any periodic filter can be used as an interleaver by matching its period to the desired channel spacing. For example, a fiber-based Mach-Zehnder interferometer is a common choice. These devices are now commercially available, and interleaving is becoming a popular approach toward realizing high channel count multiplexers and demultiplexers.

3.4 *Optical Amplifiers*

In an optical communication system, the optical signals from the transmitter are attenuated by the optical fiber as they propagate through it. Other optical components, such as multiplexers and couplers, also add loss. After some distance, the cumulative loss of signal strength causes the signal to become too weak to be detected. Before this happens, the signal strength has to be restored. Prior to the advent of optical amplifiers over the last decade, the only option was to regenerate the signal, that is, receive the signal and retransmit it. This process is accomplished by *regenerators*. A regenerator converts the optical signal to an electrical signal, cleans it up, and converts it back into an optical signal for onward transmission.

Optical amplifiers offer several advantages over regenerators. On one hand, regenerators are specific to the bit rate and modulation format used by the communication system. On the other hand, optical amplifiers are insensitive to the bit rate or signal formats. Thus a system using optical amplifiers can be more easily upgraded, for example, to a higher bit rate, without replacing the amplifiers. In contrast, in a system using regenerators, such an upgrade would require all the regenerators to be replaced. Furthermore, optical amplifiers have fairly large gain bandwidths, and as a consequence, a single amplifier can simultaneously amplify several WDM signals. In contrast, we would need a regenerator for each wavelength. Thus optical amplifiers have become essential components in high-performance optical communication systems.

Amplifiers, however, are not perfect devices. They introduce additional noise, and this noise accumulates as the signal passes through multiple amplifiers along its path due to the analog nature of the amplifier. The spectral shape of the gain, the output power, and the transient behavior of the amplifier are also important considerations for system applications. Ideally, we would like to have a sufficiently high output power to meet the needs of the network application. We would also like the gain to be flat over the operating wavelength range and to be insensitive to variations in input power of the signal. We will study the impact of optical amplifiers on the physical layer design of the system in Chapters 4 and 5. Here we explore their principle of operation.

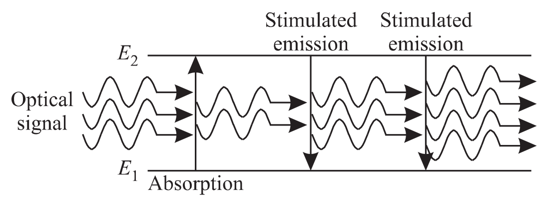

Figure 3.33 Stimulated emission and absorption in an atomic system with two energy levels.

We will consider three different types of amplifiers: *erbium-doped fiber amplifiers, Raman amplifiers,* and *semiconductor optical amplifiers*.

3.4.1 Stimulated Emission

In all the amplifiers we consider, the key physical phenomenon behind signal amplification is *stimulated emission* of radiation by atoms in the presence of an electromagnetic field. (This is not true of fiber Raman or fiber Brillouin amplifiers, which make use of fiber nonlinearities, but we do not treat these here.) This field is an optical signal in the case of optical amplifiers. Stimulated emission is the principle underlying the operation of lasers as well; we will study lasers in Section 3.5.1.

According to the principles of quantum mechanics, any physical system (for example, an atom) is found in one of a discrete number of energy levels. Accordingly, consider an atom and two of its energy levels, E_1 and E_2, with $E_2 > E_1$. An electromagnetic field whose frequency f_c satisfies $hf_c = E_2 - E_1$ induces transitions of atoms between the energy levels E_1 and E_2. Here, h is Planck's constant (6.63×10^{-34} J s). This process is depicted in Figure 3.33. Both kinds of transitions, $E_1 \rightarrow E_2$ and $E_2 \rightarrow E_1$, occur. $E_1 \rightarrow E_2$ transitions are accompanied by *absorption* of photons from the incident electromagnetic field. $E_2 \rightarrow E_1$ transitions are accompanied by the *emission* of photons of energy hf_c, the same energy as that of the incident photons. This emission process is termed *stimulated emission* to distinguish it from another kind of emission called *spontaneous emission*, which we will discuss later. Thus if stimulated emission were to dominate over absorption—that is, the incident signal causes more $E_2 \rightarrow E_1$ transitions than $E_1 \rightarrow E_2$ transitions—we would have a net increase in the number of photons of energy hf_c and an amplification of the signal. Otherwise, the signal will be attenuated.

It follows from the theory of quantum mechanics that the rate of the $E_1 \rightarrow E_2$ transitions per atom *equals* the rate of the $E_2 \rightarrow E_1$ transitions *per atom*. Let this common rate be denoted by r. If the populations (number of atoms) in the energy levels E_1 and E_2 are N_1 and N_2, respectively, we have a net increase in power (energy per unit time) of $(N_2 - N_1)rhf_c$. Clearly, for amplification to occur, this must be positive, that is, $N_2 > N_1$. This condition is known as *population inversion*. The reason for this term is that, at thermal equilibrium, lower energy levels are more highly populated, that is, $N_2 < N_1$. Therefore, at thermal equilibrium, we have only absorption of the input signal. In order for amplification to occur, we must *invert* the relationship between the populations of levels E_1 and E_2 that prevails under thermal equilibrium.

Population inversion can be achieved by supplying additional energy in a suitable form to pump the electrons to the higher energy level. This additional energy can be in optical or electrical form.

3.4.2 Spontaneous Emission

Before describing the operation of the different types of amplifiers, it is important to understand the impact of spontaneous emission. Consider again the atomic system with the two energy levels discussed earlier. Independent of any external radiation that may be present, atoms in energy level E_2 transit to the lower energy level E_1, emitting a photon of energy hf_c. The spontaneous emission rate per atom from level E_2 to level E_1 is a characteristic of the system, and its reciprocal, denoted by τ_{21}, is called the *spontaneous emission lifetime*. Thus, if there are N_2 atoms in level E_2, the rate of spontaneous emission is N_2/τ_{21}, and the spontaneous emission power is hf_cN_2/τ_{21}.

The spontaneous emission process does not contribute to the gain of the amplifier (to first order). Although the emitted photons have the same energy hf_c as the incident optical signal, they are emitted in random directions, polarizations, and phase. This is unlike the stimulated emission process, where the emitted photons not only have the same energy as the incident photons but also the same direction of propagation, phase, and polarization. This phenomenon is usually described by saying that the stimulated emission process is *coherent*, whereas the spontaneous emission process is *incoherent*.

Spontaneous emission has a deleterious effect on the system. The amplifier treats spontaneous emission radiation as another electromagnetic field at the frequency hf_c, and the spontaneous emission also gets amplified, in addition to the incident optical signal. This *amplified spontaneous emission* (ASE) appears as noise at the output of the amplifier. The implications of ASE for the design of optical communication systems are discussed in Chapters 4 and 5. In addition, in some amplifier designs, the ASE can be large enough to *saturate* the amplifier. Saturation effects are explored in Chapter 5.

3.4.3 Erbium-Doped Fiber Amplifiers

An erbium-doped fiber amplifier (EDFA) is shown in Figure 3.34. It consists of a length of silica fiber whose core is doped with ionized atoms (ions), Er^{3+}, of the rare earth element erbium. This fiber is pumped using a pump signal from a laser, typically at a wavelength of 980 nm or 1480 nm. In order to combine the output of the pump laser with the input signal, the doped fiber is preceded by a wavelength-selective coupler.

At the output, another wavelength-selective coupler may be used if needed to separate the amplified signal from any remaining pump signal power. Usually, an isolator is used at the input and/or output of any amplifier to prevent reflections into the amplifier. We will see in Section 3.5 that reflections can convert the amplifier into a laser, making it unusable as an amplifier.

A combination of several factors has made the EDFA the amplifier of choice in today's optical communication systems: (1) the availability of compact and reliable high-power semiconductor pump lasers, (2) the fact that it is an all-fiber device, making it polarization independent and easy to couple light in and out of it, (3) the simplicity of the device, and (4) the fact that it introduces no crosstalk when amplifying WDM signals. This last aspect is discussed later in the context of semiconductor optical amplifiers.

Principle of Operation

Three of the energy levels of erbium ions in silica glass are shown in Figure 3.35 and are labeled E_1, E_2, and E_3 in order of increasing energy. Several other levels in Er^{3+} are not shown. Each energy level that appears as a discrete line in an isolated

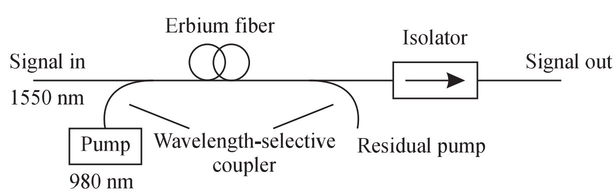

Figure 3.34 An erbium-doped fiber amplifier.

ion of erbium is split into multiple energy levels when these ions are introduced into silica glass. This process is termed *Stark splitting*. Moreover, glass is not a crystal and thus does not have a regular structure. Thus the Stark splitting levels introduced are slightly different for individual erbium ions, depending on the local surroundings seen by those ions. Macroscopically, that is, when viewed as a collection of ions, this has the effect of spreading each discrete energy level of an erbium ion into a continuous *energy band*. This spreading of energy levels is a useful characteristic for optical amplifiers since they increase the frequency or wavelength range of the signals that can be amplified. Within each energy band, the erbium ions are distributed in the various levels within that band in a nonuniform manner by a process known as *thermalization*. It is due to this thermalization process that an amplifier is capable of amplifying several wavelengths simultaneously. Note that Stark splitting denotes the phenomenon by which the energy levels of free erbium ions are split into a number of levels, or into an energy band, when the ion is introduced into silica glass. Thermalization refers to the process by which the erbium ions are distributed within the various (split) levels constituting an energy band.

Recall from our discussion of the two-energy-level atomic system that only an optical signal at the frequency f_c satisfying $hf_c = E_2 - E_1$ could be amplified in that

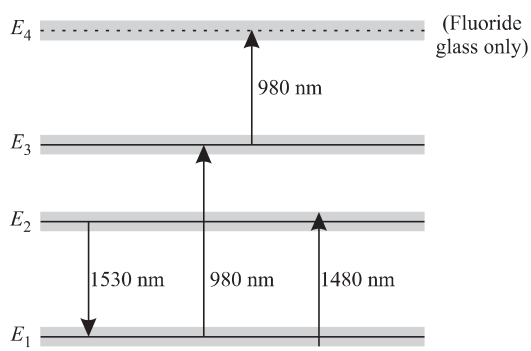

Figure 3.35 Three energy levels E_1, E_2, and E_3 of Er^{3+} ions in silica glass. The fourth energy level, E_4, is present in fluoride glass but not in silica glass. The energy levels are spread into bands by the Stark splitting process. The difference between the energy levels is labeled with the wavelength in nm of the photon corresponding to it. The upward arrows indicate wavelengths at which the amplifier can be pumped to excite the ions into the higher energy level. The 980 nm transition corresponds to the band gap between the E_1 and E_3 levels. The 1480 nm transition corresponds to the gap between the bottom of the E_1 band to the top of the E_2 band. The downward transition represents the wavelength of photons emitted due to spontaneous and stimulated emission.

case. If these levels are spread into bands, all frequencies that correspond to the energy difference between some energy in the E_2 band and some energy in the E_1 band can be amplified. In the case of erbium ions in silica glass, the set of frequencies that can be amplified by stimulated emission from the E_2 band to the E_1 band corresponds to the wavelength range 1525–1570 nm, a bandwidth of 50 nm, with a peak around 1532 nm. By a lucky coincidence, this is exactly one of the low-attenuation windows of standard optical fiber that optical communication systems use.

Denote ionic population in level E_i by N_i, $i = 1, 2, 3$. In thermal equilibrium, $N_1 > N_2 > N_3$. The population inversion condition for stimulated emission from E_2 to E_1 is $N_2 > N_1$ and can be achieved by a combination of absorption and spontaneous emission as follows. The energy difference between the E_1 and E_3 levels corresponds to a wavelength of 980 nm. So if optical power at 980 nm—called the *pump power*—is injected into the amplifier, it will cause transitions from E_1 to E_3 and vice versa. Since $N_1 > N_3$, there will be a net absorption of the 980 nm power. This process is called *pumping*.

The ions that have been raised to level E_3 by this process will quickly transit to level E_2 by the spontaneous emission process. The lifetime for this process, τ_{32}, is about 1 μs. Atoms from level E_2 will also transit to level E_1 by the spontaneous emission process, but the lifetime for this process, τ_{21}, is about 10 ms, which is much larger than the E_3 to E_2 lifetime. Moreover, if the pump power is sufficiently large, ions that transit to the E_1 level are rapidly raised again to the E_3 level only to transit to the E_2 level again. The net effect is that most of the ions are found in level E_2, and thus we have population inversion between the E_2 and E_1 levels. Therefore, if simultaneously a signal in the 1525–1570 nm band is injected into the fiber, it will be amplified by stimulated emission from the E_2 to the E_1 level.

Several levels other than E_3 are higher than E_2 and, in principle, can be used for pumping the amplifier. But the pumping process is more efficient, that is, uses less pump power for a given gain, at 980 nm than these other wavelengths. Another possible choice for the pump wavelength is 1480 nm. This choice corresponds to absorption from the bottom sublevel of the E_1 band to the top sublevel of the E_2 band itself. Pumping at 1480 nm is not as efficient as 980 nm pumping. Moreover, the degree of population inversion that can be achieved by 1480 nm pumping is lower. The higher the population inversion, the lower the noise figure of the amplifier. Thus 980 nm pumping is preferred to realize low-noise amplifiers. However, higher-power pump lasers are available at 1480 nm, compared to 980 nm, and thus 1480 nm pumps find applications in amplifiers designed to yield high output powers. Another advantage of the 1480 nm pump is that the pump power can also propagate with low loss in the silica fiber that is used to carry the signals. Therefore, the pump laser can be located remotely from the amplifier itself. This feature is used in some systems to avoid placing any active components in the middle of the link.

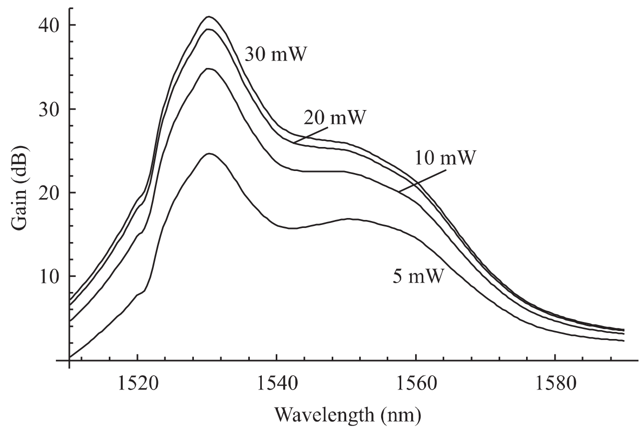

Figure 3.36 The gain of a typical EDFA as a function of the wavelength for four different values of the pump power, obtained through simulations. The length of the doped fiber is taken to be 15 m and 980 nm pumping is assumed.

Gain Flatness

Since the population levels at the various levels within a band are different, the gain of an EDFA becomes a function of the wavelength. In Figure 3.36, we plot the gain of a typical EDFA as a function of the wavelength for different values of the pump power. When such an EDFA is used in a WDM communication system, different WDM channels undergo different degrees of amplification. This is a critical issue, particularly in WDM systems with cascaded amplifiers, and is discussed in Section 5.5.2.

One way to improve the flatness of the amplifier gain profile is to use fluoride glass fiber instead of silica fiber, doped with erbium [Cle94]. Such amplifiers are called erbium-doped fluoride fiber amplifiers (EDFFAs). The fluoride glass produces a naturally flatter gain spectrum compared to silica glass. However, there are a few drawbacks to using fluoride glass. The noise performance of EDFFAs is poorer than EDFAs. One reason is that they must be pumped at 1480 nm and cannot be pumped at 980 nm. This is because fluoride glass has an additional higher energy level E_4 above the E_3 level, as shown in Figure 3.35, with the difference in energies between these two levels corresponding to 980 nm. This causes the 980 nm pump power to be absorbed for transitions from the E_3 to E_4 level, which does not produce useful gain. This phenomenon is called *excited state absorption*.

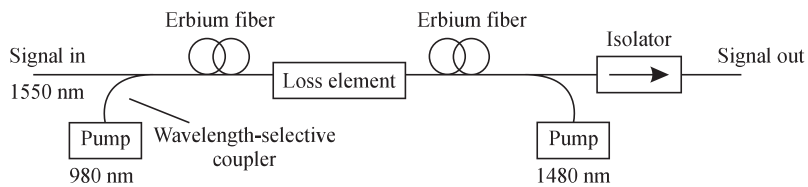

Figure 3.37 A two-stage erbium-doped fiber amplifier with a loss element inserted between the first and second stage.

In addition to this drawback, fluoride fiber itself is difficult to handle. It is brittle, difficult to splice with conventional fiber, and susceptible to moisture. Nevertheless, EDFFAs are now commercially available devices.

Another approach to flatten the EDFA gain is to use a filter inside the amplifier. The EDFA has a relatively high gain at 1532 nm, which can be reduced by using a notch filter in that wavelength region inside the amplifier. Some of the filters described in Section 3.3 can be used for this purpose. Long-period fiber gratings and dielectric thin-film filters are currently the leading candidates for this application.

Multistage Designs

In practice, most amplifiers deployed in real systems are more complicated than the simple structure shown in Figure 3.34. Figure 3.37 shows a more commonly used two-stage design. The two stages are optimized differently. The first stage is designed to provide high gain and low noise, and the second stage is designed to produce high output power. As we will see in Problem 4.5 in Chapter 4, the noise performance of the whole amplifier is determined primarily by the first stage. Thus this combination produces a high-performance amplifier with low noise and high output power. Another important consideration in the design is to provide redundancy in the event of the failure of a pump, the only active component of the amplifier. The amplifier shown in the figure uses two pumps and can be designed so that the failure of one pump has only a small impact on the system performance. Another feature of the two-stage design that we will address in Problem 4.5 is that a loss element can be placed between the two stages with negligible impact on the performance. This loss element may be a gain-flattening filter, a simple optical add/drop multiplexer, or a dispersion compensation module used to compensate for accumulated dispersion along the link.

L-Band EDFAs

So far, we have focused mostly on EDFAs operating in the C-band (1530–1565 nm). Erbium-doped fiber, however, has a relatively long tail to the gain shape extending well beyond this range to about 1605 nm. This has stimulated the development of systems in the so-called L-band from 1565 to 1625 nm. Note that current L-band EDFAs do not yet cover the top portion of this band from 1610 to 1625 nm.

L-band EDFAs operate on the same principle as C-band EDFAs. However, there are significant differences in the design of L- and C-band EDFAs. The gain spectrum of erbium is much flatter intrinsically in the L-band than in the C-band. This makes it easier to design gain-flattening filters for the L-band. However, the erbium gain coefficient in the L-band is about three times smaller than in the C-band. This necessitates the use of either much longer doped fiber lengths or fiber with higher erbium doping concentrations. In either case, the pump powers required for L-band EDFAs are much higher than their C-band counterparts. Due to the smaller absorption cross sections in the L-band, these amplifiers also have higher amplified spontaneous emission. Finally, many of the other components used inside the amplifier, such as isolators and couplers, exhibit wavelength-dependent losses and are therefore specified differently for the L-band than for the C-band. There are several other subtleties associated with L-band amplifiers; see [Flo00] for a summary.

As a result of the significant differences between C- and L-band amplifiers, these amplifiers are usually realized as separate devices rather than as a single device. In a practical system application, the C- and L-band wavelengths on a fiber are first separated by a demultiplexer, then amplified by separate amplifiers, and recombined together afterward.

3.4.4 Raman Amplifiers

In Section 2.5.3, we studied stimulated Raman scattering (SRS) as one of the nonlinear impairments that affect signals propagating through optical fiber. The same nonlinearity can be exploited to provide amplification as well. As we saw in Figure 2.17, the Raman gain spectrum is fairly broad, and the peak of the gain is centered about 13 THz below the frequency of the pump signal used. In the near-infrared region of interest to us, this corresponds to a wavelength separation of about 100 nm. Therefore, by pumping a fiber using a high-power pump laser, we can provide gain to other signals, with a peak gain obtained 13 THz below the pump frequency. For instance, using pumps around 1460–1480 nm provides Raman gain in the 1550–1600 nm window.

A few key attributes distinguish Raman amplifiers from EDFAs. Unlike EDFAs, we can use the Raman effect to provide gain at any wavelength. An EDFA provides

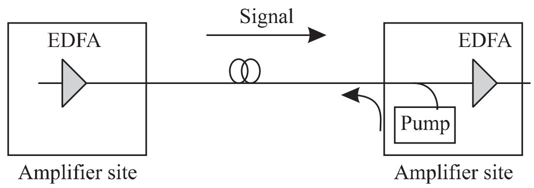

Figure 3.38 Distributed Raman amplifier using a backward propagating pump, shown operating along with discrete erbium-doped fiber amplifiers.

gain in the C- and L-bands (1528–1605 nm). Thus Raman amplification can potentially open up other bands for WDM, such as the 1310 nm window, or the so-called S-band lying just below 1528 nm. Also, we can use multiple pumps at different wavelengths and different powers simultaneously to tailor the overall Raman gain shape.

Second, Raman amplification relies on simply pumping the same silica fiber used for transmitting the data signals, so that it can be used to produce a *lumped* or *discrete* amplifier, as well as a *distributed* amplifier. In the lumped case, the Raman amplifier consists of a sufficiently long spool of fiber along with the appropriate pump lasers in a package. In the distributed case, the fiber can simply be the fiber span of interest, with the pump attached to one end of the span, as shown in Figure 3.38.

Today the most popular use of Raman amplifiers is to complement EDFAs by providing additional gain in a distributed manner in ultra-long-haul systems. The biggest challenge in realizing Raman amplifiers lies in the pump source itself. These amplifiers require high-power pump sources of the order of 1 W or more, at the right wavelength. We will study some techniques for realizing these pump sources in Section 3.5.5.

The noise sources in Raman amplifiers are somewhat different from EDFAs. The Raman gain responds instantaneously to the pump power. Therefore fluctuations in pump power will cause the gain to vary and will appear as crosstalk to the desired signals. This is not the case with EDFAs. We will see in Section 3.4.6 that the response time of the gain is much slower—on the order of milliseconds—in those devices. Therefore, for Raman amplifiers, it is important to keep the pump at a constant power. Having the pump propagate in the opposite direction to the signal helps dramatically because fluctuations in pump power are then averaged over the propagation time over the fiber. To understand this, first consider the case where the pump propagates along with the signal in the same direction. The two waves travel at approximately the same velocity. In this case, when the pump power is high at the input, the signal sees high gain, and when the power is low, the signal sees a lower gain. Now consider the case when the signal and pump travel in opposite directions. To keep things simple, suppose that the pump power varies between two states: high and low. As the signal propagates through the fiber, whenever it overlaps with the pump signal in the high power state, it sees a high gain. When it overlaps with the pump signal in the low power state, it sees a lower gain. If the pump fluctuations are relatively fast compared to the propagation time of the signal across the fiber, the gain variations average out, and by the time the signal exits the fiber, it has seen a constant gain.

Another major concern with Raman amplifiers is crosstalk between the WDM signals due to Raman amplification. A modulated signal at a particular wavelength depletes the pump power, effectively imposing the same modulation on the pump signal. This modulation on the pump then affects the gain seen by the next wavelength, effectively appearing as crosstalk on that wavelength. Again, having the pump propagate in the opposite direction to the signal dramatically reduces this effect. For these reasons, most Raman amplifiers use a counterpropagating pump geometry.

Another source of noise is due to the back-reflections of the pump signal caused by Rayleigh scattering in the fiber. Spontaneous emission noise is relatively low in Raman amplifiers. This is usually the dominant source of noise because, by careful design, we can eliminate most of the other noise sources.

3.4.5 *Semiconductor Optical Amplifiers*

Semiconductor optical amplifiers (SOAs) actually preceded EDFAs, although we will see that they are not as good as EDFAs for use as amplifiers. However, they are finding other applications in switches and wavelength converter devices. Moreover, the understanding of SOAs is key to the understanding of semiconductor lasers, the most widely used transmitters today.

Figure 3.39 shows the block diagram of a semiconductor optical amplifier. The SOA is essentially a *pn*-junction. As we will explain shortly, the depletion layer that is formed at the junction acts as the *active region*. Light is amplified through stimulated emission when it propagates through the active region. For an amplifier, the two ends of the active region are given an antireflection (AR) coating to eliminate ripples in the amplifier gain as a function of wavelength. Alternatively, the facets may also be angled slightly to reduce the reflection. In the case of a semiconductor laser, there would be no AR coating.

SOAs differ from EDFAs in the manner in which population inversion is achieved. First, the populations are not those of ions in various energy states but of *carriers*—*electrons* or *holes*—in a semiconductor material. Holes can also be thought of as charge carriers similar to electrons except that they have a positive charge. A semiconductor consists of two bands of electron energy levels: a band of

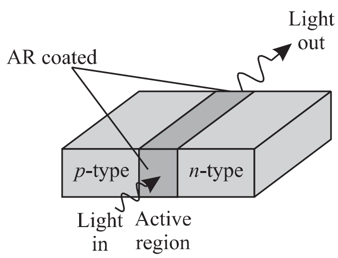

Figure 3.39 Block diagram of a semiconductor optical amplifier. Amplification occurs when light propagates through the active region. The facets are given an antireflective coating to prevent undesirable reflections, which cause ripple in the amplifier gain.

low-mobility levels called the *valence band* and a band of high-mobility levels called the *conduction band*. These bands are separated by an energy difference called the *bandgap* and denoted by E_g. No energy levels exist in the bandgap. Consider a p-type semiconductor material. At thermal equilibrium, there is only a very small concentration of electrons in the conduction band of the material, as shown in Figure 3.40(a). With reference to the previous discussion of EDFAs, it is convenient to think of the conduction band as the higher energy band E_2, and the valence band as the lower energy band E_1. The terms *higher* and *lower* refer to the electron energy in these bands. (Note that if we were considering an n-type semiconductor, we would be considering hole energies rather than electron energies, the conduction band would be the lower energy band E_1, and the valence band, the higher energy band E_2.) In the population inversion condition, the electron concentration in the conduction band is much higher, as shown in Figure 3.40(b). This increased concentration is such that, in the presence of an optical signal, there are more electrons transiting from the conduction band to the valence band by the process of stimulated emission than there are electrons transiting from the valence band to the conduction band by the process of absorption. In fact, for SOAs, this condition must be used as the defining one for population inversion, or optical gain.

Population inversion in an SOA is achieved by forward-biasing a pn-junction. A *pn-junction* consists of two semiconductors: a p-type semiconductor that is doped with suitable impurity atoms so as to have an excess concentration of holes, and an n-type semiconductor that has an excess concentration of electrons. When the two semiconductors are in juxtaposition, as in Figure 3.41(a), holes diffuse from the p-type semiconductor to the n-type semiconductor, and electrons diffuse from the n-type semiconductor to the p-type semiconductor. This creates a region with net negative charge in the p-type semiconductor and a region with net positive

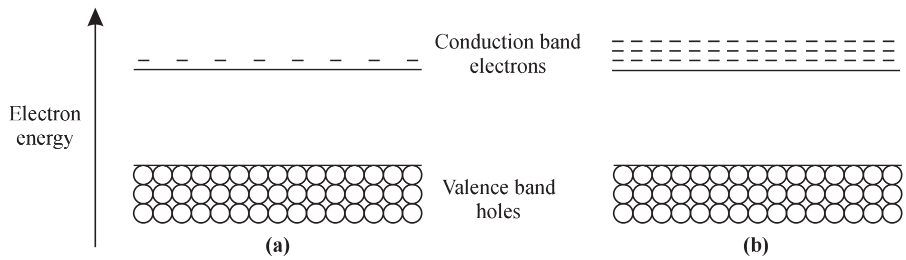

Figure 3.40 The energy bands in a *p*-type semiconductor and the electron concentration at (a) thermal equilibrium and (b) population inversion.

charge in the *n*-type semiconductor, as shown in Figure 3.41(b). These regions are devoid of free charge carriers and are together termed the *depletion region*. When no voltage (bias) is applied to the *pn*-junction, the minority carrier concentrations (electrons in the *p*-type region and holes in the *n*-type region) remain at their thermal equilibrium values. When the junction is *forward biased*—positive bias is applied to the *p*-type and negative bias to the *n*-type—as shown in Figure 3.41(c), the width of the depletion region is reduced, and there is a drift of electrons from the *n*-type region to the *p*-type region. This drift increases the electron concentration in the conduction band of the *p*-type region. Similarly, there is a drift of holes from the *p*-type to the *n*-type region that increases the hole concentration in the valence band of the *n*-type region. When the forward-bias voltage is sufficiently high, these increased minority carrier concentrations result in population inversion, and the *pn*-junction acts as an optical amplifier.

In practice, a simple *pn*-junction is not used, but a thin layer of a different semiconductor material is sandwiched between the *p*-type and *n*-type regions. Such a device is called a *heterostructure*. This semiconductor material then forms the *active region* or *layer*. The material used for the active layer has a slightly smaller bandgap and a higher refractive index than the surrounding *p*-type and *n*-type regions. The smaller bandgap helps to confine the carriers injected into the active region (electrons from the *n*-type region and holes from the *p*-type region). The larger refractive index helps to confine the light during amplification since the structure now forms a dielectric waveguide (see Section 2.3.4).

In semiconductor optical amplifiers, the population inversion condition (stimulated emission exceeds absorption) must be evaluated as a function of optical frequency or wavelength. Consider an optical frequency f_c such that $hf_c > E_g$, where E_g is the bandgap of the semiconductor material. The lowest optical frequency (or largest wavelength) that can be amplified corresponds to this bandgap. As the

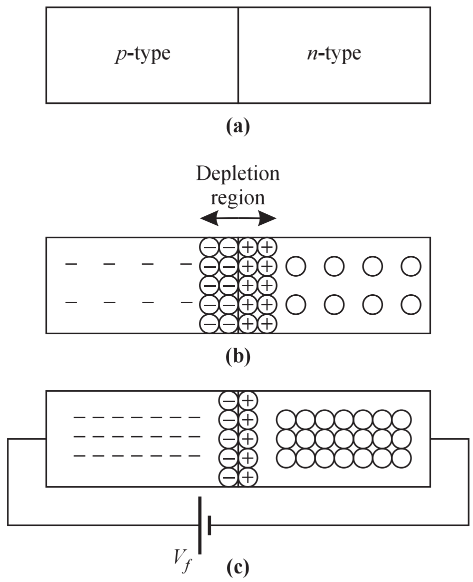

Figure 3.41 A forward-biased *pn*-junction used as an amplifier. (a) A *pn*-junction. (b) Minority carrier concentrations and depletion region with no bias voltage applied. (c) Minority carrier concentrations and depletion region with a forward-bias voltage, V_f.

forward-bias voltage is increased, the population inversion condition for this wavelength is reached first. As the forward bias voltage increases further, the electrons injected into the *p*-type region occupy progressively higher energy levels, and signals with smaller wavelengths can be amplified. In practice, bandwidths on the order of 100 nm can be achieved with SOAs. This is much larger than what is achievable with EDFAs. Signals in the 1.3 and 1.55 μm bands can even be simultaneously amplified using SOAs. Nevertheless, EDFAs are widely preferred to SOAs for several reasons. The main reason is that SOAs introduce severe crosstalk when they are used in WDM systems. This is discussed next. The gains and output powers achievable with EDFAs are higher. The coupling losses and the polarization-dependent losses are also lower with EDFAs since the amplifier is also a fiber. Due to the higher input coupling loss, SOAs have higher *noise figures* relative to EDFAs. (We will discuss noise figure in Section 4.4.5. For our purposes here, we can think of it as a measure of the noise introduced by the amplifier.) Finally, the SOA requires very high-quality antireflective coatings on its facets (reflectivity of less than 10^{-4}), which is not easy to achieve. Higher values of reflectivity create ripples in the gain spectrum and cause gain variations due to temperature fluctuations. (Think of this device as a Fabry-Perot filter with very poor reflectivity, and the spectrum as similar to the one plotted in Figure 3.17 for the case of poor reflectivity.) Alternatively, the SOA facets can be angled to obtain the desired reflectivities, at the cost of an increased polarization dependence.

3.4.6 Crosstalk in SOAs

Consider an SOA to which is input the sum of two optical signals at different wavelengths. Assume that both wavelengths are within the bandwidth of the SOA. The presence of one signal will deplete the minority carrier concentration by the stimulated emission process so that the population inversion seen by the other signal is reduced. Thus the other signal will not be amplified to the same extent and, if the minority carrier concentrations are not very large, may even be absorbed! (Recall that if the population inversion condition is not achieved, there is net absorption of the signal.) Thus, for WDM networks, the gain seen by the signal in one channel varies with the presence or absence of signals in the other channels. This phenomenon is called *crosstalk*, and it has a detrimental effect on system performance.

This crosstalk phenomenon depends on the spontaneous emission lifetime from the high-energy to the low-energy state. If the lifetime is large enough compared to the rate of fluctuations of power in the input signals, the electrons cannot make the transition from the high-energy state to the lower-energy state in response to these fluctuations. Thus there is no crosstalk whatsoever. In the case of SOAs, this lifetime is on the order of nanoseconds. Thus the electrons can easily respond to fluctuations in power of signals modulated at gigabit/second rates, resulting in a major system impairment due to crosstalk. In contrast, the spontaneous emission lifetime in an EDFA is about 10 ms. Thus crosstalk is introduced only if the modulation rates of the input signals are less than a few kilohertz, which is not usually the case. Thus EDFAs are better suited for use in WDM systems than SOAs.

There are several ways of reducing the crosstalk introduced by SOAs. One way is to operate the amplifier in the small signal region where the gain is relatively independent of the input power of the signal. Another is to *clamp* the gain of the amplifier using a variety of techniques, so that even at high signal powers, its gain remains relatively constant, independent of the input signal. Also, if a sufficiently large number of signals at different wavelengths are present, although each signal varies in power, the total signal power into the amplifier can remain fairly constant.

The crosstalk effect is not without its uses. We will see in Section 3.8.2 that it can be used to make a *wavelength converter*.

3.5 Transmitters

We will study many different types of light sources in this section. The most important one is the laser, of which there are many different types. Lasers are used as transmitters as well as to pump both erbium-doped and Raman amplifiers.

When using a laser as a light source for WDM systems, we need to consider the following important characteristics:

1. Lasers need to produce a reasonably high output power. For WDM systems, the typical laser output powers are in the 0–10 dBm range. Related parameters are the threshold current and slope efficiency. Both of these govern the efficiency of converting electrical power into optical power. The *threshold current* is the drive current at which the laser starts to emit optical power, and the *slope efficiency* is the ratio of output optical power to drive current.
2. The laser needs to have a narrow *spectral width* at a specified operating wavelength so that the signal can pass through intermediate filters and multiple channels can be placed close together. The side-mode suppression ratio is a related parameter, which we will discuss later. In the case of a tunable laser, the operating wavelength can be varied.
3. Wavelength stability is an important criterion. When maintained at constant temperature, the wavelength drift over the life of the laser needs to be small relative to the wavelength spacing between adjacent channels.
4. For lasers that are modulated, chromatic dispersion can be an important limiting factor that affects the link length. We will see in Chapter 5 that the dispersion limit can be stated in terms of a penalty as a function of the total accumulated dispersion along the link.

Pump lasers are required to produce much higher power levels than lasers used as WDM sources. Pump lasers used in erbium-doped fiber amplifiers put out 100–200 mW of power, and pump lasers for Raman amplifiers may go up to a few watts.

3.5.1 Lasers

A laser is essentially an optical amplifier enclosed within a reflective cavity that causes it to oscillate via positive feedback. *Semiconductor lasers* use semiconductors as the gain medium, whereas *fiber lasers* typically use erbium-doped fiber as the gain medium. Semiconductor lasers are by far the most popular light sources for optical communication systems. They are compact, usually only a few hundred micrometers

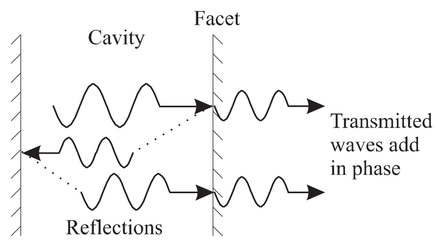

Figure 3.42 Reflection and transmission at the facets of a Fabry-Perot cavity.

in size. Since they are essentially *pn*-junctions, they can be fabricated in large volumes using highly advanced integrated semiconductor technology. The lack of any need for optical pumping, unlike fiber lasers, is another advantage. In fact, a fiber laser typically uses a semiconductor laser as a pump! Semiconductor lasers are also highly efficient in converting input electrical (pump) energy into output optical energy.

Both semiconductor and erbium fiber lasers are capable of achieving high output powers, typically between 0 and 20 dBm, although semiconductor lasers used as WDM sources typically have output powers between 0 and 10 dBm. Fiber lasers are used mostly to generate periodic trains of very short pulses (by using a technique called mode locking, discussed later in this section).

Principle of Operation

Consider any of the optical amplifiers described, and assume that a part of the optical energy is reflected at the ends of the amplifying or *gain medium,* or *cavity,* as shown in Figure 3.42. Further assume that the two ends of the cavity are plane and parallel to each other. Thus the gain medium is placed in a *Fabry-Perot cavity* (see Section 3.3.5). Such an optical amplifier is called a *Fabry-Perot amplifier.* The two end faces of the cavity (which play the role of the mirrors) are called *facets.*

The result of placing the gain medium in a Fabry-Perot cavity is that the gain is high only for the resonant wavelengths of the cavity. The argument is the same as that used in the case of the Fabry-Perot filter (Section 3.3.5). After one pass through the cavity, as shown in Figure 3.42, part of the light leaves the cavity through the right facet, and part is reflected. Part of the reflected wave is again reflected by the left facet to the right facet. For the resonant wavelengths of the cavity, all the light waves transmitted through the right facet *add in phase.* As a result of in-phase addition, the amplitude of the transmitted wave is greatly increased for these resonant wavelengths compared to other wavelengths. Thus, when the facets are at least partially reflecting, the gain of the optical amplifier becomes a function of the wavelength.

If the combination of the amplifier gain and the facet reflectivity is sufficiently large, the amplifier will start to "oscillate," or produce light output, even in the absence of an input signal. For a given device, the point at which this happens is called its *lasing threshold*. Beyond the threshold, the device is no longer an amplifier but an oscillator or *laser*. This occurs because the stray spontaneous emission, which is always present at all wavelengths within the bandwidth of the amplifier, gets amplified even without an input signal and appears as the light output. This process is quite similar to what happens in an electronic oscillator, which can be viewed as an (electronic) amplifier with positive feedback. (In electronic oscillators, the thermal noise current due to the random motion of electrons serves the same purpose as spontaneous emission.) Since the amplification process is due to stimulated emission, the light output of a laser is *coherent*. The term *laser* is an acronym for *light amplification* by *stimulated emission* of *radiation*.

Longitudinal Modes

For laser oscillation to occur at a particular wavelength, two conditions must be satisfied. First, the wavelength must be within the bandwidth of the gain medium that is used. Thus, if a laser is made from erbium-doped fiber, the wavelength must lie in the range 1525–1560 nm. The second condition is that the length of the cavity must be an integral multiple of half the wavelength in the cavity. For a given laser, all the wavelengths that satisfy this second condition are called the *longitudinal modes* of that laser. The adjective "longitudinal" is used to distinguish these from the waveguide modes (which should strictly be called spatial modes) that we studied in Section 2.2.

The laser described earlier is called a *Fabry-Perot laser* (FP laser) and will usually oscillate simultaneously in several longitudinal modes. Such a laser is termed a *multiple-longitudinal mode* (MLM) laser. MLM lasers have large spectral widths, typically around 10 nm. A typical spectrum of the output of an MLM laser is shown in Figure 3.43(a). We saw in Section 2.4 that for high-speed optical communication systems, the spectral width of the source must be as narrow as possible to minimize the effects of chromatic dispersion. Similarly, a narrow spectral width is also needed to minimize crosstalk in WDM systems (see Section 3.3). Thus it is desirable to design a laser that oscillates in a single-longitudinal mode (SLM) only. The spectrum of the output of an SLM laser is shown in Figure 3.43(b). Single-longitudinal mode oscillation can be achieved by using a filtering mechanism in the laser that selects the desired wavelength and provides loss at the other wavelengths. An important

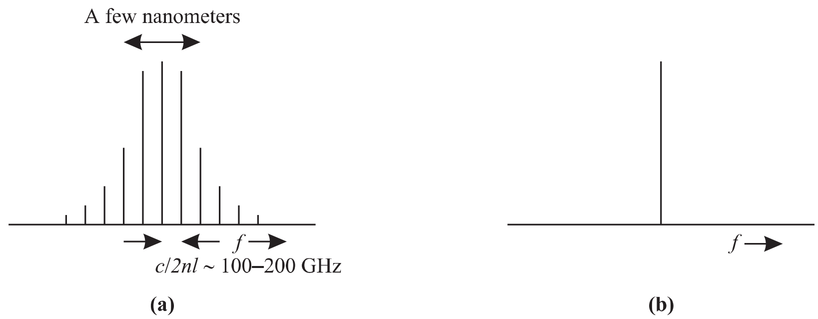

Figure 3.43 The spectrum of the output of (a) an MLM laser and (b) an SLM laser. The laser cavity length is denoted by l, and its refractive index by n. The frequency spacing between the modes of an MLM laser is then $c/2nl$.

attribute of such a laser is its *side-mode suppression ratio,* which determines the level to which the other longitudinal modes are suppressed, compared to the main mode. This ratio is typically more than 30 dB for practical SLM lasers. We will now consider some mechanisms that are commonly employed for realizing SLM lasers.

Distributed-Feedback Lasers

In the Fabry-Perot laser described earlier, the feedback of the light occurs from the reflecting facets at the ends of the cavity. Thus the feedback can be said to be *localized* at the facets. Light feedback can also be provided in a *distributed* manner by a series of closely spaced reflectors. The most common means of achieving this is to provide a periodic variation in the width of the cavity, as shown in Figure 3.44(a) and (b).

In the corrugated section of the cavity, the incident wave undergoes a series of reflections. The contributions of each of these reflected waves to the resulting transmitted wave from the cavity add in phase if the period of the corrugation is an integral multiple of half the wavelength in the cavity. The reasoning for this condition is the same as that used for the Fabry-Perot cavity. This condition is called the Bragg condition and was discussed in Section 3.3.3. The Bragg condition will be satisfied for a number of wavelengths, but the strongest transmitted wave occurs for the wavelength for which the corrugation period is *equal* to half the wavelength, rather than some other integer multiple of it. Thus this wavelength gets preferentially amplified at the expense of the other wavelengths. By suitable design of the device, this effect can be used to suppress all other longitudinal modes so that

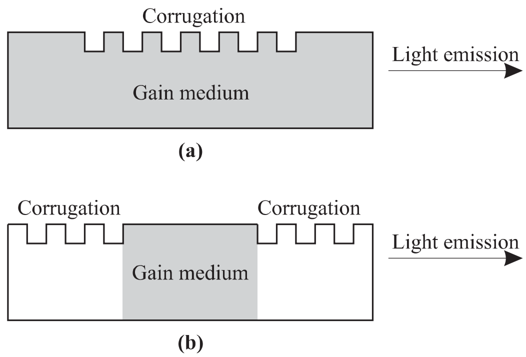

Figure 3.44 The structure of (a) a DFB laser and (b) a DBR laser. In a DFB laser, the gain and wavelength selection are obtained in the same region, whereas in a DBR laser, the wavelength selection region is outside the gain region.

the laser oscillates in a single-longitudinal mode whose wavelength is equal to twice the corrugation period. By varying the corrugation period at the time of fabrication, different operating wavelengths can be obtained.

Any laser that uses a corrugated waveguide to achieve single-longitudinal mode operation can be termed a distributed-feedback laser. However, the acronym *DFB laser* is used only when the corrugation occurs within the gain region of the cavity, as shown in Figure 3.44(a). When the corrugation is outside the gain region, as in Figure 3.44(b), the laser is called a *distributed Bragg reflector* (DBR) laser. The main advantage of DBR lasers is that the gain region is decoupled from the wavelength selection region. Thus it is possible to control both regions independently. For example, by changing the refractive index of the wavelength selection region, the laser can be tuned to a different wavelength without affecting its other operating parameters. Indeed, this is how many of the tunable lasers that we will study in Section 3.5.3 are realized.

DFB lasers are inherently more complex to fabricate than FP lasers and thus relatively more expensive. However, DFB lasers are required in almost all high-speed transmission systems today. FP lasers are used for shorter-distance data communication applications.

Reflections into a DFB laser cause its wavelength and power to fluctuate and are prevented by packaging the laser with an isolator in front of it. The laser is also usually packaged with a thermoelectric (TE) cooler and a photodetector attached to its rear facet. The TE cooler is necessary to maintain the laser at a constant operating temperature to prevent its wavelength from drifting. The temperature sensitivity of

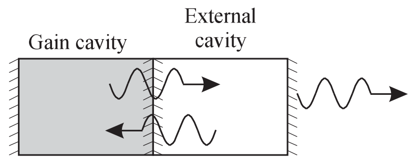

Figure 3.45 The structure of an external cavity laser.

a semiconductor DFB laser operating in the 1.55 μm wavelength region is about 0.1 nm/°C. The photodetector monitors the optical power leaking out of the rear facet, which is proportional to the optical power coming out of the laser.

The packaging of a DFB laser contributes a significant fraction of the overall cost of the device. For WDM systems, it is very useful to package multiple DFB lasers at different wavelengths inside a single package. This device can then serve as a multiwavelength light source or, alternatively, as a tunable laser (only one of the lasers in the array is turned on, depending on the desired wavelength). These lasers can all be grown on a single substrate in the form of an array. Four- and eight-wavelength laser arrays have been fabricated in research laboratories, but have not quite progressed to volume manufacturing. The primary reason for this is the relatively low yield of the array as a whole. If one of the lasers doesn't meet specifications, the entire array will have to be discarded.

External Cavity Lasers

Suppression of oscillation at more than one longitudinal mode can also be achieved by using another cavity—called an *external cavity*—following the primary cavity where gain occurs. This is illustrated in Figure 3.45. Just as the primary cavity has resonant wavelengths, so does the external cavity. This effect can be achieved, for example, by using reflecting facets for the external cavity as well. The net result of having an external cavity is that the laser is capable of oscillating only at those wavelengths that are resonant wavelengths of *both* the primary and external cavity. By suitable design of the two cavities, it can be ensured that only one wavelength in the gain bandwidth of the primary cavity satisfies this condition. Thus the laser oscillation can be confined to a single-longitudinal mode.

Instead of another Fabry-Perot cavity, as shown in Figure 3.45, we can use a diffraction grating (see Section 3.3.1) in the external cavity, as shown in Figure 3.46. Such a laser is called a *grating external cavity* laser. In this case, the facet of the gain cavity facing the grating is given an antireflection coating. The wavelengths reflected by the diffraction grating back to the gain cavity are determined by the

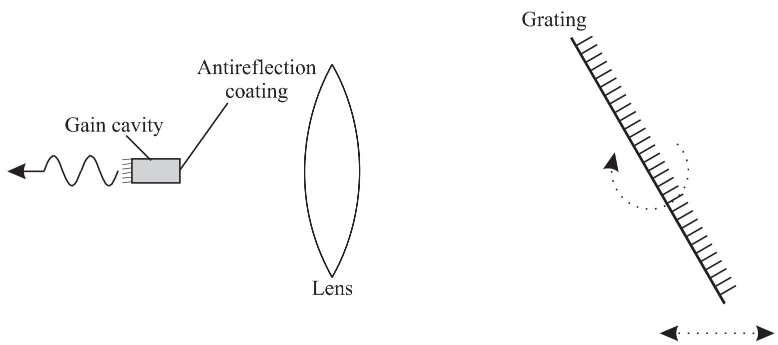

Figure 3.46 The structure of a grating external cavity laser. By rotating the grating, we can tune the wavelength of the laser.

pitch of the grating (see Section 3.3.1) and its tilt angle (see Figure 3.46) with respect to the gain cavity. An external cavity laser, in general, uses a *wavelength-selective mirror* instead of a wavelength-flat mirror. (A highly polished and/or metal-coated facet used in conventional lasers acts as a wavelength-flat mirror.) The reflectivity of a wavelength-selective mirror is a function of the wavelength. Thus only certain wavelengths experience high reflectivities and are capable of lasing. If the wavelength-selective mirror is chosen suitably, only one such wavelength will occur within the gain bandwidth, and we will have a single-mode laser.

Several of the filters discussed in Section 3.3 can be used as wavelength-selective mirrors in external cavity lasers. We have already seen the use of the diffraction grating (Section 3.3.1) and Fabry-Perot filter (Section 3.3.5) in external cavity lasers. These laser structures are used today primarily in optical test instruments and are not amenable to low-cost volume production as SLM light sources for transmission systems. One version of the external cavity laser, though, appears to be particularly promising for this purpose. This device uses a fiber Bragg grating in front of a conventional FP laser with its front facet AR coated. This device then acts as an SLM DBR laser. It can be fabricated at relatively low cost compared to DFB lasers and is inherently more temperature stable in wavelength due to the low temperature-coefficient of the fiber grating.

One disadvantage of external cavity lasers is that they cannot be modulated directly at high speeds. This is related to the fact that the cavity length is large.

Vertical Cavity Surface-Emitting Lasers

In this section, we will study another class of lasers that achieve single-longitudinal mode operation in a slightly different manner. As we saw in Figure 3.43, the frequency