Kirk-Othmer Encyclopedia of Chemical Technology.

LIGHT GENERATION SEMICONDUCTOR LASERS

Diode lasers have evolved from a laboratory curiosity to devices used in a variety of everyday applications, from consumer products such as the compact disc (CD) audio players to high speed telecommunication networks based on optical fibers (see Fiber optics). This evolution was made possible by a series of rapid advances in materials preparation, laser design, and fabrication. Progress in materials growth and device design has been particularly impressive (1–4). Most of the diode lasers emitting in the red and near-infrared (nir) parts of the electromagnetic spectrum are based on compound semiconductors (qv) composed of elements of Group 13 (III) and 15 (V) of the Periodic Table. Some of the newer green and blue emitting lasers (qv) use compound semiconductors composed of Group 12 (IIB) and 16 (VI) of the Periodic Table. Compounds containing as many as four different elements are often used in order to control the laser wavelength, electrical properties, and other desired characteristics.

1. Semiconductors

Many of the optical and electrical characteristics of semiconductors are conveniently described by energy level diagrams. Electrons in atoms are restricted to sets of discrete energy states and these states are separated by gaps in which electrons are not allowed. In solids formed by bringing isolated atoms together, the allowed energy levels of discrete atoms spread into essentially continuous energy bands. Two such bands are of particular importance in semiconductors: the low lying valence band (VB) and the upper band, known as the conduction band (CB). The conduction and valence bands are separated from each other by an energy gap, E_g. In an ideal undoped semiconductor there are no energy states within the energy gap. At a temperature of absolute zero all of the states in the valence band are occupied by electrons, which thus cannot contribute to conductivity, and all of the conduction band states are empty. Thermal excitation of electrons across the energy gap becomes possible at higher temperatures establishing a net electron concentration in the conduction band as illustrated in Figure 1**a.** The excited electrons leave behind empty states in the valence band called holes which behave like positively charged electrons. Both the electrons in the conduction band and holes in the valence band are free to move in space and participate in electrical conductivity.

In semiconductors of interest for light-emitting sources the band gap energy is considerably larger than room temperature thermal energy ($kT = 0.025$ eV) and the electrical conductivity resulting from the thermally excited electrons is negligibly small. In order to obtain appreciable conductivities these materials must be intentionally doped. Doping is accomplished by adding small amounts of impurity atoms having either more or fewer valence electrons than required for bonding into the host crystal lattice. The impurity atoms deficient in electrons tend to attract an electron and become ionized. The impurity atoms having an excess electron tend to donate it to the lattice and become ionized. These two different kinds of impurity atoms are known, respectively, as acceptors and donors. The regions of semiconductor doped with the donor impurities contain more ionized donors than acceptors and are called n-type. The p-type regions have more acceptor than donor impurities. The boundary region between the p-type and n-type regions is called a p–n junction. Such a region

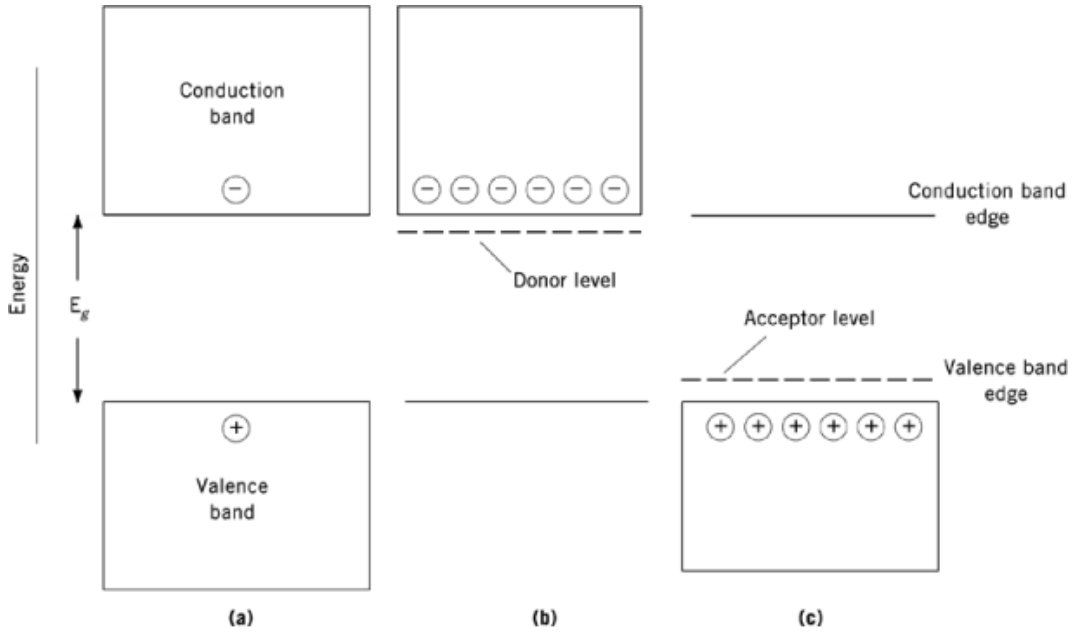

Fig. 1. Band-edge energy diagram where the energy of electrons is higher in the conduction band than in the valence band: (**a**) an undoped semiconductor having a thermally excited carrier; (**b**) *n*-type doped semiconductor having shallow donors; and (**c**) a *p*-type doped semiconductor having shallow acceptors.

behaves electrically like a rectifying diode in which current flow with applied voltage is not symmetric with respect to voltage polarity. All of the so-called injection lasers are based on *p–n* diodes.

The impurity atoms used to form the *p–n* junction form well-defined energy levels within the band gap. These levels are shallow in the sense that the donor levels lie close to the conduction band (Fig. 1**b**) and the acceptor levels are close to the valence band (Fig. 1**c**). The thermal energy at room temperature is large enough for most of the dopant atoms contributing to the impurity levels to become ionized. Thus, in the *p*-type region, some electrons in the valence band have sufficient thermal energy to be excited into the acceptor level and leave mobile holes in the valence band. Similar excitation occurs for electrons from the donor to conduction bands of the *n*-type material. The electrons in the conduction band of the *n*-type semiconductor and the holes in the valence band of the *p*-type semiconductor are called majority carriers. Likewise, holes in the *n*-type, and electrons in the *p*-type semiconductor are called minority carriers.

At thermal equilibrium characterized by temperature, *T,* the distribution of electrons over the allowed band of energies is given by a Fermi-Dirac distribution function:

$$f(E) = [1 + e^{(E-E_f)/kT}]^{-1}$$

where $k = 8.62 \times 10^{-5}$ eV/K is the Boltzmann's constant. The function, $f(E)$, describes the probability of a state with an energy, E, being occupied at a temperature, T. The quantity, E_f (the Fermi level), denotes the energy level with the occupation probability of $1/2$. At $T = 0$ all the available states below E_f are filled with electrons, and all the states above E_f are empty. The function, $f(E)$, is symmetrical about E_f for all temperatures. Thus $f(E)$ is an expression for an average number of fermions in a single orbital of energy, E. The Fermi level is then the chemical potential μ.

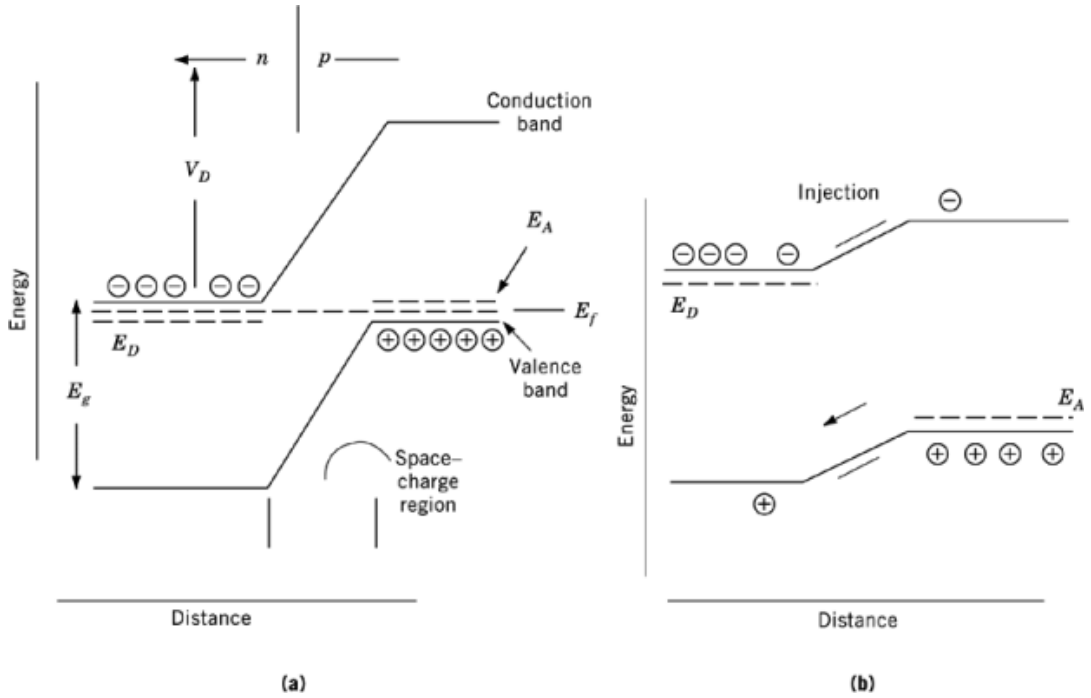

Fig. 2. Representation of the band edges in a semiconductor p–n junction where shallow donor, acceptor energies, and the Fermi level are labeled E_D, E_A, and E_f, respectively. (**a**) Without external bias; V_D is the built-in potential of the p–n junction; (**b**) under an applied forward voltage V_B.

Instead of plotting the electron distribution function in a band energy level diagram, it is convenient to indicate the Fermi level. For instance, it is easy to see that in p-type semiconductors the Fermi level lies near the valence band.

The spatial distribution of carriers in the immediate vicinity of the p–n junction is very important. Some of the majority carriers at the p–n junction neutralize each other resulting in a thin region depleted of free carriers, known as the space–charge region. In this region the fixed negative charges on the p-side repel the mobile electrons from the n-side of the junction. Similarly, the mobile holes from the p-side are repelled by the fixed positive charges at the p-side. The result is a built-in electric field which inhibits carrier diffusion across the p–n junction. The potential V_D associated with this field bends the conduction and valence bands in the space–charge region by an amount called the barrier height, as illustrated in Figure 2**a.** The potential difference, V_D, is also called the contact potential.

The current flow across the p–n junction can be accomplished only by the application of an external voltage bias which opposes the V_D and reduces the barrier height. This bias, called a forward bias, supplies electrons at the n-contact and holes at the p-contact. These carriers flow, as majority carriers, toward the p–n junction. Because the barrier height is similar in magnitude to the band gap energy, the external voltage bias needs to be fairly small, also on the order of the band gap energy divided by the electron charge. That is, the applied voltage needs to be less than the numerical value of the band gap. As electrons move up on the energy band diagram their energy increases. Holes increase their energy when moving down.

Under a forward bias the majority carriers cross the p–n junction and become minority carries, eg, holes on the n-side of the junction. These holes rapidly come into thermal equilibrium with the lattice of the semiconductor and reach the energy approximately equal to the energy of the valence band edge on the n-side

of the junction. The minority electrons reach an equilibrium energy at the p-side of the junction. The minority carriers are not in thermodynamic equilibrium with the majority carriers and must give up their excess energy. The hole on the n-side is said to recombine with a majority electron which then loses energy about equal to E_g. The minority electron loses a similar energy and recombines with a majority hole. In semiconductors of interest herein the energy given up by the minority carriers results in emission of light. This process is known as the radiative recombination or spontaneous emission. Recombination can also occur through a range of competing nonradiative processes, which occur without the emission of light. These depend on defects and imperfections in the host lattice and must be carefully controlled in order to assure efficient emission of light in forward biased p–n junctions.

The distributions of excess, or injected, carriers are indicated in band diagrams by so-called quasi-Fermi levels for electrons, F_n; or holes, F_p. These functions describe steady-state concentrations of excess carriers in the same form as the equilibrium concentration. In equilibrium $F_n = F_p = F_f$.

Light, ie, a photon, is emitted in the transition from a higher to a lower energy state. The energy difference between the two states, ΔE, is converted to a photon having wavelength, λ, of

$$\lambda = hc/\Delta E$$

where h is Planck's constant and c is the velocity of light. A practical conversion when the wavelength of light, λ, is in units of μm and ΔE is in eV is $\lambda = 1.24/\Delta E$.

2. Direct and Indirect Band Gap Semiconductors

Semiconductors can be divided into two groups: direct and indirect band gap materials. In direct semiconductors the minimum energy in the conduction band and the maximum in the valence band occur for the same value of the electron momentum. This is not the case in indirect materials. The difference has profound consequences for the transitions of electrons across the band gap in which light is emitted, the radiative transitions, of interest here.

In direct band gap materials, recombination of an electron and a hole occurs without the need for momentum transfer making it possible to achieve high rates of radiative recombination across the band gap. In addition, in band-to-band transitions the energy of the photon needed to stimulate the transition is the same as that emitted in the transition. In indirect semiconductors the conduction–valence band transition can occur only if momentum is conserved. This can be achieved by the simultaneous emission and absorption of phonons, ie, quanta of lattice vibration. However, the rates for indirect radiative recombination processes are low, and competing nonradiative transitions can be much more important. The radiative rates can be somewhat adjusted by the intentional incorporation of deep levels within the band gap, a process useful in some types of light-emitting diodes. The energy of the photon emitted in this process is considerably lower than that of the band gap. Indirect band gap materials cannot be used in lasers owing to their low efficiency of converting injected carriers into light (low quantum efficiency).

3. Semiconductor Laser

The operation of a laser is based on stimulated emission. A photon generated in the forward biased junction, in the process of spontaneous emission, can be absorbed in the semiconductor or cause an additional transition by stimulating an electron to make a transition to a lower energy state. In this latter process another photon is emitted. The rate of stimulated emission is governed by a detailed balance between absorption and spontaneous and stimulated emission rates. That is, stimulated emission occurs when the probability of a photon

causing a transition of an electron from the conduction to the valence band, with the probability emission of another photon, is greater than for the upward transition of an electron from the valence to conduction band upon absorption of the photon. These rates are commonly described in terms of Einstein's theory of A and B coefficients (5, 6). For semiconductors, a simple condition describing the carrier density necessary for stimulated emission has been derived (7). Lasing can start when the density of electrons injected into the conduction band exceeds the hole density in the valence band. This is a condition of population inversion which occurs when the separation of the quasi-Fermi levels for the holes, F_{p}, and the electrons, F_{e}, is greater than the energy of the emitted photon,

$$F_e - F_p > hv$$

and the photon energy hv must be at least equal to the band gap energy. Thus in semiconductor lasers stimulated emission occurs between distributions of states in the conduction and valence bands. In most other lasers (qv), such as gas or glass lasers, this transition occurs between discrete energy levels.

The distributions of states in conduction and valence bands are commonly described by the effective density of states. The concentration of electrons, n, in the conduction band can be calculated as

$$n = \int_{E_c}^{\infty} f(E)N(E)dE$$

where $N(E)dE$ is the density of states (in cm^{-3}) in the energy range dE. A simpler way of calculating n is to represent all of the electron states in the conduction band by an effective density of states, N_{c}, at the energy, E_{c} (band edge). The electron density is then simply

$$n = N_c\, f(E_c)$$

Population inversion is easier to achieve when the effective density of states in the conduction band is low.

The processes contributing to laser action are illustrated in Figure 3 which shows absorption, emission, and the gain spectra of a semiconductor (8). Significant absorption occurs only for the photon energies greater than that of the band gap. The spectrum of spontaneous emission is determined by the absorption spectrum, at the low energy side, and the electron energy distribution at the high energy side. The gain spectrum represents the difference between the spontaneous emission and absorption. Lasing occurs at the energy of the gain peak.

A semiconductor laser requires a means of generating spatially localized high concentrations of minority carriers, a medium to provide the gain, and a means of providing some feedback to the stimulated emission. The medium is the semiconductor structure arranged in a way which helps to confine the carriers and light. Light is generated in this structure by means of a p–n junction which injects electrons from the valence to the conduction band and thus provides the population inversion. This is followed by recombination of electrons with holes and emission of light. Further recombination can be stimulated (stimulated emission) by light already present in the medium. This optical feedback is carefully arranged by forming a cavity which has two mirrors parallel to each other. Light generated within the cavity is then partially reflected back into the crystal. Such mirrors can be formed in most compound semiconductors lasers by cleaving two ends of a waveguide.

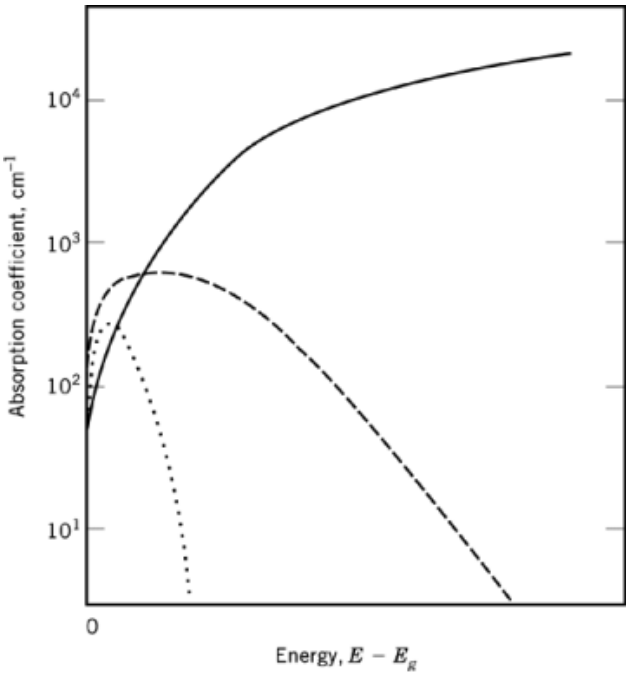

Fig. 3. Spectra showing absorption coefficient as a function of the photon energy in a direct band gap semiconductor where (—) represents absorption, (— — —) the spontaneous emission, and (...) gain.

4. Heterostructures

In the *p–n* junction illustrated in Figure 2 both sides are made of the same type of semiconductor and have the same energy gap. Such a junction is called a homojunction. The disadvantage of this type of junction is that the minority carriers are free to move away from the junction. It is thus very difficult to achieve the high carrier density needed in a laser. Higher carrier densities can be obtained by introducing an energy barrier at, or very near to, the *p–n* junction. The energy barrier arises when two semiconductors, having different band gaps, are joined together (9, 10). Barriers in the conduction band (CB) and valence band (VB), the band-edge discontinuities ΔE_c and ΔE_v, respectively, can be formed by changing the composition of the semiconductor. A junction of a small and large band gap semiconductor is called a single heterojunction. It confines one type of minority carrier (electrons or holes) to the *p–n* junction region.

A more effective carrier confinement is offered by a double heterostructure in which a thin layer of a low band gap material (the active layer) is sandwiched between larger band gap layers. The physical junction between two materials of different band gaps, and chemical compositions, is called a heterointerface. A schematic representation of the band diagram of such a structure is shown in Figure 4. Electrons injected under forward bias across the *p–N* junction into the lower band gap material encounter a potential barrier, ΔE_c, at the *p–P* junction which inhibits their motion away from the junction. The holes see a potential barrier of ΔE_v at the *N–p* heterointerface which prevents their injection into the *N* region. The result is that the injected minority carriers are confined to the thin narrow band gap region. If this region is thinner than the average distance the carrier can move before recombination (called diffusion length), on the order of a few micrometers in most compound semiconductors, a very high density of injected carriers can be obtained in a forward biased diode. In laser diodes this thin low band gap layer is called the active layer. In most of the conventional double

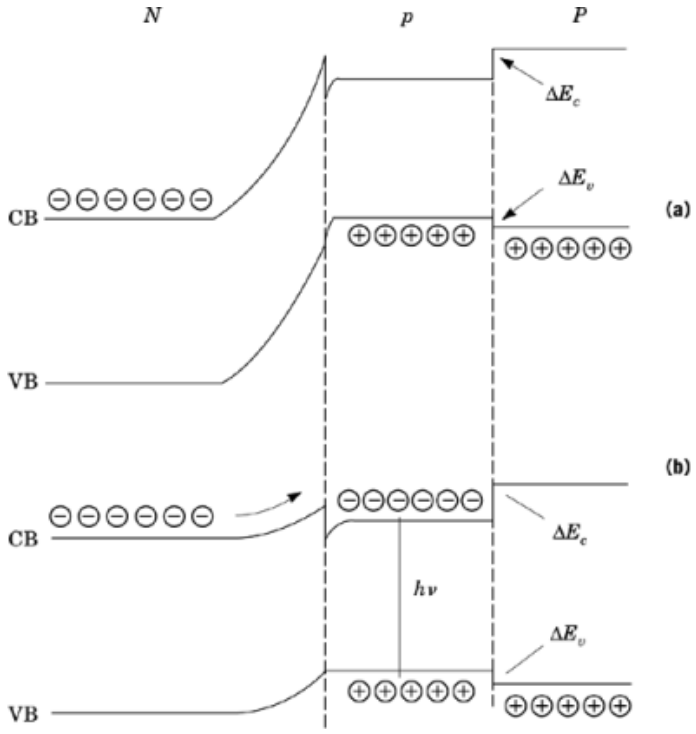

Fig. 4. Schematic cross section and the band diagram of a double heterostructure showing the band-edge discontinuities, ΔE_c and ΔE_v, used to confine carriers to the smaller band gap active layer. (**a**) Without and (**b**) with forward bias. See text.

heterostructure lasers the thickness of this layer is about 0.1 μm. The thinner the active layer, the easier it is to achieve the carrier density needed for stimulated emission. However, as a matter of practice, the thickness of the active layer may be limited by the precision of the epitaxial growth technique.

In addition to the carrier confinement, the active region is crucial in providing light confinement and waveguiding because the index of refraction of this low band gap layer is larger than the indexes of the surrounding layers. Such light confinement is absent in homostructure or single heterostructure lasers. In most modern laser structures special waveguide layers having a band gap, and thus the index of refraction, intermediate to that of the active and confining layers, are included in order to assure adequate light confinement. In such structures the carriers are confined to the narrow section of the active layer having the lowest band gap and the highest index of refraction, and light is confined to a wider region defined by the waveguides on each side of the active layer. This type of structure is known as a separate confinement heterostructure.

Lasing occurs whenever the gain arising from stimulated emission exceeds the cavity losses. Internal losses, α_i, result from absorption and scattering of light. The reflectivity, R, of the mirror facet must be <1 and this contributes a loss term of $(1/L)\ln(1/R)$, where L is the cavity length. At threshold, the gain, g, is equal to losses and

$$g\Gamma = \alpha_i + \frac{1}{L}\ln\frac{1}{R}$$

where Γ is the confinement factor, a fraction of the light intensity within the active region. It is typically <1 because the active layer (a region of carrier confinement) is thinner than the waveguide layers (regions of light confinement). The threshold current density, J_{th}, in A/cm^2 can be written (6) as

$$J_{th} = 4.5 \times 10^3 (d/\eta) + (20d/\eta\Gamma)\left\{\alpha_i + \frac{1}{L}\ln\left(\frac{1}{R}\right)\right\}$$

where d, the thickness of the active region written, is in μm, and η is the efficiency with which current is converted to photons.

Large threshold current density of homostructure lasers results mainly from poor carrier confinement and the resulting large active region thickness d. This is because the diffusion length of electrons is fairly long in most semiconductors discussed herein, on the order of many micrometers. For $d = 3$ μm and $L = 500$ μm the threshold current density is as large as 15 kA/cm^2. In practice threshold current densities are much larger because the homostructure also suffers from small Γ and high α_i. Very high threshold current density limits operation to short pulses and cryogenic temperatures.

In a double heterostructure laser having a gallium arsenide, GaAs, active region, $d \sim 0.1$ μm, and $Al_xGa_{1-x}As$ waveguide layers, threshold current density drops to less than 0.5 kA/cm^2. Good light confinement results in Γ close to 1 and the internal losses are less than 30 cm^{-1}. Such lasers readily achieve continuous operation at room temperature and are capable of high power output.

5. Quantum Wells

Epitaxial crystal growth methods such as molecular beam epitaxy (MBE) and metalorganic chemical vapor deposition (MOCVD) have advanced to the point that active regions of essentially arbitrary thicknesses can be prepared (see Thin films, film deposition techniques). Most semiconductors used for lasers are cubic crystals where the lattice constant, the dimension of the cube, is equal to two atomic plane distances. When the thickness of this layer is reduced to dimensions on the order of 0.01 μm, between 20 and 30 atomic plane distances, quantum mechanics is needed for an accurate description of the confined carrier energies (11). Such layers are called quantum wells and the lasers containing such layers in their active regions are known as quantum well lasers (12).

The uncertainty principle, according to which either the position of a confined microscopic particle or its momentum, but not both, can be precisely measured, requires an increase in the carrier energy. In quantum wells having abrupt barriers (square wells) the carrier energy increases in inverse proportion to its effective mass (the mass of a carrier in a semiconductor is not the same as that of the free carrier) and the square of the well width. The confined carriers are allowed only a few discrete energy levels (confined states), each described by a quantum number, as is illustrated in Figure 5. Stimulated emission is allowed to occur only as transitions between the confined electron and hole states described by the same quantum number.

The two-dimensional carrier confinement in the wells formed by the conduction and valence band discontinuities changes many basic semiconductor parameters. The parameter important in the laser is the density of states in the conduction and valence bands. The density of states is greatly reduced in quantum well lasers (11, 12). This makes it easier to achieve population inversion and thus results in a corresponding reduction in the threshold carrier density. Indeed, quantum well lasers are characterized by threshold current densities as low as $100 - 150$ A/cm^2, dramatically lower than for conventional lasers. In the quantum well lasers, carriers are confined to the wells which occupy only a small fraction of the active layer volume. The internal loss owing to absorption induced by the high carrier density is very low, as little as $\alpha_i \sim 2$ cm^{-1}. The output efficiency of such lasers shows almost no dependence on the cavity length, a feature useful in the preparation of high power lasers.

The incorporation of quantum wells into the material has other subtle consequences. The valence band in direct bulk semiconductors is degenerate at the maximum energy point (valence band maximum). At this point there are two types of holes of the same energy, the heavy and light ones. The effective mass of the light hole is

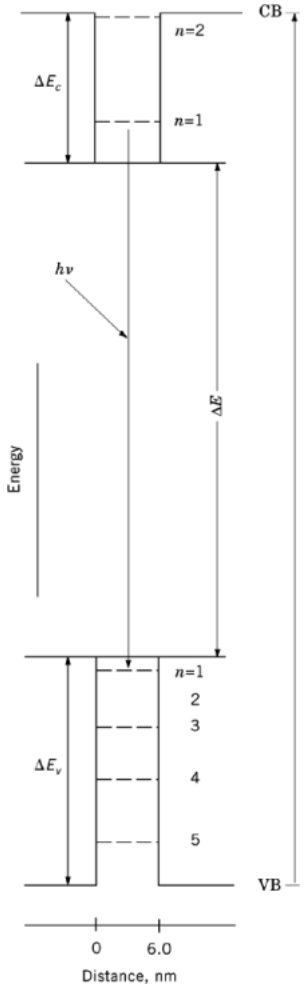

Fig. 5. Energy levels of electrons and heavy holes confined to a 6-nm wide quantum well, $In_{0.53}Ga_{0.47}As$, with InP valence band, ΔE_v, and conduction band, ΔE_c, barriers. In this material system approximately 60% of the band gap discontinuity lies in the valence band. Lasing occurs between the confined electron from the upper level and heavy hole states in the lower level.

similar to that of the electron, whereas that of the heavy hole is about 10 times larger than that of the electron. In quantum wells the two hole bands become separated in energy, and it is said that the degeneracy is lifted (12). The energy shift owing to quantum well confinement is larger for the light holes. The lasing transition is more likely to occur between the confined electron and heavy hole states, as is illustrated in Figure 5.

All the layers of conventional heterostructure lasers, that is, lasers not based on quantum wells, must precisely replicate the lattice structure of the substrate. These layers are fairly thick on the order of 0.1 μm, and any lattice mismatch large enough to alter electronic properties invariably results in generation of defects. Strain can be introduced in quantum wells, with the thickness on the order of 0.01 μm or less, without generating defects (13). In such thin layers strain is accommodated without any change in the in-plane lattice constant, as long as the thickness of the well is lower than a certain critical thickness (14). Depending on the material system, strain in the quantum well can be either compressive, with the well's lattice constant trying

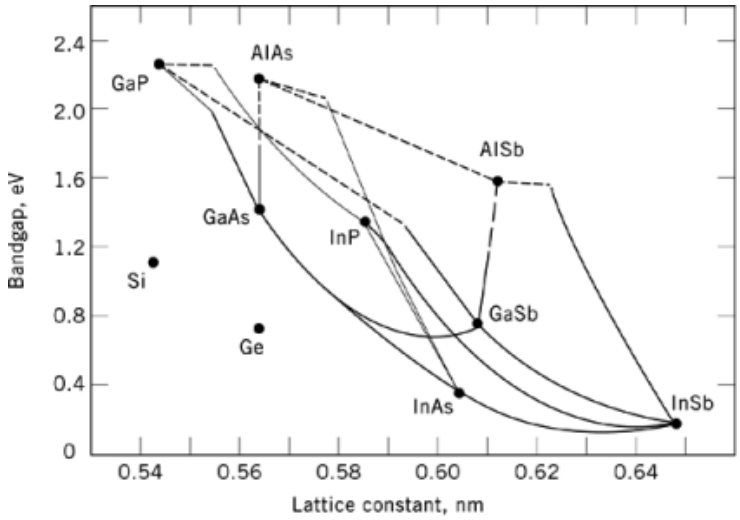

Fig. 6. Band gap versus lattice constant for Group 13–15 (III–V) semiconductors where (—) denotes direct gap and (_) indirect. Lines joining the binary semiconductors indicate possible compositions of ternary compounds. The quaternary compound $Ga_xIn_{1-x}As_{1-y}P_y$ may have any band gap and lattice constant that lie in the region of the plane bound by InP, GaP, GaAs, and InAs. Only the compositions where $_{x=2.2y}$ are lattice matched to InP.

to be larger than that of the substrate, or tensile. Compressive strain increases the separation of the light and heavy hole bands already occurring because of quantum confinement. The tensile strain initially reduces the light and heavy hole band separation and thus counteracts the quantum confinement. However, as the tensile strain is increased further the two hole bands can cross and the light hole band becomes the lowest energy band in the valence band. Laser transition then occurs between the confined electrons and light holes. Both of these strain directions can result in lower density of states and the threshold current density even lower than that of lattice matched quantum well lasers (15). The detailed manipulation of the electronic states of the band gap is now a subject of research.

6. Materials

Several compound semiconductor systems permit the growth of high quality, thin-layered crystal structures that have large and abrupt changes in the band gap and the index of refraction required in heterostructure lasers. The layer-to-layer changes result from changes in composition. Several of the Group 13–15 (III–V) material systems used for the preparation of laser structures are shown in the band gap–lattice constant plots of Figure 6. All of these materials have cubic structures known as zinc blende. The special nature of the binary semiconductor compounds such as gallium arsenide [1303-00-0], GaAs, or indium phosphide [22398-80-7], InP, arises from their availability as the substrate materials needed for the growth of lasers.

Two of the materials systems shown in Figure 6 are of particular importance. These are the ternary compounds formed from the Group 13 (III) elements such as Al and Ga in combination with As (6) and quaternary compounds formed from Ga and In in combination with As and P (16–18). The former, aluminum gallium arsenide, $Al_xGa^{1-x}As$, grown on GaAs, is the best known of the general class of compounds $A^{III}{}_xB^{III}{}_{1-x}C^V$. The latter, gallium indium arsenide phosphide, $Ga_xIn_{1-x}As_{1-x}As_{1-y}P_y$, grown on InP, is of the general class $A^{III}{}_xB^{III\,1-x}C^V{}_yD^V{}_{1-y}$. In general, the lattice constants, band gaps, indexes of refraction, and a number of other

Table 1. Composition of a Double Heterostructure Laser[a,b]

Layer	Width, μm	Composition	Comment
contact	0.2	InGaAs	$p = 2 \times 10^{19}$
cladding layer	2	InP	$p = 2 \times 10^{18}$
waveguide	0.05	InGaAsP	$E_g = 1.15$ μm
active	0.075	InGaAsP	$E_g = 1.30$ μm
buffer	1	InP	$n = 2 \times 10^{18}$

[a] Based on InGaAsP/InP system.
[b] n-type InP substrate.

parameters of these materials depend on the values of x and y. These properties are intermediate between those of the parent binary compounds.

$Al_xGa_{1-x}As$ grown on GaAs was the first material system used for the preparation of injection lasers and as of this writing, a great majority of the lasers in use are based on it. These include the popular compact disc (CD) lasers and most of the high power semiconductor lasers. The lattice constants, a, of GaAs and AlAs are almost identical, being 0.565 and 0.566 nm, respectively. Thus aluminum atoms can be substituted for Ga atoms in the GaAs lattice to form $Al_xGa_{1-x}As$ without any significant change in the lattice constant. It is possible to vary the bandgap from 1.43 eV, the E_g of GaAs to 2.16 eV, and the E_g of AlAs, simply by adjusting the Al fraction x in the epitaxial layer. This feature of $Al_xGa_{1-x}As$ is quite unique among the compound semiconductors. The band gap of the ternary alloy changes from direct to indirect for Al concentrations where $x > 0.43$. Compounds having lower Al concentrations can be used in active layers. Lasers having a large range of emission wavelengths, from nearly visible (red) to infrared, can be prepared in this way.

An even wider range of wavelength, toward the infrared, can be covered with quantum well lasers. In the $Al_xGa_{1-x}As$ system, compressively strained wells of $Ga_xIn_{1-x}As$ are used. This ternary system is indicated in Figure 6 by the line joining GaAs and InAs. In most cases the Al fraction is quite small, $x < 0.2$. Such wells are under compressive strain and their thickness must be carefully controlled in order not to exceed the critical layer thickness. Lasers prepared in this way are characterized by unusually low threshold current density, as low as ca 50 A/cm^2 (19).

The usual acceptor and donor dopants for $Al_xGa_{1-x}As$ compounds are from Groups 2 and 12 (II), 14 (IV), and 16 (VI) of the Periodic Table. Group 2 and 12 elements are acceptors and Group 16 elements are donors. Depending on the growth conditions and the growth method, Si and Ge (Group 14) can be either donors or acceptor, ie, be amphoteric, which is of special interest in light-emitting diodes. Doping levels in lasers range from mid-10^{17} cm^{-3} in the buffer and cladding layers to mid-10^{19} cm^{-3} in the contact layers. In quantum well-based lasers it is also desirable to vary the doping level very abruptly, ie, by two to three orders of magnitude, over dimensions as small as 0.01 μm. This requires dopants which do not move significantly (diffuse) at the growth temperature.

Another principal material system is $Ga_xIn_{1-x}As_{1-y}P_y$ grown on InP. A summary of compositions and doping levels of a laser heterostructure using such layers is presented in Table 1. Quaternary compounds are not naturally lattice matched to that substrate. In order to avoid generation of defects care must be taken in adjusting the ratio of x and y to maintain the lattice matched composition of $x = 2.2y$. The available band gaps range from that of InP to that of $In_{0.53}Ga_{0.47}As$, the ternary compound shown in Figure 6 by a line joining GaAs and InAs. The particular advantage of this system lies in the excellent match of available band gaps with the optical properties of fused silica optical fibers (17).

Quantum well lasers in this system typically use ternary $In_{0.53}Ga_{0.47}As$ wells and binary InP barriers. All quaternary lasers, ie, lasers in which both the wells and barriers are formed by quaternary compounds, are also being developed. These structures can be lattice matched or strained.

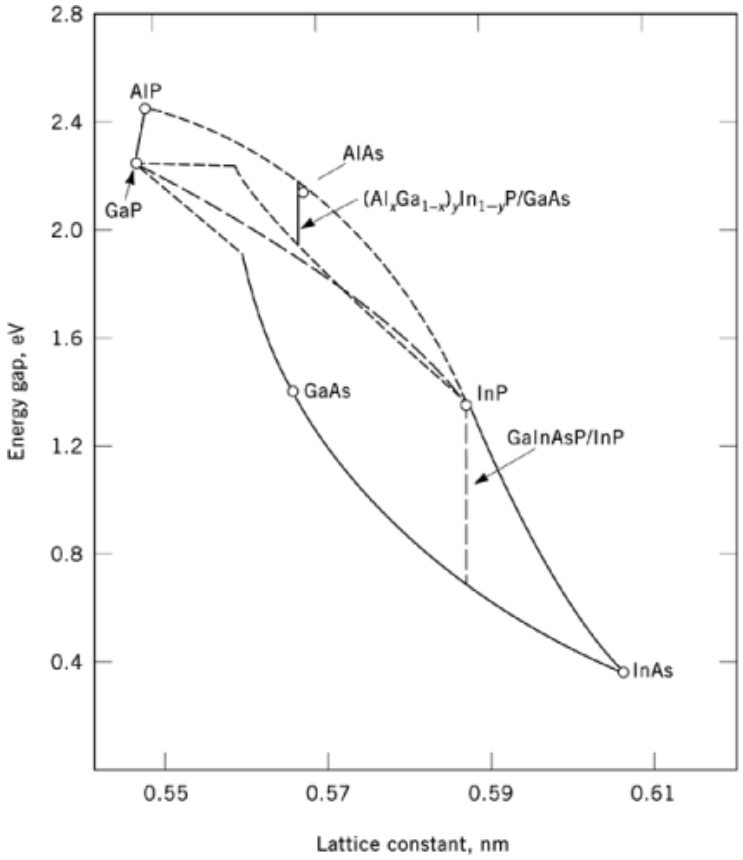

Fig. 7. Band gap versus lattice constant for the $(Al_xGa_{1-x})_yIn_{1-y}P$ system.

The materials discussed yield lasers operating in the infrared and near visible spectral ranges. Many applications of lasers, such as printing or high density memories, require as short a wavelength as possible. The III–V system most suitable for short wavelength visible operation is the $(Al_xGa_{1-x})_yIn_{1-y}P$ system (Fig. 7) (18). Direct band gap compositions of $(Al_xGa_{1-x})_{0.5}In_{0.5}P$ are lattice matched to GaAs and span the range from 2.25 eV (for $x = 0.3$), to 1.92 eV for $Ga_{0.5}In_{0.5}P$. Room temperature lasing up to 580 nm is possible, the shortest emission wavelength of any lattice matched III–V heterostructure. Because in a double heterostructure some of the bandgap must be used for carrier confinement, the shortest room temperature emission wavelength achieved as of this writing in continuous wave (CW) operation is 633 nm which matches that of a HeNe laser.

More exotic Group 13–15 (III–V) materials such as gallium nitride [25617-97-4] are being investigated for shorter lasing wavelengths. One of the prime candidates is the InGaN/GaN system in which blue ($\sim$410 nm) light-emitting diodes have been prepared (20). The search is on for other large band gap III–V-like semiconductor systems.

Injection lasers based on Group 2 or 12–16 (II–VI) compounds have also been demonstrated (21). These large band gap compounds make green lasers a reality and promise an extension of the lasing wavelength into the blue. The first green lasers were based on a cadmium zinc selenide, $Cd_{0.2}Zn_{0.8}Se$ quantum well sandwiched between cladding layers consisting of zinc sulfur selenide, $ZnS_{0.07}Se_{0.93}$, and zinc selenide [1315-09-9], ZnSe, waveguide layers. The band gap–lattice constant plots for these materials are shown in Figure 8. The lattice

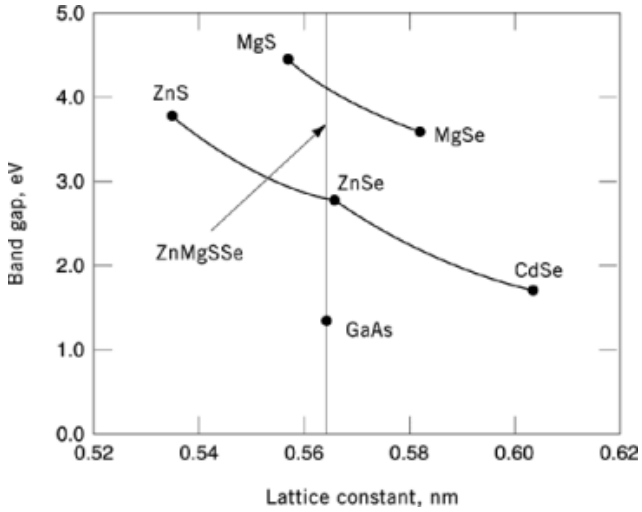

Fig. 8. A plot of band gap versus lattice constant for Groups 2–16 and 12–16 (II–VI) semiconductors used for the preparation of green and blue lasers.

constant of the ternary $ZnS_{0.07}Se_{0.93}$, but not that of ZnSe, matches that of (100) GaAs which serves as a substrate for this structure. The first lasers operated at cryogenic temperatures (77 K) and had an emission wavelength of 490 nm. Pulsed room temperature operation, and CW operation for very brief times (seconds) was also reported at 509 nm (22).

The absence of lattice match between the ZnSe waveguide and the substrate is a serious problem. In order to be effective as a buffer or waveguide this layer must be fairly thick, on the order of 0.2–1 μm. The strain is then too large to be accommodated elastically and the layer relaxes. The relaxation occurs through the generation of defects, ie, misfit dislocations, at the ZnSe/$ZnS_{0.07}Se_{0.93}$ interface. The presence of such defects significantly increases the threshold current density. This problem can be solved by using a quaternary compound, zinc magnesium sulfur selenide, $Zn_{1-x}Mg_xS_ySe_{1-y}$, also shown in Figure 8 (23). Cladding and waveguide layers can be grown lattice matched to GaAs, and their composition adjusted to yield a desired sequence of the band gap steps. Larger band gap range available for the cladding layer allows for increased band gap of the active layer and a shorter emission wavelength. Room temperature laser operation having a threshold current density as low as 500 A/cm^2 has been reported at 516 nm (24). This quaternary II–VI compound appears to have the flexibility analogous to that already exploited in InGaAsP.

The potential importance of II–VI compounds as visible laser materials has been recognized for some time. Injection lasers are difficult to prepare because of the inability to introduce dopants, and particularly *p*-type dopants. A convenient *n*-type dopant is chlorine, which can be introduced from $ZnCl_2$ sources. Fairly high doping levels, ca $10^{18}/cm^3$, can be readily achieved. Historically, the most common *p*-type dopants were lithium, Li, and nitrogen, N, the latter most often from NH_3. These dopant elements can be introduced into the semiconductor but only a small fraction becomes electrically active, ie, contributes to electrical conductivity. The *p*-type concentrations in large band gap materials had not even reached ca $10^{17}/cm^3$, a level too low to form *p*–*n* junctions or contact layers with low electrical resistance, until the impasse was broken by the discovery that nitrogen free-radicals produced by a high frequency plasma source could be incorporated up to a level of $p \sim 10^{18}/cm^3$ (25) (see Plasma technology). This allowed for successful preparation of diode lasers.

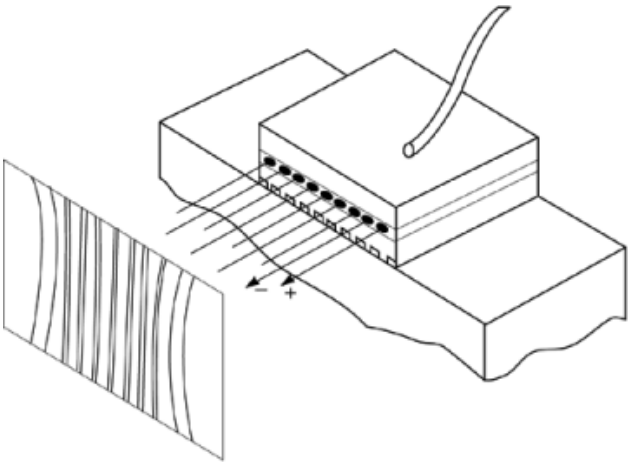

Fig. 9. High power array of phase coupled $_{\mathrm{GaAs/AlGaAs}}$ lasers mounted p-side down on a thermal heat sink. The $_{\pi/2}$ shift of the neighboring lasers is indicated by the $_{+}$ and $_{-}$ signs. The output pattern consists of two dominant peaks, each associated with the lasers of the same phase, and much weaker peaks in between.

7. Laser Structures

Early injection lasers were small rectangular parallelepipeds made by cutting a wafer of GaAs. Feedback was provided by mirrors polished on two edges or by cleaving. The wafer had a p–n junction incorporated into it and broad area or stripe contacts were provided. Laser structures have since evolved to satisfy a wide range of application specific requirements.

$Al_xGa_{1-x}As$ lasers show outstanding performance in two areas: high power lasers and vertical cavity surface emitting lasers. High power lasers are based on phase coupled arrays (26) of the type illustrated in Figure 9. The individual lasers are formed simply by providing a narrow stripe contact, 5–10 μm in width. Proton implantation, which turns GaAs highly resistive, is used to confine the current to a narrow region close to the stripe (see Ion implantation). These lasers have no lateral index step to guide the light and are said to be gain guided. Lasing occurs in the areas under and very near the stripe in which the carrier density necessary for threshold is reached. The array is formed by placing a number of individual lasers close enough, 5–10 μm from each other, so that their radiation fields influence each other, hence the name phase coupled. The advantage is a very narrow angle of emission in the plane of the array and high power output. The narrow stripe of individual lasers results in a fairly wide angle of in-plane emission, approximately 14° for a 5-μm wide stripe (6). An array of interacting lasers may operate as a spatially coherent unit. The wide aperture of such a unit results in a very low divergence beam. Because the electric field must go through zero between lasers, it changes its phase from laser to laser by $\pi/2$; the most likely emission pattern consists of two narrow beams each from all the lasers operating in phase. These two beams are emitted at a small angle from each other. The typical values are ca 1–1.5° for the width of each beam and about 4° for the separation between beams, respectively. The exact values depend strongly on the dimensions of the array and the current drive.

The large emission aperture and tight current confinement in laser arrays are important for high power operation. In GaAs lasers the output power is limited by catastrophic facet damage, the damage caused by a high light flux at the facet. It must be realized that the submicrometer active layer thickness of double heterostructure and the narrow stripe width result in radiative power fluxes on the order of $\mathrm{MW/cm^2}$. Facet damage is known to occur at a flux of approximately 10 $\mathrm{MW/cm^2}$ (27). The large aperture of laser arrays is very beneficial in this respect. Power output exceeding 100 mW/facet was achieved in a coupled array consisting of

five lasers in 1978. The output powers have been vastly increased. For instance, a large array of 1716 lasers, arranged in a two-dimensional pattern, reached a peak power output of 1 kW at a current drive of 60 A (28). The power output of such large arrays is a matter of thermal management and proper heat sinking more than the individual laser properties. Finally, in these arrays the current is confined by a narrow stripe and the lasing filament is well defined. It stays reasonably well defined at current values many times greater than the threshold current. The result is a light output pattern stable as a function of power level.

GaAs-based lasers also excel in new applications directed toward optical computing and optical data processing (see also Information storage materials). Such applications require two-dimensional arrays of lasers having a very low threshold current, high beam quality, and high speed of operation. These requirements are well served by vertical cavity surface emitting lasers (VCSELs). In the lasers discussed previously light propagates in the plane of the wafer, in a carefully constructed waveguide, and is emitted through a cleaved facet. The edge-emitting lasers are quite long, the typical length exceeds 250 μm, and the facet reflectivity is about 3%, the natural reflectivity of the semiconductor–air interface arising from the differences in respective indexes of refraction. It is also possible to rearrange the geometry of the laser to produce light emission in the direction perpendicular to the plane of the p–n junction, surface emission (29, 30). In such lasers the cavity length is quite small (<1 μm) and the lower gain must be compensated for by much higher facet reflectivity.

VCSELs are formed by placing the active layer between two highly reflecting mirrors, as illustrated in Figure 10. The Bragg mirrors, as they are called, are formed from alternating layers of semiconductor having high and low indexes of refraction, such as GaAs and AlAs. The thickness of each layer is chosen to be equal to one-fourth of the wavelength of light in the semiconductor, ie, $\lambda/(4n)$ where n is the index of refraction at the lasing wavelength. Because the index changes with the wavelength this condition can be strictly satisfied only at a single wavelength. The choice of thicknesses of individual layers results in a maximum reflectivity for each pair of layers at the lasing wavelength. Because the composition-induced index variation is not large, as many as 30 pairs must be used in order to produce a mirror having the needed reflectivity greater than 99%. The total active layer thickness is chosen to be λ/n is the effective index of refraction. This condition allows for a standing wave to be formed between the lower and the upper mirrors. VCSELs emit light in a narrow mode at the wavelength at which a single-λ cavity is formed. Many of their properties are thus determined by the interplay between the complicated reflectivity spectra of the mirrors and the gain spectrum. This may result in very low temperature dependence of the threshold current and other desirable characteristics. These lasers show very low (<1 mA) threshold currents, and modulation bandwidths, ie, the speed with which they can be turned on and off, can possibly reach 70 GHz (31). High reflectivity of the facet limits the light output to a few mW. The exact value depends strongly on the laser diameter.

Lasers fabricated from alloys of InGaAsP operate in the 1.3–1.55 μm wavelength range and are used in fiber optic communications (17, 18) (see Fiber optics). These wavelengths correspond, respectively, to the minimum dispersion and loss spectral windows of fused silica fibers. The lasers are designed to be coupled to optical fibers and must have excellent beam properties, narrow spectral width, and very high speed. All of these characteristics can be achieved at the same time but with a considerable increase in the structural complexity and difficulty of preparation. An example of a device designed for such demanding application is a buried heterostructure laser shown in Figure 11. This laser structure is known as capped mesa buried heterostructure (CMBH). It is prepared in three (or more) epitaxial growth cycles, all typically carried out by metalorganic chemical vapor deposition (MOCVD). The planar base structure containing the active, waveguide, and confining layers is grown first. It may be based on a conventional bulk active layer design or quantum wells, or contain a distributed feedback structure. The active layer mesa is formed by masking a narrow stripe using a dielectric layer, such as SiO_2, and etching off the remainder of the first growth. The width of the mesa is kept at approximately 0.8 μm in order to assure low threshold current and a narrow angle of emission. The narrow mesa width can be achieved using conventional wet etching only if the thickness of the InP cladding layer is <1 μm. The current blocking layers are grown in the second epitaxial growth cycle in the exposed areas of the wafer but not on the active mesas which are protected by a dielectric mask. The blocking structure

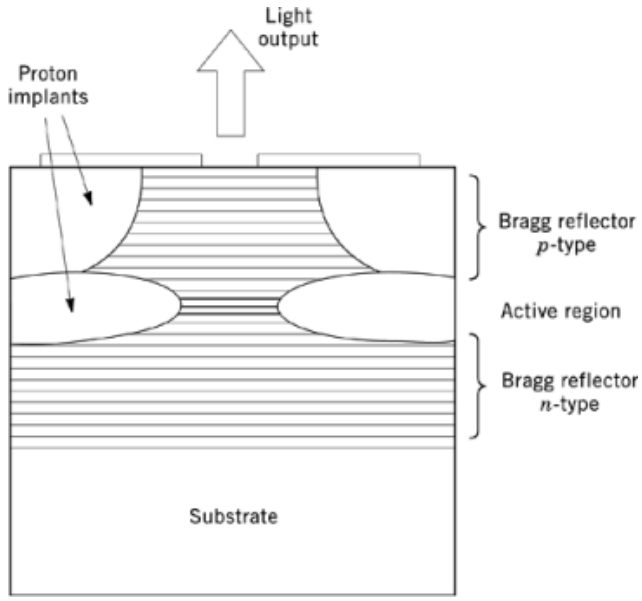

Fig. 10. Cross-sectional drawing of a vertical cavity surface emitting laser (VCSEL). Proton implantation is used to channel the current through a small active region. Light is emitted in the direction perpendicular to the plane of the wafer. This makes preparation of two-dimensional arrays quite easy.

typically consists of a *n–i–n–p* doped sequence of layers of InP. Here the *i*-layer is semi-insulating, prepared by doping with iron which acts as a deep acceptor. The *p*-type InP layers on each side of the *i*-layer are intended to prevent injection of holes and electrons into the semi-insulating layer even at high injection levels or high operating temperatures. In the third cycle of growth the stripe mask is removed and *p*-type InP and GaInAs layers are grown over the entire wafer. The *p*-type layer of InP increases the cladding layer thickness to the final desired value and the *p*-type GaInAs contact layer assures low contact resistance.

The complex current blocking structure is effective in confining the current to the narrow active layer but it adds a parasitic capacitance, a feature undesirable in high speed operation. A fourth growth step may be added at this point. Most of the current blocking structure is removed by etching, leaving only a 20–50 μm wide ridge. The removed regions are then backfilled with Fe-doped InP in order to maintain planarity of the device.

The resulting laser offers excellent performance. The threshold currents are usually much lower than 10 mA and the light output is linear at least up to 20 mW. Specialized high power devices are capable of producing power levels in excess of 250 mW and having excellent beam characteristics. The modulation bandwidth of these lasers exceeds 25 GHz. Another feature of the buried heterostructure lasers fabricated from InGaAsP is their outstanding long-term stability. These lasers are reliably used in fiber optic communication cables spanning the oceans.

Most of the lasers discussed operate in a small number of discrete longitudinal modes, the Fabry-Perot modes. The individual modes are very narrow, much less than 0.01 nm, but are separated by spectral distances of ca 1.0 nm. Thus the overall width of the laser spectrum may exceed 4–5 nm. This is unacceptable in communication systems operating at 1.55 μm, the minimum loss region of optical fibers, over distances on the order of 100 km and data rates as high as 10 gigabites per second (Gb/s). This is because light of different wavelengths travels in optical fibers at slightly different speeds. This property of optical fibers results in time spreading of optical pulses traveling down the fiber and carrying information in digital form. Eventually it is

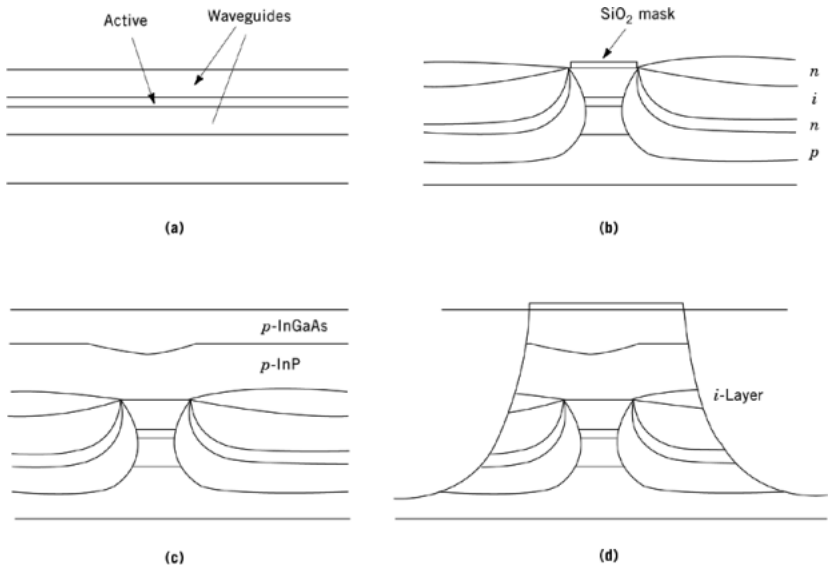

Fig. 11. Schematic diagram illustrating the preparation sequence of an $_{\text{InGaAsP/InP-based}}$ buried heterostructure laser. (**a**) The active and waveguide layers; (**b**) the SiO_2 mask; (**c**) the *p*-type regions; and (**d**) the injection layer. See text.

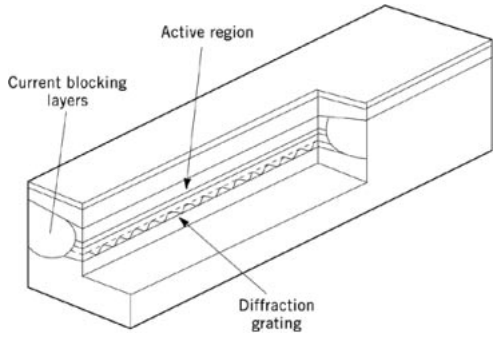

Fig. 12. Cut-out drawing of a distributed feedback (DFB) laser showing the active region and a diffraction grating, under the active layer, which produces the feedback.

not possible to distinguish between the individual pulses. The only way to avoid this problem is to work with lasers emitting light in a single and very narrow mode. The most effective of such devices is the distributed feedback (DFB) laser shown in a cross section of Figure 12.

Selection of a single longitudinal mode is accomplished by the introduction of a diffraction grating in one of the waveguide layers of a separate confinement cavity (17, 32). The grating extends throughout the length of the cavity and produces feedback, hence the name of this laser, at a single wavelength. The wavelength depends on the period of the grating and the effective index of the cavity. Lasing at other wavelengths would require much higher gain and thus cannot be reached. The grating, which requires a fairly fine pitch and a period of ca 230 nm, is usually prepared by holographic methods (see Holography). A successful addition of such a grating requires a great deal of sophistication in the crystal growth. The grating must be carefully preserved in the overgrowth. At the same time the overgrown layer must become planar after less than 0.1–0.2 μm of growth. If this is not accomplished the resulting undulations in the active layer introduce high loss at all wavelengths.

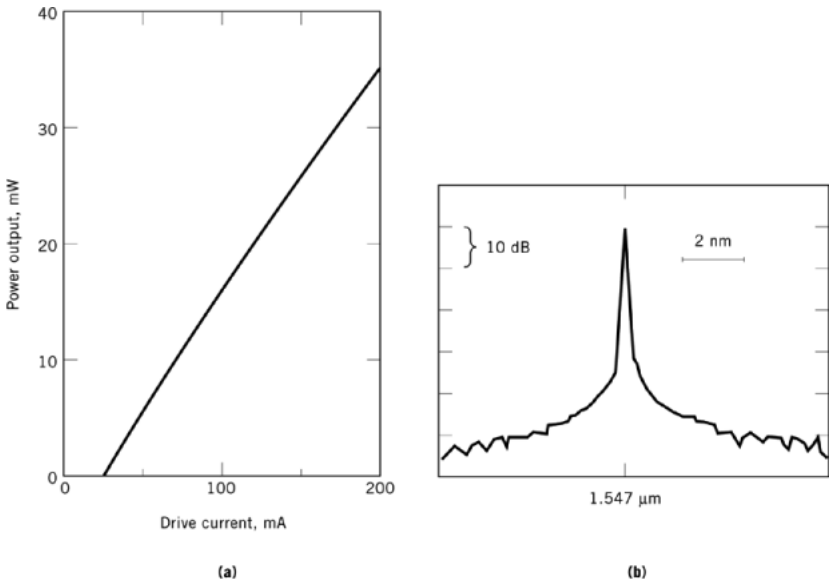

Fig. 13. Characteristics of a 50-μm long DFB laser. (**a**) Light-current properties, (**b**) spectral intensity plotted on a logarithmic scale to better illustrate suppression of undesired modes.

What is desired is a periodic perturbation in the index of refraction of the active layer and not the width of the layer itself. After the DFB active layer structure is prepared it is then incorporated into a CMBH laser of the type discussed above.

When properly done the incorporation of a grating changes only the laser spectrum. The light-current characteristics of a DFB laser are illustrated in Figure 13**a.** The laser has low threshold current and the output power is highly linear to at least 50 mW. The spectrum, shown in Figure 13**b,** is characterized by a single narrow line at the feedback frequency of the grating. All other modes are suppressed; their intensities are at least a thousand times weaker than the main line. In DFB lasers based on strained quantum wells, the width of the laser line is so narrow that it must be measured in units of frequency. In the best lasers of this type the linewidth is lower than 250 kHz or approximately 2×10^{-6} nm.

BIBLIOGRAPHY

"Light-Emitting Diodes and Semiconductor Lasers" in *ECT* 3rd ed., Vol. 14, pp. 269–294, by M. B. Parish, Bell Laboratories.

Cited Publications

1. A. S. Grove, *Physics and Technology of Semiconductor Devices*, John Wiley & Sons, Inc., New York, 1967.
2. S. M. Sze, *Physics of Semiconductor Devices*, John Wiley & Sons, Inc., New York, 1981.
3. B. G. Streetman, *Solid State Electronic Devices*, Prentice Hall, Englewood Cliffs, N.J., 1990.
4. M. Shur, *Physics of Semiconductor Devices*, Prentice Hall, Englewood Cliffs, N.J., 1990.
5. A. Einstein, *Phys. Z.* **18**, 121 (1917).
6. H. C. Casey, Jr. and M. B. Panish, *Heterostructure Lasers*, Academic Press, Inc., New York, 1978.
7. M. G. A. Bernard and G. Duraffourg, *Phys. Stat. Solidi.* **1**, 699 (1961).
8. C. H. Henry, in Y. Yamamoto, ed., *Coherence, Amplification, and Quantum Effects in Semiconductor Lasers*, John Wiley

& Sons, Inc., New York, 1991.
9. H. Kroemer, *Proc. IEEE* **51**, 1782 (1963).
10. USSR Pat. 181,737 (1963), Zh. I. Alferov and R. F. Kazarinov.
11. G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures*, Halsted Press, New York, 1988.
12. P. S. Zory, Jr., ed., *Quantum Well Lasers*, Academic Press, Inc., Boston, 1993.
13. G. C. Osbourn, *Phys. Rev.* **B27**, 5126 (1983).
14. J. W. Matthews and A. E. Blakeslee, *J. Cryst. Growth* **27**, 118 (1974).
15. E. Yablonovitch and E. O. Kane, *J. Lightwave. Technol.* **LT-4**, 504 (1986).
16. T. P. Pearsall, ed., *GaInAsP Alloy Semiconductors*, John Wiley & Sons, Inc., New York, 1982.
17. G. P. Agrawal and N. K. Dutta, *Semiconductor Lasers*, van Nostrand Reinhold, New York, 1993.
18. M. B. Panish and H. Temkin, *Gas Source Molecular Beam Epitaxy*, Springer Verlag, Berlin, 1994.
19. N. Chand and co-workers, *Appl. Phys. Lett.* **58**, 1704 (1991).
20. S. Nakamura, M. Senoh, and T. Mukai, *Appl. Phys. Lett.* **62**, 2390 (1993).
21. M. A. Haase and co-workers, *Appl. Phys. Lett.* **59**, 1272 (1991).
22. A. Salokatve and co-workers, *Electron. Lett.*, (Nov. 1993).
23. H. Okuyama and co-workers, *J. Cryst. Growth* **117**, 139 (1992).
24. J. M. Gaines and co-workers, *Appl. Phys. Lett.* **62**, 2462 (1993).
25. R. M. Park and co-workers, *Appl. Phys. Lett.* **57**, 2127 (1990).
26. D. Scifres, R. D. Burnham, and W. Streifer, *Appl. Phys. Lett.* **33**, 1015 (1978).
27. C. H. Henry and co-workers, *J. Appl. Phys.* **50**, 3721 (1979).
28. A. Rosen and co-workers, *IEEE Photonics Techn. Lett.* **1**, 43 (1989).
29. K. Iga, F. Koyama, and S. Kinoshita, *IEEE J. Quant. Electron.* **24**, 1845 (1988).
30. J. Jewell and co-workers, *IEEE J. Quant. Electron.* **27**, 1332 (1991).
31. D. Tauber and co-workers, *Appl. Phys. Lett.* **62**, 325 (1993).
32. H. Temkin and co-workers, *Appl. Phys. Lett.* **57**, 1295 (1990).

HENRYK TEMKIN
Colorado State University

Related Articles

Thin films, film formation; Semiconductors; Light generation, light emitting diodes; Ion implantation