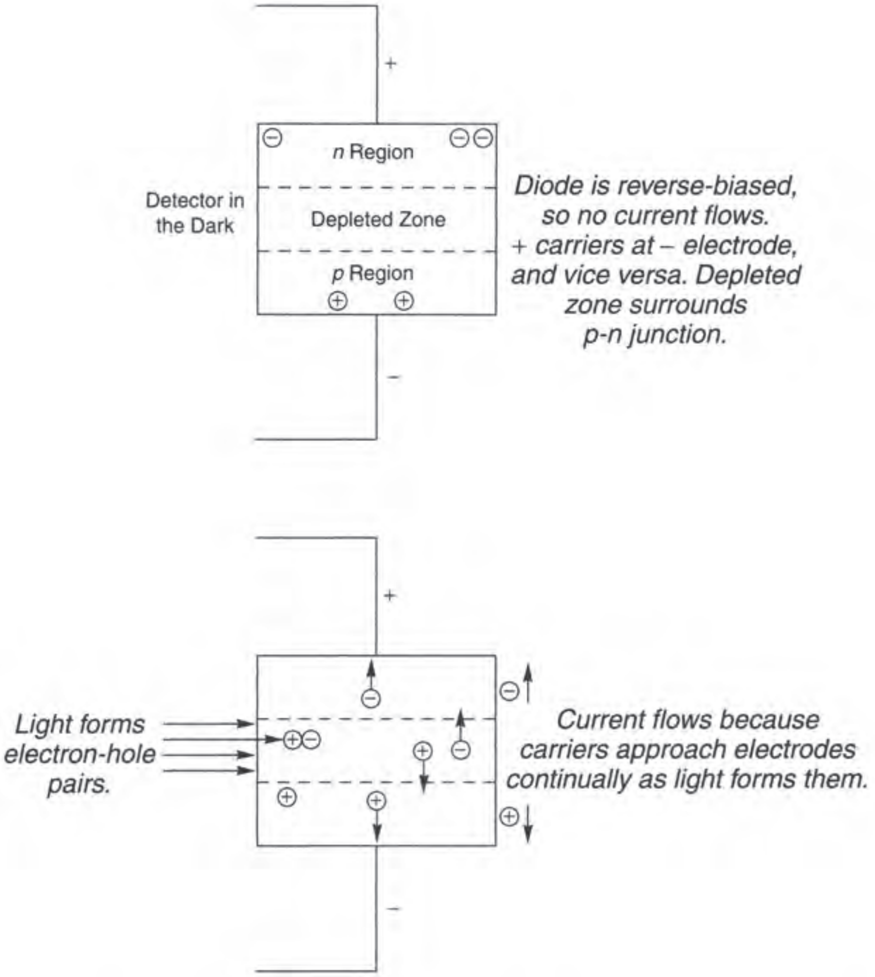
The detectors used in fiber-optic communications are semiconductor diodes that produce electron-hole pairs when illuminated by light. They are often called *photodiodes* or *photodetectors.* The incoming photons must have enough energy to raise electrons from the valence band to the conduction band. Because this bandgap energy depends on the material, the wavelength response also depends on the material. The diode also must be designed and operated so the current carriers can produce a signal that can be transmitted to other electronic devices.

The simplest photodetectors are solar cells, diodes in which internal structures generate electromagnetic fields that separate the electrons and holes produced by incident light. This separation generates a voltage without requiring any input electrical power. Solar cells work well for producing electricity, but are too slow and inefficient for use as detectors in communications.

Photodiodes are much faster and more sensitive if electrically reverse-biased, as shown in Figure 11.4. (Recall that LEDs and lasers are forward-biased to emit light.) The reverse bias draws current-carrying electrons and holes out of the junction region, creating a depleted region, which stops current from passing through the diode. Photons with enough energy can create electron-hole pairs in this region by raising an electron from the valence band to the conduction band, leaving a hole behind. The bias voltage causes these current carriers to move quickly away from the junction region, so a current flows proportional to the light illuminating the detector. Several types of detectors can be used in fiber-optic systems, as described below.

Semiconductor photodiodes are reverse-biased to detect light; they produce a current proportional to the illumination level.

FIGURE 11.4
Photodetector operation.



Composition and structure combine to determine the operational characteristics of photodetectors. We will start by looking at detector materials, then consider common structures. As you will see, the actual structures used are more complex than the simple example of Figure 11.4.

Detector Materials

The wavelengths at which detectors respond depend on their composition.

Photodetectors can be made of silicon, germanium, gallium arsenide, indium gallium arsenide, or other semiconductors. The wavelengths at which they respond to light depend on their composition. To produce a photocurrent, photons must have enough energy to raise an electron from the valence band to the conduction band—that is, their energy must equal or exceed the bandgap energy. The need to have at least this minimum energy means that photodetector sensitivity tends to drop steeply at the long-wavelength, low-energy end of the spectrum. Other effects, such as light absorption in other parts of the device, cause the response to drop more gradually for more energetic photons at shorter wavelengths.

Table 11.1 Typical operating ranges of important detectors

Material	Wavelengths (nm)
Silicon	400–1100
Germanium	600–1600
GaAs	400–900
InGaAs	900–1700
InGaAsP	800–1600

Table 11.1 lists approximate spectral ranges for important detector materials used in fiber-optic systems. Note that two of the most important semiconductor materials, silicon and gallium arsenide, are not sensitive at the 1280- to 1650-nm wavelengths used in long-distance fiber-optic systems. The band gaps in InGaAs and InGaAsP depend on the material composition, so the range of operating wavelengths may vary between devices. For example, while most fiber-optic InGaAs detectors are sensitive at 900 to 1700 nm, some InGaAs detectors respond to wavelengths longer than 2200 nm.

The response of each material also varies with wavelength. Figure 11.5 shows the relative response of silicon, germanium, GaAs, and InGaAs across the range of wavelengths normally used in fiber-optic systems. Silicon and GaAs have good response at the 650-nm

Silicon and GaAs have good response at 650 and 850 nm; InGaAs is used at 1250 to 1700 nm.

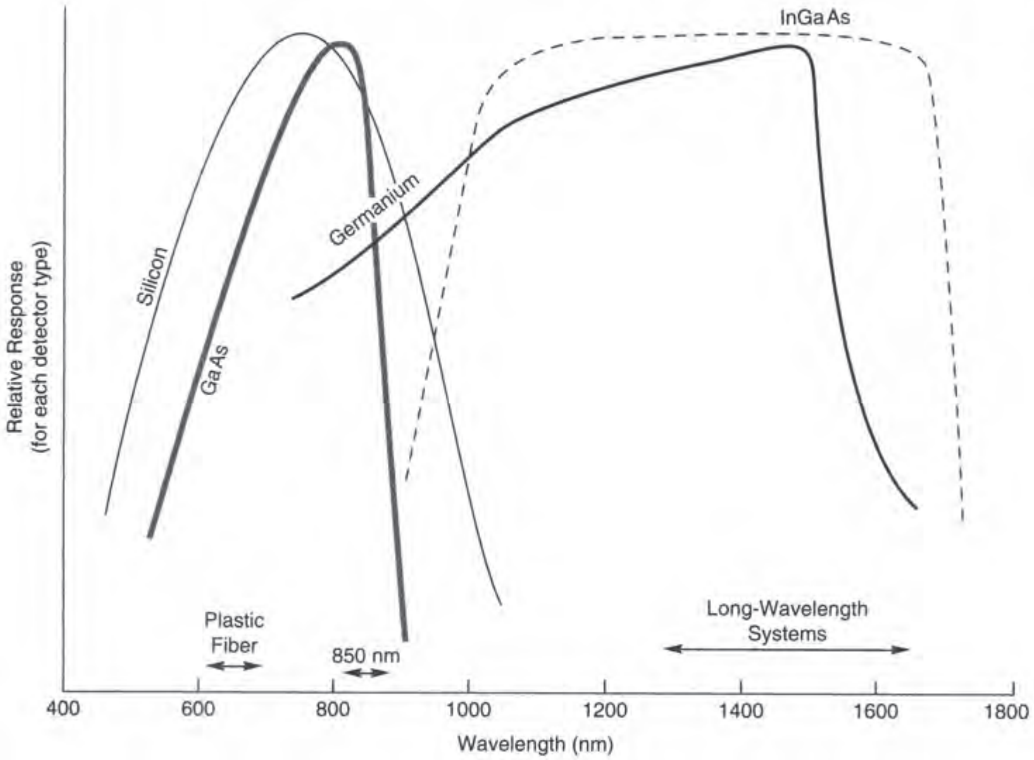


FIGURE 11.5

Detector response curves, showing quantum efficiency relative to the peak value for the material.

wavelength used in plastic fibers and the 850-nm wavelength used for short-distance transmission. Germanium has a very broad range, but response drops near 1550 nm and it suffers much higher noise levels than do other materials. InGaAs has response through the entire 1250 to 1700 nm range, making it the most common material for long-wavelength detectors. InGaAsP responds to similar wavelengths, depending on composition, but is not as widely used.

All these semiconductor materials also can be used in electronic circuits, allowing integration of detectors with amplifiers and other signal processing electronics in receivers. Detectors also can be combined with other electronic components to make hybrid receiver circuits.

pn and *pin* Photodiodes

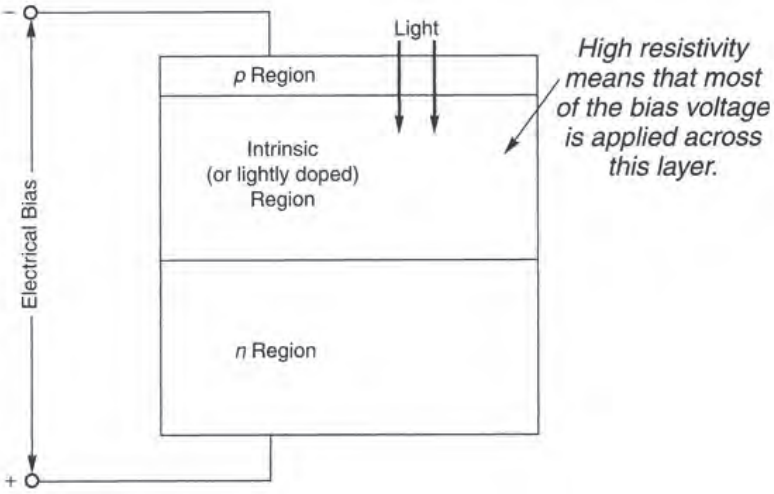
Photodiode sensitivity is improved by sandwiching an undoped intrinsic region between the *p* and *n* regions.

You saw earlier that photodetectors are reverse-biased diodes that generate current signals. They often are said to operate in a *photoconductive* mode because light changes the effective resistance of the device. However, they are not purely resistive devices because they contain *pn* junctions and function as diodes, which conduct current differently depending on the bias voltage. Purely resistive photoconductive detectors exist in which light produces current carriers that change the resistivity of a bulk semiconductor without a junction layer, but they are slow and not used as fiber-optic detectors.

Reverse biasing draws current carriers out of the central depleted region, blocking current flow unless light frees electrons and holes to carry current. The amount of current increases with the amount of light absorbed, and the light absorption increases with the thickness of the depleted region. Depletion need not rely entirely on the bias voltage. The same effect can be obtained if a lightly doped or undoped intrinsic semiconductor region is between the *p*- and *n*-doped regions shown in Figure 11.6. In a sense, such *pin* (*p*-intrinsic-*n*) photodiodes come predepleted because the intrinsic region lacks the impurities that can generate current carriers in the dark. This design has other practical advantages. By concentrating absorption in the intrinsic region, it avoids the noise and slow response that occur when the *p* region of ordinary *pn* photodiodes absorbs some light. The

FIGURE 11.6

A simple pin photodiode.



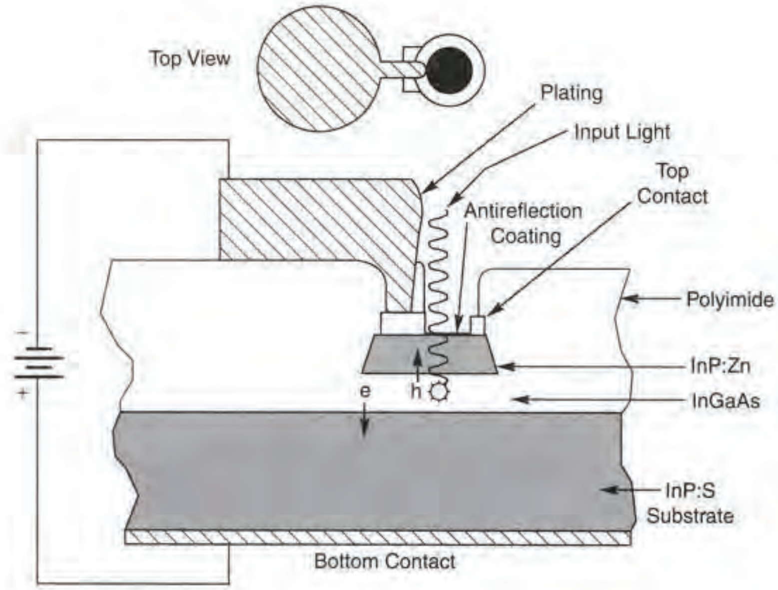


FIGURE 11.7

Structure of a multigigahertz pin *detector. (Copyright © 1993 Hewlett-Packard Company. Reproduced with permission.)*

bias voltage is concentrated across the intrinsic semiconducting region because it has higher resistivity than the rest of the device, helping raise speed and reduce noise.

The speed of *pin* photodiodes is limited by variations in the time it takes electrons to pass through the device. This time spread can be reduced in two ways—by increasing the bias voltage and/or by decreasing the thickness (and width) of the intrinsic layer. Reducing intrinsic layer thickness must be traded off against detector sensitivity because this reduces the fraction of the incident light absorbed. Typical biases are 3 to 20 V, although some devices have specified maximum bias above 100 V. Typical response times range from a few nanoseconds to about 5 ps. Sensitivity of *pin* detectors is measured as amperes of current generated per watt of light. For silicon, the peak sensitivity is about 0.7 A/W at 800 nm; InGaAs has a peak of around 1 A/W near 1600 nm. An important attraction of *pin* photodiodes is a large dynamic range; their output-current characteristics can be linear over 50 dB.

●

pin detectors can have response times well under 1 ns and dynamic ranges of 50 dB.

The speed and sensitivity of *pin* photodiodes are more than adequate for most fiber-optic applications, and they are widely used even in high-performance systems. Sending their electrical output directly to an electronic preamplifier can boost sensitivity.

●

pin photodiodes are widely used because of their high speed and good sensitivity.

The designs of actual *pin* photodiodes are more complex than this simple example, particularly in fast devices, like the multigigahertz detector in Figure 11.7. Light signals at 1200 to 1600 nm pass through the antireflection coating and upper layer of InP (which are transparent at those wavelengths) and are absorbed in the intrinsic InGaAs layer. Other designs direct light through the InP substrate, which is also transparent to the signal wavelengths.

Phototransistors

●

Phototransistors are used in low-cost, low-speed systems.

Some detectors have internal amplification, which increases their sensitivity. The simplest of these is a special type of transistor called a *phototransistor*, in which the incoming light

Table 11.2 Typical detector characteristics

Device	Responsivity	Rise Time	Dark Current
Phototransistor (Si)	18 A/W	2.5 μs	25 nA
Photodarlington (Si)	500 A/W	40 μs	100 nA
pin photodiode (Ge)	0.4 A/W	0.1–1 ns	100 nA
pin photodiode (Si)	0.5 A/W	0.1–5 ns	1–10 nA
pin photodiode (InGaAs)	0.8 A/W	0.005–5 ns	0.1–3 nA
Avalanche photodiode (Ge)	(voltage-dependent)	0.3–1 ns	400 nA (voltage-dependent)
Avalanche photodiode (Si)	10–125 A/W (voltage-dependent)	0.1–2 ns	10–250 nA (voltage-dependent)
Avalanche photodiode (InGaAs)	7–9 A/W (voltage-dependent)	0.1–0.5 ns	6–160 nA (voltage-dependent)

passes through the emitter region to illuminate the base, where it produces a photocurrent. The phototransistor amplifies the current generated in the base, just as an ordinary transistor amplifies its input base current.

The internal amplification stage gives the phototransistor much higher responsivity than a simple photodiode. However, this increase comes at a steep price in response time and linearity, and at some cost in noise, as you can see in Table 11.2. Most phototransistors are limited to speeds below the megahertz range. Commercial phototransistors normally are made of silicon. Their major uses are in sensors and in inexpensive fiber links.

Output of a phototransistor is fed to the base of a second transistor in a *photodarlington*. This is a simple, integrated darlington amplifier, which adds a second transistor to the phototransistor to increase the responsivity. This has the trade-offs of lowering speed and increasing noise. That constrains the photodarlington's uses even more narrowly than those of the phototransistor, but it offers higher responsivity for low-cost, low-speed applications.

Avalanche Photodiodes

Avalanche photodiodes are fast detectors with internal amplification.

The *avalanche photodiode* (APD) is another photodiode that provides internal amplification. It operates at much higher speeds, making it more useful for fiber-optic communications.

You can think of an APD as a two-stage device. The first stage is a conventional photodiode in which light generates current carriers. The second is an internal amplification stage based on avalanche multiplication in which a strong electric field accelerates the light-produced electrons so much that they can knock valence electrons out of the semiconductor lattice. At high voltage, the result is a near-avalanche of carriers—thus the name—that is still proportional to the amount of incident light. However, a further increase of the

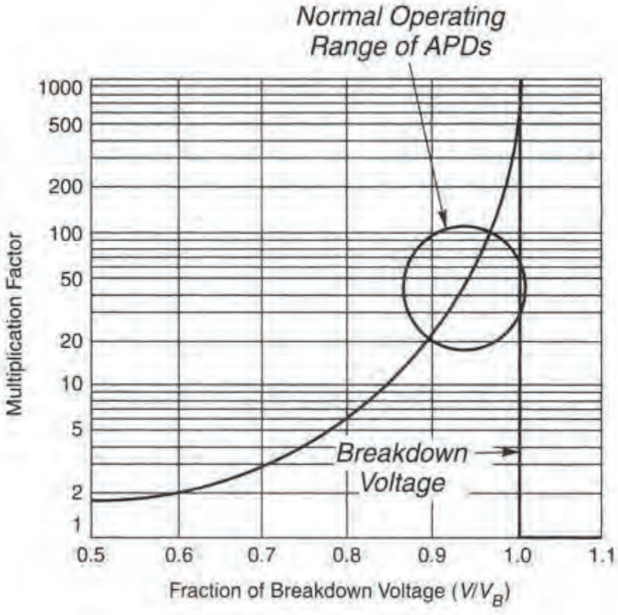


FIGURE 11.8
Increase of multiplication factor in APD.

voltage to a point called the *breakdown voltage,* V_B, causes current to flow freely through the semiconductor without regard to the amount of incident light; this can damage the device.

The factor by which each initial carrier is multiplied, called the *multiplication factor,* typically is 30 to 100. Multiplication factor M is defined as

$$M = \frac{1}{1 - \left(\frac{V}{V_B}\right)^n}$$

where V is the operating voltage, V_B is the breakdown voltage, and n is a number between 3 and 6, depending on the device characteristics. The multiplication factor can become very large as the operating voltage approaches the breakdown voltage. Care must be taken when increasing the operating voltage because reaching or exceeding the breakdown voltage can damage the device. A representative plot of multiplication factor as a fraction of breakdown voltage is shown in Figure 11.8. Note that multiplication factors of 100 require bias voltages within a few percent of breakdown. Typical APD operating voltages are 150 to 400 V in silicon but only 20 to 60 V in InGaAs, which has an inherently lower breakdown voltage.

The multiplication factor increases as bias voltage approaches the breakdown voltage.

APDs are fast, but the uneven nature of multiplication introduces noise. Avalanche gain is an average; not all photons are multiplied by the same factor. Signal power increases with roughly the square of the multiplication factor M, and at moderate values M increases faster than noise. However, as M reaches high levels, noise increases roughly as $M^{2.1}$, faster than M^2. As a result, APDs have an optimum multiplication factor to control noise; for silicon devices it is typically 30 to 100, but for InGaAs it is much lower.

Avalanche photodiodes require much higher operating voltages than the few volts normally used for *pin* photodiodes or other semiconductor electronics. The need for special circuits to provide this drive voltage, and to compensate for the temperature sensitivity of APD characteristics, makes APD receivers more complex than *pin* types. As Table 11.2 shows, APDs also suffer from higher dark currents and cannot match the rise times of the fastest *pin* detectors.

pin-FETs and Integrated Receivers

The distinction between detectors and the receivers that contain them sometimes can be hazy, especially when detectors are integrated on the same substrate with electronic circuits. One example is the *pin*-FET, which integrates a *pin* photodiode with an electronic amplifier circuit based on field-effect transistors. These devices are really *detector-amplifiers*, but they may be lumped together with detectors or simply called "detectors" themselves.

Performance Considerations

The factors that affect detector and receiver performance are complex and often interrelated. These considerations generally apply to both detectors and receivers, but some apply more to the detector. Remember that overall performance depends on the entire receiver system, and can be changed dramatically by replacing a detector or adding an optical preamplifier.

We can divide performance considerations into several broad categories. Four directly measure receiver performance: the strength of electrical signal generated in response to an optical input, internal noise levels, linearity, and the speed of the response. Signal coding and modulation format also play critical roles. Finally, the quality and power level of the optical input signal strongly influence performance. Let's look at each of these in turn.

Sensitivity

Sensitivity measures output signal produced for a given optical input.

Sensitivity measures the response to an optical input signal as a function of its intensity, in units such as amperes (of output signal intensity) per watt (of input light). Although sensitivity may sound like a simple concept, setting a precise definition can be tricky.

Detector sensitivity can be measured in two subtly different units. *Responsivity* is the ratio of electrical output from the detector to the input optical power. If the output current varies proportionally to the input, this is measured as amperes per watt (A/W). In practice, input powers usually are in the microwatt range, so responsivity is sometimes given as microamperes per microwatt ($\mu A/\mu W$), which is equivalent.

Quantum efficiency measures the fraction of incoming photons that generate electrons at the detector:

$$\text{Quantum efficiency} = \frac{\text{Electrons out}}{\text{Photons in}}$$

This sounds like the same thing as responsivity, but the two are not equivalent. Recall that the energy of a photon depends on the wavelength, so a 400-nm photon carries twice as much energy as an 800-nm photon. Suppose a detector generates one electron from every photon that reachs it at either wavelength. Then the input power at 400 nm would be twice the level at 800 nm, but the output current would be the same at both wavelengths. Thus responsivity—measured relative to power—at the shorter wavelength would be only half the level at the longer wavelength.

Both responsivity and quantum efficiency depend on the input wavelength, but their differences mean that the two curves are not interchangeable. The shape of the curve depends largely on the detector material, but the height also depends on the detector type and structure. Some curves, such as the one in Figure 11.5, show quantum efficiency at different wavelengths relative to the maximum value, rather than in absolute terms. You should always check the scales you're looking at.

Responsivity and quantum efficiency both depend on input wavelength and the detector material.

Typical values of responsivity from a *pin* detector alone are 0.4 to 1 A/W. InGaAs detectors typically have responsivity of 0.5 to 1 A/W in their long-wavelength range, and silicon has responsivity of 0.4 to 0.5 A/W at shorter wavelengths. (Note that according to the difference in wavelength, a silicon detector used at 800 nm should have about half the responsivity of an InGaAs detector used at 1600 nm.) Quantum efficiency of a detector with no internal amplification can approach 100%. Internal or external amplification can give much higher values by multiplying the number of electrons effectively generated by each incoming photon. Detector response also depends on operating temperature.

Receiver sensitivity is a different specification, the minimum optical input signal, usually in microwatts or dBm (decibels relative to one milliwatt) needed to operate at the required performance level. This quantity depends on the detector type as well as the receiver circuitry. For example, one receiver has sensitivity of −23 dBm with a *pin* detector, but −32 dBm with an avalanche photodiode as the detector. That reflects how internal amplification in the APD allows a receiver to process much weaker signals.

Dark Current and Noise-Equivalent Power

The electronic signal emerging from a detector includes noise as well as signal. The noise comes from many sources, including the optical input, the detector itself, and the amplification electronics. Electromagnetic interference can add noise by inducing currents in conductors in the receiver. The full complexities of noise are beyond the scope of this book, but you should understand two key concepts: dark current and noise-equivalent power.

An ideal detector generates an output signal that depends only on the input light, so in the dark it should produce no signal at all. Nature isn't that kind. Any detector generates some output current when it is operated in the normal manner but receives no light at all. This *dark current* measures the electrical noise inherent within the detector, which also is present when the detector is exposed to light. It's analogous to the low level of hiss you can hear during silent intervals on an analog tape cassette. It sets a floor on the minimum detectable signal, because a signal must produce more current than the dark current in order to be detected. Dark current depends on operating temperature, bias voltage, and the type of detector.

Dark current is the noise a detector generates in the dark.

Noise-equivalent power (NEP) is the input power needed to generate an electrical current equal to the root-mean-square noise from the detector (or receiver). This more directly measures the minimum detectable signal because it compares noise directly to optical power. NEP depends on the frequency of the modulated signal, the bandwidth over which noise is measured, the detector area, and the operating temperature. It's measured in the peculiar

Noise-equivalent power is the optical power needed to generate average detector noise.

units of watts divided by the square root of frequency (in hertz), or W/Hz$^{1/2}$. Specified values normally are measured with a 1-kHz modulation frequency and a 1-Hz bandwidth.

Speed and Bandwidth

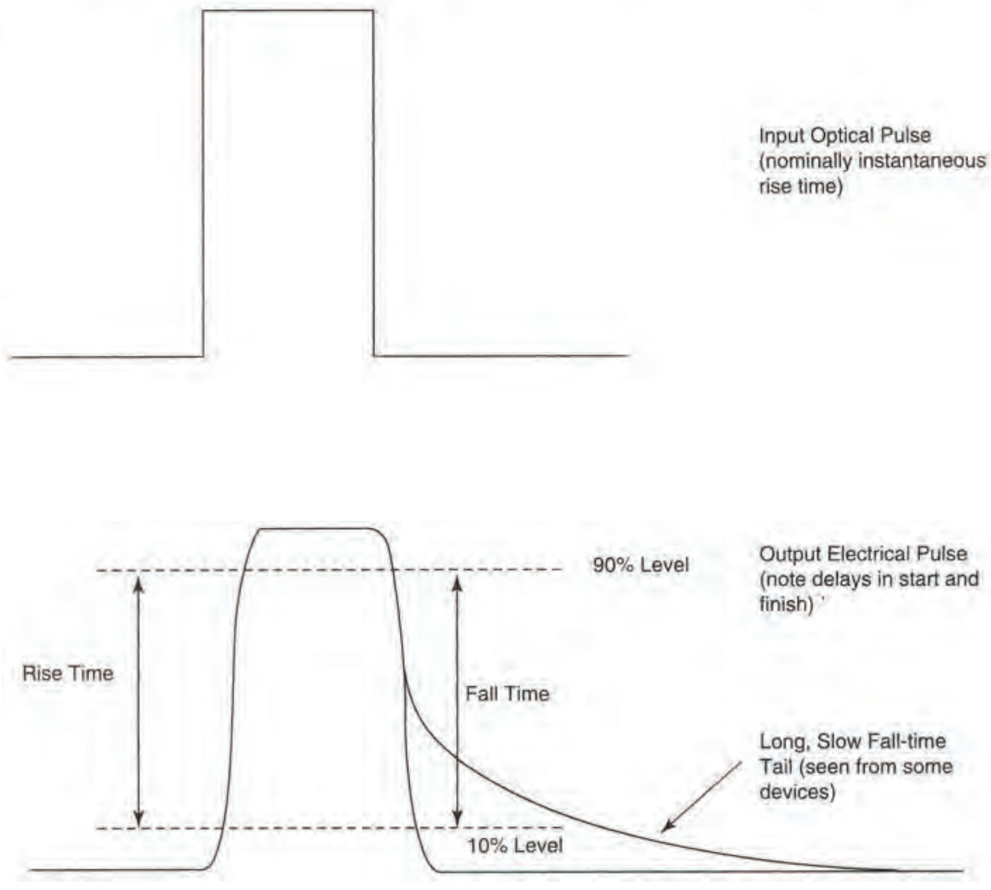
Detector bandwidth depends on rise time.

You can think of the speed of a detector in two ways. One is the time it takes to convert an input light signal into an output electronic signal. This measures how long the signal takes to pass through the detector, and has little impact on system operation because it does not change the shape of the pulses passing through. However, how fast the detector output responds to a change in the optical input changes the shape of the signal pulse, which does impact system performance. This can be measured as *rise time* or *bandwidth.*

Rise time is the time the output signal takes to rise from 10% to 90% of the final level after the input is turned on instantaneously, as shown in Figure 11.9. The *fall time* is how long the output takes to drop from 90% to 10% of the peak value after the input turns off abruptly. The two are not always symmetrical; some devices have long, slow-falling tails. Both are shown in Figure 11.9. Very slow fall times are undesirable because they can limit the detector's response speed more severely than can slow rise times.

Rise and fall times are a measure of bandwidth, the range of frequencies that a detector can reproduce in a signal. Detectors themselves can reproduce extremely low frequencies,

FIGURE 11.9
Rise time of a pulse.



although receiver electronics may limit low-frequency response. The most important limitation in fiber optics is the way detector response falls off at higher frequencies. If you study electronics, you will learn that the highest frequencies are responsible for the sharp edges on a square-wave function like the input digital pulse of Figure 11.9. Lose those high frequencies, and the edges of the signal pulse become rounded, and the pulse takes longer to rise and fall, as shown in the output of Figure 11.9.

Detector bandwidth usually is defined as the frequency at which the output signal has dropped to 3 dB (50%) below the power at a low frequency. This means that only half as much signal is getting through the detector at the higher frequency. Frequencies higher than the upper bound of the bandwidth are attenuated even more. As bandwidth decreases, the higher frequencies fade away, and the pulses become more rounded.

Loss of high frequencies rounds the edges of digital pulses.

If rise and fall times are equal, the 3-dB bandwidth can be estimated from the rise time using the formula

$$\text{Bandwidth} = \frac{0.35}{\text{Rise time}}$$

Thus a detector with 1-ns rise time has a 3-dB bandwidth of 350 MHz, while one with 10-ps rise time has a 3-dB bandwidth of 35 GHz. You can also flip the formula to estimate rise time if you know bandwidth:

$$\text{Rise time} = \frac{0.35}{\text{Bandwidth}}$$

Device geometry, material composition, electrical bias, and other factors all combine to determine bandwidth and rise time. For relatively slow devices, the rise time is proportional to the RC time constant—the photodiode capacitance multiplied by the sum of the load resistance and the diode's internal resistance. Reducing the equivalent capacitance can increase speeds. At higher speeds, two other factors limit rise time, the diffusion of current carriers and the time needed for carriers to cross the depletion region.

Linearity of Response

In an ideal detector, the output current would be a constant multiplied by the input power. This is called a *linear response*, and in practice it is available only over a limited range of input powers, called the *dynamic range*. Once the input exceeds that range, the output does not increase as rapidly, and the signal becomes distorted.

A detector has a linear response over a limited dynamic range.

As you can see in Figure 11.10, an increase of Δp in the input light produces an increase of Δi in the output current for a receiver or a detector. In the lower portion of the curve, where response is linear, $\Delta i/\Delta p$ is a constant. Once the power exceeds the detector's dynamic range, each additional photon no longer produces as many electrons in the output current, so the device starts to saturate. This happens every time the signal reaches this level, which could be every pulse in a digital system.

The result is distortion, much as when an audio speaker is driven with more power than it can handle. Exceeding the dynamic range in a digital receiver increases the bit error rate. In either case, inserting attenuators in front of the detector can reduce average signal intensity to a level within the dynamic range so the detector responds linearly.

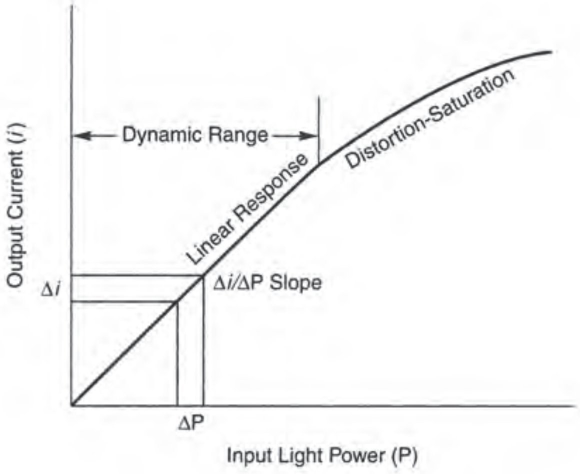


FIGURE 11.10
Receiver output as a function of input light power. Both the detector and the receiver electronics affect dynamic range.

Receiver circuits also impact dynamic range, so you should consider their properties as well as those of the detector.

Detectors also need to receive a minimum level of input power, so the signal levels will exceed the background noise or dark current mentioned earlier.

Signal Coding and Modulation

Signal formats also affect detector and receiver performance. Analog modulation looks simple, because it only requires replicating the input signal. It also uses bandwidth efficiently. However, perfectly replicating the signal is difficult because detectors are not perfectly linear. Other complications include noise from the signal and the detector itself, and signal intensity. Figure 11.11 shows how distortion can affect analog signals.

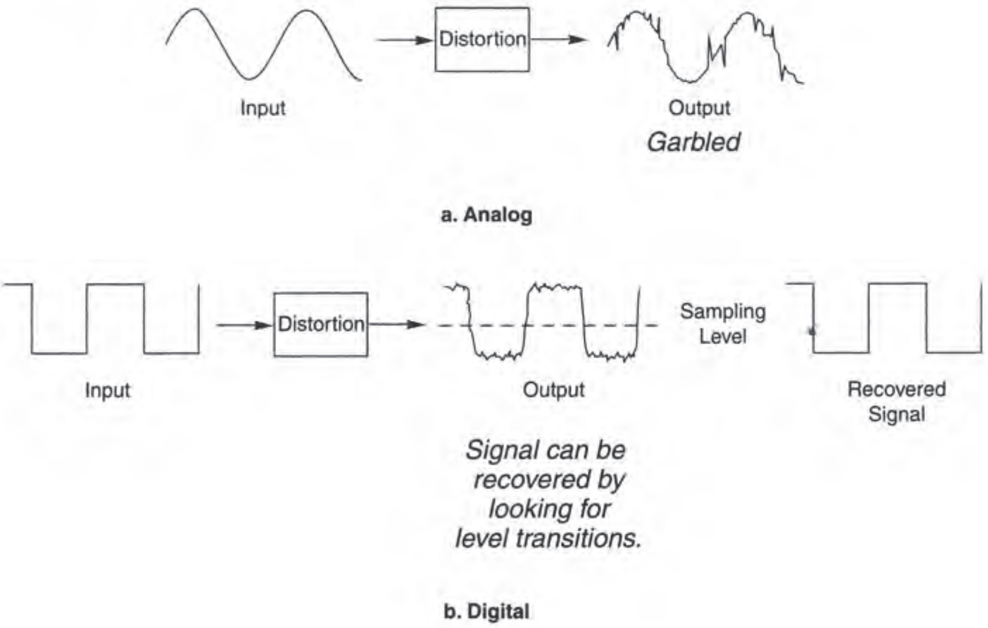


FIGURE 11.11
Effects of distortion on (a) analog and (b) digital signals.

Digital signals can withstand the effects of noise and distortion better than analog signals. Conventional digital signals are binary codes in which the light is either "on" or "off"—corresponding to digital 1s and 0s. To interpret this type of input, a digital receiver needs only to tell whether the signal is "off" or "on," not to reproduce the input signal level. Digital receivers do this by deciding if pulses exceed a minimum discrimination level, as shown in Figure 11.11. Pulses that cross this threshold level are considered "on"; those that don't are considered "off." The decision is done by circuits in the receiver electronics.

Digital signals are more robust than analog signals, and can be extracted from noise.

Figure 11.3 shows that a digital receiver actually consists of an analog receiver followed by a set of circuits that clean up the digital output. First the circuits decide at what point the power crosses the threshold of being "on" or "off." Then they adjust the timing of the pulse to fit it into the right time interval. Sometimes called *regeneration* or *pulse recovery*, this process produces a fresh new signal for the next stage of transmission.

The analog receiver has to reproduce an analog input signal very accurately or the output will be distorted, as shown in Figure 11.11(a). The analog stages of the digital receiver don't have to be that good. A digital receiver is like a sound monitor that only needs to be sensitive enough to tell if someone is talking in a room. An analog receiver is like a sound monitor that lets you understand conversations in the room. If you've ever tried to understand the announcements in an airport, rail station, or bus terminal, you know there's a big difference between sound and intelligible speech. In a sense, a digital receiver is really a cheap and dirty analog receiver with a digital back end.

Analog receivers must reproduce signals much more accurately than digital receivers.

It's worth noting that the "off" state is not completely off in many fiber-optic systems, because many transmitters generate a very weak optical signal for 0 and a much stronger signal for a 1. That would pose a problem if the digital signal had to be reproduced precisely. However, decision circuits can easily tell if binary signals are in either an "on" or "off" state. In that way, noise can be removed from digital signals, something that is impossible with analog signals.

Digital transmission does not have to be binary. Some digital systems may have a series of steps in energy level, so the pulse may have, for example, four rather than two possible levels. That can carry more information, but requires telling the difference between four rather than two energy levels. Such systems have been demonstrated in the laboratory, but are not in practical use.

Signal Quality and Power

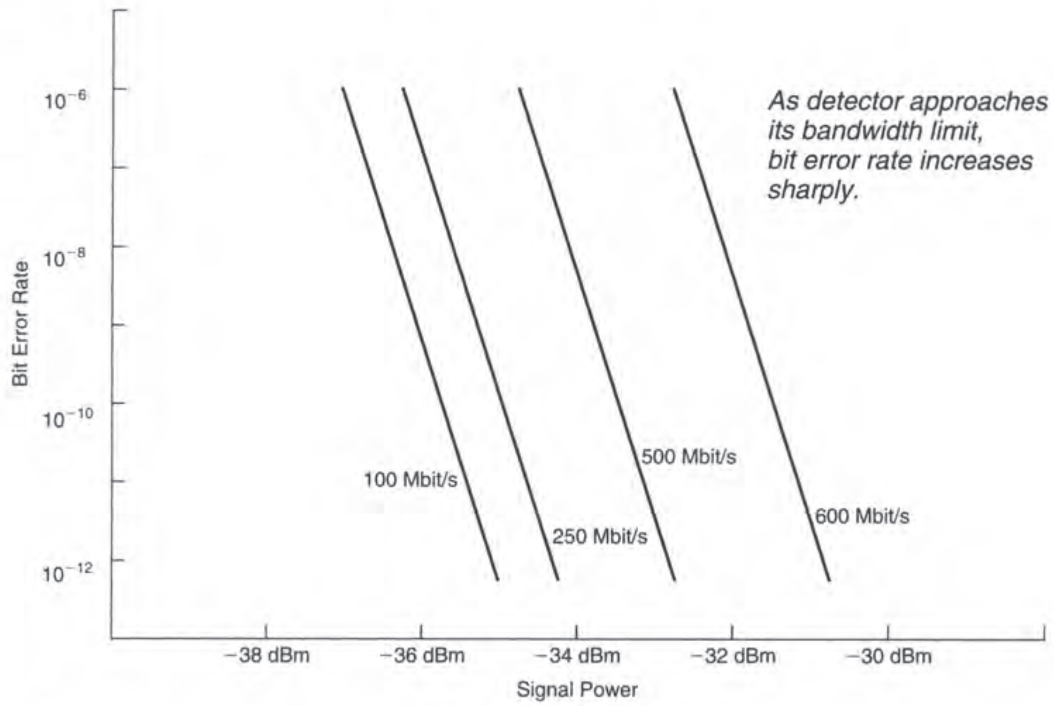
Power and quality of the input optical signal are important in detector and receiver response. The input consists of two components—the signal and the noise. The larger the difference between them, the better the signal quality. Usually the difference between the signal and the noise matters more than the absolute values of either. The ringing of someone else's cell phone is annoying if you're listening to a quiet passage of classical music, but you'd never hear the same ring if a heavy-metal band was playing full blast.

Signal-to-noise ratio measures analog signal quality.

Analog signal quality is measured as *signal-to-noise* ratio, usually in decibels. Normally average powers are used for both signal and noise. A 30-dB signal-to-noise ratio (or S/N ratio) means the signal amplitude is 1000 times higher than that of the noise. Whether or not that is acceptable depends on the situation, the type of signal, and who's measuring or listening. Many systems specify a minimum acceptable signal-to-noise ratio—for example, at certain points in a cable television network.

FIGURE 11.12

Bit error rate increases as average power drops and data rate increases.



Bit error rate measures quality of digital signals.

Digital signal quality normally is measured as *bit error rate,* the fraction of bits received incorrectly. Specified values depend on the application. For telephone and digital data transmission, the target normally is 10^{-12}, or no more than one incorrect bit in every trillion. Other applications may allow more errors.

The error rate depends on the number of photons received in a pulse. As the number of photons decreases, the chance of error increases. As you would expect, this means that the error rate increases as power decreases. This effect can be quite dramatic in certain power ranges. The plots in Figure 11.12 show how a 1-dB decrease in received power causes the bit error rate to jump from 10^{-12} to 10^{-9}.

Photons received per bit decreases for shorter pulses.

Increases in signal speed also can decrease the number of photons received per bit because shorter pulses have less time to deliver photons. Signal power measures the number of photons delivered per unit time, so the shorter the pulse, the fewer photons delivered during it. Suppose, for example, your laser transmitter delivers an average power of one microwatt at a wavelength of 1 μm, which equals roughly 5×10^{12} (five trillion) photons a second. At a data rate of 1 Gbit/s, that comes to roughly 5000 photons per bit interval. Raise the data rate to 10 Gbit/s and you have an average of only 500 photons per bit.

That's an average, and since there are both "on" and "off" bits, it might sound pretty easy to tell the bits apart: No photons should arrive during an "off" bit, so 1000 photons should arrive during an "on" bit. However, detectors don't sense every photon; they sense noise during the "off" pulses, and the goal is to have very low error rates, in the range of 10^{-12}—one error in a trillion bits. The more the power is reduced, the more likely an error is to occur. You can see this effect in Figure 11.12, where a small increase in data rate or power causes a large increase in error rate. The effect is strongest as the detector approaches its bandwidth limit.

Electronic Functions

Converting an optical signal into electrical form is only the first part of a receiver's job. The raw electrical signal generally requires some further processing before it can serve as input to a terminal device at the receiver end. Typically, photodiode signals are weak current signals that require amplification and conversion to voltage. In addition, they may require such cleaning up as squaring off digital pulses, regenerating clock signals for digital transmission, or filtering out noise introduced in transmission. The major electronic functions are as follows:

Detector output must be processed before other equipment can use the electronic signal.

1. Preamplification
2. Amplification
3. Equalization
4. Filtering
5. Discrimination or decision
6. Timing

If you're familiar with audio or other electronics, you will recognize some of these functions. Not all are required in every receiver, and even some of those included may not be performed by separate, identifiable devices. A phototransistor, for example, both detects and amplifies. And many moderate-performance digital systems don't need special timing circuits. Nonetheless, each of these functions may appear on block diagrams such as in Figure 11.3. Their operation is described briefly below.

Preamplification and Amplification

Typical optical signals reaching a fiber-optic receiver are 1 to 10 μW and sometimes lower. If a *pin* photodiode with 0.6 to 0.8 A/W responsivity detects such signals, its output current is in the microampere range and must be amplified for most uses. In addition, most electronics require input signals as voltage, not current. Thus, detector output must be amplified and converted.

Microampere-level *pin* detector outputs must be converted to logic-level voltages.

Receivers may include one or more amplification stages. Often the first is called preamplification because it is a special low-noise amplifier designed for weak input. (An optical amplifier placed in front of a detector also may be called a preamplifier.) In some cases, as mentioned earlier, the preamplifier may be packaged with the detector. The preamplifier output often goes into an amplifier, much as the output of a tape-deck or CD preamplifier goes to an audio amplifier that can produce the power needed to drive speakers.

Equalization

Detection and amplification can distort the received signal. For example, high and low frequencies may not be amplified by the same factor. The equalization circuit evens out these differences, so the amplified signal is closer to the original. Much the same is done in analog high-fidelity equipment, where standard equalization circuits process signals from tape heads and phonograph cartridges so they more accurately represent the original music.

Filtering blocks noise while transmitting the signal.

Filtering

Filtering helps increase the S/N ratio by selectively attenuating noise. This can be important when noise is at particular frequencies (e.g., a high-frequency hiss on analog audio tapes). It is most likely to be used in fiber optics to remove undesired frequencies close to the desired signal, such as harmonics.

Discrimination circuits generate digital pulses from an analog input.

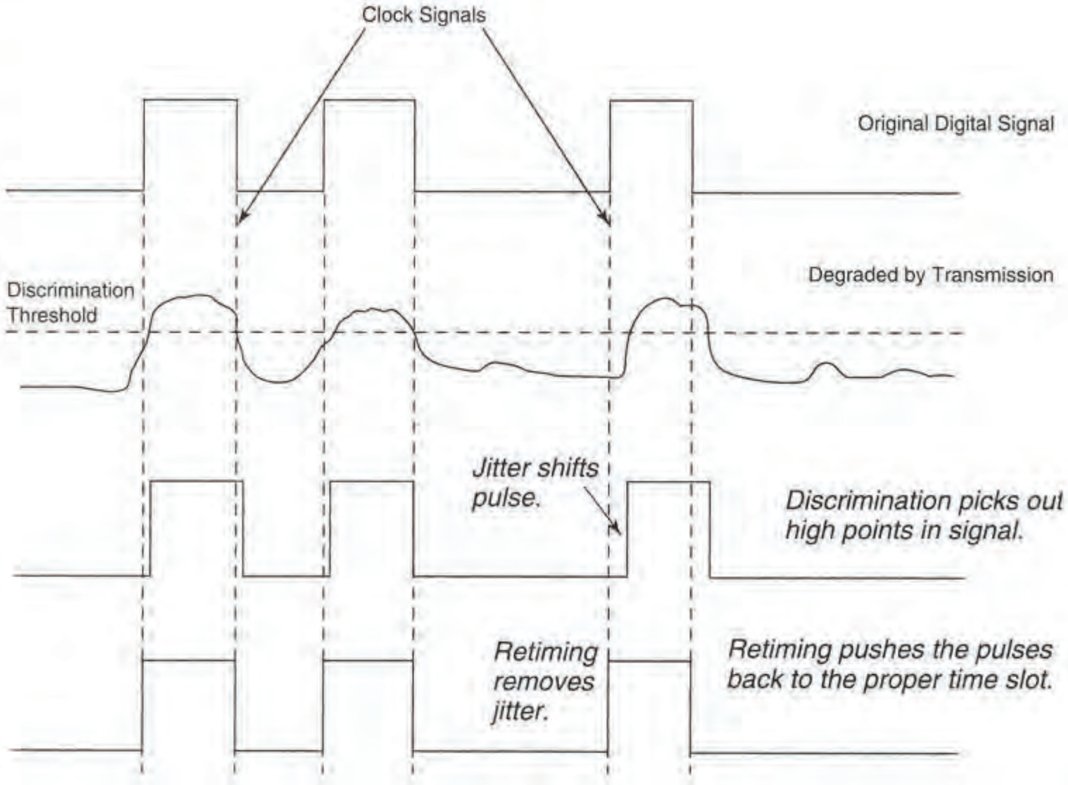
Discrimination

Pulse regeneration requires circuitry to decide if the input is on or off.

So far, the functions you've looked at are needed to reproduce the original waveform for both analog and digital receivers. However, a further stage is needed to turn a received analog signal back into a series of digital pulses—decoding and discrimination. Rectangular pulses that started with sharp turn-on and turn-off edges have been degraded into unboxy humps, as shown in Figure 11.13. Dispersion may have blurred the boundaries between pulses.

As mentioned earlier, this rounding of square pulses represents loss of high frequencies, which make up the sharp rising and falling edges of the pulse, so if they are lost, the pulses lose their square edges. The remaining low-frequency components contain most of the

FIGURE 11.13
Discrimination and retiming regenerate digital pulses.



information needed, but they are not clean enough to serve as input to other electronic devices. Regeneration of clean pulses requires circuitry that decides whether or not the input is in the on or off state by comparing it to an intermediate threshold level. The decision circuit generates an "on" pulse if the power is above the threshold; otherwise it produces an "off" signal. Care must be taken in selecting this threshold level to avoid misinterpreting input; too low a threshold, for example, could turn noise spikes in the off state into signal pulses.

Timing

Another essential task in many receivers, particularly in high-performance systems, is resynchronizing the signal. Digital signals are generated at a characteristic clock rate, such as once every nanosecond for a 1-Gbit/s data stream. You can see in Figure 11.13 that discrimination circuits do not necessarily spot the exact times the pulses start and end. These random errors, called *jitter*, can cause the signal to drift from the clock rate, introducing errors.

Timing of digital pulses often must be resynchronized.

Timing synchronization recreates the clock signal (the dashed vertical lines in Figure 11.13) and puts the regenerated pulses in the right time slots. It is an essential part of cleaning up signals at a sophisticated receiver.

Packaging Considerations

As with transmitters, packaging is important for receivers. The basic requirements are electronic, mechanical, and optical interfaces that are simple and easy to use. The main mechanical issues are mounts. Electronic interfaces must allow for input of bias voltage and amplifier power (where needed) and for output of signals in the required format. Details can vary significantly.

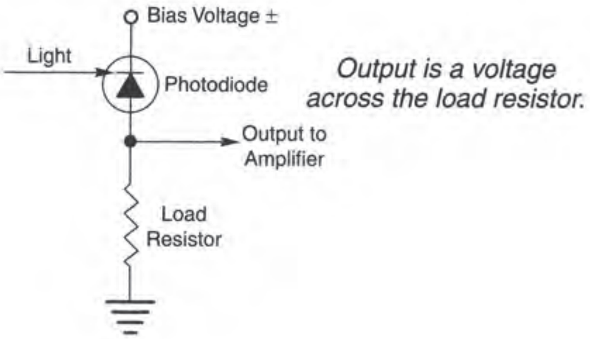
Optical interface requirements are simpler than for transmitters because mechanical tolerances for aligning fibers with detectors generally are looser. The active areas of detectors are larger than the cores of single-mode fibers. Larger-core multimode fibers transmit signals at slower speeds, so they normally are used only with slower detectors with larger active areas. In practice, receivers are assembled with integral fiber pigtails or connectors that collect light from the input fiber and deliver it to the detector.

In general, packaged receivers look very much like transmitters, and often the two are packaged together as a transceiver. You may have to read the labels to tell them apart. Detector modules are packaged inside receivers just as light-source modules are put inside transmitters. Internal design constraints become increasingly severe at high frequencies because of the problems inherent in high-frequency electronic transmission.

Sample Receiver Circuits

Details of receiver circuitry vary widely with the type of detector used and with the purpose of the receiver. For purposes of this book, I will show only a few simple circuits for important devices and avoid detailed circuit diagrams.

FIGURE 11.14
Basic circuit for photoconductive pin *or* pn *photodiode.*



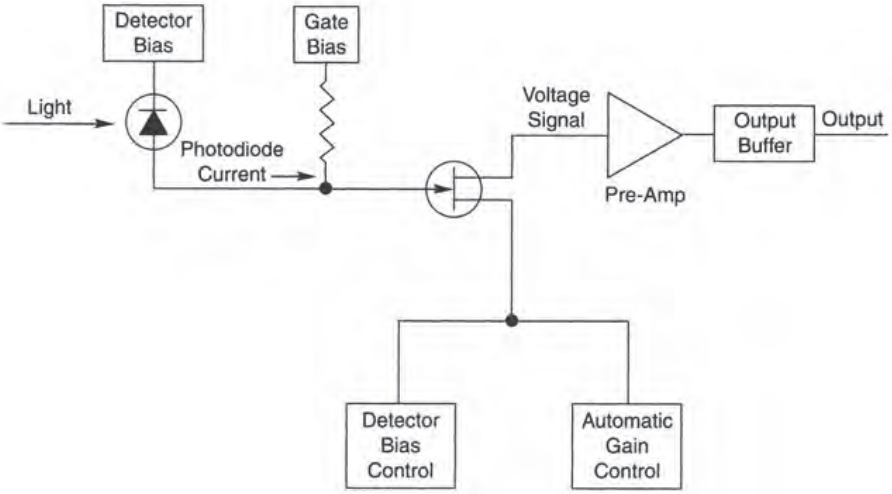
Photoconductive Photodiodes

Photoconductive photodiodes have a load resistor in series with the bias voltage.

The typical *pin* or *pn* photodiode used in a fiber-optic receiver is used in a circuit with a reverse-bias voltage applied across the photodiode and a series load resistor, such as that shown in Figure 11.14. In this mode, the photodiode is photoconductive because the photocurrent flowing is proportional to the nominal resistance of the illuminated photodiode. This simple circuit converts the photocurrent signal from a photodiode into a voltage signal.

The division of the bias voltage between the photodiode and the fixed resistor depends on illumination level. The higher the illumination of the photodiode, the more current it will conduct and, thus, the larger the voltage drop across the load resistor. In the simple circuit shown, the signal voltage is the drop across the load resistor. Most circuits are more complex, with amplification stages beyond the load resistor, as in *pin*-FET and detector-preamplifier circuits. Figure 11.15 is a block diagram of one circuit that includes automatic gain control, which can increase dynamic range by turning down the amplification factor before any other components are overloaded.

FIGURE 11.15
Block diagram of pin-FET receiver circuit.



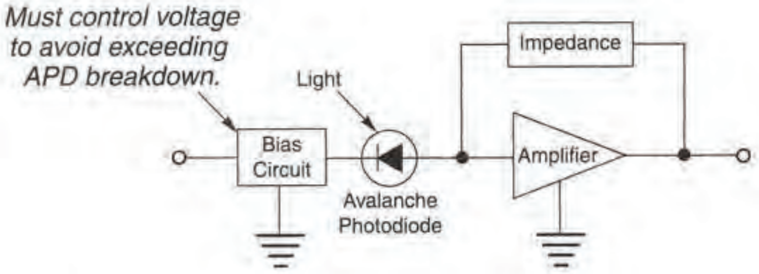


FIGURE 11.16 *Basic receiver circuit for avalanche photodiode.*

Avalanche Photodiode Circuits

The circuits used for avalanche photodiodes are conceptually similar to those used for photoconductive *pin* photodiodes. However, because of the high bias voltages required and the sensitivity of the photodiode to bias voltage, care must be taken to assure stable bias voltage. This adds to circuit complexity, as shown in the block diagram of Figure 11.16.

10-Gbit/s Receiver Module

Designs for 10-Gbit/s receivers are highly modularized, as shown in the block diagram of Figure 11.17. Optical input is directed to a *pin* photodiode, then amplified in a preamplifier before being transmitted (in electronic form) to an automatic gain-control circuit. Outputs from that circuit go to a clock and regeneration circuit and to a separate circuit that detects service outages and triggers external alarms. The clock and regeneration chip interfaces with a voltage-controlled oscillator and delivers two output signals: one containing the data, the other a clock signal. That clock can be used by electronic demultiplexer circuits (not shown here) to step the electronic data rate to slower speeds.

The receiver module shown in Figure 11.17 is a simple one that converts an optical input at 10 Gbit/s into a raw electronic signal at 10 Gbit/s. It does not include such electronic functions as forward error correction or demultiplexing to lower speeds for signal distribution. Such modules can be combined with transmitters to make the transceivers you learned about in Chapter 10.

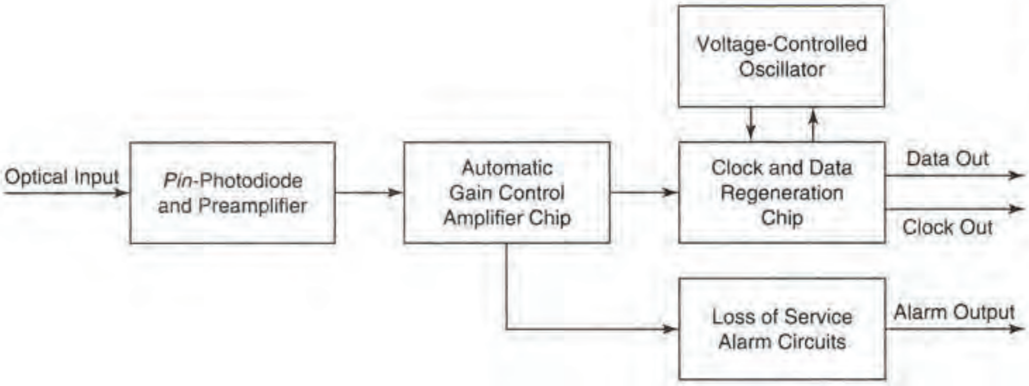


FIGURE 11.17 *Elements of a 10-Gbit/s receiver.*

What Have You Learned?

1. A receiver detects optical signals and converts them to electronic form.
2. A digital receiver includes an optical detector, an electronic amplifier, and circuits for retiming and regenerating the original digital signal. An analog receiver is similar, but lacks the retiming and regeneration circuits.
3. Wavelength-division demultiplexing must be performed before the optical detector because the detector is color-blind. One detector is needed for each optical channel.
4. Electronic post-processing corrects errors and decodes the sequence of incoming bits after the receiver module converts them to electronic form.
5. When a photon of sufficient energy strikes a semiconductor detector, it excites an electron from the valence band into the conduction band, producing an electron and a hole. This produces a current that flows through a reverse-biased semiconductor photodiode. Electronic circuits in the receiver convert this current signal into a voltage signal.
6. Detectors respond to photons that have enough energy to raise electrons from the valence band to the conduction band. The wavelengths at which they respond depend on composition. Silicon responds at 400 to 1000 nm. InGaAs responds at 800 to 1700 nm.
7. A high-resistivity intrinsic layer between the *p* and *n* layers of a *pin* photodiode absorbs light and improves its sensitivity. *pin* photodiodes are widely used for their high speed and sensitivity.
8. A high bias current creates an internal cascade of electrons in an avalanche photodiode, multiplying its electrical output to higher levels than those produced by *pin* photodiodes. However, APDs are slower than *pin* photodiodes and have higher noise.
9. Detector sensitivity is the output signal produced for a given input power. Responsivity is the current per watt of optical input. Quantum efficiency is the fraction of incoming photons that generate electrons.
10. Dark current and noise-equivalent power measure noise in detectors.
11. The bandwidth of a detector depends on its rise time in response to a signal. Loss of high frequencies reduces rise time and rounds sharp-cornered signal pulses.
12. Detectors operate best over a limited dynamic range, where their output depends linearly on the input signal.
13. Digital signals are more robust than analog signals because receivers do not have to reproduce digital signals as accurately.
14. Signal-to-noise ratio measures the quality of analog signals. Bit error rate measures the quality of digital signals.
15. A receiver's ability to detect pulses depends on the number of photons in the pulse.

What's Next?

In Chapter 12, I move on to optical amplifiers and to electro-optic repeaters and regenerators, which combine the functions of receivers and transmitters.

Further Reading

S. R. Forrest, "Optical detectors for lightwave communication," pp. 569–599 in Stewart E. Miller and Ivan P. Kaminow, eds., *Optical Fiber Telecommunications II* (Academic Press, 1988)

David A. Johnson, *Handbook of Optical Through the Air Communications* (*http://www.imagineeringezine.com/ttaoc/detector.html*)

Gerd Keiser, *Optical Fiber Communications* (McGraw-Hill, 2000), See Chapter 6 "Photodetectors" and Chapter 7 "Optical Receiver Operation".

Jim Rue and Bouchiab Nessar, "High speed avalanche photodiode optical receivers," *Fiberoptic Product News* (November 1999)

Questions to Think About

1. The input signal at a receiver is -30 dBm (1 μW), which is too low for your *pin* photodiode detector. What alternatives are there to increase receiver sensitivity?
2. An input signal is -30 dBm (1 μW). If its data rate is 1 Gbit/s, how much energy does each pulse contain, remembering that a power of one watt equals one joule per second? If the signal is at 1.5 μm, how many photons does that correspond to, using the following equation?

$$E = \frac{1.989 \times 10^{-19}(\text{joules}/\mu\text{m})}{\lambda(\mu\text{m})}$$

3. If a detector has a response of 1 μA/μW and the input is -30 dBm, what is the output current? How many electrons per second does this correspond to, recalling that $1 \text{ A} = 6.24 \times 10^{18}$ electron charges per second? How many electrons would be contained in a 1-nanosecond pulse?
4. A silicon *pin* detector has peak sensitivity of 0.7 A/W at 800 nm. If the input signal is -20 dBm, how many electrons does a 1-ns pulse produce? Use the equations and conversion factors from Questions 2 and 3.
5. The rise time of an InGaAs *pin* photodiode is 0.005 ns. The rise time of an InGaAs avalanche photodiode at the same wavelength is 0.1 ns. All other things being equal, how much larger is the bandwidth of the *pin* photodiode? If the rise time is all that limits the bandwidth of the two devices, what are their bandwidths?

6. The breakdown voltage of a silicon APD is 100 V. What voltage should it be operated at to have an electron multiplication factor of 20?
7. The dark current in a germanium *pin* photodiode is 100 nA, compared to 1 nA in an InGaAs *pin* photodiode. Suppose that all other things are equal, and the germanium detector has a sensitivity of −15 dBm, limited by dark current. What would be the sensitivity of the InGaAs detector?

Chapter Quiz

1. How many separate receivers are required for a 16-channel wavelength-division multiplexed system with an optical preamplifier that provides 15-dB gain on all channels?
   a. 1
   b. 4
   c. 8
   d. 16
   e. 32
2. When would an optical preamplifier be used at the receiver end of a system?
   a. always
   b. when input power is below 1 mW
   c. when the input signal contains multiple optical channels
   d. when input power is below receiver sensitivity
   e. never
3. What is present in a digital receiver that is not present in an analog receiver?
   a. nothing
   b. a detector
   c. thresholding and retiming circuits
   d. amplification circuits
   e. wavelength-division multiplexing
4. Photodiodes used as fiber-optic detectors normally are
   a. reverse-biased.
   b. thermoelectrically cooled.
   c. forward-biased.
   d. unbiased to generate a voltage like a solar cell.
   e. none of the above
5. Silicon detectors are usable at wavelengths of
   a. 800 to 900 nm.
   b. 1300 nm.

c. 1550 nm.

d. all of the above

6. Which detector material is most often used in the 1550-nm window?

a. silicon

b. InGaAs

c. GaAs

d. germanium

e. all of the above

7. A *pin* photodiode is a

a. point-contact diode detector in which a pin makes contact with the semiconductor.

b. semiconductor detector with an undoped intrinsic region between *p* and *n* materials.

c. circuit element used in receiver amplification.

d. photovoltaic detector.

e. hybrid detector-amplifier.

8. A phototransistor

a. has an internal amplification stage based on avalanche multiplication of electrons.

b. has an external amplification stage containing a single transistor.

c. generates a photocurrent in the base of a transistor, which amplifies the signal.

d. is an ordinary transistor that generates an optical signal under bright lights.

e. is the same as a photodarlington.

9. An avalanche photodiode

a. has an internal amplification stage based on avalanche multiplication of electrons.

b. has an external amplification stage containing a single transistor.

c. generates a photocurrent in the base of a transistor, which amplifies the signal.

d. is an ordinary transistor that generates an optical signal under bright lights.

e. is the same as a photodarlington.

10. What type of photodetector could have a responsivity of 20 amperes per watt?

a. silicon *pin* photodiode

b. silicon avalanche photodiode

c. InGaAs *pin* photodiode

d. InGaAs avalanche photodiode

e. none

11. What bit error rate is most often specified for digital telecommunications systems?

a. 40 dB
b. 10^{-4}
c. 10^{-6}
d. 10^{-12}
e. 10^{-18}

12. Noise equivalent power is

a. optical input power required to generate a signal equal to the noise.
b. noise required to equal the signal intensity.
c. the power of the current generated when a detector is in the dark.
d. noise present in the electrical output signal.

13. What happens when you increase the bias voltage above the breakdown voltage in an avalanche photodiode?

a. You stabilize the output current.
b. You stop the avalanche of electrons produced inside the semiconductor.
c. You get in trouble because that can damage the APD.
d. You increase signal-to-noise ratio to infinity.

14. What's the maximum value of quantum efficiency possible in a *pin* photodiode?

a. 0
b. 0.5
c. 0.9
d. 1.0
e. 100

CHAPTER

12

Amplification, Regeneration, and Wavelength Conversion

About This Chapter

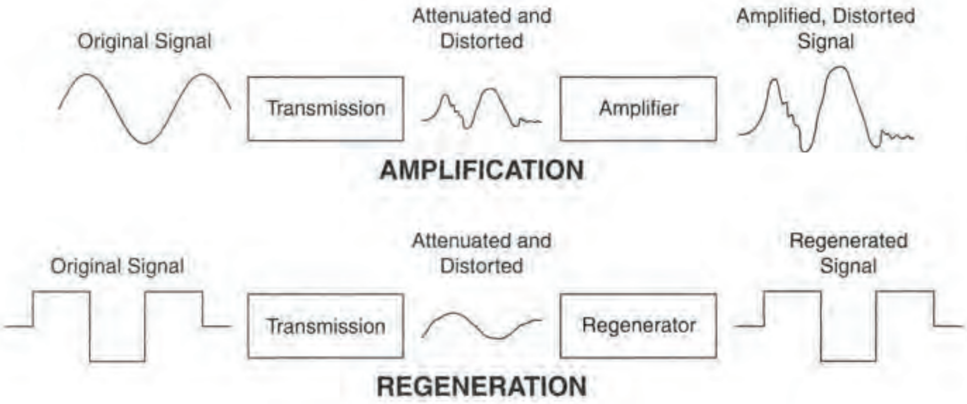
The previous three chapters have shown how transmitters and receivers send and receive signals in a fiber-optic system. Sometimes signals need a boost or special processing somewhere along the line to reach the final receiver in intelligible form. Optical amplifiers, repeaters, and regenerators give them that needed boost. Signals also may need to be converted from one wavelength to another as they are transmitted through a system.

In this chapter, you will learn about the differences among repeaters, regenerators, optical amplifiers, and wavelength converters; the functions they perform; and the technologies they use. Repeaters and regenerators are electro-optic devices, closely related to transmitters and receivers because they convert optical signals to electronic form. Optical amplifiers are purely optical devices, related to the lasers described in Chapter 9. Optical regeneration also is possible, but not widely used. Wavelength conversion is a separate function that may be performed optically or electro-optically.

Amplification and Regeneration

Signals suffer noise, distortion, and attenuation when traveling through any transmission medium. The further the signal goes, the more noise accumulates, the more the signal is distorted, and the more the signal fades in strength. As you learned in Chapter 11,

FIGURE 12.1
Amplification and regeneration.



receivers can regenerate the original signal, but only if the output signal is reasonably strong and of good quality. If a digital signal fades below a certain level, the bit error rate rises rapidly. If an analog signal fades too far, it's lost in the noise.

Amplification increases signal amplitude but does not reshape the signal.

Communication systems can avoid this problem by amplifying or regenerating the signals along the way. The two processes are different, as you can see in Figure 12.1. *Amplification* multiplies the amplitude of a weak input signal but doesn't clean it up or reshape it. It's like turning up the volume if you're listening to a distant AM radio station or a cassette tape that was recorded too faintly. The signal is loud, but it's still distorted. In practice, amplification itself can add some noise and distortion. *Regeneration* processes a weak input signal to regenerate the original input, like the digital receiver described in Chapter 11. The output is a clean signal. Regeneration also amplifies the signal to the proper strength. The operation of a regenerator depends on the signal being digital. (It's very hard to remove noise and distortion from analog signals.)

Regeneration reproduces the original signal, removing noise and distortion.

Each amplifier can increase signal strength by only a certain amount. If more amplification is needed to span a long distance, multiple amplifiers can be spaced along the length of the system to amplify the signal repeatedly. Each time the signal strength drops a certain amount, another amplifier is added. The amplification must be done before noise or distortion start to overwhelm the signal. Once the quality of a digital signal drops too much, regeneration is needed as well as amplification.

Optical and Electronic Transmission

The principles of amplification and regeneration are the same for optical and electronic signals (although optical signals can go much longer distances through fibers than electronic signals can go through copper cables). Technical details and terminology do differ, however.

Noise and distortion as well as attenuation generally are lower in optical fibers than in copper cables, so amplifiers are spaced much farther apart in optical systems. In practice, an optical signal may pass through several amplifiers before it must be regenerated. For example, an optical signal may pass through four optical amplifiers 80 kilometers apart before requiring regeneration. If more amplifiers are spaced closer together, noise and distortion can be diminished and regenerator spacing can be increased further. The spacing of video amplifiers on coaxial cables typically is one kilometer or less. Actual amplification and regeneration requirements depend on the system design.

Functions and Terminology

The terminology has evolved considerably over the years, and can be confusing because it is not applied consistently. We've already seen that amplification and regeneration have different meanings, so you may be wondering where repeaters come in. To understand that, you need a very quick history lesson.

Repeaters were first used to extend telegraph transmission.

Repeaters were first used in long-distance electrical telegraph systems to detect faint input signals and repeat them automatically for transmission through another length of wire. The signals were dots and dashes, so the repeater both amplified the signal (by generating a new one) and cleaned it up (by replacing it with a fresh signal). When telephone systems began spanning long distances, they borrowed the name "repeater" for electronic amplifiers that amplified weak input signals.

The first fiber-optic systems also used repeaters, which consisted of a receiver and a transmitter placed back to back. The receiver converted the input signal to electronic form and amplified it, then the transmitter took the electronic input and generated a new optical signal to span the next length of fiber. The term "repeater" was still accurate because the receiver-transmitter pair *repeated* the input signal on a new length of fiber.

Optical amplifiers are based on the laser principle.

The first practical *optical amplifiers* were invented in the late 1980s. They rely on the laser principle to amplify optical input. The weak input signal stimulates the emission of light, which amplifies the signal strength. An optical amplifier is inherently an analog device, so its output signal is what you put in, but amplified in strength and with a dash of noise added. If the signal is noisy to start with, the optical amplifier multiplies the input noise by the same factor as the input signal. Yet optical amplifiers have a compelling simplicity, and—because they leave signals in optical form—they can amplify light at a range of wavelengths passing through the fiber. They can simultaneously amplify separate optical signals carried at different wavelengths without scrambling the signals, so one optical amplifier can boost the strength of many optical channels transmitted by the same fiber. As you will learn later in this chapter, there are several different types of optical amplifiers.

Although repeaters and optical amplifiers perform similar functions, they have different meanings in fiber optics. A repeater is specifically an electro-optic (or opto-electronic) device that converts input light into electronic form for processing, then generates an output optical signal. An optical amplifier is an *all-optical* device that never converts the optical signal into electronic form. Figure 12.2 illustrates the difference.

Optical amplifiers can amplify separate signals at different wavelengths.

Recall that an optical receiver can process only one optical channel at a time; and while one optical amplifier can amplify signals at several separate wavelengths, separate repeaters are needed for each optical channel. This has proved to be a compelling advantage, so optical repeaters are considered obsolete for long-distance transmission. (You will hear of "repeaters" used in submarine fiber-optic cable, but that's an anachronism. Undersea cables use optical amplifiers, but they're called "repeaters" because they are packaged in the same type of case used for electro-optic repeaters.)

Repeaters have sometimes been called "regenerators," and the term "regenerative repeater" has been used for repeaters that actually regenerated the original signal, as on electrical