Module 6 - Device Physics

Dr. Alan Kost

Associate Research Professor of Optical Sciences, University of Arizona

Dr. Alan Kost is an Associate Research Professor of Optical Sciences in the University of Arizona. His areas of research include optical data storage, optical communications, integrated optics, nano-optics.

Email: akost@arizona.edu

Introduction

An optical fiber communication systems is comprised of four essential components: a source that generates the optical signal that is transmitted through the system, an optical waveguide (fiber) that carries the signal, an optical amplifier that periodically boosts the signal, and a receiver that detects the signal at the end of the system. Optical fiber manufacture and design have been treated in prior course modules and amplifiers and receivers will be topics of later sections of this course. In the following section, we will discuss both the basic theory and specific properties of optical sources for fiber communications systems.

6.1 Radiative Processes

Energy Levels

Every atom or molecule has a collection of discrete energy levels that may be irregularly spaced. **Figure 6.1** shows six levels in a hypothetical collection. Each of the energy levels may be comprised of multiple sublevels of the same energy, in which case we say that the energy level is degenerate. For the time being we will ignore degeneracy. As pictured in **Figure 6.1**, the atom or molecule is in level 1.

The authors would like to acknowledge support from the National Science Foundation through CIAN NSF ERC under grant #EEC-0812072.

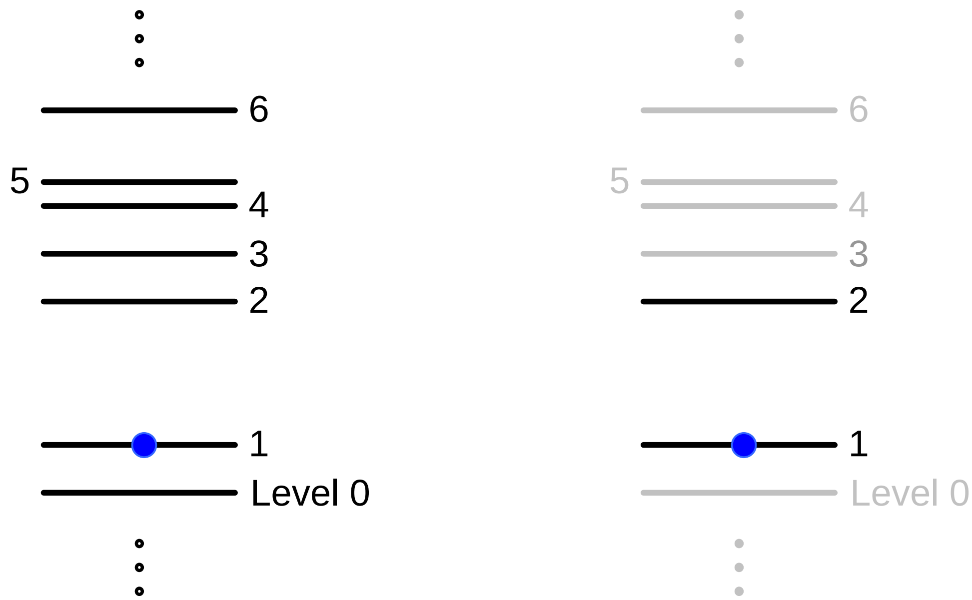


Figure 6.1: *Energy levels for an atom or molecule. Frequently we will discuss the interaction of light with just two of the levels.*

Light generated by a laser source will be of a single frequency (or of a very narrow band of frequencies). As a result, we will call such light monochromatic. The light energy emitted by the source will be in the form of electromagnetic waves (as discussed previously) and we may describe discrete packets of light energy, where $E = h\nu$, as photons. Monochromatic light generated by a laser, and incident on an atom, molecule or collection of atoms or molecules, may interact strongly with only two of the system's energy levels, say levels 1 and 2, if the photon energy equals the difference in the energy of the two levels (i.e. $h\nu = E_2 - E_1$). In this case we say that the light interacts resonantly with levels 1 and 2, and we often focus our attention on these levels and, at least temporarily, ignore the rest. From here on we will refer to an atom or molecule as a particle.

We almost always deal with a collection of particles, each of which may be in a different energy level, as pictured in **Figure 6.2**. Because the number of particles may be very large, it is common to use a visual shorthand and represent the collection with just two levels as in **Figure 6.3**.

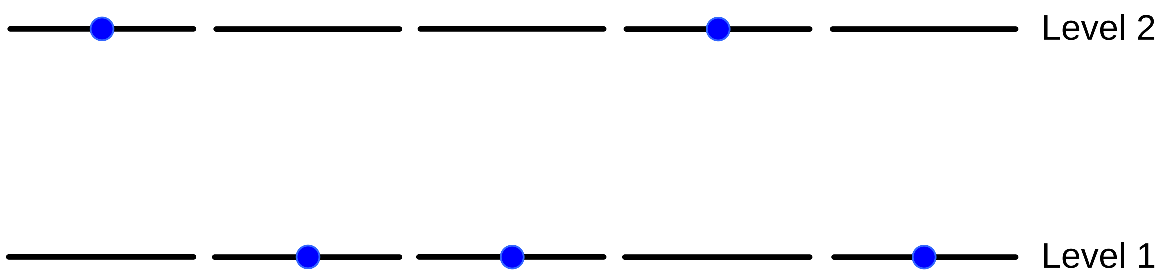


Figure 6.2: *A collection of five particles with energies distributed between the energy levels.*

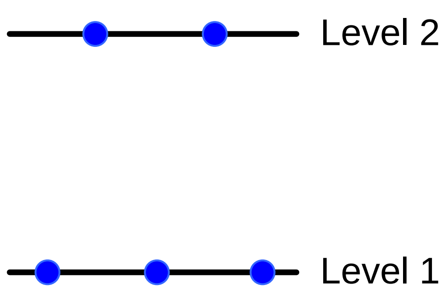


Figure 6.3: *A shorthand version of the diagram in Figure 6.2.*

In order to describe the operation of a laser and other optoelectronic devices, we enumerate three radiative processes involving the interaction of a photon with two energy levels in a particle: absorption, stimulated emission, and spontaneous emission.

Absorption

A photon resonant with two energy levels can cause an electron to transition to a higher energy as pictured in **Figure 6.4**. The process is called photon absorption, and the end result is that an additional particle occupies energy level 2 and there is one less photon in the system. In the language of quantum electrodynamics we say that the photon has been destroyed.

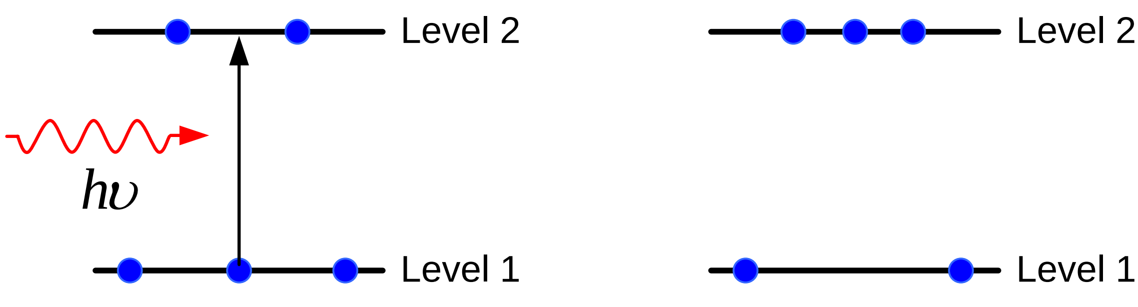


Figure 6.4: *A photon is absorbed as a particle is excited from energy level 1 to energy level 2.*

The strength of the interaction between the photon and the levels is proportional to the dipole moment of the two levels:

$$\mu \equiv \int_{\substack{\text{All}\\\text{Space}}} dr^3 q\, u_2(\vec{r})\,\vec{r}\; u_1(\vec{r}), \qquad \text{(Equation 6.1)}$$

where u_1 and u_2 are the wavefunctions for electrons in levels 1 and 2 respectively. In some cases, it we may find that $\mu = 0$. Then we say that the transition is not allowed. This is an example of a "selection rule" for an optical transition. The strength of all three radiative processes considered in this section are proportional to the dipole moment, μ.

Stimulated Emission

A photon resonant with two energy levels can also stimulate a transition to lower energy as pictured in **Figure 6.5**. The process, called stimulated emission, creates an additional photon.

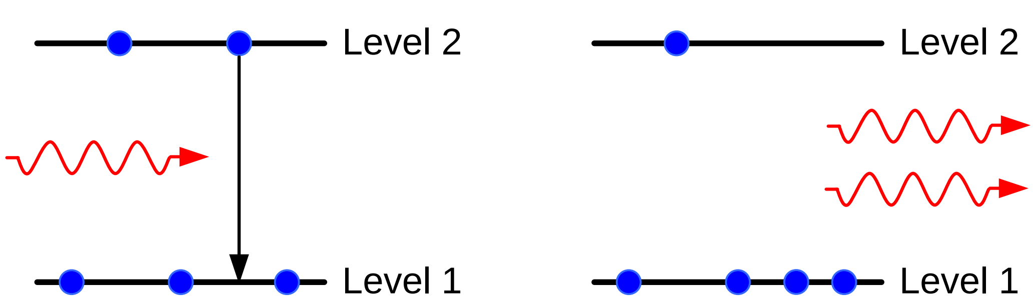


Figure 6.5: *A Photon stimulates a particle to a lower energy level, creating a second photon.*

Spontaneous Emission

Unlike absorption and emission, the third radiative process, spontaneous emission, is not initiated by a photon. It is possible for a particle to undergo a spontaneous transition to lower energy while emitting a photon. Quantum electrodynamics describes this process, pictured in **Figure 6.6**, as a transition initiated by fluctuations of the electromagnetic vacuum.

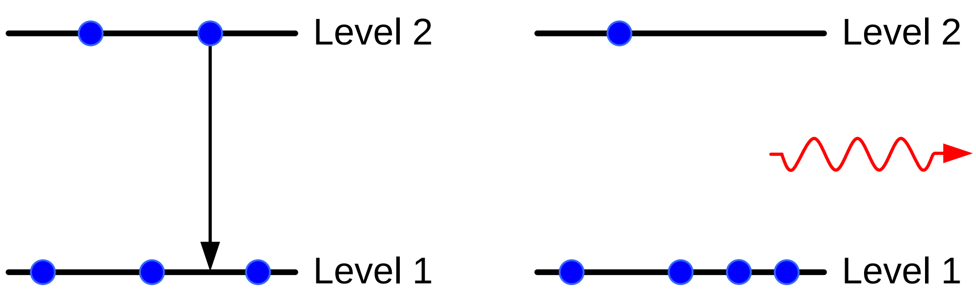


Figure 6.6: ***A particle spontaneously emits a photon.***

Population Inversion

When resonant photons are incident on a collection of particles with a distribution of energies in both level 1 and level 2, we expect all three radiative processes (absorption, stimulated emission, spontaneous emission) to take place. If the flux of incident photons is large, the stimulated transitions (i.e. absorption and stimulated emission) will occur most rapidly and we can neglect spontaneous emission.

If, in addition, more particles are initially in level 2 than in level 1 (**Figure 6.7**), then stimulated emission dominates absorption and a resonant light beam will gain intensity as it traverses the volume occupied by the particles. This is the opposite of what is typically observed when a light beam traverses a medium – absorption. Unless they have been suitably prepared, most media have larger populations in lower energy levels. It is for this reason that we call the condition that produces optical gain a population inversion. To be more quantitative, we say we have a population inversion provided

$$N_2 - N_1 > 0, \qquad \text{(Equation 6.2)}$$

where N_2 is the number of particles per unit volume in level 2 and N_1 is the number of particles per unit volume in level 1. If the energy levels are degenerate (i.e. contain sublevels), then the condition for population conversion and optical gain is

$$g_1N_2 - g_2N_1 > 0, \qquad \text{(Equation 6.3)}$$

where g_1 is the number of sublevels in level 1 and g_2 is the number of sublevels in level 2.

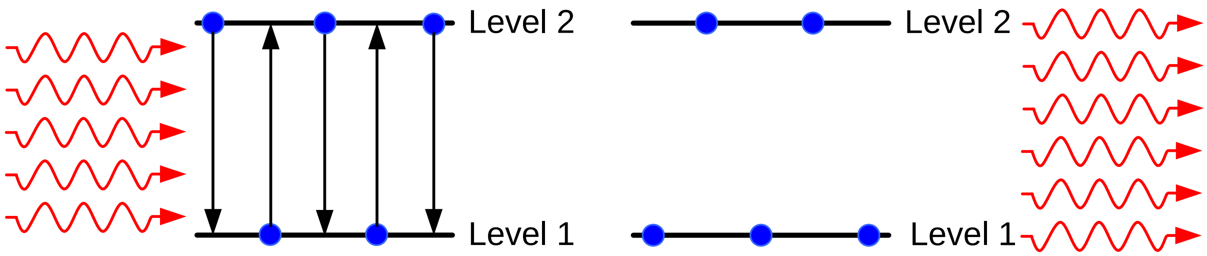


Figure 6.7. *An "inverted" population amplifies light.*

6.2 Semiconductors

Energy Levels

Semiconductors are also characterized by a collection of energy levels. However, the energy levels of interest are grouped into bands of very closely spaced levels (**Figure 6.8**). Bands are approximately one eV in width and contain on the order of 10^8 levels. Each level is doubly degenerate, containing a sublevel that can be filled by a "spin up" electron and a sublevel that can be occupied by a "spin down" electron. Within a band, levels, or states as they are often called, are labeled by a wavenumber k. The quantity $\hbar$k represents the linear momentum of the electron or, more precisely, its crystal momentum.

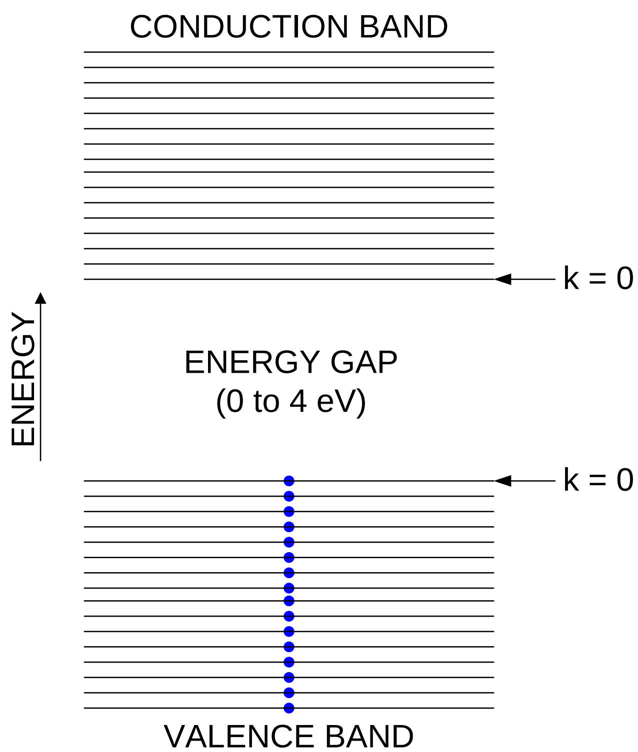


Figure 6.8: ***Energy bands in a semiconductor***

Of particular interest is a pair of bands - a valence band, with energy levels completely (or very nearly completely) filled with electrons, and an empty (or very nearly empty) conduction band that lies just above the valence band **(Figure 6.8**). The "band gap" that lies between the valence and conduction band may have a width that ranges anywhere from 0 to 4 eV, but band gaps of most technological interest are on the order of 1 eV. For energy bands relevant to the operation of a semiconductor laser, the states at the top and bottom of the valence and conduction band, respectively, have k = 0.

Figure 6.9 shows the energy bands of **Figure 6.8** as curves of energy versus electron wavenumber. The curves in **Figure 6.9** give a better feel for how the energy of electrons vary with the value k, and these type of curves are most often used for describing radiative processes in semiconductors. **Figure 6.9** also shows a missing electron in the valence band that is referred to as a hole. In many ways, a hole acts much like an electron but with an opposite electric charge. States in the band structure of a semiconductor that contain "free" or "excess" electrons are often called *n-type* states while those that contain holes (e.g. missing valence electrons) are called *p-type* states.

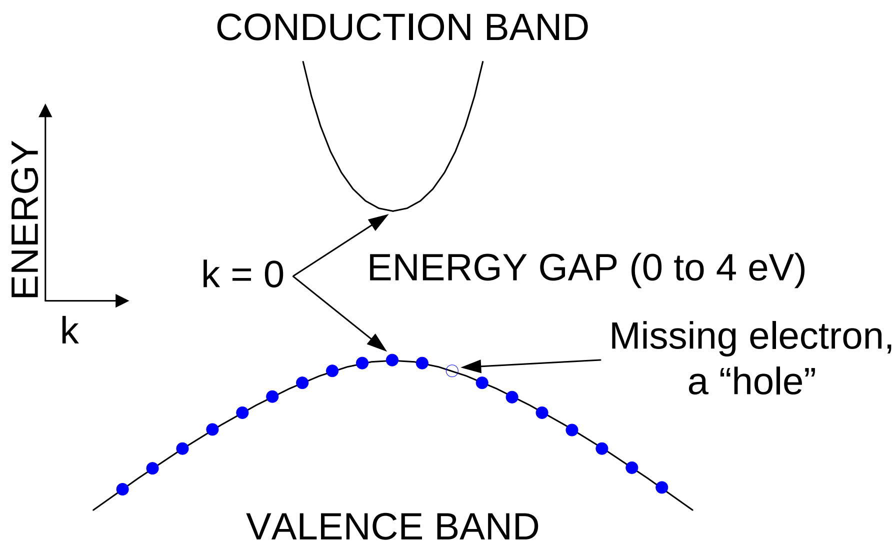


Figure 6.9: ***Valence and conduction bands in the energy versus momentum picture.***

Another advantage of the E versus k curves is that they are better suited for displaying the multiple overlapping valence bands that are present in all semiconductors of interest. **Figure 6.10** shows an E vs. k picture of a conduction band and two overlapping valence bands. The separate valence bands are easily distinguished by their curvature. Interactions with the crystal lattice causes electrons and holes in semiconductors to move as if they have an effective mass that differs from the free mass of the electron. The effective mass is determined by the curvature of the energy band – the greater the curvature, the smaller the effective mass. Thus we call the highly curved valence band in **Figure 6.10** the light-hole band and the less curved valence band the heavy-hole band.

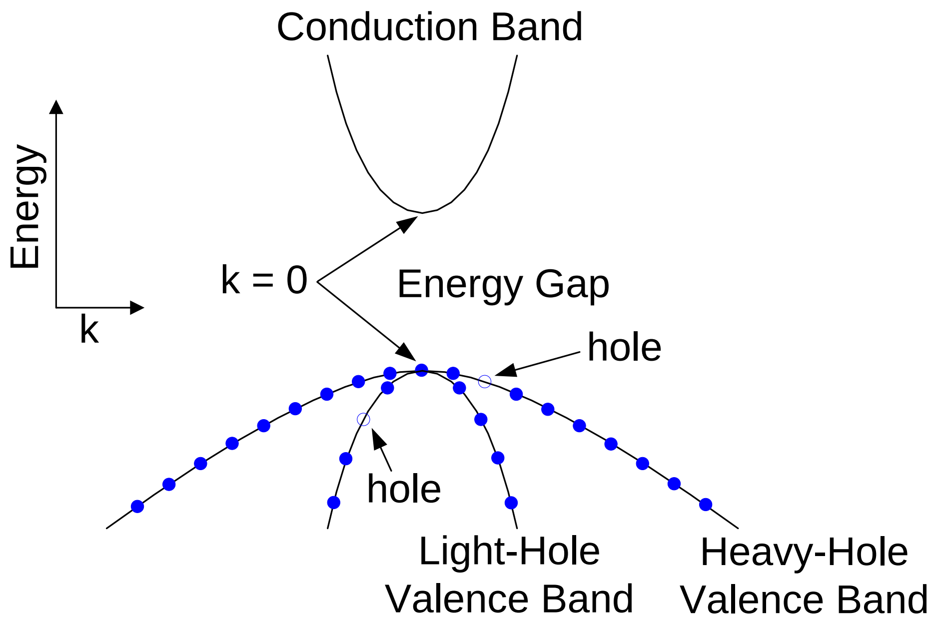


Figure 6.10: ***Energy versus momentum picture with two valence bands.***

Optical Transitions

As with atoms and molecules, radiative processes in semiconductors also include absorption, stimulated emission, and spontaneous emission. The strength of an optical interaction is also determined by the dipole moment between initial and final states, but it is common to describe selection rules that emphasize the importance of conservation of linear momentum. An optical transition between two energy levels in a semiconductor can take place only if

$$\hbar k_i = \hbar k_f \pm \hbar k_{photon}, \qquad \textit{(Equation 6.4)}$$

where $\hbar k_i$ is the linear momentum of the electron in the initial state, $\hbar k_f$ is the electron momentum in the final state, and $\hbar k_{photon}$ is the momentum of the photon involved in the transition. The plus sign in the equation is used when the optical interaction is one of spontaneous or stimulated emission of a photon, while the minus sign is used for photon absorption.

Equivalently we can write momentum conservation in terms of wavenumbers:

$$k_i = k_f \pm k_{photon}. \qquad \textit{(Equation 6.5)}$$

Momentum and wavenumber conservation is illustrated in **Figure 6.11** for the absorption of a photon and transition of an electron from a valence band to a conduction band.

Note that on the scale of the drawing, the wavenumber of the photon is very small so that the arrow in the drawing that connects the initial and final electron states is almost vertical. Speaking approximately, we say that the transition is vertical, and we almost always make the approximation

$$k_i \approx k_f .$$

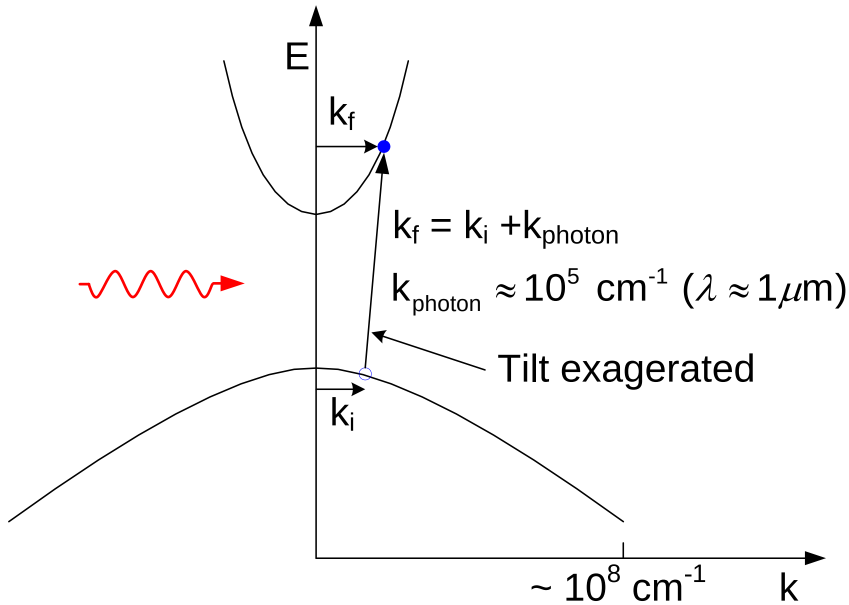


Figure 6.11: *Optical transitions in semiconductors are "vertical".*

Population Inversion

Figure 6.12 is a simplified picture of a population inversion in a semiconductor with just two bands. All valence band states within a range of k values centered about k = 0 are empty, and an

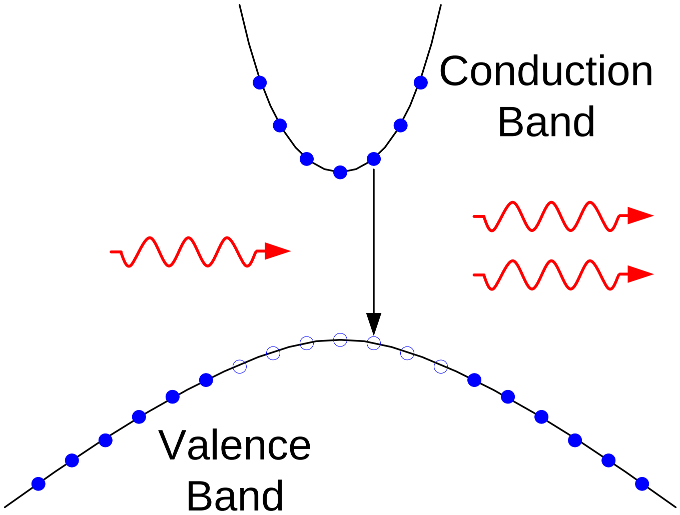


Figure 6.12: *A simple picture of population inversion in a semiconductor.*

equal population of conduction band states about k = 0 are filled. Photons that are resonant with these states stimulate emission of additional photons, and light beams that contain these photons experience optical gain.

A more precise picture of population inversion in semiconductors must include the concept of quasi-Fermi levels (**Figure 6.13**). When there are electrons in the conduction band, the probability that a state of energy E is occupied is given by the Fermi distribution

$$f_c\left(E_c\right)=\frac{1}{1+e^{\frac{(E_c-F_c)}{kT}}},$$ ***(Equation 6.6)***

where F_c is the quasi-Fermi level for the conduction band. Similarly, when there are holes in the valence bands, the probability that a state of energy E is occupied by a hole is

$$f_v\left(E_v\right)=\frac{1}{1+e^{\frac{(F_v-E_v)}{kT}}},$$ ***(Equation 6.7)***

where F_v is the quasi-Fermi level for the valence band.

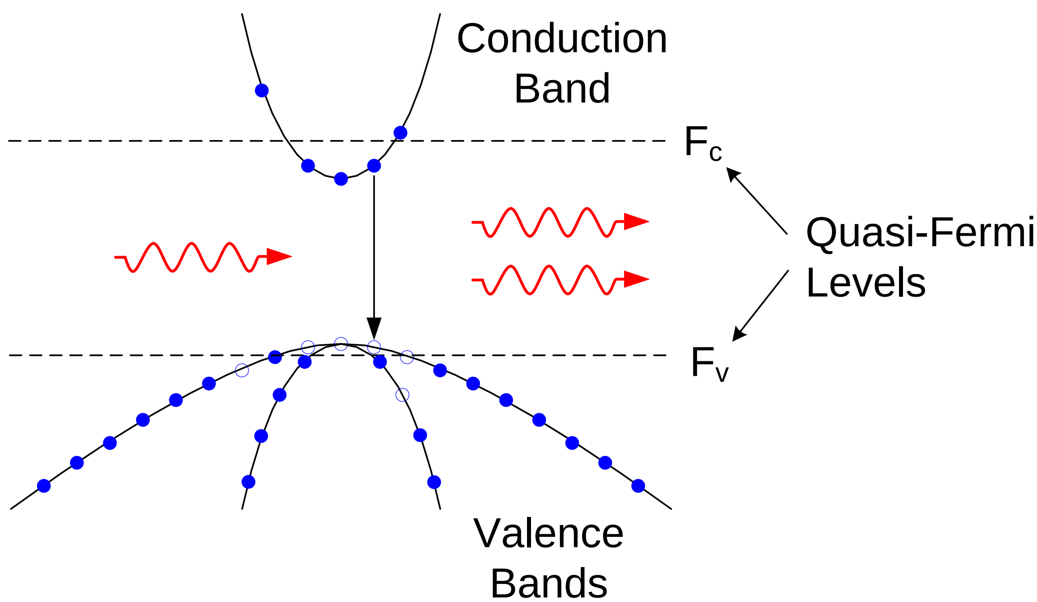


Figure 6.13: *A more precise picture of population inversion in semiconductors includes quasi-Fermi levels.*

The condition for optical gain for light resonant with energy levels E_c and E_v (i.e. $h\nu = E_c - E_v$) is

$$\underbrace{f_c\left(E_c\right)f_v\left(E_v\right)}_{\text{Proportional to the rate of stimulated emission}}-\underbrace{\left(1-f_c\left(E_c\right)\right)\left(1-f_v\left(E_v\right)\right)}_{\text{Proportional to the rate of photon absorption}}>0\cdot$$ ***(Equation 6.8)***

It is not difficult to show that the condition for optical gain implies

$$F_c - F_v > E_c - E_v = h\nu \,. \qquad \textit{(Equation 6.9)}$$

Photons with a wide range of optical wavelengths can interact resonantly with a semiconductor, so whether or not we have optical gain depends on the photon energy. We say we have a population inversion for those photons that experience gain.

The expression

$$F_c - F_v > h\nu \qquad \textit{(Equation 6.10)}$$

is known as the Bernard-Duraffourg relation.

Impurity and Defect States

The behavior of semiconductors is dependent upon their band structures which, in turn, are strongly tied to the crystal structures of the materials. Numerous defects can, and do, exist in semiconductor crystals, and these will impact the band structure and the resulting optical properties of devices made from the materials. Defects in crystal structure can be unintentional, arising from impurities, or intentional, resulting from doping of the material and their effect on the crystal lattice will depend on the type of defect and its charge.

- Substitutional defects: These defects are characterized by the substitution of an impurity or dopant atom for a semiconductor atom in the lattice site of the crystal lattice. If substitutional impurities are present, there are three possibilities, depending on their charge or valence: (1) same valence substitutions cause little change in the band structure; (2) substitutional defects with more electrons become donors; (3) substitutional impurities with fewer electrons become acceptors. If the material contains a relatively small concentration of impurities, the defects form isolated defect states. At higher concentrations, they form defect bands.
- Interstitial defects: In this case the impurity or dopant atom resides in a location in the lattice that rests between existing lattice sites. Regardless of the valence of the interstitial defect, the outer shell electrons are available for excitation to "free" states, so they form donor defects.
- Vacancy defects: The absence of constituents corresponds to missing electrons and forms acceptor defects.
- Frenkel defects: These defects are formed by paired vacancies and interstitials; consequently, the charges are added up to determine whether the defects are donors or acceptors.

Acceptor defects form p-type states and generate excess holes. Donor defects form n-type states and generate excess electrons. Thus, p-type semiconductors have excess holes and n-type semiconductors have excess electrons.

Electrical Pumping

Contact between metals and semiconductors and between different semiconductors results in the flow of free carriers across the interface, since the electron affinity varies from material to material. Semiconductor devices are typically formed by bringing into contact oppositely doped semiconductors, forming an interface called a p-n junction. A p-n junction is used to create a population inversion in the "gain region" of a semiconductor laser or light-emitting diode (LED). **Figure 6.14** diagrams a typical p-n junction for a semiconductor laser. The p-n junction is forward biased by applying a voltage between metal electrodes on p and n-type semiconductor layers. The forward biased diode structure conducts current by injecting holes from the p-type semiconductor, and electrons from the n-type semiconductor, into the undoped gain region, placed between the p and n-type layers.

It is common to use a semiconductor with a relatively large band gap for the p and n-type layers and a semiconductor with a smaller band gap for the gain region. This combination forms a potential well that helps to confine electrons and holes to the gain region. The interface between two types of semiconductors is called a heterojunction. The arrangement of semiconductors in **Figure 6.14** is called a double heterostructure, and a semiconductor laser (or light-emitting diode) that uses the structure is called a double heterostructure laser (or light-emitting diode).

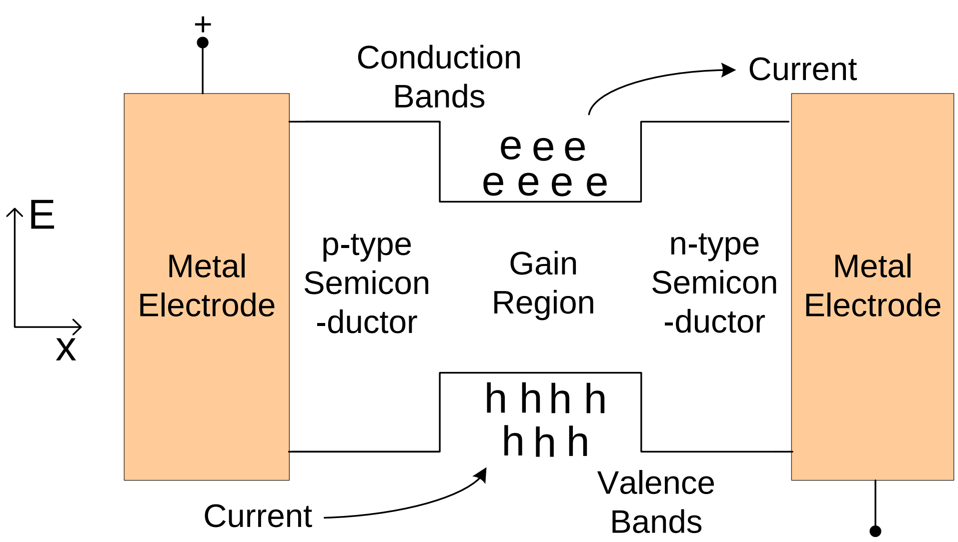


Figure 6.14: *A double heterostructure, p-n junction for creating a pop-ulation inversion in a semiconductor.*

6.3 Non-Radiative Processes

In addition to the radiative processes considered above, there are non-radiative processes that also impact the operation of optical sources such as light emitting diodes and semiconductors lasers. In particular, non-radiative processes decrease optical source efficiency by converting pump energy into unusable and detrimental heat. In this section we will discuss the two most important non-radiative processes and introduce the concept of electron-hole lifetime.

Shockley-Read-Hall Recombination

Shockley-Read-Hall (SRH) recombination is called an "extrinsic" process because it requires the presence of impurity states in the energy gap that would be absent in an ideal crystal. SRH recombination proceeds in the manner illustrated in **Figure 6.15**. A charged particle, say an electron, encounters a trap state in the middle of the forbidden gap of a semiconductor. There the electron waits for a hole to pass nearby. When the hole encounters the electron, they recombine. SRH recombination may produce infrared photons, but it is common to refer to this as a non-radiative process because any photons produced are outside of the region of interest.

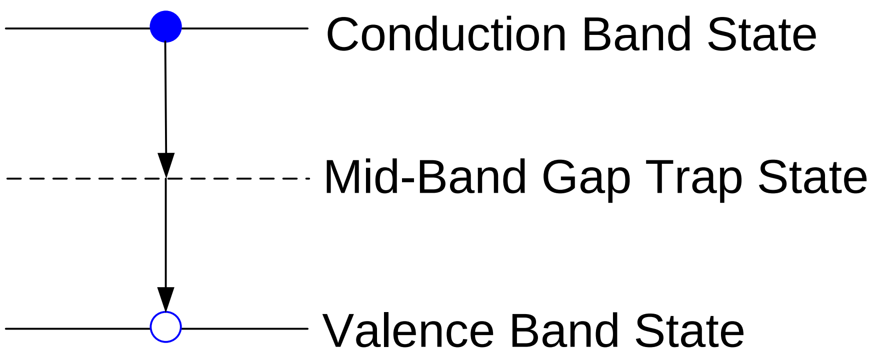


Figure 6.15: ***Shockley-Read-Hall type recombination of an electron with a hole.***

It often turns out that the SRH recombination rate is proportional to density of electrons or the density of holes. A condition of charge neutrality forces equality of the density of electrons and holes (i.e. n=p) in the gain region of a semiconductor laser. Under these conditions it is possible to write:

$$\frac{dn}{dt} = -A_{nr}n \quad \left(\text{or equivalently } \frac{dp}{dt} = -A_{nr}p \right). \qquad \textit{(Equation 6.11)}$$

This relation implies an exponential decay of n in time with a time constant

$$\tau_{SRH} = \frac{1}{A_{nr}} \qquad \textit{(Equation 6.12)}$$

that we will call the Shockley-Read-Hall lifetime. A typical value for SHR lifetime in a high-quality III-V semiconductor is on the order of a nanosecond.

Auger Recombination

Auger recombination is an "intrinsic" recombination mechanism that can occur in an ideal crystal without imperfections. There are several types of Auger recombination. The process illustrated in **Figure 6.16** is called a CCCH Auger process because it involves three electron states in a conduction band and a hole in a valence band.

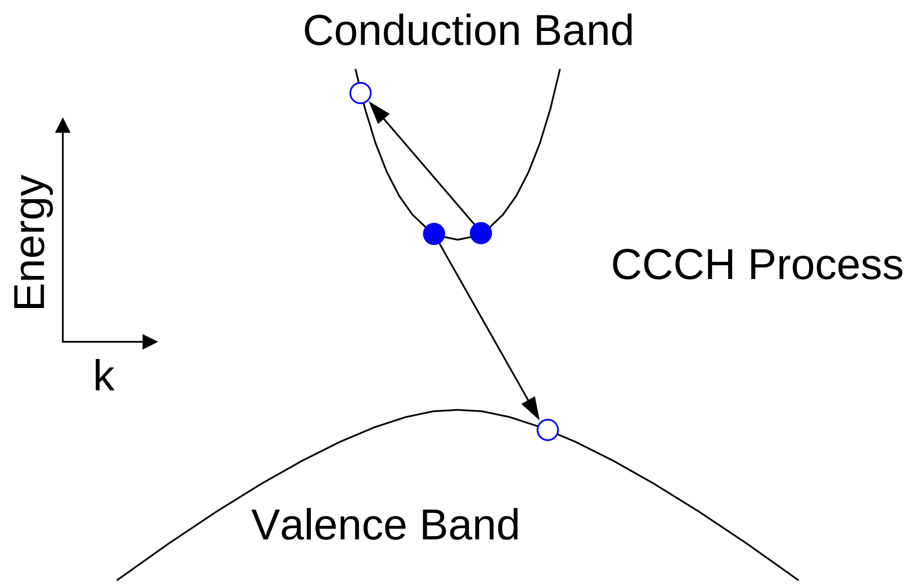


***Figure 6.16**: A CCCH Auger recombination process.*

In the CCCH process an electron in the conduction band recombines with a hole. However,

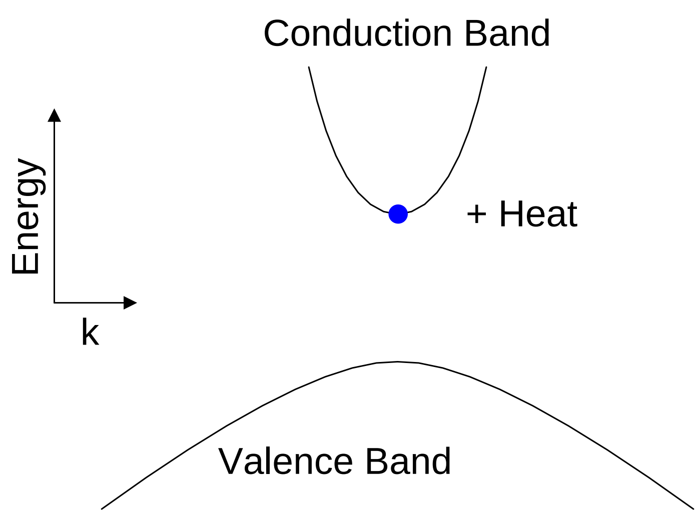


***Figure 6.17:** A CCCH Auger process removes an electron-hole pair from the semiconductor and converts the energy of the pair into heat.*

instead of producing a photon, the energy from the recombination goes to promote a second electron in the conduction band to an energy that is higher in the band. The excited electron makes its way back to the bottom of the conduction band, emitting phonons (heat) on its way. The end result is pictured in **Figure 6.17.**

Auger processes are "three body" processes, and the rate is given by

$$\frac{dn}{dt} = -Cn^3 \qquad \textit{(Equation 6.13)}$$

when n=p. The decay of the population density n due to Auger processes in not exponential, but we can define a density dependent Auger lifetime

$$\tau_{Auger} = \frac{1}{n^2 C}. \qquad \textit{(Equation 6.14)}$$

The last relation shows that the Auger rate increases rapidly and the lifetime drops rapidly with increasing population density. This is why Auger recombination is particularly important for the operation of light emitting diodes and semiconductor lasers with large electron-hole populations injected in the gain region. Auger recombination also tends to be larger for semiconductors with smaller band gaps such as those used for optical communication sources.

Figures of Merit for Laser Pumping Efficiency

Spontaneous emission is a "two body" process with a rate

$$\frac{dn}{dt} = -Bn^2 \qquad \textit{(Equation 6.15)}$$

and a corresponding electron-hole lifetime

$$\tau_{se} = \frac{1}{Bn}. \qquad \textit{(Equation 6.16)}$$

The electron-hole lifetime when all three recombination processes contribute simultaneously is

$$\tau_{tot} = \frac{1}{A_{nr} + Bn + Cn^2}. \qquad \textit{(Equation 6.17)}$$

One can define a figure of merit for pumping efficiency for a light emitting diode based on the electron-hole recombination rates and lifetimes. A suitable figure of merit compares the rate of recombination by spontaneous emission, which produces optical output, with the total rate of electron-hole recombination, which contains processes that produce only heat. The figure of merit is

$$\eta = \frac{Bn^2}{A_{nr}n + Bn^2 + Cn^3} = \frac{\tau_{tot}}{\tau_{se}}. \quad \textit{(Equation 6.18)}$$

For a laser above threshold, one must include stimulated emission which produces photons at a rate

$$\frac{dn}{dt} = -g_m F, \quad \textit{(Equation 6.19)}$$

where g_m is the gain for the light in the laser (discussed in more detail in the module "Sources I") and F is the photon flux in number per unit area. Including stimulated emission the figure of merit for pumping efficiency is

$$\eta = \frac{\beta Bn^2 + g_n F}{A_{nr}n + Bn^2 + Cn^3 + g_n F}, \quad \textit{(Equation 6.20)}$$

where β is called the spontaneous emission factor and is the fraction of spontaneous emission that is in the same direction as the laser light.

6.4 Three Main Components of a Laser

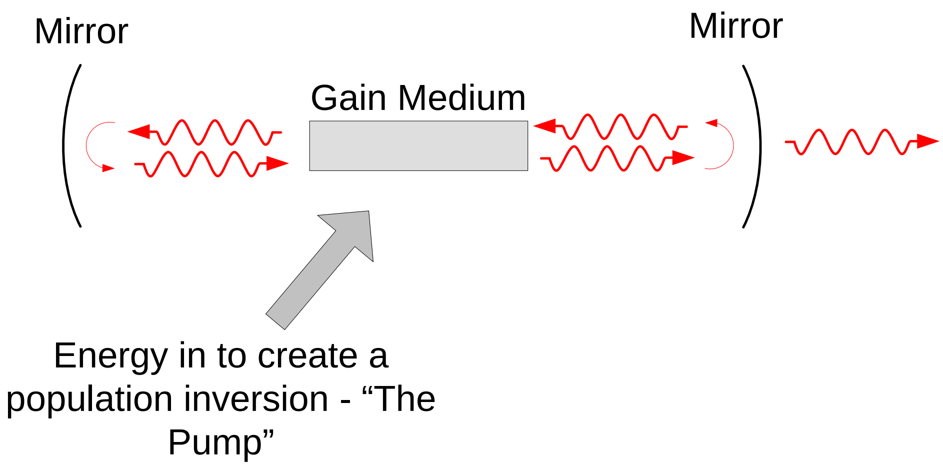


***Figure 6.18**: Lasers have a gain medium, a pump, and mirrors.*

Active Gain Medium

Every laser has a gain medium that emits and amplifies light (**Figure 6.18**). The laser is almost always named after the gain medium. For example, a gas laser with a combination of helium and neon gases as the active medium is called a helium-neon laser. The gain medium can be either gaseous, liquid or solid-state.

Pump

The active gain medium must be supplied energy to produce a population inversion so that medium can amplify light. Energy may be supplied in a variety of forms including optical, electrical, or chemical. We call the device that delivers energy to the gain medium (e.g. a flashlamp) the pump.

Resonator

In lasers, mirrors form a "resonator" or "cavity" that is used to circulate light so that the light passes multiple times through the gain medium, experiencing amplification on each pass. The mirrors are essential for some of the most important properties of a laser such as monochromaticity and directionality. One of the mirrors, the "output coupler" must be partially transmissive in order to allow circulating laser light to escape.

Basic Conditions for Lasing

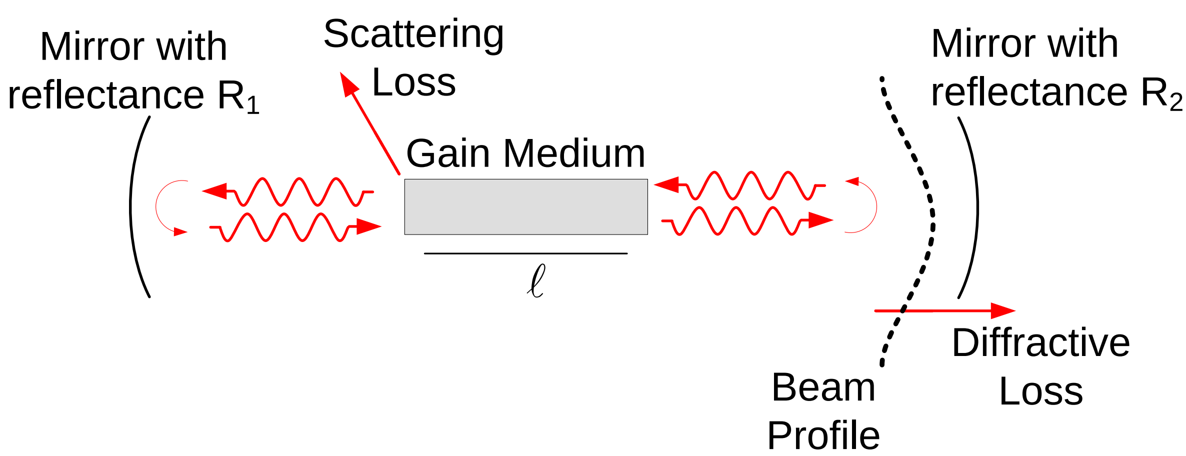


Figure 6.19: ***Gain and loss mechanisms for a laser.***

Lasing occurs when the gain from the gain medium balances optical loss in the laser resonator. For a laser with discrete components, such as the laser pictured in **Figure 6.19,** loss mechanisms include scattering at surfaces, optical loss at the mirrors, which are never 100% reflective, and leakage of the circulating beam out of the resonator when the beam profile extends beyond the edges of a mirror. The quantitative condition for lasing and the balance between gain and loss is

$$G(2\ell)R_1R_2(1-L_i)^2 = 1, \qquad \textbf{(Equation 6.21)}$$

where $G(2\ell)$ is the optical gain after two passes through the gain medium (i.e. $P_{final} = G(2\ell)P_{initial}$), R1 is the reflectance of the rear mirror, R2 is the reflectance of the front mirror, and L_i is the additional optical loss after a single (one way) pass of the resonator.

It is common to define a gain coefficient g with

$$G(z) \equiv e^{gz}. \qquad \textit{(Equation 6.22)}$$

The gain coefficient has units of inverse length and is a property of a material medium that, like the absorption coefficient, is independent of the dimensions of the medium. In terms of g, the basic balance condition for lasing is

$$e^{2g\ell} R_1 R_2 (1 - L_i)^2 = 1 \qquad \textit{(Equation 6.23)}$$

6.5 Semiconductor Lasers

Loss Mechanisms

As pictured in **Figure 6.20**, the gain medium for the most common types of semiconductor lasers extends the entire length of the optical resonator. The mirrors (i.e. reflectors) that circulate light in the semiconductor are the surfaces (or "facets") of the semiconductor. An uncoated semiconductor surface reflects approximately 30 % of incident light. Dielectric coatings are added to the semiconductor surfaces to increase or decrease reflectance.

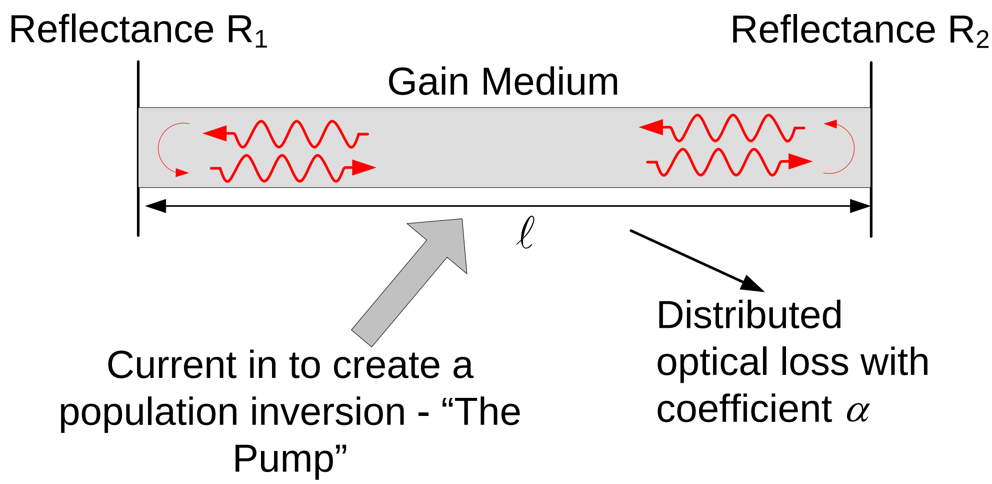


Figure 6.20: *Electrical current is the pump source for most semicon-ductor lasers. Optical loss is distributed throughout the gain medium.*

Most semiconductor lasers are pumped with electrical current. Furthermore, there are significant contributions to optical loss inside the gain medium. These losses are distributed throughout the medium and characterized with an optical loss coefficient α, which, like the gain coefficient, has units of inverse length.

Lasing Conditions

The balance equation for a semiconductor laser of the type pictured in **Figure 6.20** is

$$e^{2g\Gamma\ell} R_1 R_2 e^{-2\alpha\ell} = 1, \qquad \textit{(Equation 6.24)}$$

which can also be written

$$e^{2(g\Gamma-\alpha)\ell}R_1R_2=1 \qquad \textit{(Equation 6.25)}$$

where Γ is called the "confinement factor". The confinement factor is a number that varies between zero and one and is included to take into account the imperfect overlap between the interaction volume of the circulating optical beam within the gain medium. The quantity is considered in more detail in the following section.

The Optical Confinement Factor

Care must be taken when writing the balance equation for semiconductor lasers because it is almost always true that there is light circulating in the laser that does not overlap the gain region, as pictured in the side view of a semiconductor laser in **Figure 6.21**.

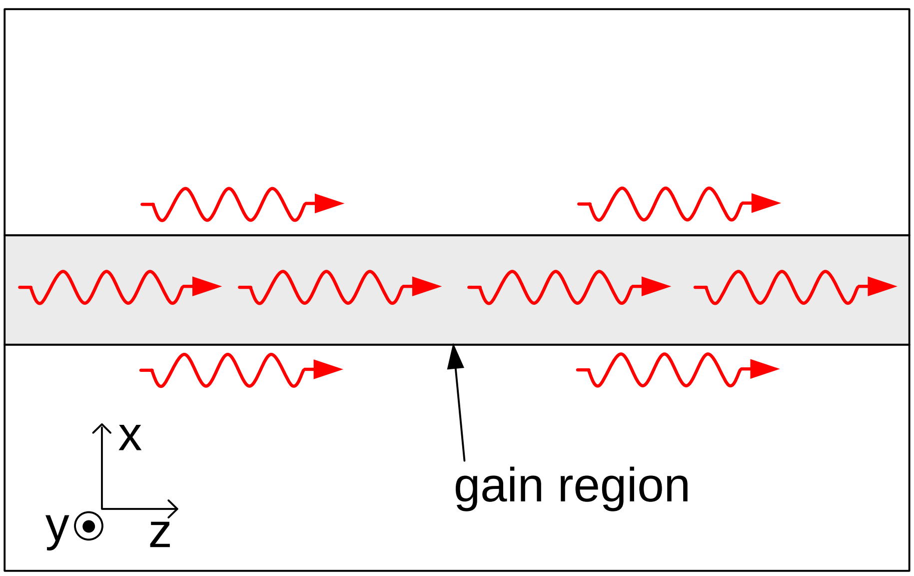


Figure 6.21: *Semiconductor lasers contain light outside of the gain region.*

The perspective of **Figure 6.22** is of an observer looking down the length of a semiconductor laser. The red ellipses indicate the profile of the optical beam in the laser with each ellipse being a curve of constant electric field strength.

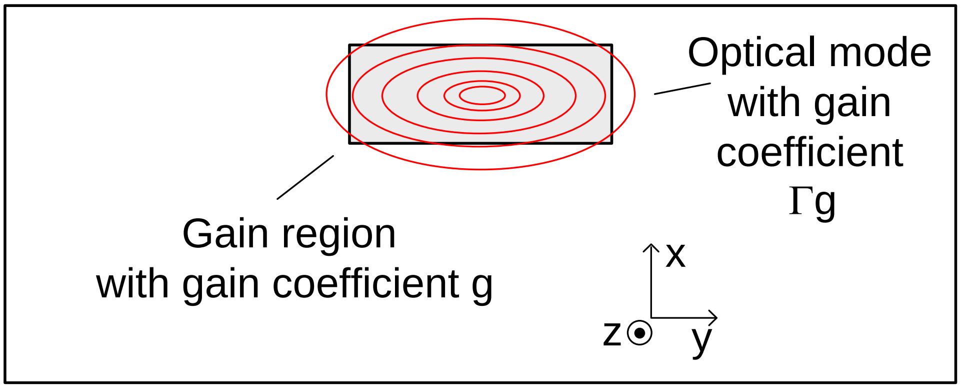


***Figure 6.22**: The overlap of an optical beam with the gain region defines the confinement factor.*

Intuitively we would surmise that the greater the overlap of the optical beam with the gain region, the greater the optical gain experienced by the beam. We can make this thought somewhat more quantitative by defining the optical confinement factor for the beam/gain region:

$$\Gamma \equiv \frac{\int_{\substack{\text{cross section}\\ \text{of gain region}}} \left|E^+\right|^2 dA}{\int_{\text{Entire plane}} \left|E^+\right|^2 dA}, \quad \textit{(Equation 6.26)}$$

where E^+ is the electric field for circulating light traveling in the positive z direction, the integral in the numerator is over a cross section of the gain region in the x-y plane, and the integral in the denominator is over the entire x-y plane.

It is also common to introduce an effective gain coefficient for light in the laser, called the modal gain coefficient, which is related to the gain coefficient in the gain region by the expression

$$g_m = \Gamma g \,. \quad \textit{(Equation 6.27)}$$

In terms of the modal gain coefficient, the laser balance equation is

$$e^{2(g_m - \alpha)\ell} R_1 R_2 = 1 \quad \textit{(Equation 6.28)}$$

Using the Balance Equation

Problem

A semiconductor laser has a length of 250 μm. We find that by injecting current into the laser that we can create a gain coefficient of 250 cm^{-1} in the gain region. The optical confinement factor is 0.3. The gain region has imperfections that contribute a distributed optical loss α = 40 cm^{-1}. The front mirror of the laser is uncoated. What must be the reflectance of the back surface to achieve lasing?

Solution

For lasing to occur, gain must balance loss. From the balance equation we find

$$R_1 = \frac{1}{R_2} e^{-2(g_m - \alpha)\ell} = \frac{1}{R_2} e^{-2(g\Gamma - \alpha)\ell} . \qquad \textbf{\textit{(Equation 6.29)}}$$

$$= \frac{1}{0.3} e^{-2(250 \cdot 0.3 - 40)250\times10^{-4}} = 58\%$$

Output Power

In this section we consider the effects of increasing the pumping current for a semiconductor laser from zero to a value greater than what is required for lasing. The results are illustrated In **Figure 6.23**. Initially, the electron density in the active region, the gain coefficient, and the light output increase roughly linearly with current. The light output is relatively low and due primarily to spontaneous emission. In this regime, the device is operating as a light emitting diode.

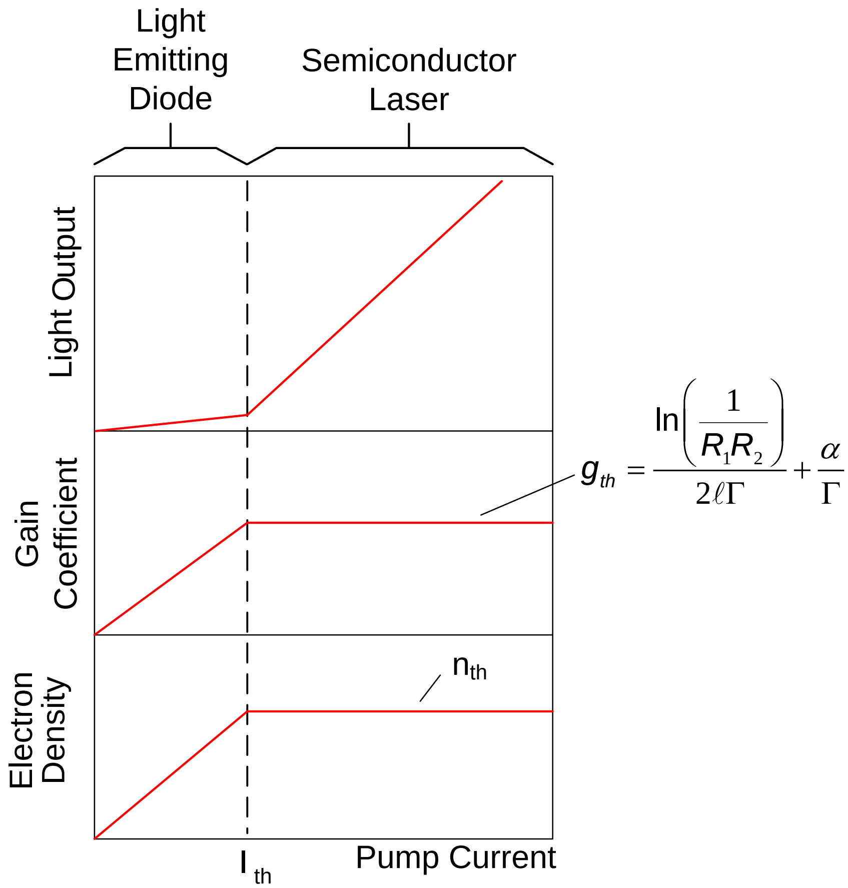


***Figure 6.23** Above threshold the electron density, n, and gain coefficient, g, are clamped at values determined by the gain-loss balance equation while light output increases linearly.*

When the pump current reaches a threshold value the gain-loss balance condition for lasing is met. Above this value for current, additional energy from the electrical pump goes into increasing output from the laser, primarily by stimulated emission. The electron density and gain coefficient are clamped at the threshold values. This last result may seem counterintuitive, but consider the following argument. If n and g were to increase beyond their threshold values, then gain would exceed loss. In this case, the light circulating in the laser would increase upon each round trip – and the optical power of the laser would continue to increase without limit. The curves in **Figure 6.23** represent steady state values, and a condition for steady state operation of the laser is that gain just balances loss.

The Phase Condition for Lasing and the Laser Spectrum

For lasing to occur, circulating light must add constructively to itself after a round trip in the resonator. We will make this statement quantitative with the aid of **Figure 6.24**.

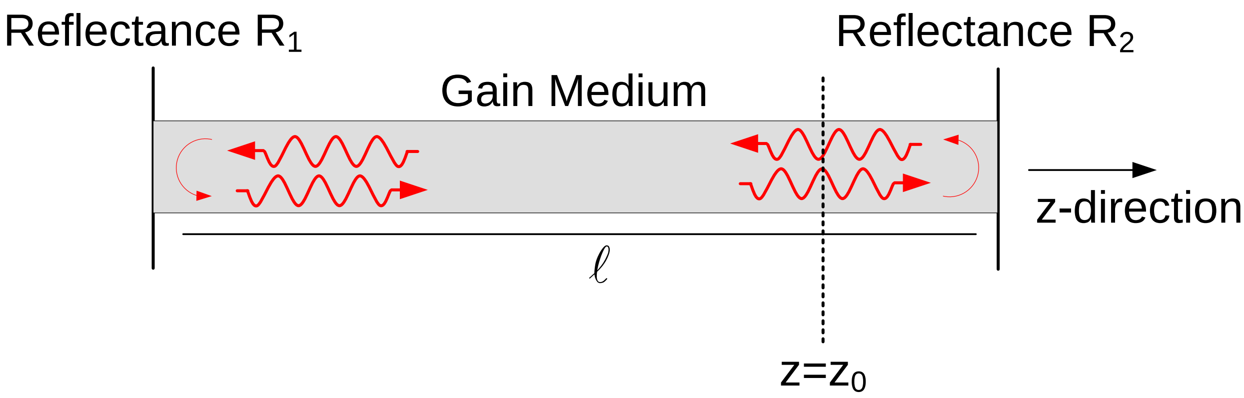


***Figure 6.24**: To establish a phase condition for lasing we examine the phase of the circulating light at the plane z = z_0.*

The electric field propagating in the forward direction in the laser is

$$E = E(x,y)e^{j\left(\frac{2\pi n_{eff}}{\lambda_0}z-\omega t\right)}, \quad \textit{(Equation 6.30)}$$

where E(x,y) is the transverse beam profile and $\frac{2\pi n_{eff}}{\lambda_0}z - \omega t$ is the phase of the wave. Because the light propagates in both the gain region and the surrounding semiconductor, the phase of the wave is determined by an effective refractive index n_{eff} that can be thought of as weighted average of the material index over the beam profile.

At the point z = z_0 and time t_0 the electric field is

$$E(z = z_0, t_0) = E(x,y)e^{j\left(\frac{2\pi n_{eff}}{\lambda_0}z_0-\omega t_0\right)}, \quad \textit{(Equation 6.31)}$$

and after one round trip in the cavity the field is

$$E(z = z_0 + 2\ell, t_0 + \Delta t) = E(x,y)e^{j\left(\frac{2\pi n_{eff}}{\lambda_0}(z_0+2\ell)-\omega(t_0+\Delta t)\right)}.$$

(Equation 6.32)

In order for the beam to add constructively to itself we must have

$$e^{j\left(\frac{2\pi n_{eff}}{\lambda_0}(z_0+2\ell)-\omega(t_0+\Delta t)\right)} = e^{j\left(\frac{2\pi n_{eff}}{\lambda_0}(z_0)-\omega(t_0+\Delta t)\right)}, \qquad \textit{(Equation 6.33)}$$

and this gives the phase condition

$$\frac{2\pi n_{eff}}{\lambda_0} 2\ell = p2\pi \qquad \textit{(Equation 6.34)}$$

where p is a positive integer. The phase condition can be rearranged to give the discrete set of wavelengths at which a laser can operate:

$$\lambda_{0,p} = \frac{2n_{eff}\ell}{p}, \quad p = 1, 2, 3 \ \cdots . \qquad \textit{(Equation 6.35)}$$

We call a circulating light beam with one of these wavelengths a longitudinal mode for the laser. A semiconductor laser will often operate in several of longitudinal modes simultaneously as illustrated in **Figure 6.25**. The frequencies for the longitudinal modes are

$$\nu_p = \frac{c}{2n_{eff}\ell} p \cdot \qquad \textit{(Equation 6.36)}$$

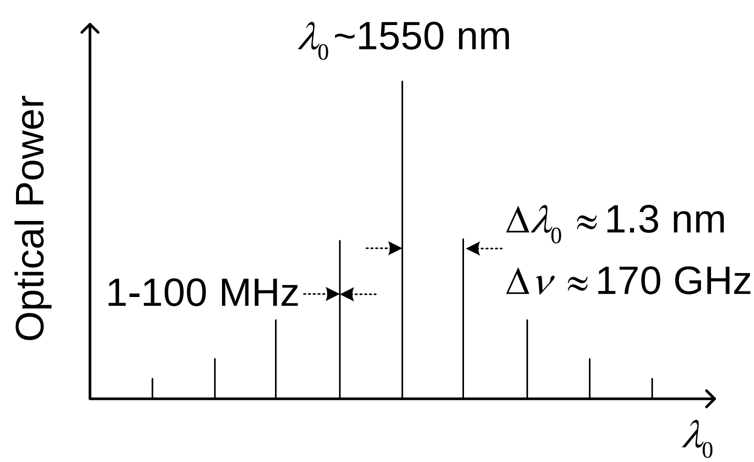


Figure 6.25: ***A semiconductor laser can operate in several longitudinal modes.***

The resonant frequencies are uniformly spaced:

$$\Delta\nu = \nu_p - \nu_{p-1} = \frac{c}{2n_{eff}\ell} \text{ (independent of p).} \qquad \textit{(Equation 6.37)}$$

Notice that the number of possible longitudinal modes for a given resonator is large (essentially infinite as p increases). However, a given gain medium can only emit light over certain, specific, spectral ranges – depending on the band structure of the gain medium. The emission band for a given medium is known as its gain bandwidth ($\Delta\nu_{gain}$). For this reason, the number of possible emission lines for a laser is determined both by the longitudinal mode spacing, as determined by the geometry of the laser resonator cavity, and also by the gain bandwidth of the gain medium. This is seen clearly in **Figure 6.25** where the number of longitudinal modes and their amplitudes are determined by the gain bandwidth envelope for the semiconductor gain medium. For a 250 μm long semiconductor laser and an effective index of 3.5, for example, the spacing of longitudinal modes is about 170 GHz. The wavelength spacing of the longitudinal modes varies slowly, and if the laser operates at approximately 1550 nm, the longitudinal modes are separated by about 1.3 nm. The spectral width of the individual modes is of the order of one to 100 megahertz. The gain bandwidth allows for the oscillation of 9 longitudinal modes, as seen in the figure.

6.6 Types of Semiconductor Lasers

Buried Heterostructure Lasers

One of the most commonly used semiconductor lasers for optical fiber communications is the buried heterostructure laser. **Figure 6.26** shows the cross section of a typical buried heterostructure device. An intrinsic (undoped) gain region is sandwiched between p and n-type layers, forming a forward biased diode during laser operation. A p-type InGaAsP layer is used to make a low resistance contact to the upper metal electrode.

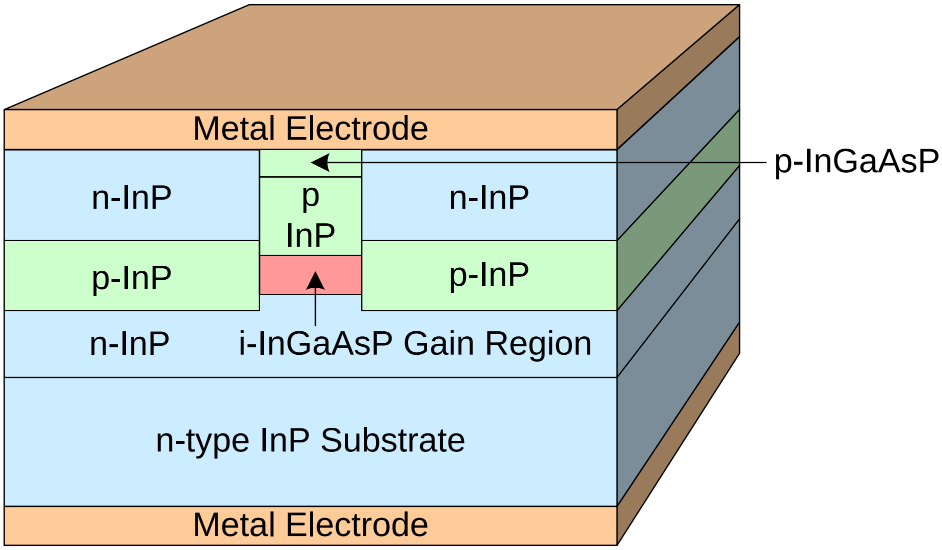


***Figure 6.26:** A buried heterostructure semiconductor laser.*

Figure 6.27 shows the flow of current injected into the diode when applying a positive voltage to the upper electrode or a negative voltage to the lower electrode. The n-InP and p-InP layers just below the top electrode on either side of the gain region are reversed biased and current blocking – concentrating the current flow through the gain region. The gain region fills with electrons in the conduction band and holes in the valence band, creating a population inversion and optical gain. Layers have been removed in the schematic of **Figure 6.28** to provide a better view of the gain region. This region is formed from InGaAsP, or a similar material, and has a refractive index that is larger than the surrounding InP. Light tends to be confined to materials of larger refractive index so the gain region is an optical waveguide in which the laser light circulates from one end to the other.

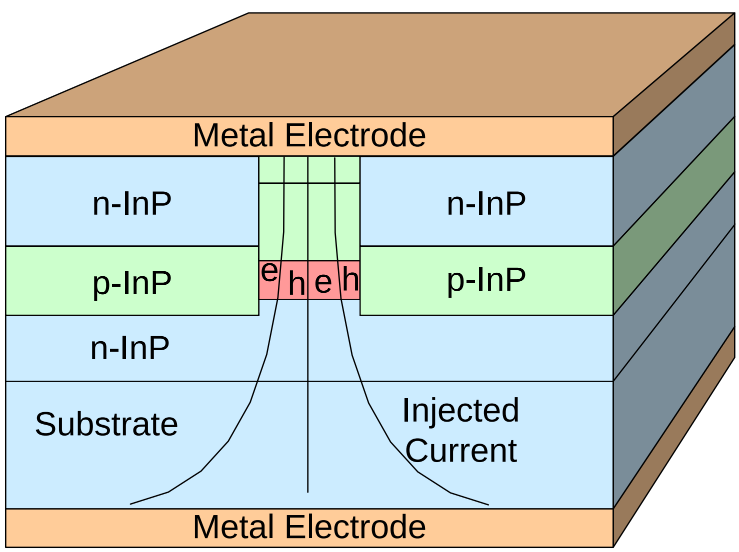


***Figure 6.27:** Current flow in a buried heterostructure semiconductor laser.*

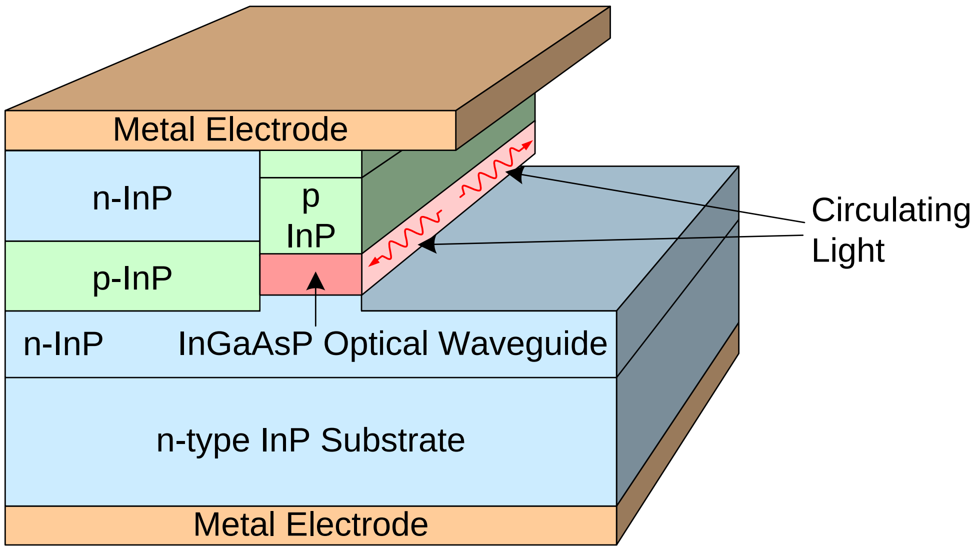


Figure 6.28: ***The InGaAsP gain region has a high refractive index than the surrounding material and forms the core of an optical waveguide that confines light that circulates from end to end.***

Distributed Feedback Lasers

The semiconductor laser pictured in **Figures 6.26-6.28** is called a "Fabry-Perot" type because light circulates between two planar reflectors separated by a medium that has a uniform refractive index along an axis perpendicular to the reflectors. A drawback of the Fabry-Perot structure is that it can lase in multiple longitudinal modes so that the output spectrum may span 10 nm or more. The wideband spectral output can lead to pulse spreading and signal degradation in optical communication links. For this reason, it is desirable to modify the Fabry-Perot structure so that the laser operates in a single longitudinal mode.

The basic idea for modifying the frequency spectrum of a semiconductor laser is to introduce a frequency selective element in the resonator as pictured in **Figure 6.29**. The selective element may or may not overlap the gain medium but it must overlap the circulating optical beam.

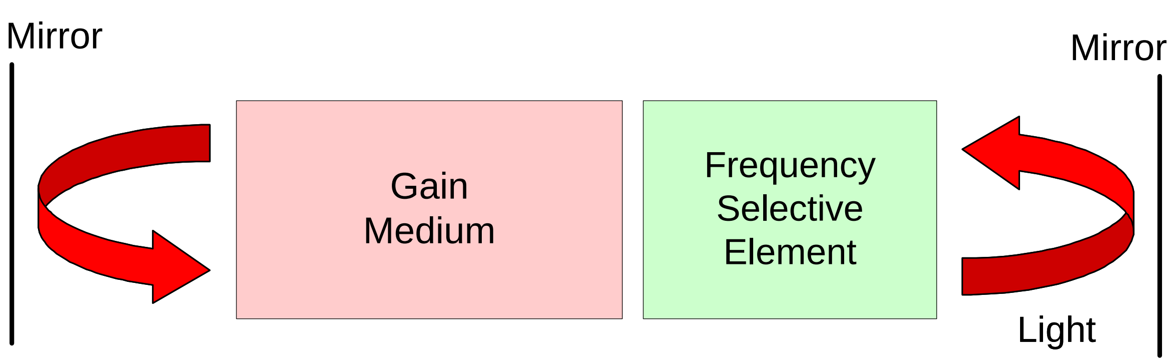


Figure 6.29: *A frequency selective element is introduced into a semiconductor laser to narrow the output spectrum.*

A widely used frequency selective element for semiconductor lasers is a periodic modulation to the refractive index of the laser material. The idea is illustrated in **Figure 6.30**. The drawing at the top left of **Figure 6.30** is for a semiconductor structure that is similar to the Fabry-Perot structure of **Figures 6.26-6.28**. Thin layers of n and p-type InGaAsP have been added to either side of the intrinsic InGaAsP active region. The purpose of these additional layers can be understood with the aid of the cross section of the laser that is pictured in **Figure 6.30**. The InP below the gain region has been etched to form a corrugation at the surface. InGaAsP is deposited on top of the corrugation. There is a periodic modulation along the length of this "grating" because InP and InGaAsP have different refractive indices. The grating is the frequency selective unit that is used to narrow that laser's spectral output.

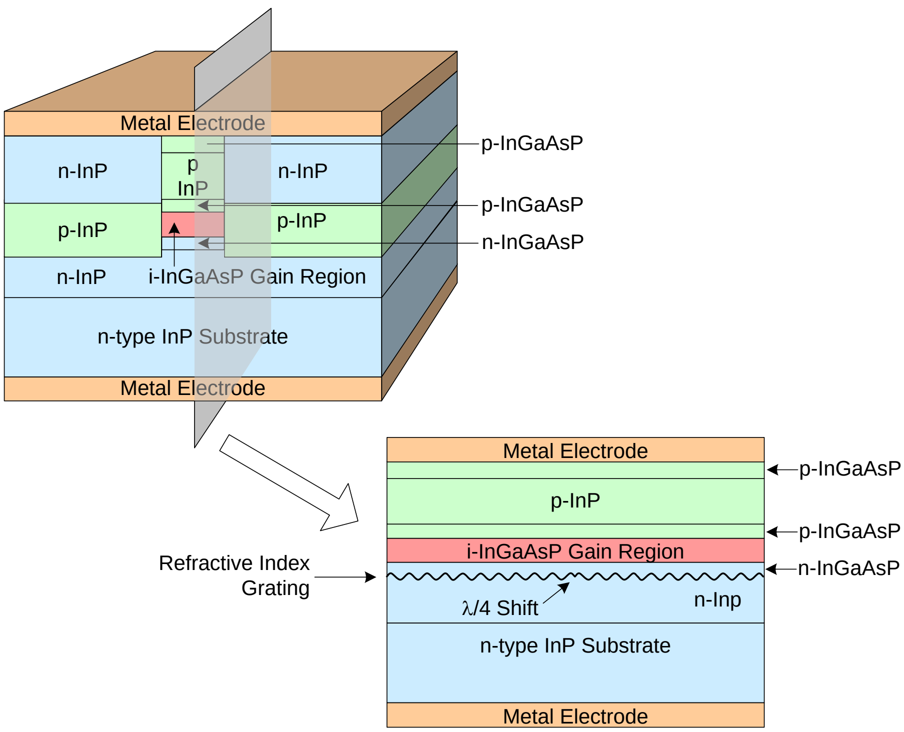


Figure 6.30 *A cross section of a distributed feedback (DFB) semiconductor laser.*

The grating is produced in the added n-InGaAsP, instead of the active region, in order to maintain high crystalline quality in the active region. Nevertheless, the circulating optical beam extends beyond the active region and overlaps the grating. The circulating light is continuously reflected back onto itself and the phase condition for constructive interference is

$$\frac{\lambda_0}{n_{eff}} = \Lambda, \qquad \textit{(Equation 6.38)}$$

where Λ is the period of the refractive index grating. With a uniform grating, the phase condition allows the laser to operate in two modes that see very slightly different effective refractive indices. In order to ensure laser operation in a single longitudinal mode and the narrowest possible output spectrum, a Λ/4 shift is added to the grating as indicated in the cross section of **Figure 6.30.**

Distributed Bragg Reflector Lasers

In a distributed Bragg reflector (DBR) laser, the gain medium is physically separated from the frequency selective element. The selective element is a stack of quarter wavelength layers of alternating refractive index with a very narrow reflectance band that serves as the rear mirror of the laser. The laser operates in the one longitudinal mode of the gain region that is reflected by the Bragg stack.

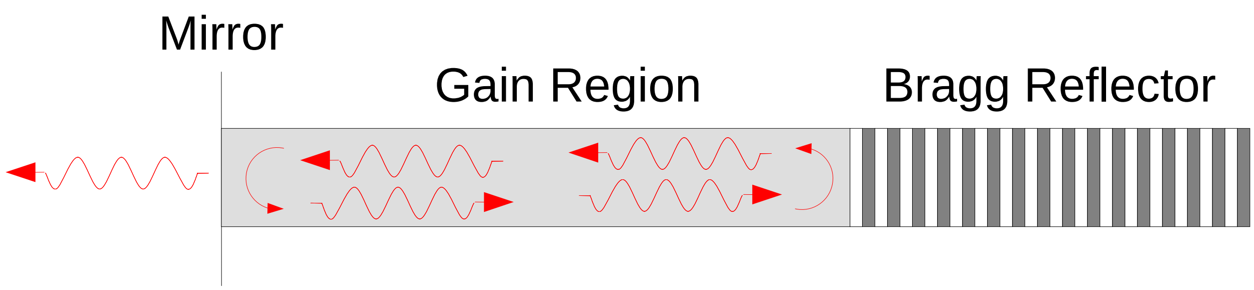


Figure 6.31: ***A single mode distributed Bragg reflector laser.***

The separate Bragg reflector used in a DBR laser is a somewhat less selective element than the grating that is integrated into the gain region of a DFB laser. For this reason, DFB lasers are the preferred choice for optical communication links when frequency stability is a primary consideration. However, the use of a separate Bragg reflector allows more flexibility in the design of a semiconductor laser, and the DBR laser is the basis for a tunable laser described in the next section.

Tunable Distributed Bragg Reflector Lasers

A tunable distributed Bragg reflector laser (**Figure 6.32**) has electrodes for injecting current into each of three segments - a gain region, a Bragg reflector, and a phase control region. The current to the gain region supplies the pump power that allows for lasing. The current to the Bragg reflector adjusts the center wavelength of the reflector, and thus, the wavelength at which the laser can operate.

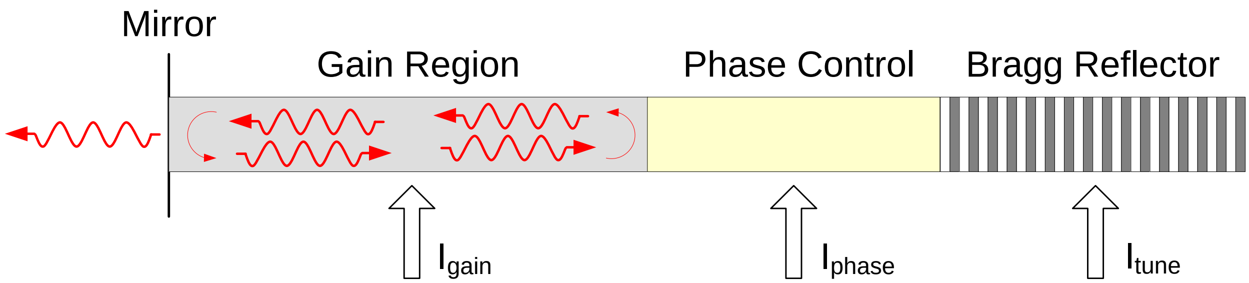


***Figure 6.32**: A tunable distributed Bragg reflector laser.*

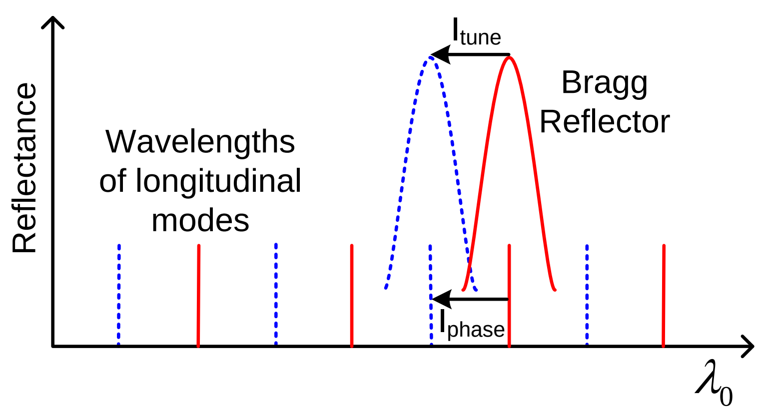
The physical mechanism that adjusts the wavelength of the Bragg reflector is a carrier density dependent refractive index. The tuning current injects charge into the Bragg reflector that decreases the refractive index of the layers by an amount

$$\Delta n_r(\lambda_0) = -\frac{q^2\lambda_0^2}{8\pi^2 n_r(\lambda_0)\varepsilon_0 c^2}\left(\frac{n}{m_e} - p\frac{m_{lh}^{1/2} + m_{hh}^{1/2}}{m_{lh}^{3/2} + m_{hh}^{3/2}}\right),$$

(Equation 6.39)

where q is the electronic charge, λ_0 is the optical wavelength, $n_r(\lambda_0)$ is the refractive index at λ_0, n and p are the density of injected electrons and holes respectively, and m_e, m_{lh}, and m_{hh} are the effective masses for electrons, light-holes and heavy-holes respectively. The center wavelength of the reflector is four times the optical thickness of the Bragg layers, so an increase in charge carrier density in the layers decreases the center wavelength.

The way in which the tuning current and the phase control work together can be made clear with the aid of **Figure 6.33**. Current to the Bragg reflector sets the position of the high reflectance band that determines the lasing wavelength. However, lasing can only take place if there is a



1

***Figure 6.33:** The phase current positions a longitudinal mode in the high reflectance band of the Bragg reflector.*

Longitudinal mode that coincides with this wavelength. The phase current modifies the optical path length of the laser resonator and thus the location of the longitudinal modes. For proper operation of the tunable laser, the phase current is adjusted to position a longitudinal mode at the Bragg reflectance maximum.

Vertical Cavity Surface Emitting Lasers (VCSELs)

The structure of a surface emitting laser is pictured in **Figure 6.34**. This particular structure is for a semiconductor laser known as a vertical cavity surface emitting laser or VCSEL (pronounced "vick sell").

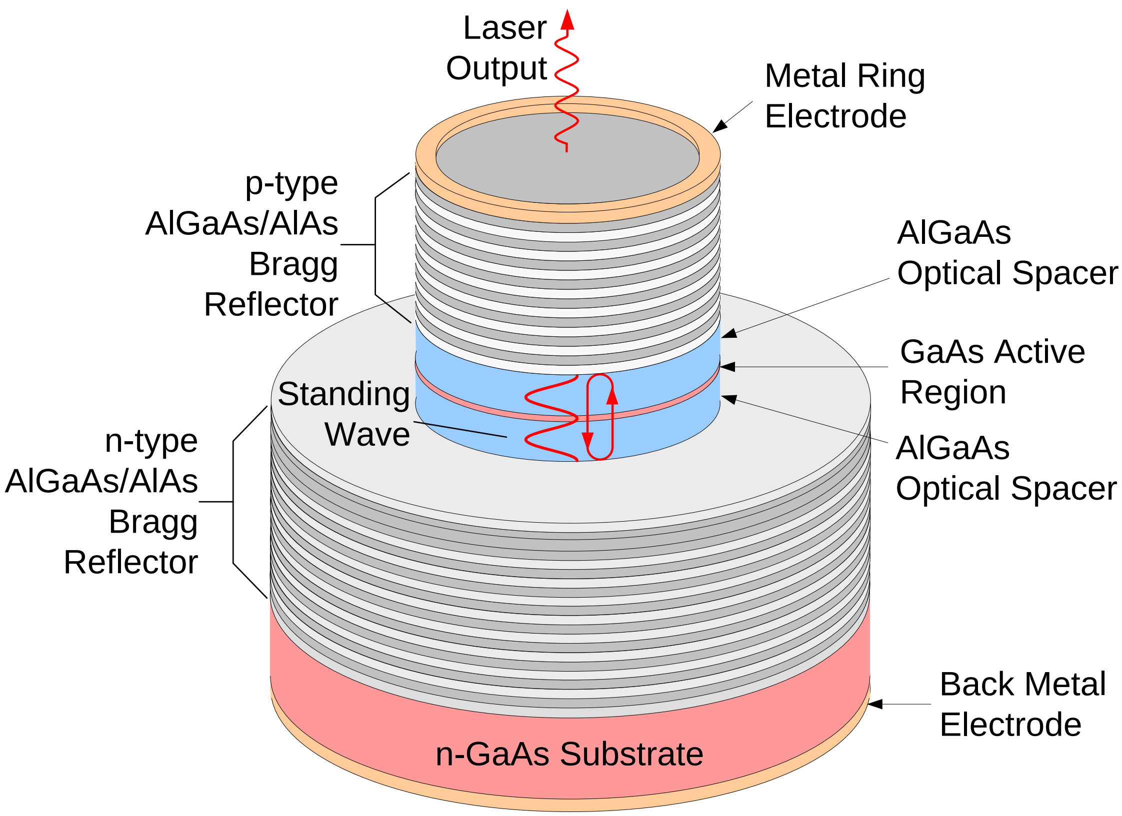


***Figure 6.34:** A GaAs surface emitting laser.*

The gain region of the VCSEL is a layer of GaAs. The mirrors for the laser are distributed Bragg reflectors formed by depositing stacks of quarter wavelength layers of AlGaAs (high index) and AlAs (low index) material. The stacks are doped p and n-type to form a p-n junction for electrical pumping of the laser. Light circulates between the two Bragg stacks to be amplified by the GaAs gain region. Optical spacers are used to adjust the optical thickness of the region between the Bragg stacks to an integer number of wavelengths. This ensures that the circulating light forms a standing wave pattern with an intensity maximum at the gain region and maximizes the interaction of the light with the gain region.

Because of the very short resonator length, VCSELs operate in a single longitudinal mode. As can be seen in **Figure 6.35,** the short resonator produces widely spaced longitudinal modes separated by approximately 100 nm. Only one of these modes can experience high gain.

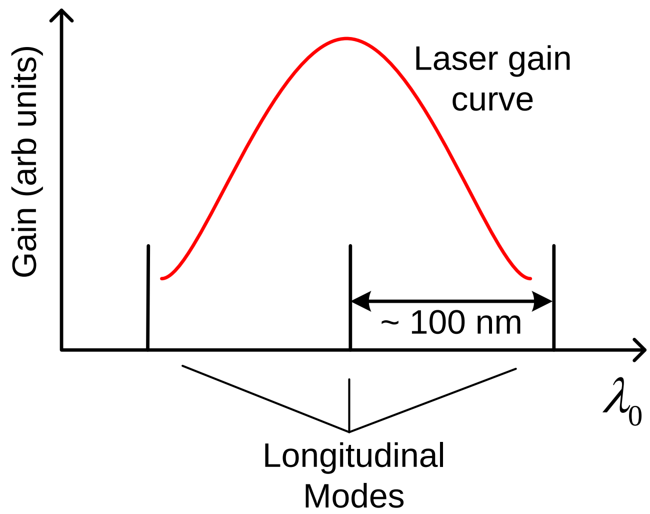


Figure 6.35: *Only one longitudinal mode falls under a VCSEL gain curve.*

In addition to high spectral purity, other advantages for VCSELs include a circular optical output profile that couples well to optical fibers and low manufacturing costs. On the other hand, the small gain volume limits the output power of VCSELs to values on the order of one mW. Furthermore, VCSELs are most easily constructed for lasing at wavelengths near 1 μm, where optical absorption in optical fibers is relatively high. For these reasons, VCSELs are most frequency used for data transmission in local area networks with short fiber spans.

Light Emitting Diodes

The simplest and least expensive light source for optical transmitters is a light emitting diode or LED. A surface emitting LED is pictured in **Figure 6.36.**

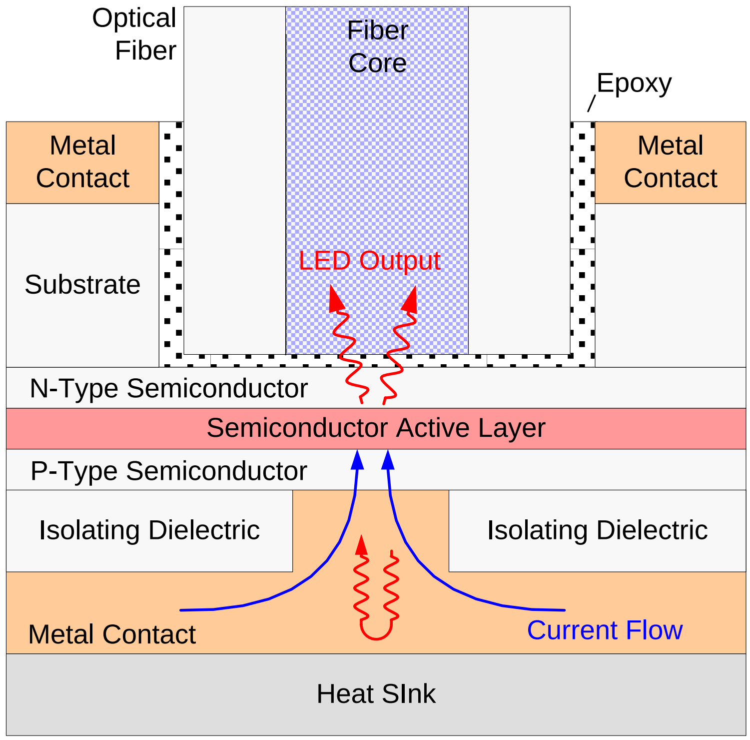


Figure 6.36*: A light emitting diode (LED).*

Electrical current is injected into an optically active semiconductor layer by forward biasing a p-n junction. The electrically excited active region emits optically incoherent light generated primarily from spontaneous emission. Light is emitted in all directions but downward emitted light is reflected upwards off a metal electrode. Light that leaves the surface of the LED can be coupled into the core of an optical fiber.

LED emission is relatively broadband, and this limits both the data rate and length of optical data links. For this reason, LED's are most often used for optical transmitters when the data rate is 200 Mbits/sec or less, and for applications such as local area networks where fiber spans are short.