CHAPTER 14

SEMICONDUCTOR LASERS

Semiconductor lasers, or diode lasers as they are often called, are by far the most ubiquitous of all lasers. Every year, diode-laser sales are a hundredfold greater than the unit sales of any other type of laser. At the beginning of the twenty-first century, diode-laser sales are exceeding sales of 100 million units per year. They are used in an enormous variety of applications, from CD players to laser printers to telecommunications systems.

A very simple semiconductor laser is shown in Fig. 14.1, and you can readily see why diode lasers are so much smaller, lighter, and more rugged than other lasers. What is shown in Fig. 14.1 is the entire package, but the laser itself is just the small crystalline block with the wire bonded to its top. The laser is formed by the junction of two dissimilar types of semiconductor, and the light emerges from the edge of the block, coming directly from the junction.

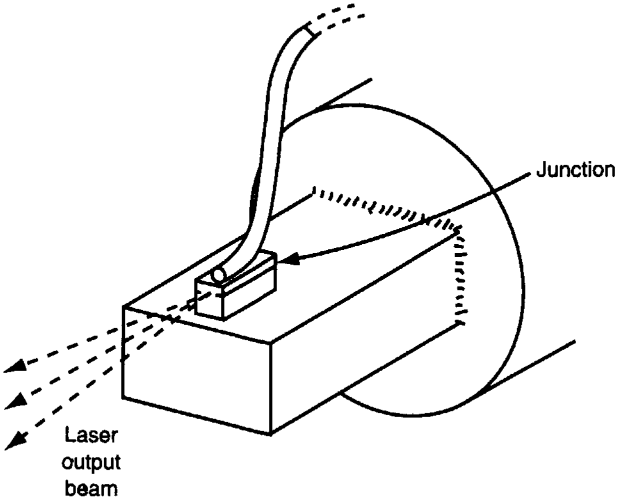


Figure 14.1 The semiconductor laser itself is the tiny, rectangular crystal on the mounting post.

Introduction to Laser Technology, Third Edition Breck Hitz, J. J. Ewing and Jeff Hecht
ISBN: 0-780-35373-0

14.1 SEMICONDUCTOR PHYSICS

To understand how the semiconductor laser in Fig. 14.1 works, you need to learn a little bit about what semiconductors are, and what conductors and insulators are. An electric current is composed of moving electrons, and any material that has "loose" electrons can be a conductor. When a voltage is applied across a copper wire, the "loose" electrons in the wire flow from one end to the other. In an insulator, on the other hand, all the electrons are held tightly in place. When a voltage is applied to an insulator, nothing happens because there are no "loose" electrons to move from one place to another.

In Chapter 6 we discussed the cloud of electrons that surrounds the nucleus of an atom. These electrons are arranged in "shells," and it's usually the electrons in the outermost shell that are most likely to shake loose and become carriers of an electrical current. In a metal such as copper, the force holding the outermost electrons to the nucleus is weak, so copper is a good electrical conductor. In a poor conductor such as silicon, however, the electrons in the outermost shell are tightly bound to the nucleus and are not readily available to conduct a current. Thus, silicon is a semiconductor.

Silicon has four electrons in its outermost shell, and they're all necessary to form bonds with adjacent atoms in the silicon crystal, as illustrated in Fig. 14.2a. An arsenic atom is more or less the same size as a silicon atom, but it has five electrons in its outermost shell. If a silicon crystal is *doped* with arsenic—that is, a small amount of arsenic is added as an impurity to the silicon crystal as it is grown—then arsenic atoms will replace silicon atoms at some locations in the crystal. This is shown schematically in Fig. 14.2b.

Four of the arsenic atom's outermost-shell electrons are used to form bonds with adjacent atoms. But the fifth electron is not tightly attached to the nucleus and can, under the right conditions, become available to conduct a

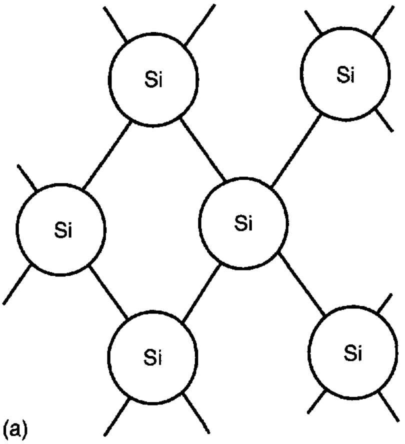


Figure 14.2a In a normal silicon crystal, all four outermost shell electrons are used, forming bonds with neighbors. There are no "loose" electrons, and normal silicon is not an electrical conductor.

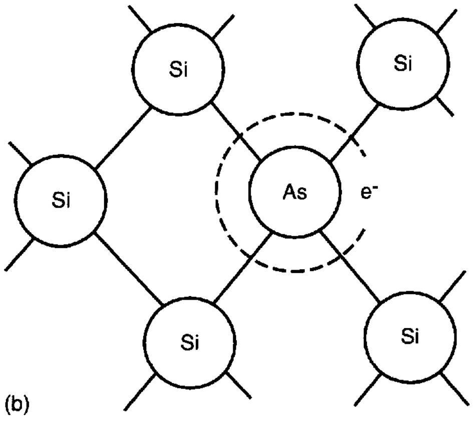


Figure 14.2b When a silicon crystal is doped with arsenic (which has five electrons in the outermost shell), the extra electron is "loose" and turns the crystal into a semiconductor.

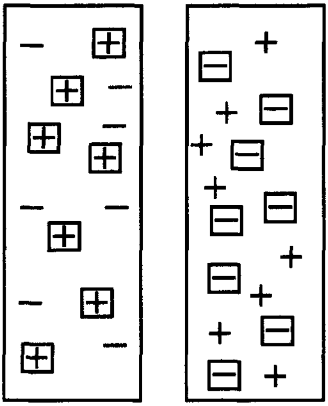
current. The arsenic thereby enhances the conductivity of the silicon semiconductor.

But there is a variation on this theme. Instead of doping the silicon with arsenic, which has five outermost-shell electrons, you could dope it with boron, an atom that has only three outermost-shell electrons. Then there would be a "hole" where the fourth electron is supposed to be. And it turns out that such "holes" can conduct electricity just the way electrons can. Of course, you understand that when a hole moves from point A to point B, what *really* happens is that an electron moves from point B to point A. But it turns out that the physics is easier to deal with if you think of holes being the things that move, rather than electrons.

Thus, there are two fundamental types of semiconductors. In an n-doped semiconductor, there are extra electrons that are the charge carriers, and in a p-doped semiconductor, there are extra holes that are the charge carriers.

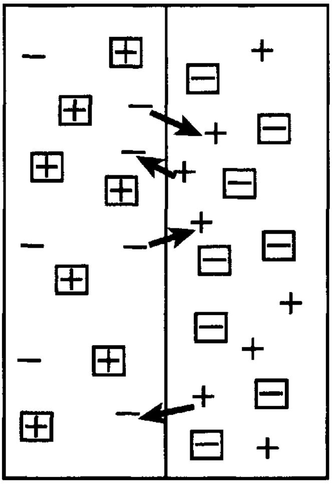
In a diode laser, laser action occurs at the junction of an n- and a p-doped material. Let's consider what happens when you bring an n-doped material into contact with a p-doped material, as illustrated in Fig. 14.3. The p-doped material has loosely bound positive charges (holes), and the n-doped material has loosely bound negative charges (electrons). Hence, initially, some of the p-doped material's holes move across the junction and combine with some of the n-doped material's electrons, and vice versa.

It's very important to understand what happens when an electron and a hole combine. Remember that a hole is a place where there's supposed to be an electron, so when an electron and hole combine, the electron literally falls into the hole, filling it. And since the electron loses energy as it falls, at least part of that energy can be liberated as a photon. That's where the photons come from in a semiconductor laser: they're created when an electron loses energy by combining with a hole.



(a)

Figure 14.3a Both p-type (right) and n-type (left) semiconductors are electrically neutral, but p-type semiconductors have mobile positive carriers (holes), and n-type semiconductors have mobile negative carriers (electrons). In this figure, mobile carriers are represented as charges without boxes around them.



(b)

Figure 14.3b When a junction is formed between an n- and a p-type semiconductor, some of the holes and electrons move across the junction to combine with opposite charges in the other material.

(When an electron and a hole combine, the energy isn't always released as a photon. Sometimes it is released in the form of heat. In particular, in the silicon crystal we discussed earlier, the energy takes the form of heat, because the total momentum of the electron and hole changes when they combine. Momentum must be conserved during the process, and that's not possible if the energy is emitted as a photon. However, there are other crystals, gallium arsenide [GaAs] for example, in which the total momentum is unchanged when the electron and hole combine. So in GaAs, the energy can be emitted as light. Semiconductor lasers are made of GaAs and related materials, not silicon.)

Let's go back to the junction we've created between an n- and a p-doped material. We saw that initially a number of charge carriers moved across the junction to combine with opposite charges in the other material. But soon, as shown in Fig. 14.4a, there is an excess of positive charge on the n-doped ma-

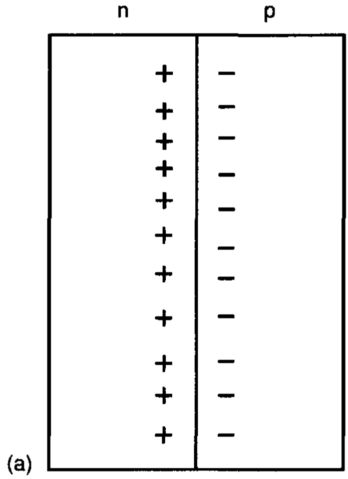


Figure 14.4a After an initial cascade of charge carriers across the junction, a potential barrier builds up and prevents further charge migration.

terial side, and an excess of negative charge on the p-doped material side. This excess charge stops more carriers from moving across the junction.

Thus, when you form a junction between an n- and a p-type semiconductor, you get an initial shower of photons as mobile charges move across the junction and combine with opposite charges, but the show is quickly over. To prolong it, you can inject additional mobile charges, as shown in Fig. 14.4b (i.e., you pass an electric current through the junction). By doing this, you can create a steady-state condition in which photons are continuously being emitted from the junction.

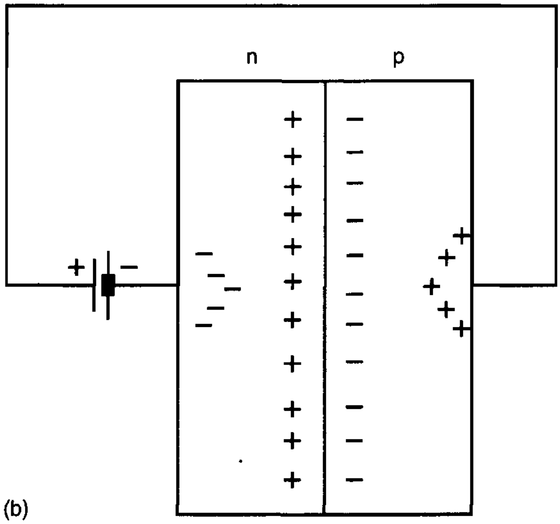


Figure 14.4b By injecting charge carriers into the junction as shown here, a steady current across the junction can be maintained.

In fact, the device illustrated in Fig. 14.4b is a light-emitting diode (LED). To turn this LED into a diode laser, you have to recall from Chapter 6 the two things required for laser action: a population inversion and feedback. To create a population inversion, you must have more population at an upper level than a lower level. This usually requires a very high current density. And to provide feedback, you can place a reflector at the ends of the junction. In many cases, the reflectivity of the polished crystal face is adequate.

Interestingly, diode lasers are generally more efficient than LEDs, because an LED depends on spontaneous emission, and, therefore, a given atom must wait for the entire spontaneous lifetime before it can emit its photon. While it's waiting, the energy can be lost to competing, nonradiative decay mechanisms. In a diode laser, on the other hand, the photons are created by stimulated emission, so the energy doesn't have to wait as long in the upper laser level.

14.2 MODERN DIODE LASERS

The first diode lasers, developed in the early 1960s, looked something like the device illustrated in Fig. 14.1. But semiconductor lasers have evolved enormously since then. The laser in Fig. 14.1 requires very high current flow to maintain a population inversion, and the heat generated by the steady-state current would quickly destroy the device.

To reduce the current and heat while maintaining a population inversion, modern diode lasers pack the stimulated emission into a small region. Hence, the current density is still great enough to maintain a population inversion, but the total current does not overheat the laser. There are two approaches to increasing the density of stimulated emission: increasing the density of charge carriers and increasing the density of intracavity optical power.

Both techniques invoke the sophisticated semiconductor fabrication techniques that have evolved over the past 30 years. These techniques allow the complex structures illustrated in Figs. 14.5 through 14.9 to be grown, literally one molecule at a time, from the basic raw materials. Today, methods such as molecular-beam epitaxy and metal-organic chemical vapor deposition allow the creation of semiconductor structures that are only several atoms thick.

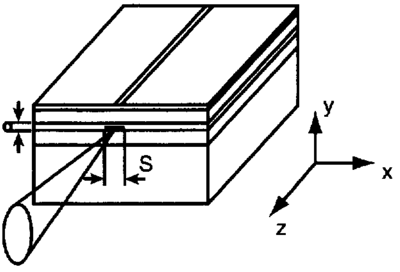


Figure 14.5 This laser has a stripe electrode on the top to restrict the current flow to a narrow region, and a double heterostructure (see text) to confine the photons.

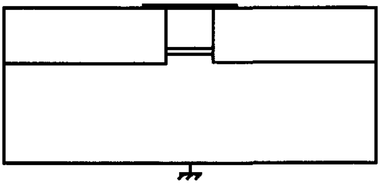


Figure 14.6 The lower refractive index of InP in the blocking regions prevents the laser photons from spreading outside the micrometer-wide active region.

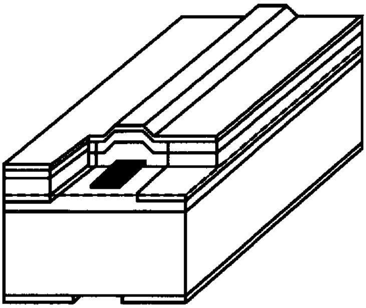


Figure 14.7 Frequency-selective feedback is provided by the optical gratings at the ends of this "distributed-feedback" laser.

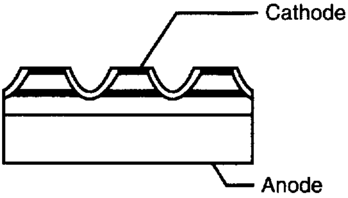


Figure 14.8 A one-dimensional diode-laser array. The original layered structure was grown on the substrate, and then the individual lasers were etched out chemically.

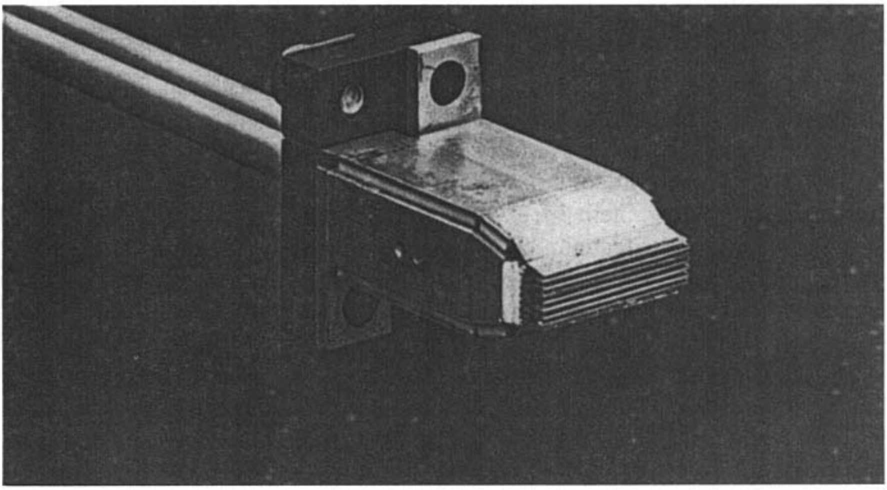


Figure 14.9 A mechanical "stack" of one-dimensional arrays finds numerous commercial applications (Photo courtesy of SDL Inc).

One way to increase the density of charge carriers is to use a stripe electrode, as illustrated in Fig. 14.5. Instead of injecting the current over a wide area of the diode's surface, current is injected only along a narrow stripe, resulting in a much higher concentration of charge carriers inside the diode.

The laser in Fig. 14.5 confines the current to a small region with its stripe electrode. This is called *current confinement* in the plane of the junction, and it also confines the generated photons perpendicular to the plane of the junction with its "double heterostructure" design. Note that there are several junctions of dissimilar material in this laser, not just one junction, as in Fig. 14.1. The electrons and holes combine in the narrow region shown as having thickness *d*, and the material there has a higher refractive index than the material above or below it. That means that photons are reflected off the interface between the materials and, therefore, are confined to the region of thickness *d*. (Total internal reflection is discussed in Chapter 3.)

A more sophisticated method of increasing the density of charge carriers involves the quantum-mechanical nature of carriers in very thin regions. If the dimension *d* in Fig. 14.5 is very tiny, a few tens of nanometers, quantum mechanical effects become important and the charge carriers are trapped in the thin region. Such superthin structures are called *quantum wells* and are employed frequently in modern semiconductor lasers.

Figure 14.5 illustrates two techniques of confining charge carriers to a small region: the stripe electrode and a quantum well. It also illustrates one technique for confining the photons: total internal reflection from the interface above and below the active region. But the photons are still free to spread out sideways. Index guiding structures, like the one shown in Fig. 14.6, stop this spreading. Here, the laser has been fabricated with a low-index material on both sides of the active regions, as well as above and below the active region. Now the photons cannot spread out in any direction. Note that Fig. 14.7 also illustrates the stripe electrode, a few micrometers thick; a narrow quantum well; and a double heterostructure.

We've talked about confining the generated photons in three-dimensional space, but confining them in frequency can also be beneficial. Chapter 10 discussed some of the techniques of reducing a laser's bandwidth, and the same principles apply to diode lasers. In particular, the best way to reduce a laser's bandwidth is usually to reduce the bandwidth of the laser's feedback. In Chapter 10 we discussed using gratings for this purpose, and small, built-in gratings are often used with semiconductor lasers. Figure 14.7 shows a semiconductor laser with a grating fabricated into the structure. Such lasers are sometimes called *distributed-feedback* (DFB) lasers because the feedback, or reflectivity, is distributed over the length of the grating, rather than occurring all at once at a mirror. The wavelength that is fed back is determined by the period of the grating. Usually, a DFB laser has a grating fabricated into the entire length of the laser. A variation referred to as a distributed Bragg reflector has a distinct grating fabricated into the substrate on each side of the active area.

The end faces of an "external resonator" diode laser aren't fabricated into mirrors. Rather, the laser diode is placed in a separate resonator like those described in Chapter 10. Such lasers are capable of very narrow bandwidth and good frequency stability.

Another important development was the evolution of laser-diode arrays—multiple lasers in a single device. A one-dimensional diode array is illustrated in Fig. 14.8. Although the beam produced by this device is of dubious quality for many applications, the output power is several times greater than can be obtained from a single laser. (Many applications of high-power arrays do not require high beam quality. Optical pumping of solid-state lasers is one example.) In some cases, the optical phase of adjacent lasers can be locked, resulting in improved beam quality and stability.

Even greater power can be obtained by stacking one-dimensional arrays on top of each other, in essence creating a two-dimensional array. Several manufacturers produce high-power diode-laser "stacks" whose output exceeds 100 W. An example of diode-laser stacks is shown in Fig. 14.10. Although the beam quality from such stacks is relatively low, these devices are ideally suited for applications that require efficient delivery of high power, such as optical pumping of other lasers, as well as for many industrial and medical applications.

Although they aren't readily available commercially, semiconductor lasers have also been fabricated in monolithic, two-dimensional arrays. These arrays offer the possibility of combining the beams from hundreds, or even thousands, of individual lasers into a single, high-quality, powerful beam. A proto-

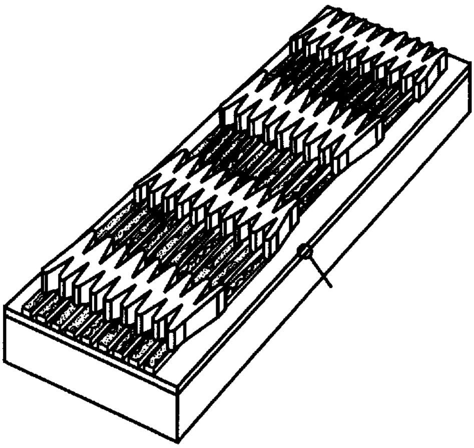


Figure 14.10 A monolithic, two-dimensional array. This cut-away view of the laser shows its internal structure.

type device is illustrated in Fig. 14.10. Here, the Y-like structures are the individual diode lasers, and the gratings between the Y-like structures provide the (distributed) feedback. Each leg of each Y-like structure is an individual resonator, and all the resonators are coupled in the chain of Y-like structures.

It's interesting to understand how the output coupling of this laser works. The gratings between the Y-like structure are designed such that a small portion of the light incident on them is diffracted *straight up,* out of the surface of the device. Thus, a tiny beam is projected upward from each of the gratings. These beams can combine at some distance from the laser to produce a single beam whose power is the sum of all the individual beams.

14.2.1 Wavelength of Diode Lasers

The wavelength of a diode laser depends on the amount of energy released in the form of a photon when an electron and a hole combine. In the argot of semiconductor physics, this is the semiconductor's *bandgap energy.* For semiconductors composed of two elements (*binary*), the bandgap is fixed at a given value. The bandgap energy of GaAs, for example, corresponds to a photon wavelength of 870 nm. However, for semiconductors composed of three (*ternary*) or four (*quaternary*) elements, the bandgap energy depends on the relative concentration of different elements. The ternary semiconductor gallium aluminum arsenide (GaAlAs) has a bandgap energy corresponding to photons from 900 to 620 nm. That is, the wavelength of a GaAlAs laser can be changed by altering the relative amounts of gallium and aluminum and arsenic in the crystal. The range of wavelengths varies from 620 to 900 nm in theory, or from about 750 to about 850 nm in practical devices. Shorter wavelengths—in the blue and even the ultraviolet—are obtainable with gallium-nitride compounds.

You can create ternary and quaternary semiconductors with a wide range of bandgap energies, corresponding to photons from the midinfrared (several microns) through the visible and into the ultraviolet. But you cannot build semiconductor lasers with many of these compounds, because the lattice spacing of the compound must be very close to that of the substrate. (Remember that a diode laser is a complex assemblage of many layers of semiconductors. The substrate is the bottom layer of a semiconductor device, the "table" on which you build the device itself.) Gallium arsenide and indium phosphide are two of the best substrates, and the requirement to match the lattice spacing of one of these crystals limits the wavelengths available from diode lasers.

In addition to GaAlAs, commercial semiconductor lasers are made from AlGaInP (whose wavelength ranges from about 650 to about 680 nm), InGaAsP (about 1.1 to about 1.65 μm), and InGaAsSb (about 1.7 to about 4.3 μm). More recently, GaN compounds have produced semiconductor lasers in the blue and even UV ranges. Other semiconductor lasers, based on the so-called lead-salt compounds, include PbSnTe (about 6 to about 25 μm) and PbEuSeTe (about 2 to about 4 μm).

14.2.2 Vertical Cavity, Surface-Emitting Lasers

One of the most fundamental problems with the diode lasers we've discussed so far is the highly divergent, elliptical beam. In Chapter 5 we explained that the divergence of a laser beam is inversely proportional to the beam size at the source—the smaller the source, the larger the divergence. But the laser illustrated in Fig. 14.5, for example, is an extremely small light source, perhaps 1 μm wide. And it is much smaller in the vertical direction than in the horizontal. Thus, the beam that emerges from the device in Fig. 14.5 diverges much more rapidly in the vertical direction than in the horizontal. The vertical divergence is typically tens of degrees, and the horizontal divergence is several degrees.

The highly divergent, elliptical beam can be corrected—to an extent—with a cylindrical lens, but the inherent problem of a small, elliptical source can never be completely rectified. And the diode lasers we've discussed so far have other fundamental limitations. Although they are extremely small, their resonators are still hundreds of micrometers in length—long enough to support multiple longitudinal modes. Unless the laser's bandwidth is artificially reduced (e.g., by fabricating a DFB structure), mode hopping among these modes produces an instability in both the amplitude and frequency of the laser's output.

Because the output beam emerges from the edge of the cleaved crystal, and the crystal is not cleaved until the end of the manufacturing process, it is not possible to test the devices optically during manufacture. This limitation tends to drive up the price of manufacturing.

Finally, although monolithic two-dimensional arrays can be fabricated, as illustrated in Fig. 14.10, these devices are extremely difficult to manufacture, and it is questionable whether they can ever be priced low enough to find widespread commercial use.

The vertical cavity, surface-emitting laser (VCSEL, pronounced like *vixel*) avoids these shortcomings. In conventional diode lasers, as we've already discussed, the cavity, or resonator, is in the horizontal plane. In a vertical-cavity laser, the cavity is along the vertical direction. The two approaches are illustrated conceptually in Fig. 14.11. In the vertical-cavity laser, the mirrors are located above and below the population inversion, instead of on either side. The horizontal-cavity laser is an edge emitter, while the vertical-cavity laser is a surface emitter.

You can readily see the advantage of such a design. For one thing, it immediately eliminates the problem of a divergent, elliptical beam caused by a small, irregular emitting surface of an edge emitter. The emitting area of a surface emitter is round, and many times larger.

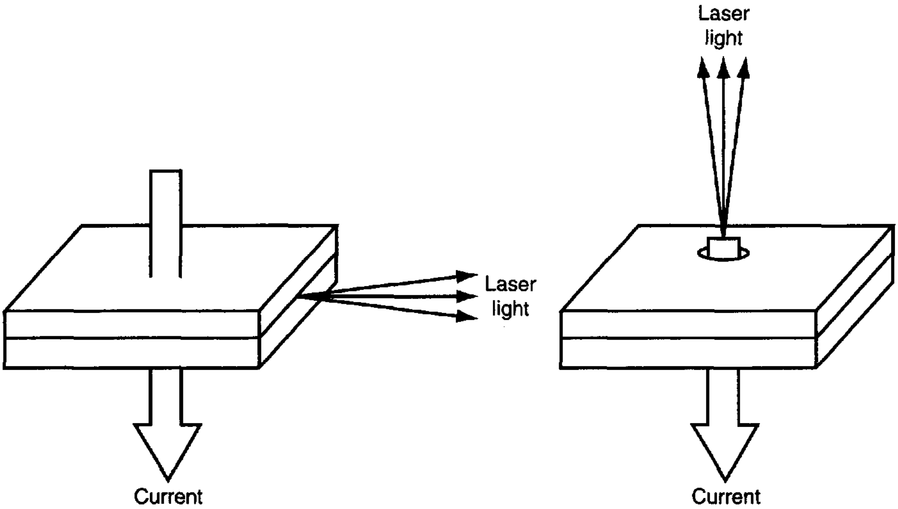


Figure 14.11 In a VCSEL (right) the resonator is vertical, and the light emerges from the surface of the laser rather than from the edge of a conventional diode laser (left).

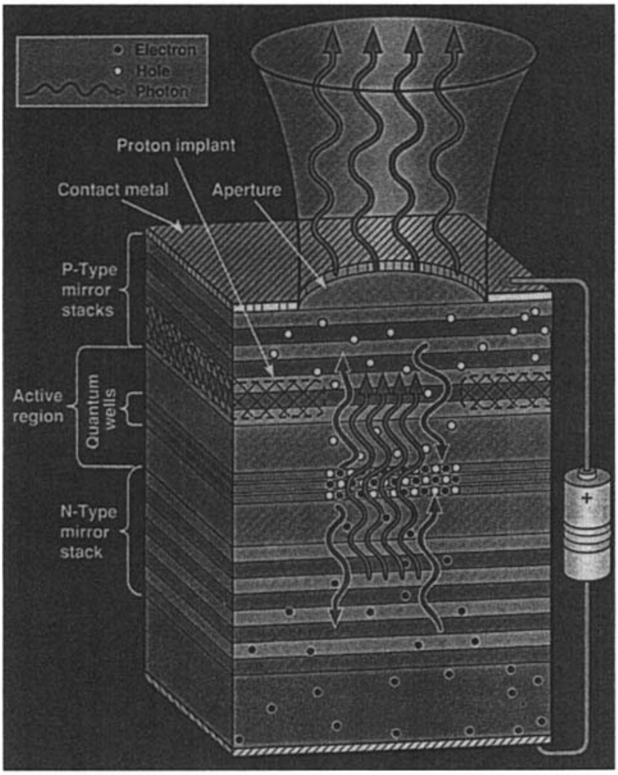


Figure 14.12 A structural view of a VCSEL. The mirrors for the vertical resonator are formed by layers of semiconductors above and below the active region.

Figure 14.11 makes it clear that the resonator of a VCSEL is shorter than a conventional laser's. In fact, it is so short that the spacing between longitudinal modes is too great for more than one mode to oscillate. (Longitudinal modes are spaced by $c/2\ell$, as discussed in Chapter 9.) Thus, the mode-hopping instability of conventional lasers is eliminated.

Moreover, the manufacturing difficulties are reduced because it is possible to subject a waferfull of VCSELs to optical testing during the manufacturing process. Very high densities, tens of millions of diodes per wafer, can be achieved, further driving down the cost of individual diodes. And it seems much more straightforward to fabricate two-dimensional arrays. Figure 14.12 shows the detailed structure of a VCSEL. The mirrors are thin layers of semiconductors, fabricated to the proper thickness for constructive interference of the light reflected from different interfaces.