UNIT- V

8. SEMICONDUCTOR PHYSICS

INTRODUCTION:

Semiconductors are materials whose electronic properties are intermediate between those of conductors and insulators. These electrical properties of a solid depend on its band structure. A semiconductor has two bands of importance (neglecting bound electrons as they play no part in the conduction process) the valence and the conduction bands. They are separated by a forbidden energy gap. At OK the valence band is full and the conduction band is empty, the semiconductor behaves as an insulator. Semiconductor has both positive (hole) and negative (electron) carriers of electricity whose densities can be controlled by doping the pure semiconductor with chemical impurities during crystal growth.

At higher temperatures, electrons are transferred across the gap into the conduction band leaving vacant levels in the valence band. It is this property that makes the semiconductor a material with special properties of electrical conduction.

Generally there are two types of semiconductors. Those in which electrons and holes are produced by thermal activation in pure Ge and Si are called intrinsic semiconductors. In other type the current carriers, holes or free electrons are produced by the addition of small quantities of elements of group III or V of the periodic table, and are known as extrinsic semiconductors. The elements added are called the impurities or dopants.

Intrinsic semiconductors:

A pure semiconductor which is not doped is termed as intrinsic semiconductor. In Si crystal, the four valence electrons of each Si atom are shared by the four surrounding Si atoms. An electron which may break away from the bond leaves deficiency of one electron in the bond. The vacancy created in a bond due to the departure of an electron is called a hole. The vacancy may get filled by an electron from the neighboring bond, but the hole then shifts to the neighboring bond which in turn may get filled by electron from another bond to whose place the hole shifts, and so on thus in effect the hole also undergoes displacement inside a crystal. Since the hole is associated with deficiency of one electron, it is equivalent for a positive charge of unit magnitude. Hence in a semiconductor, both the electron and the hole act as charge carriers.

In an intrinsic semiconductor, for every electron freed from the bond, there will be one hole created. It means that, the no of conduction electrons is equal to the no of holes at any given temperature. Therefore there is no predominance of one over the other to be particularly designated as charge carriers.

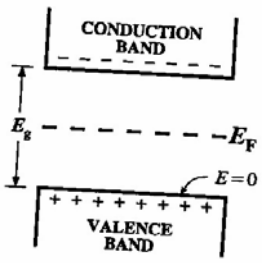
Carriers Concentration in intrinsic semiconductors:

A broken covalent bond creates an electron that is raised in energy, so as to occupy the conduction bond, leaving a hole in the valence bond. Both electrons and holes contribute to overall conduction process.

In an intrinsic semiconductor, electrons and holes are equal in numbers. Thus

$$n = p = n_i$$

Where n is the number of electrons in the conduction band in a unit volume of the material (concentration), p is the number of holes in the valance band in a unit volume of the material. And n_i, the number density of charge carriers in an intrinsic semi conductor. It is called intrinsic density.



FERMI LEVEL IN INTRINSIC SEMICONDUCTOR

For convenience, the top of the valence bond is taken as a zero energy reference level arbitrarily.

The number of electrons in the conduction bond is

$$n = N\,P(Eg)$$

Where P(Eg) is the probability of an electron having energy Eg. It is given by Fermi Dirac function eqn., and N is the total number of electrons in both bands.

Thus,

$$n = \frac{N}{1 + \exp[(Eg - E_F)/KT]}$$

Where E_F is the Fermi Level

The probability of an electron being in the valence bond is given by putting Eg = 0 in eqn. Hence, the number of electrons in the valence bond is given by

$$n_v = \frac{N}{1 + \exp(-E_F/KT)}$$

The total number of electrons in the semiconductor. N is the sum of those in the conduction band n and those in the valence bond n_v. Thus,

$$N = \frac{N}{1 + \exp[(Eg - EF)/KT]} + \frac{N}{1 + \exp(-E_F/KT)}$$

For semiconductors at ordinary temperature, Eg >> KT as such in equation one may be neglected when compared with $\exp\left(\frac{Eg - E_F}{RT}\right)$ Then

$$1 = \frac{1}{\exp\left(\frac{Eg - E_F}{RT}\right)} + \frac{1}{1 + \exp\left(\frac{-E_F}{RT}\right)}$$

Rearranging the terms, we get

$$\exp\left(\frac{-E_g + E_F}{RT}\right) = \frac{\exp(-E_F/RT)}{1 + \exp(-E_F/RT)}$$

$$\approx \exp\left(\frac{-E_F}{RT}\right)$$

or $$\exp\left(\frac{2E_F - E_g}{RT}\right) = 1$$

This leads to

$$E_F = E_g/2$$

Thus in an intrinsic semiconductor, the Fermi level lies mid way between the conduction and the valence bonds. The number of conduction electrons at any temperature T is given by

$$n = \frac{N}{1 + \exp(E_g/2KT)} \quad (\therefore E_F = E_g/2)$$

In eqn may be approximated as

$$n \approx N \exp(-E_g / 2RT)$$

From the above discussion, the following conclusions may be drawn.

a) The number of conduction electrons and hence the number of holes in an intrinsic semiconductor, decreases exponentially with increasing gap energy Eg this accounts for lack of charge carries in insulator of large forbidden energy gap.
b) The number of available charge carries increases exponentially with increasing temperature.

The above treatment is only approximate as we have assumed that all states in a bond have the same energy. Really it is not so. A more rigorous analysis must include additions terms in eqn.

The no of conduction bond, in fact is given by

$$n = \int S(E)\, P(E)\, dE$$

Where S(E) is the density of available states in the energy range between E and E + dE, and P(E) is the probability, that an electron can occupy a state of energy E.

$$S(E) \frac{8\sqrt{2}\,\pi\, m^{3/2}}{n^3} E^{1/2}$$

Inclusion of S(E) and integration over the conduction bond leads to

$$n = N_e \exp[(-Eg - E_F)/RT]$$

In a similar way, we arrive at

$$p = N_V \exp[-E_F/RT]$$

If we multiply eq: we get

$$np = n_i^2 = N_e N_V \exp(-Eg/RT)$$

For the intrinsic material

$$Ni = 2 \frac{(2\pi RT)^{3/2}}{h^2} (m_e^* m_n^*)^{3/4} \exp(-Eg/2RT)$$

Notice that this expression agress with the less rigorous one derived earlier since the temperature dependence is largely controlled by the rapidly varying exponential term.

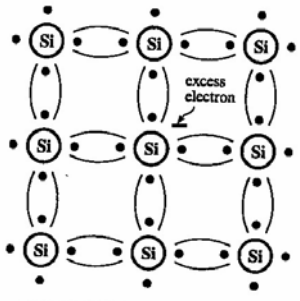
EXTRINSIC SEMICONDUCTORS:

Intrinsic Semiconductors are rarely used in semiconductor devices as their conductivity is not sufficiently high. The electrical conductivity is extremely sensitive to certain types of impurity. It is the ability to modify electrical characteristics of the material by adding chosen impurities that make extrinsic semiconductors important and interesting.

Addition of appropriate quantities of chosen impurities is called doping, usually, only minute quantities of dopants (1 part in 10^3 to 10^{10}) are required. Extrinsic or doped semiconductors are classified into main two main types according to the type of charge carries that predominate. They are the n-type and the p-type.

N-TYPE SEMICONDUCTORS:

Doping with a pentavalent impurity like phosphorous, arsenic or antimony the semiconductor becomes rich in conduction electrons. It is called n-type the bond structure of an n-type semiconductor is shown in Fig below.



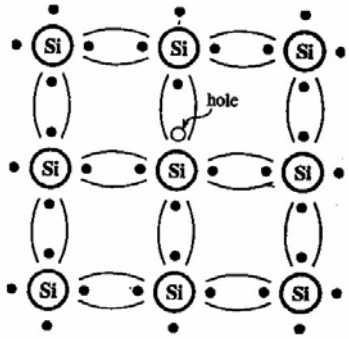
FREE ELECTRON DUE TO DOPING

Even at room temperature, nearby all impurity atoms lose an electron into the conduction bond by thermal ionization. The additional electrons contribute to the conductivity in the same way as those excited thermally from the valence bond. The essential difference beam

the two mechanisms is that ionized impurities remain fixed and no holes are produced. Since penta valent impurities denote extra carries elections, they are called donors.

P-TYPE SEMICONDUCTORS:-

p-type semiconductors have holes as majority charge carries. They are produced by doping an intrinsic semiconductor with trivalent impurities.(e.g. boron, aluminium, gallium, or indium). These dopants have three valence electrons in their outer shell. Each impurity is short of one electron bar covalent bonding. The vacancy thus created is bound to the atom at OK. It is not a hole. But at some higher temperature an electron from a neighbouring atom can fill the vacancy leaving a hole in the valence bond for conduction. It behaves as a positively charge particle of effective mass m_h*. The bond structure of a p-type semiconductor is shown in Fig below.



FORMATION OF HOLE DUE TO DOPING

Dopants of the trivalent type are called acceptors, since they accept electrons to create holes above the tope of the valence bond. The acceptor energy level is small compared with thermal energy of an electron at room temperature. As such nearly all acceptor levels are occupied and each acceptor atom creates a hole in the valence bond. In extrinsic semiconductors, there are two types of charge carries. In n-type, electrons are more than holes. Hence electrons are majority carriers and holes are minority carries. Holes are majority carries in p-type semiconductors; electrons are minority carriers.

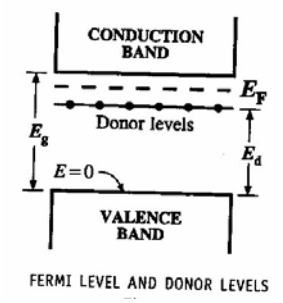
CARRIER CONCENTRATION IN EXTRINSIC SEMICONDUCTORS:

Equation gives the relation been electron and hole concentrations in a semiconductor. Existence of charge neutrality in a crystal also relates n and p. The charge neutrality may be stated as

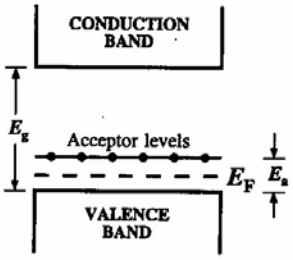
$$N_D + p = N_A + n$$

Since donors atoms are all ionized, N_D positive charge per cubic meter are contributed by N_D donor ions. Hence the total positive charge density $= N_D + p$. Similarly if N_A is the concentration of the acceptor ions, they contribute N_A negative charge per cubic meter. The total negative charges density $= N_A + n$. Since the

semiconductor is electrically neutral the magnitude of the positive charge density must be equal to the magnitude of the total negative charge density.



FERMI LEVEL AND DONOR LEVELS



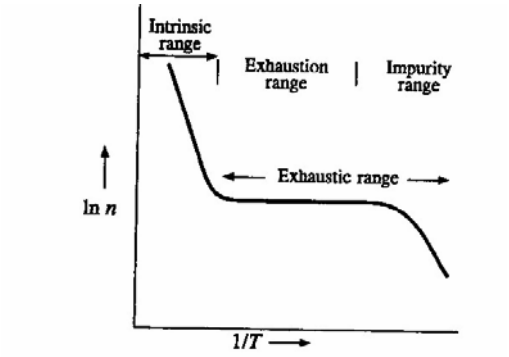
FERMI LEVEL AND ACCEPTOR LEVELS

n-type material : $N_A = 0$ Since $n >> p$, eqn reduces to $n \approx N_D$ i.e., in an n-type material the where subscript n indicates n-type material. The concentration p_n of holes in the n-type semiconductor is obtained from eqn i.e.,

$$n_n p_n = n^2_i$$

Thus $$p_n \approx \frac{n^2_i}{N_D}$$

Similarly, for a p-type semiconductor $p_p \approx N_A$ and $n_p \approx \frac{n^2_i}{N_A}$



DEPENDENCE OF ELECTRON CONCENTRATION ON TEMPERATUR

Expression for electrical conductivity:

There are two types of carries in a Semiconductor electrons and holes. Both these carries contribute to conduction. The general expression for conductivity can be written down as

$$\sigma = e(n\mu_e + \rho\mu_h)$$

Where μ_e and μ_h are motilities of electrons and holes respectively.

A. Intrinsic Semiconductor; For an intrinsic Semiconductor

$$n = p = n_i$$

eqn becomes

$$\sigma_i = en_i(\mu_n + \mu_p)$$

If the scattering is predominantly due to lattice vibrations.

$$\mu_e = AT^{3/2}$$
$$\mu_h = BT^{3/2}$$

We may put $\mu_e + \mu_h = (A + B)T^{3/2} = CT^{3/2}$

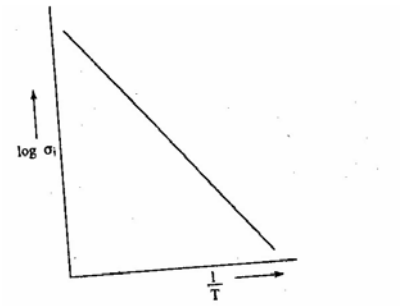
$$\sigma_i = n_i CT^{3/2}$$

Substituting for n_i from eq we get

$$\sigma_i = 2\left(\frac{2\mu RT}{h2}\right)^{3/2} CT^{3/2}(m_e^* m_h^*)^{3/4} \exp\left(\frac{-Eg}{2RT}\right)$$

$$\log \sigma_i = \log x - \frac{Eg}{2RT}$$

A graph of $\log \sigma_i$ Vs 1/T gives a straight line shown in fig below:



Variation of electrical conductivity with temperature

LIFE TIME OF MINORITY CARRIER:

In Semiconductor devices electron and hole concentrations are very often disturbed from their equilibrium values. This may happen due to thermal agitation or incidence of optical radiation. Even in a pure Semiconductor there will be a dynamic equilibrium. In a pure Semiconductor the number of holes is equal to the number of free electrons. Thermal agitation continuously produces of new EHP per unit volume per second while other EHP disappear due to recombination. On the average a hole exists for a time period of Tp while an electron exists for a time period Tn before recombination take place. This time is called the mean life time. If we are dealing with holes in an n-type Semiconductor, Tp is called the minority carrier life time. These parameters are important in Semiconductor devices as they indicate the time required for electron and hole concentration to return to their equilibrium values after they are disturbed.

Let the equilibrium concentration of electrons and holes in an n-type Semiconductor be n_0 and p_0 respectively. If the specimen is illuminated at $t = t_i$, additional EHP_S are generated throughout the specimen. The existing equilibrium is disturbed and the new equilibrium concentrations are $\bar{p}$ and $\bar{n}$. The excess concentration of holes $= \bar{p} - p_0$ – Excess concentration of electrons $= \bar{n} - n_0$ Since the radiation creates EHP_S.

$$p - p_0 = n - n_0$$

Due to incident radiation equal no of holes and electrons are created. How ever, the percentage increase of minority carriers is much more than the percentage of majority carries. In fact, the majority charge carrier change is negligibly small. Hence it is the minority charge carrier density that is important. Hence, we shall discuss the behaviour of minority carriers.

As indicated in radiation is removed at t = 0. Let us investigate how the minority carrier density returns to its original equilibrium value.

The hole concentration decreases as a result of recombination. Decrease in hole concentration per second due to recombination = p/T_p. But the increase in hole concentration per second due to thermal generation = g Since charge can neither be created nor destroyed

$$\frac{dp}{dt} = g - \frac{p}{T_p}$$

When the radiation is with drawn, the hole concentration p reaches equilibrium value p_0 Hence g = p0/Tp. Then eqn may be rewritten as

$$\frac{dp}{dt} = \frac{p_0 - p}{T_p}$$

We define excess carrier concentration. Since p is a function of time.

$p' = p - p_0 = p'(t)$
from we may write $\frac{dp}{dt} = \frac{-p'}{T_p}$

The solution to the above differential equation is given by

$$p'(t) = p(0)\, e^{-t/Tp}$$

The excess concentration decreases exponentially to zero with a time constant T_p.

DRIFT CURRENT:

In an electric field E, the drift velocity V_d of carriers superposes on the thermal velocity V_{th}. But the flow of charge carriers results in an electric current, known as the drift current. Let a field E be applied, in the positive creating drifts currents J_{nd} and J_{pd} of electrons and holes respectively.

Without E, the carriers move randomly with rms velocity V_{th}. Their mean velocity is zero. The current density will be zero. But the field E applied, the electrons have the velocity V_{de} and the holes V_{dh}.

Consider free electrons in a Semiconductor moving with uniform velocity Vde in the negative x direction due to an electric field E. Consider a smaller rectangular block of AB of length V_{de} inside the Semiconductor. Let the area of the side faces each be unity. The total charge Q in the elements AB is

Q = Volume of the element x density of partially change on each particle
$= (V_{de} \times 1 \times 1) \times n \times -q$
Thus $Q = -qnV_{de}$

Where n is the number density of electrons. The entire charge of the block will cross the face B, in unit time. Thus the drift current density Jnd due to free electrons at the face B will be.

$$J_{nd} = - q\, nV_{de}$$

Similarly for holes $J_{pd} = q\, nV_{dh}$

but $\mathbf{V_{de} = -\mu_n E}$

and $V_{dh} = \mu_p E$

hence $J_{nd} = n\, q\mu_n E$

and $J_{pd} = p\, q\mu_n E$

The total drift current due to both electrons and holes J_d is

$$J_d = J_{nd} + J_{pd} = (nq\mu_n + pq\,\mu_p)E$$

Even though electrons and holes move in opposite direction the effective direction of current flow, is the same for both and hence they get added up. Ohm's Law can be written in terms of electrical conductivity, as

$$J_d = \sigma E$$

Equating the RHS of eq we have

$$\sigma = nq\mu_n + pq\mu_p = \sigma_n + \sigma_p$$

For an intrinsic Semiconductor $n = p = n_i$

$$\sigma_i = n_i\, q(\mu_n + \mu_p)$$

DIFFUSION CURRENTS:

1. Diffusion Current: Electric current is Setup by the directed movement of charge carriers. The movements of charge carriers could be due to either drift or diffusion. Non-uniform concentration of carriers gives rise to diffusion. The first law of diffusion by Fick States that the flux F, i.e., the particle current is proportional and is directed to opposite to the concentration gradient of particles. It can be written mathematically, in terms of concentration N, as

$$F = -D\, \bar{V}\, N$$

Where D stands for diffusion constant.

In one dimension it is written as

$$F = -D \frac{\partial N}{\partial x}$$

In terms of Je and Jp the flux densities of electrons holes and their densities n and p respectively.

We get $J_e = -D_n \frac{\partial N}{\partial x}$

and $J_n = -D_p \frac{\partial p}{\partial x}$

Where D_n and D_p are the electron and hole diffusion constant constants respectively. Then the diffusion current densities become

$$J_n \text{ diff} = q D_n \frac{\partial N}{\partial x}$$

$$J_p \text{ diff} = - q D_p \frac{\partial p}{\partial x}$$

THE EINSTEIN RELATIONS:

When both the drift and the diffusion currents are present total electron and hole current densities can be summed up as

$$J_n = J_{nd} + J_n \text{ diff}$$

$$J_n = nq\mu_n E + qD_n \frac{\partial n}{\partial x}$$

$$J_p = pq\mu_p E - qD_p \frac{\partial p}{\partial x}$$

Now, let us consider a non uniformly doped n-type slab of the Semiconductor fig shown below (9) under thermal equilibrium. Let the slab be intrinsic at x = 0 while the donor concentration, increases gradually upto x = 1, beyond which it becomes a constant. Assume that the Semiconductor is non-degenerate and that all the donors are ionized. Due to the concentration gradient, electrons tend to diffuse to the left to x = l. This diffusion leaves behind a positive charge of ionized donors beyond x = 1 and accumulates electron near x = 0

plane. This charge imbalance. Sets up an electric field in which the electrons experience fill towards x = 1.

Fig shows illustrates the equilibrium potential $\phi(x)$ fig shown refers to the bond diagram of the Semiconductor.

Both E1 and EF coincide till x = 0 when n0 = ni that EF continues to be the same throughout the slab. But since the bond structure is not changed due to doping, the bond edges bond with equal separation all along. How ever, the level Ei continues to lie midway between EV and E.

In thermal equilibrium, the electrons tend to diffuse down the concentration tending to setup a current from the right to left. The presence of electric field tends to set up drift current of electrons in the opposite direction. Both the currents add upto zero. Thus we obtain

$$J_n = qD_n \frac{\partial n}{\partial x} + nq\mu_n E = 0$$

$$\text{i.e.,} \quad D_n \frac{\partial n}{\partial x} + n\mu_n E = 0$$

For a non degenerate Semi conductor.

$$n(x) = n_i \exp \frac{E_F - E_i(x)}{RT}$$

Thus relation is valid at all points in the Semiconductor further. The electronic concentration is not influenced by the small in balance of charge. Energy is defined in terms of $\phi(x)$ the potential.

$$E(x) = -q\phi(x)$$

Then $E_F - E_i(x) = E_i(0) - E_i(x) = -q[\phi(0) - \phi(x)]$

$$= q\phi(x)$$

Assuming $\phi(0) = 0$ we get $n(x) = n_i \exp \frac{q\phi(x)}{RT}$

$$\frac{dn}{} \quad -d\phi$$

Substituting ------ from eqn along with E = ----- we get

$$\frac{dx}{} \qquad \frac{dx}{}$$

$$D_n \frac{nq}{RT}\frac{d\phi}{dx} = \mu_n n \frac{d\phi}{dx}$$

Simplifying we obtain, $D_n = \frac{RT}{q}\mu_n$

Simplifying for holes $D_p = \frac{RT}{q}\mu_p$

These are known as Einstein relations and the factor (RT/q) as thermal voltage. The above relations hold good only for non degenerate Semiconductors. For the degenerate case the Einstein's relations are complex.

It is clear from the Einstein's relation that Dp μp and Dn,μn are related and they are functions of temperature also. The relation of diffusion constant D and the mobility μ confirms the fact, that both the diffusion an drift processes arise due to thermal motion and scattering of free electrons, even though they appear to be different.

EQUATION OF CONTINUITY:

If the equilibrium concentrations of carriers in a Semiconductor are disturbed, the concentrations of electrons and holes vary with time. How ever the carrier concentration in a Semiconductor is a function of both time and position.

The fundamental law governing the flow of charge is called the continuity equation. It is arrived at by assuming law to conservation of charge provided drift diffusion and recombination processes are taken into account.

Consider a small length Δx of a Semiconductor sample with area A in the Z plane fig shown above. The hole current density leaving the volume (ΔxΔ) under consideration is Jp (x + Δx) and the current density entering the volume is Jp(x). Jp (x + Δx) may be smaller or

larger than Jp(x) depending upon the generation and recombination of carriers in the element. The resulting change in hole concentration per unit time.

∂p = hole flux entering per unit time – hole flux leaving per unit

$$\frac{\partial p}{\partial t}_{x \to x+\Delta x} = Jp(x) - \frac{Jp(x+\Delta x)}{q\,\Delta x} - \frac{\delta p}{Tp}$$

Where Tp is the recombination life time. According to eqn, the rate of hole build up is equla to the rate of increase of hole concentration remains the recombination rate. As Δx approaches zero, we may write

$$\frac{\partial p}{\partial t}(x,t) = \frac{\partial\, \delta p}{\partial t} = \frac{-1\, \partial\, Jp}{q\ \partial x} - \frac{\delta p}{Tp}$$

The above is called the continuity equation for holes for electrons

$$\frac{\partial \delta n}{\partial t} = \frac{1\ \partial J_n}{q\ dx} - \frac{\delta n}{Tn}$$

If there is no drift we may write

$$J_n(diff) = qD_n \frac{\partial \delta}{\partial x}$$

Substituting the above eqn we get the following diffusion eqn for electrons.

$$\frac{\partial \delta n}{\partial t} = D_n \frac{\partial^2 \delta n}{\partial x^2} - \frac{\delta n}{T_n}$$

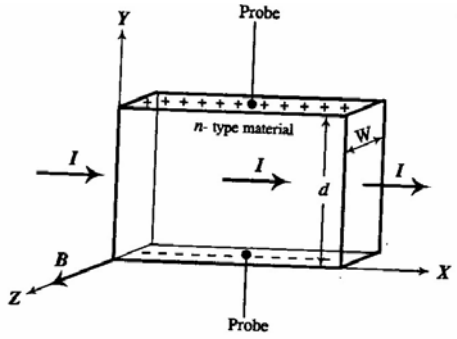
For holes we may write

$$\frac{\partial \delta p}{\partial t} = D_p \frac{\partial^2 \delta p}{\partial x^2} - \frac{\delta p}{T_p}$$

HALL EFFECT:

When a material carrying current is subjected to a magnetic field in a direction perpendicular to the direction of current, an electric field is developed across the material in a direction perpendicular to both the direction of the magnetic field and the current direction. This phenomenon is called Hall Effect.

Hall Effect finds important application in studying the electron properties of semi conductor, such as determination of carrier concentration and carrier mobility. It also used to determine whether a semi conductor is n-type, or p- type.



GENERATION OF ELECTRIC FIELD DUE TO HALL EFFECT

THEORY:

Consider a rectangular slab of an n-type Semiconductor carrying current in the positive x-direction. The magnetic field B is acting in the positive direction as indicated in fig above. Under the influence of the magnetic field, electrons experience a force F_L given by

$$F_L = - Bev \quad \text{---------- (1)}$$

Where e = magnitude of the charge of the electron

v = drift velocity

Appling the Fleming's Left Hand Rule, it indicates a force F_H acting on the electrons in the negative y-direction and electron are deflected down wards. As a consequence the lower face of the specimen gets negatively charged (due to increases of electrons) and the upper face

positively charged (due to loss of electrons). Hence a potential V_H, called the Hall voltage appears between the top and bottom faces of the specimen, which establishes an electric field E_H, called the Hall field across the conductor in negative y-direction. The field E_H exerts an upward force F_H on the electrons. It is given by

$$F_H = - eE_H \quad \text{----------(2)}$$

F_H acts on electrons in the upward direction. The two opposing forces F_L and F_H establish an equilibrium under which

$$|F_L = F_H$$

using eqns 1 and 2 $\quad -Bev = -eE_H$

$$E_H = Bv \quad \text{----------(3)}$$

If 'd' is the thickness of the Specimen

$$E_H = \frac{V_H}{d}$$

$$V_H = E_H\, d = Bvd \quad \text{from eqn (3)---------- 4}$$

If ω is the width of the specimen in z- direction.

The current density

$$J = \frac{I}{\omega d}$$

But $J = nev = \rho v$ ---------- 5

Where n = electron concentration

And ρ = charge density

$$\therefore \quad \rho v = \frac{I}{\omega d}$$

$$v = \frac{I}{\rho\omega d} \quad \text{---------- 6}$$

Substitutinf for v, from eqns 6 and 4

$$V_H = BI / \rho\omega$$

$$\rho = \frac{BI}{V_H \omega}$$

Thus, by measuring V_H, I, and ω and by knowing B, the charge density ρ can be determined.

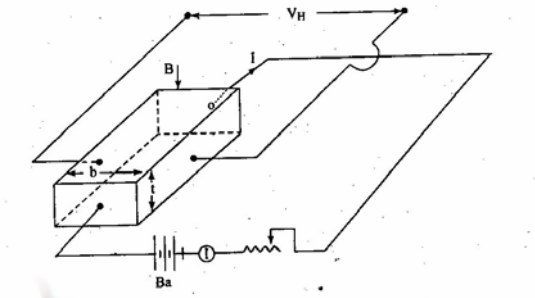


Fig. Experimental set up for the measurement of Hall voltage.

Hall Coefficient:

The Hall field E_H, for a given material depends on the current density J, and the applied field B

i.e., $E_H \propto JB$

$$E_H = R_H JB$$

Where R_H is called the Hall Coefficient

Since $V_H = \frac{BI}{\text{--------}}$

$$\rho\omega$$

$$E_H = \frac{V_H}{J\omega d}$$

$$J = \frac{I}{\omega d}$$

$$\frac{BI}{J\omega d} = R_H \frac{I}{\omega d} B$$

This leads to $R_H = \frac{I}{\rho}$

Mobility of charge carriers:

The mobility μ is given by $\mu = \frac{v}{E}$

But $J = \sigma E = nev = \rho v$

$\therefore \sigma E = \rho v$

$$\text{or } E = \frac{\rho v}{\sigma}$$

$$\Rightarrow \mu = \frac{\sigma}{\rho} = \sigma R_H \quad (\therefore 1/\rho = R_H)$$

σ is the conductivity of the semi conductor.

(C) Applications

(a) Determination of the type of Semiconductor: The Hall Coefficient R_H is negative for an n-type Semiconductor and positive for a p-type material. Thus, the sign of the Hall coefficient can be utilized to determine whether a given Semiconductor is n or p type.

(b) Determination of Carrier Concentration: Equation relates the Hall Coefficient RH and charge density is

$$R_H = \frac{1}{p} = -\frac{1}{ne} \text{ (for n-type}$$

$$= \frac{1}{pe} \text{ (for p-type}$$

$$\text{Thus } n = \frac{1}{eR_H}$$

$$\text{and } \frac{1}{eR_H}$$

(c) Determination of mobility: According to equation the mobility of charge carriers is given by

$$\mu = \sigma|R_H|$$

Determination of σ and R_H leads to a value of mobility of charge carriers.

(d) Measurement of Magnetic Induction (B):- The Hall Voltage is proportional to the flux density B. As such measurement of VH can be used to9 estimate B.

9. PHYSICS OF SEMICONDUCTOR DEVICES

To describe the basic structure and general properties of semiconductor devices. We will not be showing such devices in working circuits in this set of pages; other pages already under development will serve that purpose. However, I have received a number of inquiries on the order of "What is a MOSFET?" and "What's inside a transistor?" This set of pages is intended to answer such questions.In some cases we'll be dealing with some rather technical terms, and we will sometimes have to deal with some essential concepts involved in physics. More general explanations and definitions will also be given, so if you don't need the technical definitions, don't worry about them. The terms are present in case you actually need them. These pages will begin with basic semiconductor structure and what happens when impurities are added to a pure silicon crystal through a process known as "doping." We'll look at what happens when two or three different regions are created within a single silicon crystal. Then we'll start to look at variations: field-effect devices, devices with four and even five different regions, and finally the kinds of effects we can get when we change the amount of impurities within the crystal.

Semiconductor diodes are normally one of the following types:
1. Grown junction diode
2. Alloy type or fused junction diode
3. Diffused junction diode
4. Epitaxial grown or planar diffused diode
5. Point contact diode

Semiconductor diode fabrication types
Fabrication techniques of a P-N junction diode

1. ***Grown Junction Diode:*** Diodes of this type are formed during the *crystal pulling* process. P and N-type impurities can be alternately added to the molten semiconductor material in the crucible, which results in a P-N junction, as shown when crystal is pulled. After slicing, the larger area device can then be cut into a large number (say in thousands) of smaller-area semiconductor diodes. Though such diodes, because of larger area, are capable of handling large currents but larger area also introduces more capacitive effects, which are undesirable. Such diodes are used for low frequencies.

2. ***Alloy Type or Fused Junction Diode:*** Such a diode is formed by first placing a P- type impurity (a tiny pellet of aluminium or some other P- type impurity) into the surface of an N-type crystal and heating the two until liquefaction occurs where the two materials meet. An alloy will result that on cooling will give a P-N junction at the boundary of the alloy substrate. Similarly, an N-type impurity may be placed into the surface of a P- type crystal and the two are heated until liquefaction occurs. Alloy type diodes have a high current rating and large PIV (peak inverse voltage) rating. The junction capacitance is also large, due to the large junction area.

3. ***Diffused Junction Diode:*** Diffusion is a process by which a heavy concentration of particles diffuse into a surrounding region of lower concentration. The main difference between the diffusion and alloy process is the fact that liquefaction is not reached in the diffusion process. In the diffusion process heat is applied only to increase the activity of elements involved. For formation of such diodes, either solid or gaseous diffusion process can be employed. The process of solid diffusion starts with formation of layer of an acceptor impurity on an N- type substrate and heating the two until the impurity diffuses into the substrate to form the P-type layer, as illustrated in figure. A large P-N junction is divided into parts by cutting process. Metallic contacts are made for connecting anode and cathode leads.

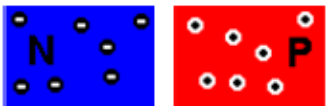
In the process of gaseous diffusion instead of layer formation of an acceptor impurity, an N-type substrate is placed in a gaseous atmosphere of acceptor impurities and then heated. The impurity diffuses into the substrate to form P- type layer on the N- type substrate. Though, the diffusion process requires more time than the alloy process but it is relatively inexpensive, and can be very accurately controlled. The diffusion technique leads itself to the simultaneous fabrication of many hundreds of diodes on one small disc of semiconductor material and is most commonly used in the manufacture of semiconductor diodes. This technique is also used in the production of transistors and ICs (integrated circuits).

4. *Epitaxial Growth or Planar Diffused Diode.* The term "epitaxial" is derived from the Latin terms *epi* meaning 'upon' and *taxis* meaning "arrangement".To construct an epitaxially grown diode, a very thin (single crystal) high impurity layer of semiconductor material (silicon or germanium) is grown on a heavily doped substrate (base) of the same material. This complete structure then forms the N- region on which P- region is diffused. $Si0_2$ layer is thermally grown on the top surface, photo-etched and then aluminium contact is made to the P- region. A metallic layer at the bottom of the substrate forms the cathode to which lead is attached.This process is usually employed in the fabrication of IC chips.

5. *Point Contact Diode.* It consists of an N-type germanium or silicon wafer about 12.5 mm square by 0.5 mm thick, one face of which is soldered to a metal base by radio-frequency heating and the other face has a phosphor bronze or tungsten spring pressed against it. A barrier layer is formed round the point contact by a pulsating current forming process. This causes a P-region to be formed round the wire and since pure germanium is N-type, a very small P-N junction in the shape of a hemisphere is formed round the point contact. The forming process cannot be controlled with precision. Because of small area of the junction, point contact diode can be used to rectify only very small currents (of the order of m A). On the other hand, the shunting capacitance of point contact diodes are very valuable in equipment operating at super high frequencies (as high as 25,000 MHz).

The PN Junction

We've seen that it is possible to turn a crystal of pure silicon into a moderately good electrical conductor by adding an impurity such as arsenic or phosphorus (for an N-type semiconductor) or aluminum or gallium (for a P-type semiconductor). By itself, however, a single type of semiconductor material isn't very useful. Useful applications start to happen only when a single semiconductor crystal contains both P-type and N-type regions. Here we will examine the properties of a single silicon crystal which is half N-type and half P-type.

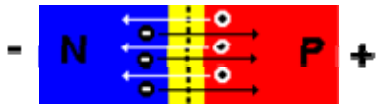


Consider the silicon crystal represented to the right. Half is N-type while the other half is P-type. We've shown the two types separated slightly, as if they were two separate crystals. The free electrons in the N-type crystal are represented by small black circles with a "-" sign inside to indicate their polarity. The holes in the P-type crystal are shown as small white circles with a "+" inside.

In the real world, it isn't possible to join two such crystals together usefully. Therefore, a practical PN junction can only be created by inserting different impurities into different parts of a single crystal. So let's see what happens when we join the N- and P-type crystals together, so that the result is one crystal with a sharp boundary between the two types.

You might think that, left to itself, it would just sit there. However, this is not the case. Instead, an interesting interaction occurs at the junction. The extra electrons in the N region will seek to lose energy by filling the holes in the P region. This leaves an empty zone, or *depletion region* as it is called, around the junction as shown to the right. This action also leaves a small electrical imbalance inside the crystal. The N region is missing some electrons so it has a positive charge. Those electrons have migrated to fill holes in the P region, which therefore has a negative charge. This electrical imbalance amounts to about 0.3 volt in a germanium crystal, and about 0.65 to 0.7 volt in a silicon crystal. This will vary somewhat depending on the concentration of the impurities on either side of the junction.
Unfortunately, it is not possible to exploit this electrical imbalance as a power source; it doesn't work that way. However, we can apply an external voltage to the crystal and see what happens in response. Let's take a look at the possibilities.

Suppose we apply a voltage to the outside ends of our PN crystal. We have two choices. In this case, the positive voltage is applied to the N-type material. In response, we see that the positive voltage applied to the N-type material attracts any free electrons towards the end of the crystal and away from the junction, while the negative voltage applied to the P-type end attracts holes away from the junction on this end. The result is that all available current carriers are attracted away from the junction, and the depletion region grows correspondingly larger. There is no current flow through the crystal because all available current carriers are attracted away from the junction, and cannot cross. (We are here considering an ideal crystal -- in real life, the crystal can't be perfect, and some leakage current does flow.) This is known as *reverse bias* applied to the semiconductor crystal.



Here the applied voltage polarities have been reversed. Now, the negative volatge applied to the N-type end pushes electrons towards the junction, while the positive voltage at the P-type end pushes holes towards the junction. This has the effect of shrinking the depletion region. As the applied voltage exceeds the internal electrical imbalance, current carriers of both types can cross the junction into the opposite ends of the crystal. Now, electrons in the P-type end are attracted to the positive applied voltage, while holes in the N-type end are attracted to the negative applied voltage. This is the condition of *forward bias*.
Because of this behavior, an electrical current can flow through the junction in the forward direction, but not in the reverse direction. This is the basic nature of an ordinary semiconductor diode.

It is important to realize that holes exist only within the crystal. A hole reaching the negative terminal of the crystal is filled by an electron from the power source and simply disappears. At the positive terminal, the power supply attracts an electron out of the crystal, leaving a hole behind to move through the crystal toward the junction again.
In some literature, you might see the N-type connection designated the *cathode* of the diode, while the P-type connection is called the *anode*. These designations come from the days of vacuum tubes, but are still in use. Electrons always move from cathode to anode inside the diode.
One point that needs to be recognized is that there is a limit to the magnitude of the reverse voltage that can be applied to any PN junction. As the applied reverse voltage increases, the depletion region continues to expand. If either end of the depletion region approaches its electrical contact too closely, the applied voltage has become high enough to generate an electrical arc straight through the crystal. This will destroy the diode.

It is also possible to allow too much current to flow through the diode in the forward direction. The crystal is not a perfect conductor, remember; it does exhibit some resistance. Heavy current flow will generate some heat within that resistance. If the resulting temperature gets too high, the semiconductor crystal will actually melt, again destroying its usefulness.

Drift-Diffusion Current Equations

The popular drift-diffusion model can be derived directly from Boltzmann's transport equation by the method of moments [104] or from the basic principles of irreversible thermodynamics [105]. In this model the electron current density is expressed as a sum of two components: The drift component which is driven by the electric field and the diffusion component caused by the gradient of the electron concentration

$$\mathbf{J} = \mathrm{q} \cdot (n \cdot \mu \cdot \mathbf{E} + D_n \cdot \mathrm{grad}\, n) \tag{3.13}$$

where μ and D_n are the mobility and the diffusivity of the electron gas, respectively. It is clear from the above reasoning that for anisotropic materials μ and D_n are all tensors of second rank and have the same form as the representative tensor σ in (3.2). They are related by the Einstein relation

$$D_n = \mu \cdot \frac{\mathrm{k_B} \cdot T_n}{\mathrm{q}} \tag{3.14}$$

where $\mathrm{k_B}$ is the Boltzmann constant and $T_n = T_\mathrm{L}$ the lattice temperature which is constant as the electron gas at drift diffusion is assumed to be in thermal equilibrium.

The current relation (3.13) is inserted into the continuity (3.11) and (3.12) to give a second order parabolic differential equation which is then solved together with POISSON's equation (3.10). More generally, according to the phenomenological equations of drift-diffusion the electron and hole current densities $\mathbf{J}_n$ and $\mathbf{J}_p$ can be expressed as

$$\mathbf{J}_n = \mathrm{q} \cdot \mu_n \cdot n \cdot \left(\mathrm{grad} \left(\frac{E_\mathrm{c}}{\mathrm{q}} - \psi \right) + \frac{\mathrm{k_B} \cdot T_\mathrm{L}}{\mathrm{q}} \cdot \frac{N_{C,0}}{n} \cdot \mathrm{grad} \left(\frac{n}{N_{C,0}} \right) \right), \tag{3.15}$$

$$\mathbf{J}_p = \mathrm{q} \cdot \mu_p \cdot p \cdot \left(\mathrm{grad} \left(\frac{E_\mathrm{v}}{\mathrm{q}} - \psi \right) - \frac{\mathrm{k_B} \cdot T_\mathrm{L}}{\mathrm{q}} \cdot \frac{N_{V,0}}{p} \cdot \mathrm{grad} \left(\frac{p}{N_{V,0}} \right) \right). \tag{3.16}$$

These current relations account for position-dependent band edge energies, E_c and E_v, and position-dependent effective masses, which are included in the effective density of states, $N_{C,0}$ and $N_{V,0}$. The index 0 indicates that $N_{C,0}$ and $N_{V,0}$ are evaluated at some (arbitrary) reference temperature, $\mathbf{T_0}$, which is constant in real space regardless of what the local values of the lattice and carrier temperatures are.

Current–voltage characteristic

A semiconductor diode's behavior in a circuit is given by its current–voltage characteristic, or I–V graph (see graph at right). The shape of the curve is determined by the transport of charge carriers through the so-called *depletion layer* or *depletion region* that exists at the p-n junction between differing semiconductors. When a p-n junction is first created, conduction band (mobile) electrons from the N-doped region diffuse into the P-doped region where there is a large population of holes (places for electrons in which no electron is present) with which the electrons "recombine". When a mobile electron recombines with a hole, both hole and electron vanish, leaving behind an immobile positively charged donor (the dopant) on the N-side and negatively charged acceptor (the dopant) on the P-side. The region around the p-n junction becomes depleted of charge carriers and thus behaves as an insulator. However, the width of the depletion region (called the depletion width) cannot grow without limit. For each electron-hole pair that recombines, a positively-charged dopant ion is left behind in the N-doped region, and a negatively charged dopant ion is left behind in the P-doped region. As recombination proceeds and more ions are created, an increasing electric field develops through the depletion zone which acts to slow and then finally stop recombination. At this point, there is a "built-in" potential across the depletion zone.If an external voltage is placed across the diode with the same polarity as the built-in potential, the depletion zone continues to act as an insulator, preventing any significant electric current flow (unless electron/hole pairs are actively being created in the junction by, for instance, light. see photodiode). This is the *reverse bias* phenomenon. However, if the polarity of the external voltage opposes the built-in potential, recombination can once again proceed, resulting in substantial electric current through the p-n junction (i.e. substantial numbers of electrons and holes recombine at the junction).. For silicon diodes, the built-in potential is approximately 0.6 V. Thus, if an external current is passed through the diode, about 0.6 V will be developed across the diode such that the P-doped region is positive with respect to the N-doped region and the diode is said to be "turned on" as it has a *forward bias*.

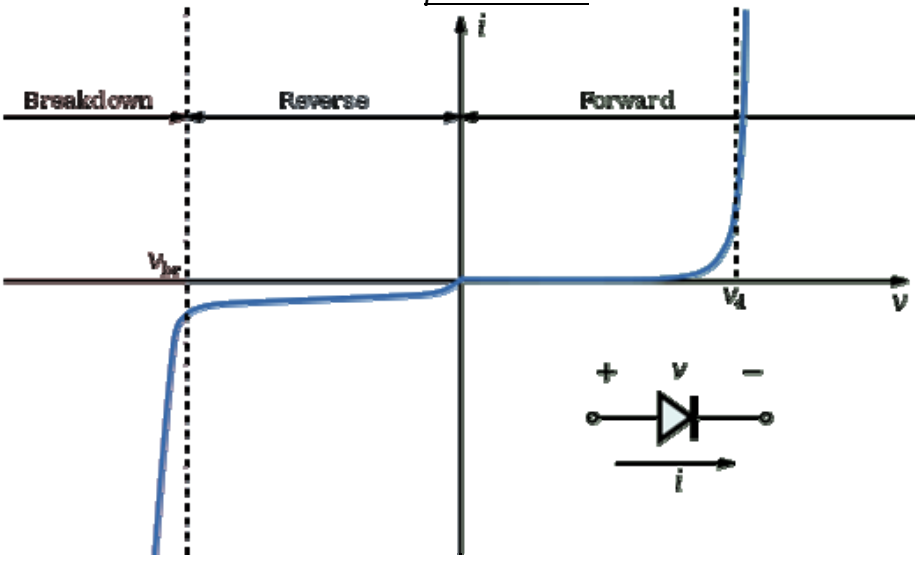


Figure 5: I–V characteristics of a P-N junction diode

A diode's I–V characteristic can be approximated by four regions of operation (see the figure at right).

At very large reverse bias, beyond the peak inverse voltage or PIV, a process called reverse breakdown occurs which causes a large increase in current (i.e. a large number of electrons and holes are created at, and move away from the pn junction) that usually damages the device permanently. The avalanche diode is deliberately designed for use in the avalanche region. In the zener diode, the concept of PIV is not applicable. A zener diode contains a heavily doped p-n junction allowing electrons to tunnel from the valence band of the p-type

material to the conduction band of the n-type material, such that the reverse voltage is "clamped" to a known value (called the *zener voltage*), and avalanche does not occur. Both devices, however, do have a limit to the maximum current and power in the clamped reverse voltage region. Also, following the end of forward conduction in any diode, there is reverse current for a short time. The device does not attain its full blocking capability until the reverse current ceases.The second region, at reverse biases more positive than the PIV, has only a very small reverse saturation current. In the reverse bias region for a normal P-N rectifier diode, the current through the device is very low (in the μA range). However, this is temperature dependent, and at suffiently high temperatures, a substantial amount of reverse current can be observed (mA or more).

The third region is forward but small bias, where only a small forward current is conducted. As the potential difference is increased above an arbitrarily defined "cut-in voltage" or "on-voltage" or "diode forward voltage drop (V_d)", the diode current becomes appreciable (the level of current considered "appreciable" and the value of cut-in voltage depends on the application), and the diode presents a very low resistance. The current–voltage curve is exponential. In a normal silicon diode at rated currents, the arbitrary "cut-in" voltage is defined as 0.6 to 0.7 volts.

The junction is biased with a voltage V_a as shown in Figure 4.2.1. We will call the junction forward-biased if a positive voltage is applied to the *p*-doped region and reversed-biased if a negative voltage is applied to the *p*-doped region. The contact to the *p*-type region is also called the anode, while the contact to the *n*-type region is called the cathode, in reference to the *anions* or positive carriers and *cations* or negative carriers in each of these regions.

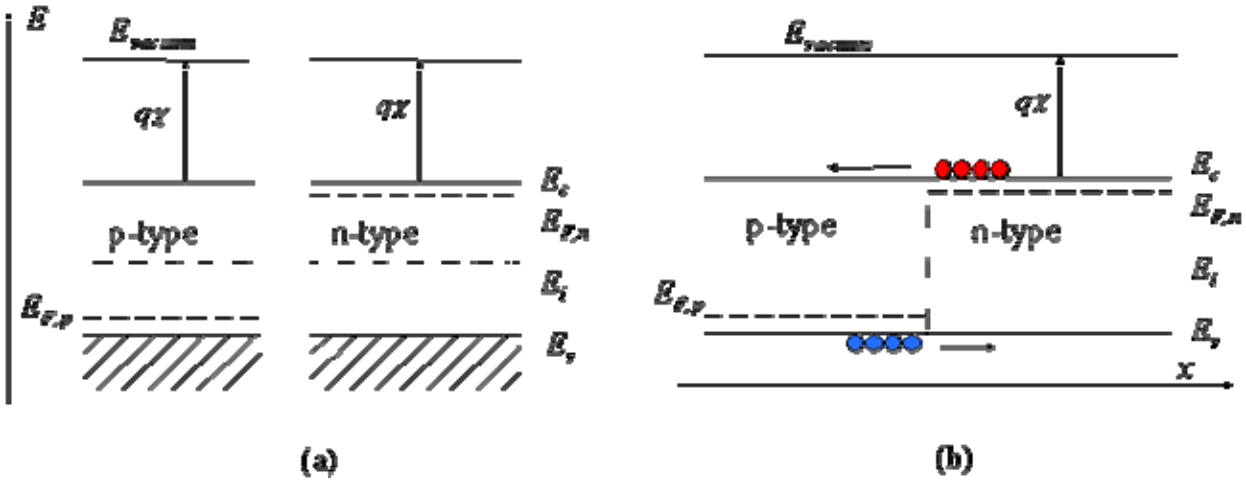


Figure 4.2.2 : Energy band diagram of a p-n junction (a) before and (b) after merging the n-type and p-type regions

Note that this does not automatically align the Fermi energies, $E_{F,n}$ and $E_{F,p}$. Also, note that this flatband diagram is not an equilibrium diagram since both electrons and holes can lower their energy by crossing the junction. A motion of electrons and holes is therefore expected before thermal equilibrium is obtained. The diagram shown in Figure 4.2.2 (b) is called a flatband diagram. This name refers to the horizontal band edges. It also implies that there is no field and no net charge in the semiconductor.

4.2.2. Thermal equilibrium

To reach thermal equilibrium, electrons/holes close to the metallurgical junction diffuse across the junction into the *p*-type/*n*-type region where hardly any electrons/holes are present. This process leaves the ionized donors (acceptors) behind, creating a region around the junction, which is depleted of mobile carriers. We call this region the depletion region, extending from $x = -x_p$ to $x = x_n$. The charge due to the ionized donors and acceptors causes an electric field, which in turn causes a drift of carriers in the opposite direction. The diffusion of carriers continues until the drift current balances the diffusion current, thereby reaching thermal equilibrium as indicated by a

constant Fermi energy. This situation is shown in Figure 4.2.3:

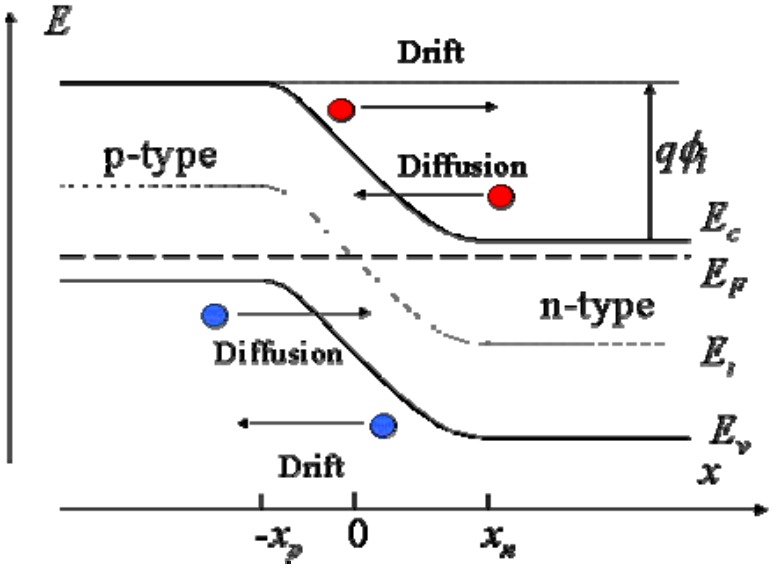


Figure 4.2.3 : Energy band diagram of a p-n junction in thermal equilibrium

While in thermal equilibrium no external voltage is applied between the *n*-type and *p*-type material, there is an internal potential, ϕ_i, which is caused by the workfunction difference between the *n*-type and *p*-type semiconductors. This potential equals the *built-in* potential, which will be further discussed in the next section.

4.2.3. The built-in potential

The built-in potential in a semiconductor equals the potential across the depletion region in thermal equilibrium. Since thermal equilibrium implies that the Fermi energy is constant throughout the p-n diode, the built-in potential equals the difference between the Fermi energies, E_{Fn} and E_{Fp}, divided by the electronic charge. It also equals the sum of the bulk potentials of each region, ϕ_n and ϕ_p, since the bulk potential quantifies the distance between the Fermi energy and the intrinsic energy. This yields the following expression for the built-in potential.

$$\phi_i = V_t \ln \frac{N_d N_a}{n_i^2} \qquad (4.2.1)$$

Example 4.1	An abrupt silicon p-n junction consists of a *p*-type region containing 2 x 10^{16} cm^{-3} acceptors and an *n*-type region containing also 10^{16} cm^{-3} acceptors in addition to 10^{17} cm^{-3} donors. a. Calculate the thermal equilibrium density of electrons and holes in the *p*-type region as well as both densities in the *n*-type region. b. Calculate the built-in potential of the p-n junction c. Calculate the built-in potential of the p-n junction at 400 K.
Solution	a. The thermal equilibrium densities are: In the *p*-type region: $p = N_a = 2 \times 10^{16}$ cm^{-3} $n = n_i^2/p = 10^{20}/2 \times 10^{16} = 5 \times 10^3$ cm^{-3} In the *n*-type region $n = N_d - N_a = 9 \times 10^{16}$ cm^{-3}

$$p = n_i 2/n = 10^{20}/(1 \times 10^{16}) = 1.11 \times 10^3 \text{ cm}^{-3}$$

b. The built-in potential is obtained from:

$$\phi_i = V_t \ln \frac{p_{n0} n_{p0}}{n_i^2} = 0.0259 \ln \frac{2\times10^{16} \times 9\times10^{16}}{10^{20}} = 0.79\,\text{V}$$

c. Similarly, the built-in potential at 400 K equals:

$$\phi_i = V_t \ln \frac{p_{n0} n_{p0}}{n_i^2} = 0.0345 \ln \frac{2\times10^{16} \times 9\times10^{16}}{(4.52\times10^{12})^2} = 0.63\,\text{V}$$

where the instrinsic carrier density at 400 K was obtained from example 2.4.b

4.2.4. Forward and reverse bias

We now consider a p-n diode with an applied bias voltage, V_a. A forward bias corresponds to applying a positive voltage to the anode (the *p*-type region) relative to the cathode (the *n*-type region). A reverse bias corresponds to a negative voltage applied to the cathode. Both bias modes are illustrated with Figure 4.2.4. The applied voltage is proportional to the difference between the Fermi energy in the *n*-type and *p*-type quasi-neutral regions.

As a negative voltage is applied, the potential across the semiconductor increases and so does the depletion layer width. As a positive voltage is applied, the potential across the semiconductor decreases and with it the depletion layer width. The total potential across the semiconductor equals the built-in potential minus the applied voltage, or:

$$\phi = \phi_i - V_a \qquad (4.2.1)$$

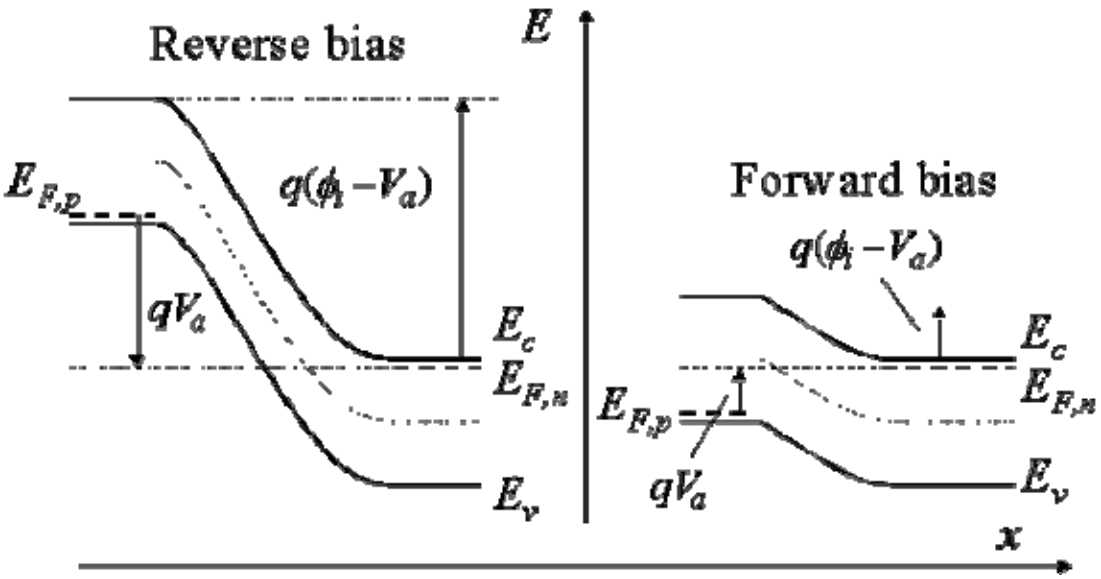


Figure 4.2.4: Energy band diagram of a p-n junction under reverse and forward bias

Introduction to Light Emitting Diodes

The past few decades have brought a continuing and rapidly evolving sequence of technological revolutions, particularly in the digital arena, which has dramatically changed many aspects of our daily lives. The developing race among manufacturers of light emitting diodes (LEDs) promises to produce, literally, the most visible and far-reaching transition to date. Recent advances in the design and manufacture of these miniature semiconductor devices may result in the obsolescence of the common light bulb, perhaps the most ubiquitous device utilized by modern society.

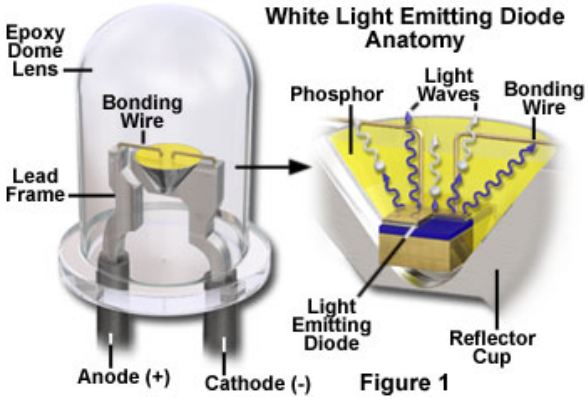


Figure 1

The incandescent lamp is the best known of Thomas Edison's major inventions, and the only one to have persisted in use (and in nearly its original form) to the present day, now more than a century after its introduction. The phonograph, tickertape, and mimeograph machines have been replaced by digital technologies in the last few decades, and recently, full-spectrum light emitting diode devices are becoming widespread, and could force incandescent and fluorescent lamps into extinction. While some applications of LED technology may be as straightforward as replacing one light bulb with another, far more visionary changes may involve dramatic new mechanisms for utilizing light. As a result of the predicted evolution, walls, ceilings, or even entire buildings could become the targets for specialized lighting scenarios, and interior design changes might be accomplished through illumination effects rather than by repainting or refurnishing. At the very least, a widespread change from incandescent to LED illumination would result in enormous energy savings. Although light emitting diodes are in operation all around us in videocassette recorders, clock radios, and microwave ovens, for example, their use has been limited mainly to display functions on electronic appliances. The tiny red and green indicator lights on computers and other devices are so familiar, the fact that the first LEDs were limited to a dim red output is probably not widely recognized. In fact, even the availability of green-emitting diodes represented a significant developmental step in the technology. In the past 15 years or so, LEDs have become much more powerful, and available in a wide spectrum of colors. A breakthrough that enabled fabrication of the first blue LED in the early 1990s, emitting light at the opposite end of the visible light spectrum from red, opened up the possibility to create virtually any color of light. More important, the discovery made it technically feasible to produce white light from the tiny semiconductor devices. An inexpensive, mass-market version of white LED is the most sought-after goal of researchers and manufacturers, and is the device most likely to end a hundred-year reliance on inefficient incandescent lamps. The widespread utilization of diode devices for general lighting is still some years away, but LEDs are beginning to replace incandescent lamps in many applications. There are a number of reasons for replacing conventional incandescent light sources with modern semiconductor alternatives. Light emitting diodes are far more efficient than incandescent bulbs at converting electricity into visible light, they are rugged and compact, and can often last 100,000 hours in use, or about 100 times longer than incandescent bulbs. LEDs are fundamentally monochromatic emitters, and applications requiring high-brightness, single-color lamps are experiencing the greatest number of applications within the current generation of improved devices. The use of LEDs is increasing for automotive taillights, turn signals, and side marker lights. As one of the first automotive applications, the high-mount brake light on cars and trucks is a particularly

appealing location for incorporating LEDs. Long LED lifespans allow manufacturers more freedom to integrate the brake light into the vehicle design without the necessity of providing for frequent (and easy) replacement, as is required when incandescent bulbs are used. Approximately 10 percent of the red traffic lights in the United States have now been replaced with LED-based lamps. The higher initial cost of the LEDs can be recovered in as little as one year, due to their higher efficiency in producing red light, which is accomplished without the need for filtering. The LEDs in a red traffic light consume about 10 to 25 watts, compared with 50 to 150 for a red-filtered incandescent light of similar brightness. The longevity of the LEDs is an obvious advantage in reducing expensive maintenance of the signals. Single-color LEDs are also being utilized as runway lights at airports and as warning lights on radio and television transmission towers. As improvements have been made in manufacturing efficiency and toward the ability to produce light emitting diodes with virtually any output color, the primary focus of researchers and industry has become the white light diode. Two primary mechanisms are being employed to produce white light from devices that are fundamentally monochromatic, and both techniques will most likely continue to be utilized for different applications. One method involves mixing different colors of light from multiple LEDs, or from different materials in a single LED, in proportions that result in light that appears white. The second technique relies on using LED emission (commonly non-visible ultraviolet) to provide energy for excitation of another substance, such as a phosphor, which in turn produces white light. Each method has both advantages and disadvantages that are likely to be in constant flux as further developments occur in LED technology.
Fundamentals of Semiconductor Diodes Details of the fundamental processes underlying the function of light emitting diodes, and the materials utilized in their construction, are presented in the ensuing discussion. The basic mechanism by which LEDs produce light can be summarized, however, by a simple conceptual description. The familiar light bulb relies upon temperature to emit visible light (and significantly more invisible radiation in the form of heat) through a process known as incandescence. In contrast, the light emitting diode employs a form of electroluminescence, which results from the electronic excitation of a semiconductor material. The basic LED consists of a junction between two different semiconductor materials (illustrated in Figure 2), in which an applied voltage produces a current flow, accompanied by the emission of light when charge carriers injected across the junction are recombined.

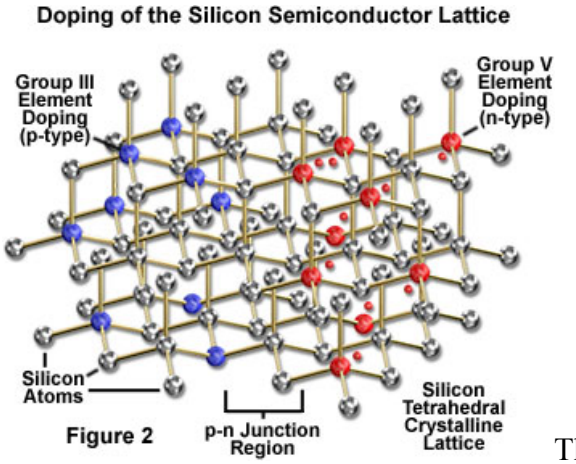
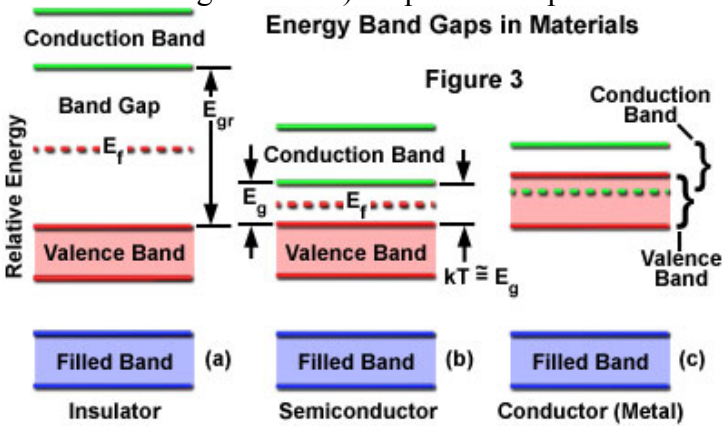


Figure 2

The fundamental element of the LED is a semiconductor chip (similar to an integrated circuit), which is mounted in a reflector cup supported by a lead frame connected to two electrical wires, and then embedded in a solid epoxy lens (see Figure 1). One of the two semiconductor regions that comprise the junction in the chip is dominated by negative charges (n-type region; Figure 2)), and the other is dominated by positive charges (p-type region). When a sufficient voltage is applied to the electrical leads, current flows and electrons move across the junction from the n region into the p region where the negatively charged electrons combine with positive charges. Each combination of charges is associated with an energy level reduction that may release a

quantum of electromagnetic energy in the form of a light photon. The frequency, and perceived color, of emitted photons is characteristic of the semiconductor material, and consequently, different colors are achieved by making changes in the semiconductor composition of the chip. The functional details of the light emitting diode are based on properties common to semiconductor materials, such as silicon, which have variable conduction characteristics. In order for a solid to conduct electricity, its resistance must be low enough for electrons to move more or less freely throughout the bulk of the material. Semiconductors exhibit electrical resistance values intermediate between those of conductors and insulators, and their behavior can be modeled in terms of the band theory for solids. In a crystalline solid, electrons of the constituent atoms occupy a large number of energy levels that may differ very little either in energy or in quantum number. The wide spectrum of energy levels tend to group together into nearly continuous energy bands, the width and spacing of which differ considerably for different materials and conditions. At progressively higher energy levels, proceeding outward from the nucleus, two distinct energy bands can be defined, which are termed the valence band and the conduction band (Figure 3). The valence band consists of electrons at a higher energy level than the inner electrons, and these have some freedom to interact in pairs to form a type of localized bond among atoms of the solid. At still-higher energy levels, electrons of the conduction band behave similarly to electrons in individual atoms or molecules that have been excited above ground state, with a high degree of freedom to move about within the solid. The difference in energy between the valence and conduction bands is defined as the band gap for a particular material. In conductors, the valence and conduction bands partially overlap in energy (see Figure 3), so that a portion of the valence electrons always resides in the conduction band. The band gap is essentially zero for these materials, and with part of the valence electrons moving freely into the conduction band, vacancies or holes occur in the valence band. Electrons move, with very little energy input, into holes in bands of adjacent atoms, and the holes migrate freely in the opposite direction. In contrast to these materials, insulators have fully occupied valence bands and larger band gaps, and the only mechanism by which electrons can move from atom to atom is for a valence electron to be displaced into the conduction band, requiring a large energy expenditure.

Semiconductors have band gaps that are small but finite, and at normal temperatures, thermal agitation is sufficient to move some electrons into the conduction band where they can contribute to electrical conduction. Resistance can be reduced by increasing the temperature, but many semiconductor devices are designed in such a manner that the application of a voltage produces the required changes in electron distribution between the valence and conduction bands to enable current flow. Although the band arrangement is similar for all semiconductors, there are large differences in the band gap (and in the distribution of electrons among the bands) at specific temperature conditions.



The element silicon is the simplest intrinsic semiconductor, and is often used as a model for describing the behavior of these materials. In its pure form, silicon does not have sufficient charge carriers, or appropriate band gap structure, to be useful in light emitting diode construction, but is widely used to fabricate other semiconductor devices. The conduction

characteristics of silicon (and other semiconductors) can be improved through the introduction of impurities in small quantities to the crystal, which serve to provide either additional electrons or vacancies (holes) in the structure. Through this process, referred to as doping, producers of integrated circuits have developed considerable ability to tailor the properties of semiconductors to suit specific applications.The process of doping to modify the electronic properties of semiconductors is most easily understood by considering the relatively simple silicon crystal structure. Silicon is a Group IV member of the periodic table, having four electrons that may participate in bonding with neighboring atoms in a solid. In pure form, each silicon atom shares electrons with four neighbors, with no deficit or excess of electrons beyond those required in the crystal structure. If a small amount of a Group III element (those having three electrons in their outermost energy level) is added to the silicon structure, an insufficient number of electrons exist to satisfy the bonding requirements. The electron deficiency creates a vacancy, or hole, in the structure, and the resulting positive electrical character classifies the material as p-type. Boron is one of the elements that is commonly utilized to dope pure silicon to achieve p-type characteristics.

Doping in order to produce the opposite type of material, having a negative overall charge character (n-type), is accomplished through the addition of Group V elements, such as phosphorus, which have an "extra" electron in their outermost energy level. The resulting semiconductor structure has an excess of available electrons over the number required for covalent silicon bonding, which bestows the ability to act as an electron donor (characteristic of n-type material).Although silicon and germanium are commonly employed in semiconductor fabrication, neither material is suitable for light emitting diode construction because junctions employing these elements produce a significant amount of heat, but only a small quantity of infrared or visible light emission. Photon-emitting diode p-n junctions are typically based on a mixture of Group III and Group V elements, such as gallium arsenide, gallium arsenide phosphide, and gallium phosphide. Careful control of the relative proportions of these compounds, and others incorporating aluminum and indium, as well as the addition of dopants such as tellurium and magnesium, enables manufacturers and researchers to produce diodes that emit red, orange, yellow, or green light. Recently the use of silicon carbide and gallium nitride has permitted blue-emitting diodes to be introduced, and combining several colors in various combinations provides a mechanism to produce white light. The nature of materials comprising p-type and n-type sides of the device junction, and the resulting energy band structure, determines the energy levels that are available during charge recombination in the junction region, and therefore, the magnitude of the energy quanta released as photons. As a consequence, the color of light emitted by a particular diode depends upon the structure and composition of the p-n junction.The fundamental key to manipulating properties of solid-state electronic devices is the nature of the p-n junction. When dissimilar doped materials are placed in contact with each other, the flow of current in the region of the junction is different than it is in either of the two materials alone. Current will readily flow in one direction across the junction, but not in the other, constituting the basic diode configuration. This behavior can be understood in terms of the movement of electrons and holes in the two material types and across the junction. The extra free electrons in the n-type material tend to move from the negatively charged area to a positively charged area, or toward the p-type material. In the p-type region, which has vacant electron sites (holes), lattice electrons can jump from hole to hole, and will tend to move away from the negatively charged area. The result of this migration is that the holes appear to move in the opposite direction, or away from the positively charged region and toward the negatively charged area (Figure 4). Electrons from the n-type region and holes from the p-type region recombine in the vicinity of the junction to form a depletion zone (or layer), in which no charge carriers remain. In the depletion zone, a static charge is established that inhibits any additional electron transfer, and no appreciable charge can flow across the junction unless assisted by an external bias voltage.

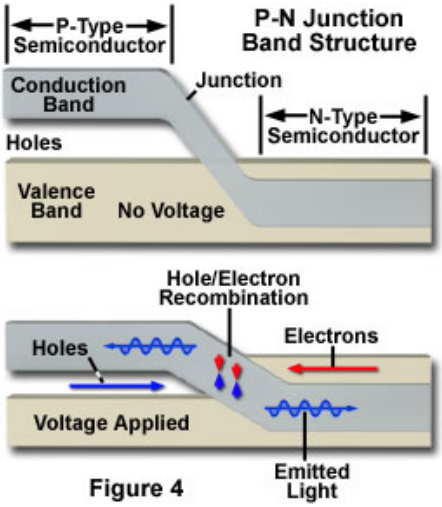


Figure 4

In a diode configuration, electrodes on opposite ends of the device enable a voltage to be applied in a manner that can overcome the effect of the depletion region. Connecting the n-type region of the diode to the negative side of an electrical circuit, and the p-type region to the positive side will cause electrons to move from the n-type material toward the p-type, and holes to move in the opposite direction. With application of a sufficiently high voltage, the electrons in the depletion region are elevated in energy to dissociate with the holes, and to begin moving freely again. Operated with this circuit polarity, referred to as forward biasing of the p-n junction, the depletion zone disappears and charge can move across the diode. Holes are driven to the junction from the p-type material and electrons are driven to the junction from the n-type material. The combination of holes and electrons at the junction allows a continuous current to be maintained across the diode.

If the circuit polarity is reversed with respect to the p-type and n-type regions, electrons and holes will be pulled in opposite directions, with an accompanying widening of the depletion region at the junction. No continuous current flow occurs in a reverse-biased p-n junction, although initially a transient current will flow as the electrons and holes are pulled away from the junction. Current flow will cease as soon as the growing depletion zone creates a potential that is equal to the applied voltage.

Light Emitting Diode Construction

Manipulation of the interaction between electrons and holes at the p-n junction is fundamental in the design of all semiconductor devices, and for light emitting diodes, the primary design goal is the efficient generation of light. Injection of carriers across the p-n junction is accompanied by a drop in electron energy levels from the conduction band to lower orbitals. This process takes place in any diode, but only produces visible light photons in those having specific material compositions. In a standard silicon diode, the energy level difference is relatively small, and only low frequency emission occurs, predominately in the infrared region of the spectrum. Infrared diodes are useful in many devices, including remote controls, but the design of visible-light emitting diodes requires fabrication with materials exhibiting a wider gap between the conduction band and orbitals of the valence band. All semiconductor diodes release some form of light, but most of the energy is absorbed into the diode material itself unless the device is specifically designed to release the photons externally. In addition, to be useful as a light source, diodes must concentrate light emission in a specific direction. Both the composition and construction of the semiconductor chip, and the design of the LED housing, contribute to the nature and efficiency of energy emission from the device. The basic structure of a light emitting diode consists of the semiconductor material (commonly referred to as a die), a lead frame on which the die is placed, and the encapsulation epoxy surrounding the assembly (see Figure 1). The LED semiconductor chip is supported in a reflector cup coined into the end of one electrode (the cathode), and, in the typical configuration, the top face of the chip is connected with a gold bonding wire to a second electrode (anode). Several junction structure designs require two bonding wires, one

to each electrode. In addition to the obvious variation in the radiation wavelength of different LEDs, there are variations in shape, size, and radiation pattern. The typical LED semiconductor chip measures approximately 0.25 millimeter-square, and the epoxy body ranges from 2 to about 10 millimeters in diameter. Most commonly, the body of the LED is round, but they may be rectangular, square, or triangular.

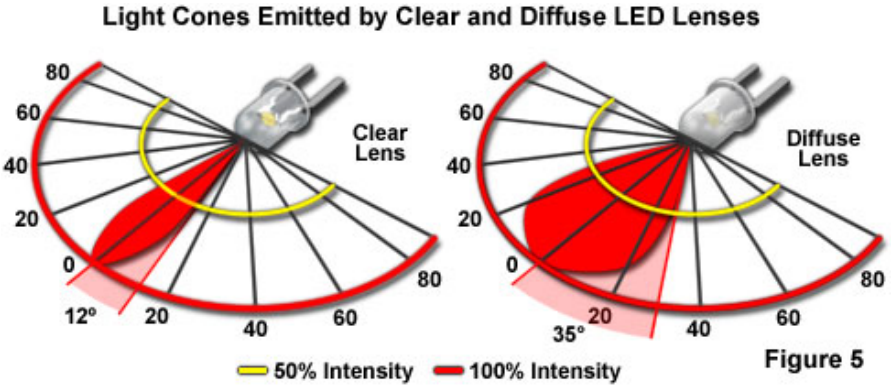


Figure 5

Although the color of light emitted from a semiconductor die is determined by the combination of chip materials, and the manner in which they are assembled, certain optical characteristics of the LED can be controlled by other variables in the chip packaging. The beam angle can be narrow or wide (see Figure 5), and is determined by the shape of the reflector cup, the size of the LED chip, the distance from chip to the top of the epoxy housing or lens, and the geometry of the epoxy lens. The tint of the epoxy lens does not determine the emission color of the LED, but is often used as a convenient indicator of the lamp's color when it is inactive. LEDs intended for applications that require high intensity, and no color in the off-state, have clear lenses with no tint or diffusion. This type produces the greatest light output, and may be designed to have the narrowest beam, or viewing angle. Non-diffused lenses typically exhibit viewing angles of plus or minus 10 to 12 degrees (Figure 5). Their intensity allows them to be utilized for backlighting applications, such as the illumination of display panels on electronic devices. For creation of diffused LED lenses, minute glass particles are embedded in the encapsulating epoxy. The diffusion created by inclusion of the glass spreads light emitted by the diode, producing a viewing angle of approximately 35 degrees on either side of the central axis. This lens style is commonly employed in applications in which the LED is viewed directly, such as for indicator lamps on equipment panels. The choice of material systems and fabrication techniques in LED construction is guided by two primary goals—maximization of light generation in the chip material, and the efficient extraction of the generated light. In the forward-biased p-n junction, holes are injected across the junction from the p region into the n region, and electrons are injected from the n region into the p region. The equilibrium charge carrier distribution in the material is altered by this injection process, which is referred to as minority-carrier injection. Recombination of minority carriers with majority carriers takes place to reestablish thermal equilibrium, and continued current flow maintains the minority-carrier injection. When the recombination rate is equal to the injection rate, a steady-state carrier distribution is established. Minority-carrier recombination can take place in a radiative fashion, with the emission of a photon, but for this to occur the proper conditions must be established for energy and momentum conservation. Meeting these conditions is not an instantaneous process, and a time delay results before radiative recombination of the injected minority carrier can take place. This delay, the minority carrier lifetime, is one of the primary variables that must be considered in LED material design. Although the radiative recombination process is desirable in LED design, it is not the only recombination mechanism that is possible in semiconductors. Semiconductor materials cannot be produced without some impurities, structural dislocations, and other crystalline defects, and these can all trap injected minority carriers. Recombinations of this type may or may not produce light photons. Recombinations that do not produce radiation are slowed by the diffusion of the carriers to

suitable sites, and are characterized by a nonradiative process lifetime, which can be compared to the radiative process lifetime. An obvious goal in LED design, given the factors just described, is to maximize the radiative recombination of charge carriers relative to the nonradiative. The relative efficiency of these two processes determines the fraction of injected charge carriers that combine radiatively compared to the total number injected, which can be stated as the internal quantum efficiency of the material system. The choice of materials for LED fabrication relies upon an understanding of semiconductor band structure and the means by which the energy levels can be chosen or manipulated to produce favorable quantum efficiency values. Interestingly, certain groups of III-V compounds have internal quantum efficiencies of nearly 100 percent, while other compounds utilized in semiconductors may have internal quantum efficiencies as low as 1 percent. The radiative lifetime for a particular semiconductor largely determines whether radiative recombinations occur before nonradiative. Most semiconductors have similar simple valence band structure with an energy peak situated around a particular crystallographic direction, but with much more variation in the structure of the conduction band. Energy valleys exist in the conduction band, and electrons occupying the lowest-energy valleys are positioned to more easily participate in recombination with minority carriers in the valence band. Semiconductors can be classified as direct or indirect depending upon the relative positioning of the conduction band energy valleys and the energy apex of the valence band in energy/momentum space. Direct semiconductors have holes and electrons positioned directly adjacent at the same momentum coordinates, so that electrons and holes can recombine relatively easily while maintaining momentum conservation. In an indirect semiconductor, the match between conduction band energy valleys and holes that would allow momentum conservation is not favorable, most of the transitions are forbidden, and the resulting radiative lifetime is long.

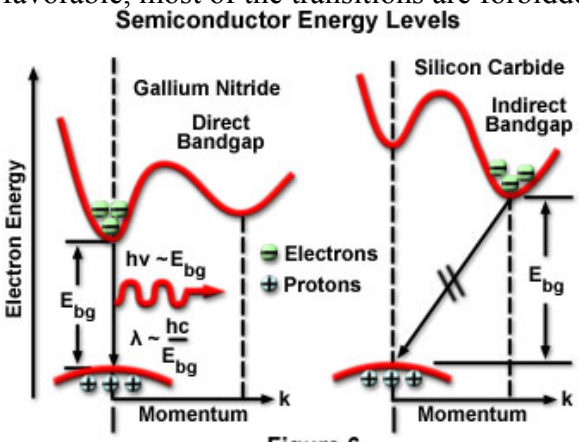


Figure 6

Silicon and germanium are examples of indirect semiconductors, in which radiative recombination of injected carriers is extremely unlikely. The radiative lifetime in such materials occurs in the range of seconds, and nearly all injected carriers combine nonradiatively through defects in the crystal. Direct semiconductors, such as gallium nitride or gallium arsenide, have short radiative lifetimes (approximately 1 to 100 nanoseconds), and materials can be produced with sufficiently low defect density that radiative processes are as likely as nonradiative. For a recombination event to occur in indirect gap materials, an electron must change its momentum before combining with a hole, resulting in a significantly lower recombination probability for the occurrence of a band-to-band transition. The quantum efficiencies exhibited by LEDs constructed of the two types of semiconductor material clearly reflect this fact. Gallium nitride LEDs have quantum efficiencies as high as 12 percent, compared to the 0.02 percent typical of silicon carbide LEDs. Figure 6 presents an energy band diagram for direct band gap GaN and indirect band gap SiC that illustrates the nature of the band-to-band energy transition for the two types of material. The wavelength (and color) of light emitted in a radiative recombination of carriers injected across a p-n junction is determined by the difference in energy between the recombining electron-hole pair of the valence and conduction bands. The approximate energies of the carriers correspond to the upper energy level of the valence band and the

lowest energy of the conduction band, due to the tendency of the electrons and holes to equilibrate at these levels. Consequently, the wavelength (λ) of an emitted photon is approximated by the following expression:

$\lambda = hc/E_{bg}$ where h represents Planck's constant, c is the velocity of light, and E(bg) is the band gap energy. In order to change the wavelength of emitted radiation, the band gap of the semiconducting material utilized to fabricate the LED must be changed. Gallium arsenide is a common diode material, and may be used as an example illustrating the manner in which a semiconductor's band structure can be altered to vary the emission wavelength of the device. Gallium arsenide has a band gap of approximately 1.4 electron-volts, and emits in the infrared at a wavelength of 900 nanometers. In order to increase the frequency of emission into the visible red region (about 650 nanometers), the band gap must be increased to approximately 1.9 electron-volts. This can be achieved by mixing gallium arsenide with a compatible material having a larger band gap. Gallium phosphide, having a band gap of 2.3 electron-volts, is the most likely candidate for this mixture. LEDs produced with the compound GaAsP (gallium arsenide phosphide) can be customized to produce band gaps of any value between 1.4 and 2.3 electron-volts, through adjustment of the content of arsenic to phosphorus. As previously discussed, maximization of light generation in the diode semiconductor material is a primary design goal in LED fabrication. Another requirement is the efficient extraction of the light from the chip. Because of total internal reflection, only a fraction of the light that is generated isotropically within the semiconductor chip can escape to the outside. According to Snell's law, light can travel from a medium of higher refractive index into a medium of lower refractive index only if it intersects the interface between the two media at an angle less than the critical angle for the two media. In a typical light-emitting semiconductor having cubic shape, only about 1 to 2 percent of the generated light escapes through the top surface of the LED (depending upon the specific chip and p-n junction geometry), the remainder being absorbed within the semiconductor materials.

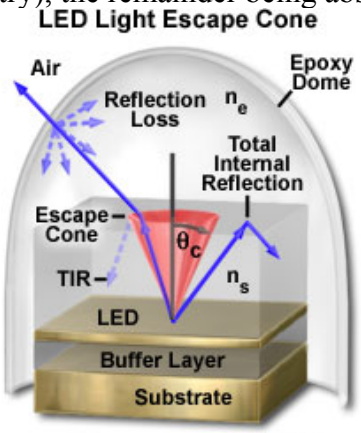


Figure 7

Figure 7 illustrates the escape of light from a layered semiconductor chip of refractive index n(s) into epoxy of lower index (n(e)). The angle subtended by the escape cone is defined by the critical angle, θ(c), for the two materials. Light rays emerging from the LED at angles less than θ(c) escape into the epoxy with minimal reflection loss (dashed ray lines), while those rays propagating at angles greater than θ(c) undergo total internal reflection at the boundary, and do not escape the chip directly. Because of the curvature of the epoxy dome, most light rays leaving the semiconductor material meet the epoxy/air interface at nearly right angles, and emerge from the housing with little reflection loss.

The proportion of light emitted from an LED chip into the surroundings is dependent upon the number of surfaces through which light can be emitted, and how effectively this occurs at each surface. Nearly all LED structures rely on some form of layered arrangement in which epitaxial growth processes are utilized to deposit several lattice-matched materials on top of one another to tailor the properties of the chip. A wide variety of structures is employed, with each material system requiring different layer architecture in order to optimize performance properties.

Most of the LED structural arrangements rely on a secondary growth step to deposit a single-crystal layer on top of a single-crystal bulk-grown substrate material. Such a multilayering approach enables designers to satisfy seemingly contradictory or inconsistent requirements. A common feature of all of the structural types is that the p-n junction, where the light emission occurs, is almost never located in the bulk-grown substrate crystal. One reason for this is that bulk-grown material generally has a high defect density, which lowers the light generation efficiency. In addition, the most common bulk-grown materials, including gallium arsenide, gallium phosphide, and indium phosphide, do not have the appropriate band gap for the desired emission wavelengths. Another requirement in many LED applications is for a low series resistance that can be met by appropriate substrate choice, even in cases in which the low doping required in the p-n junction region would not provide adequate conduction.
The techniques of epitaxial crystal growth involve deposition of one material on another, which is closely matched in atomic lattice constants and thermal expansion coefficient to reduce defects in the layered material. A number of techniques are in use to produce epitaxial layers. These include Liquid Phase Epitaxy (LPE), Vapor Phase Epitaxy (VPE), Metal-Organic Epitaxial Chemical Vapor Deposition (MOCVD), and Molecular Beam Epitaxy (MBE). Each of the growth techniques has advantages in particular materials systems or production environments, and these factors are extensively discussed in the literature.
The details of the various epitaxial structures employed in LED fabrication are not presented here, but are discussed in a number of publications. Generally, however, the most common categories of such structures are grown and diffused homojunctions, and single confinement or double confinement heterojunctions. The strategies behind the application of the various layer arrangements are numerous. These include structuring of p and n regions and reflective layers to increase the internal quantum efficiency of the system, graded-composition buffer layers to overcome lattice mismatch between layers, locally varying energy band gap to accomplish carrier confinement, and lateral constraint of carrier injection to control light emission area or to collimate the emission. Even though it does not typically contain the p-n junction region, the LED substrate material becomes an integral part of the function, and is chosen to be appropriate for deposition of the desired epitaxial layers, as well as for its light transmission and other properties. As previously stated, the fraction of generated light that is actually emitted from an LED chip is a function of the number of surfaces that effectively transmit light. Most LED chips are categorized as absorbing substrate (AS) devices, where the substrate material has a narrow band gap and absorbs all emission having energy greater than the band gap. Therefore, light traveling toward the sides or downward is absorbed, and such chips can only emit light through their top surfaces. The transparent substrate (TS) chip is designed to increase light extraction by incorporating a substrate that is transparent to the wavelength of emitted light. In some systems, transparency in the upper epitaxial layers will allow light transmitted toward the side surfaces, within certain angles, to be extracted as well. Hybrid designs, having substrate properties intermediate between AS and TS devices, are also utilized, and significant increases in extraction efficiency can be achieved by employment of a graded change in refractive index from the LED chip to air. There remain numerous other absorption mechanisms in the LED structure that reduce emission and are difficult to overcome, such as the front and back contacts on the chip, and crystal defects. However, chips made on transparent, as opposed to absorbing, substrates can exhibit a nearly-fivefold improvement in extraction efficiency.

Dyes are another suitable type of wavelength converter for white diode applications, and can be incorporated into the epoxy encapsulant or in transparent polymers. The commercially available dyes are generally organic compounds, which are chosen for a specific LED design by consideration of their absorption and emission spectra. The light generated by the diode must match the absorption profile of the converting dye, which in turn emits light at the desired longer wavelength. The quantum efficiencies of dyes can be near 100 percent, as in phosphor conversion, but they have the disadvantage of poorer long-term operational stability than phosphors. This is a serious drawback, as the molecular instability of the dyes causes them to

lose optical activity after a finite number of absorptive transitions, and the resulting change in light emitting diode color will limit its lifetime. White light LEDs based on semiconductor wavelength converters have been demonstrated that are similar in principle to the phosphor conversion types, but which employ a second semiconductor material that emits a different wavelength in response to the emission from the primary source wafer. These devices have been referred to as photon recycling semiconductors (or PRS-LEDs), and incorporate a blue-emitting LED die bonded to another die that responds to the blue light by emitting light of a complementary wavelength. The two wavelengths then combine to produce white. One possible structure for this type of device utilizes a GaInN diode as a current-injected active region coupled to an AlGaInP optically-excited active region. The blue light emitted by the primary source is partially absorbed by the secondary active region, and "recycled" as reemitted photons of lower energy. The structure of a photon recycling semiconductor is illustrated schematically in Figure 11. In order for the combined emission to produce white light, the intensity ratio of the two sources must have a specific value that can be calculated for the particular dichromatic components. The choice of materials and the thickness of the various layers in the structure can be modified to vary the color of the device output.

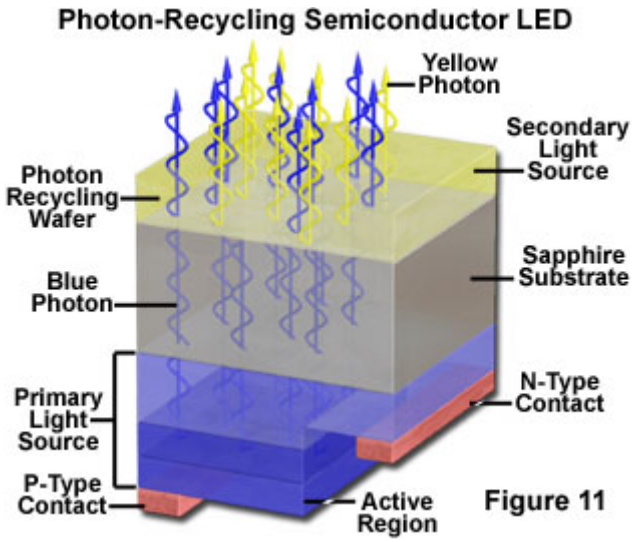
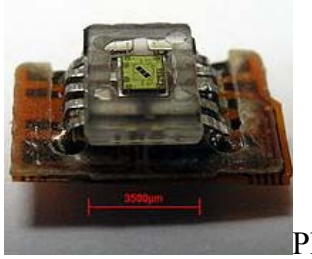


Figure 11

Because white light can be created by several different mechanisms, utilizing white LEDs in a particular application requires consideration of the suitability of the method employed to generate the light. Although the perceived color of light emitted by various techniques may be similar, its effect on color rendering, or the result of filtration of the light, for example, may be entirely different. White light created through broadband emission, through mixing of two complementary colors in a dichromatic source, or by mixing of three colors in a trichromatic source, can be located at different coordinates on the chromaticity diagram and have different color temperatures with respect to illuminants designated as standards by the CIE.

Photodiode


Photodetector from a CD-ROM Drive. 3 photodiodes are visible.

A **photodiode** is a type of photodetector capable of converting light into either current or voltage, depending upon the mode of operation.

Photodiodes are similar to regular semiconductor diodes except that they may be either exposed (to detect vacuum UV or X-rays) or packaged with a window or optical fiber connection to allow light to reach the sensitive part of the device. Many diodes designed for use specifically as a photodiode will also use a PIN junction rather than the typical PN junction.

Polarity

Photodiode schematic symbol

Some photodiodes will look like the picture to the right, that is, similar to a light emitting diode. They will have two leads, or wires, coming from the bottom. The shorter end of the two is the cathode, while the longer end is the anode. See below for a schematic drawing of the anode and cathode side. Under forward bias, conventional current will pass from the anode to the cathode, following the arrow in the symbol. Photocurrent flows in the opposite direction.

Principle of operation

A photodiode is a PN junction or PIN structure. When a photon of sufficient energy strikes the diode, it excites an electron, thereby creating a mobile electron and a positively charged electron hole. If the absorption occurs in the junction's depletion region, or one diffusion length away from it, these carriers are swept from the junction by the built-in field of the depletion region. Thus holes move toward the anode, and electrons toward the cathode, and a photocurrent is produced.

Photovoltaic mode When used in zero bias or photovoltaic mode, the flow of photocurrent out of the device is restricted and a voltage builds up. The diode becomes forward biased and "dark current" begins to flow across the junction in the direction opposite to the photocurrent. This mode is responsible for the photovoltaic effect, which is the basis for solar cells—in fact, a solar cell is just a large area photodiode.

Photoconductive mode

In this mode the diode is often reverse biased, dramatically reducing the response time at the expense of increased noise. This increases the width of the depletion layer, which decreases the junction's capacitance resulting in faster response times. The reverse bias induces only a small amount of current (known as saturation or back current) along its direction while the photocurrent remains virtually the same. The photocurrent is linearly proportional to the illuminance. Although this mode is faster, the photoconductive mode tends to exhibit more electronic noiseThe leakage current of a good PIN diode is so low (< 1nA) that the Johnson–Nyquist noise of the load resistance in a typical circuit often dominates.

Applications

P-N photodiodes are used in similar applications to other photodetectors, such as photoconductors, charge-coupled devices, and photomultiplier tubes.

Photodiodes are used in consumer electronics devices such as compact disc players, smoke detectors, and the receivers for remote controls in VCRs and televisions.

In other consumer items such as camera light meters, clock radios (the ones that dim the display when it's dark) and street lights, photoconductors are often used rather than photodiodes, although in principle either could be used.Photodiodes are often used for accurate measurement of light intensity in science and industry. They generally have a better, more linear response than photoconductors.They are also widely used in various medical applications, such as detectors for computed tomography (coupled with scintillators) or instruments to analyze samples (immunoassay). They are also used in pulse oximeters.

PIN diodes are much faster and more sensitive than ordinary p-n junction diodes, and hence are often used for optical communications and in lighting regulation.P-N photodiodes are not used to measure extremely low light intensities. Instead, if high sensitivity is needed,

avalanche photodiodes, intensified charge-coupled devices or photomultiplier tubes are used for applications such as astronomy, spectroscopy, night vision equipment and laser rangefinding.

Comparison with photomultipliers Advantages compared to photomultipliers:
Excellent linearity of output current as a function of incident light
Spectral response from 190 nm to 1100 nm (silicon), longer wavelengths with other semiconductor materials Low noise Ruggedized to mechanical stress Low cost
Compact and light weight Long lifetime High quantum efficiency, typically 80%
No high voltage required

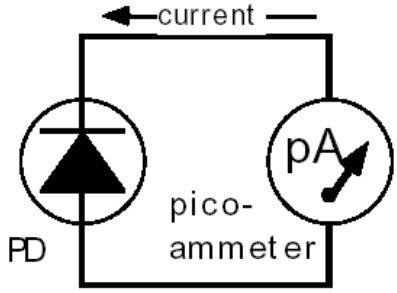
Disadvantages compared to photomultipliers:
Small area No internal gain (except avalanche photodiodes, but their gain is typically 10^2–10^3 compared to up to 10^8 for the photomultiplier) Much lower overall sensitivity
Photon counting only possible with specially designed, usually cooled photodiodes, with special electronic circuits Response time for many designs is slower

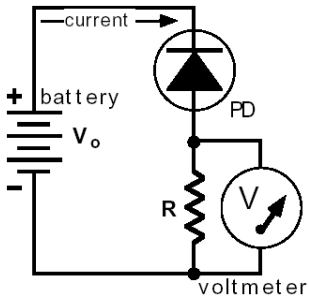
P-N vs. P-I-N photodiodes

Due to the intrinsic layer, a PIN photodiode must be reverse biased (Vr). The Vr increases the depletion region allowing a larger volume for electron-hole pair production, and reduces the capacitance thereby increasing the bandwidth. The Vr also introduces noise current, which reduces the S/N ratio. Therefore, a reverse bias is recommended for higher bandwidth applications and/or applications where a wide dynamic range is required. A PN photodiode is more suitable for lower light applications because it allows for unbiased operation.

A simple model of a photodiode

The PD can be connected directly to an ammeter, permitting the most sensitive measurement of the light power.





The diode can also be reverse biased to measure higher powers at faster speeds.
In either mode of operation, the photocurrent I is proportional to the power P of the light illuminating the photodiode. $I = K_{PD} P$, where K_{PD} is the responsivity of the photodiode. K_{PD} depends on the diode construction, the wavelength of the light, and on the temperature. The photodiode is a quantum device. Almost every incident photon generates an electron-hole charge pair. We therefore have $I = e N Q$ where Q is the quantum efficiency and e is the magnitude of the electron's charge. Q is less than or equal to 100 %. The optical power is equal to the number of photons per second, N, times the energy per photon, $E = hc/\lambda$.
$P = N E = Nhc/\lambda$ The PD responsivity therefore is given by $K_{PD} = Qe\lambda/(hc)$.

Introduction to Liquid Crystal Displays

The most common application of liquid crystal technology is in liquid crystal displays (LCDs). From the ubiquitous wrist watch and pocket calculator to an advanced VGA computer screen, this type of display has evolved into an important and versatile interface.A liquid crystal display consists of an array of tiny segments (called pixels) that can be manipulated to present information. This basic idea is common to all displays, ranging from simple calculators to a full color LCD television.Why are liquid crystal displays important? The first factor is size. As will be shown in the following sections, an LCD consists primarily of two glass plates with some liquid crystal material between them. There is no bulky picture tube. This makes LCDs practical for applications where size (as well as weight) are important. In general, LCDs use much less power than their cathode-ray tube (CRT) counterparts. Many LCDs are reflective, meaning that they use only ambient light to illuminate the display. Even displays that do require an external light source (i.e. computer displays) consume much less power than CRT devices.Liquid crystal displays do have drawbacks, and these are the subject of intense research. Problems with viewing angle, contrast ratio, and response time still need to be solved before the LCD replaces the cathode-ray tube. However with the rate of technological innovation, this day may not be too far into the future.

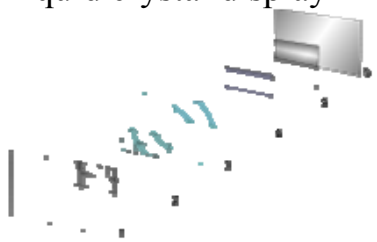
We will restrict this discussion to traditional nematic LCDs since the major technological advances have been developed for this group of devices. Other LC applications, such as those employing polymer stabilization of LC structure, are discussed in the appropriate section covering those materials.

Liquid Crystal

Ordinary fluids are isotropic in nature: they appear optically, magnetically, electrically, etc. to be the same from any perspective. Although the molecules which comprise the fluid are generally anisometric in shape, this anisometry *generally* plays little role in anisotropic macroscopic behavior (aside from viscosity). Nevertheless, there exists a large class of *highly* anisometric molecules which gives rise to unusual, fascinating, and potentially technologically relevant behavior. There are many interesting candidates for study, including polymers, micelles, microemulsions, and materials of biological significance, such as DNA and membranes. Although at times we have investigated all of these materials, our primary effort centers on liquid crystals. Liquid crystals are composed of moderate size organic molecules which tend to be elongated and shaped like a cigar, although we have studied, and the literature is full of variety of other, highly exotic shapes as well. Because of their elongated shape, under appropriate conditions the molecules can exhibit orientational order, such that all the axes line up in a particular direction. In consequence, the bulk order has profound influences on the way light and electricity behave in the material. For example, if the direction of the orientation varies in space, the orientation of the light (i.e., the polarization) can follow this variation. A well-known application of this phenomenon is the ubiquitous liquid crystal display, now comprising a $15b annual industry world-wide. Under other conditions the molecules may form a stack of layers along one direction, but remain liquid like (in terms of the absence of translational order) within the layers. As the system changes from one of these phases to another, a variety of physical parameters such as susceptibility and heat capacity, will exhibit "pretransitional behavior." Based solely on symmetry, this behavior may be related to other physical systems, such as superconductivity, magnetism, or superfluidity; this is the so-called "universality" of these phase transitions.

Using a battery of optical techniques, in addition to dielectric and certain surface probes, our research centers on the role of symmetry on liquid crystalline phases and phase transitions, how these systems behave in the presence of intense magnetic and electric fields, and the effects of confining these materials in spaces not much larger than the molecules themselves. By observing this behavior, we learn not only about the particular material under consideration, but about the global properties of anisotropic fluids and their relationships to other physical systems. Finally, we should point out that although our research is primarily fundamental in nature, determining critical exponents, surface potentials, induced polarizations, etc., a small but important component of our effort involves technology. For example, we have developed a new liquid crystal display architecture which is being developed for commercialization by American industry. This is a symbiotic approach to research, and has been an intellectual stimulation to our effort.

Liquid crystal display



Reflective twisted nematic liquid crystal display.
Polarizing filter film with a vertical axis to polarize light as it enters.
Glass substrate with ITO electrodes. The shapes of these electrodes will determine the shapes that will appear when the LCD is turned ON. Vertical ridges etched on the surface are smooth.
Twisted nematic liquid crystal. Glass substrate with common electrode film (ITO) with horizontal ridges to line up with the horizontal filter. Polarizing filter film with a horizontal axis to block/pass light. Reflective surface to send light back to viewer. (In a backlit LCD, this layer is replaced with a light source.)

A **liquid crystal display** (**LCD**) is a thin, flat panel used for electronically displaying information such as text, images, and moving pictures. Its uses include monitors for computers, televisions, instrument panels, and other devices ranging from aircraft cockpit displays, to every-day consumer devices such as video players, gaming devices, clocks, watches, calculators, and telephones. Among its major features are its lightweight construction, its portability, and its ability to be produced in much larger screen sizes than are practical for the construction of cathode ray tube (CRT) display technology. Its low electrical power consumption enables it to be used in battery-powered electronic equipment. It is an electronically-modulated optical device made up of any number of pixels filled with liquid crystals and arrayed in front of a light source (backlight) or reflector to produce images in color or monochrome. The earliest discovery leading to the development of LCD technology, the discovery of liquid crystals, dates from 1888. By 2008, worldwide sales of televisions with LCD screens had surpassed the sale of CRT units.

LCD alarm clock
Each pixel of an LCD typically consists of a layer of molecules aligned between two transparent electrodes, and two polarizing filters, the axes of transmission of which are (in most of the cases) perpendicular to each other. With no actual liquid crystal between the polarizing filters, light passing through the first filter would be blocked by the second

(crossed) polarizer.The surface of the electrodes that are in contact with the liquid crystal material are treated so as to align the liquid crystal molecules in a particular direction. This treatment typically consists of a thin polymer layer that is unidirectionally rubbed using, for example, a cloth. The direction of the liquid crystal alignment is then defined by the direction of rubbing. Electrodes are made of a transparent conductor called Indium Tin Oxide (ITO). Before applying an electric field, the orientation of the liquid crystal molecules is determined by the alignment at the surfaces of electrodes. In a twisted nematic device (still the most common liquid crystal device), the surface alignment directions at the two electrodes are perpendicular to each other, and so the molecules arrange themselves in a helical structure, or twist. This reduces the rotation of the polarization of the incident light, and the device appears grey. If the applied voltage is large enough, the liquid crystal molecules in the center of the layer are almost completely untwisted and the polarization of the incident light is not rotated as it passes through the liquid crystal layer. This light will then be mainly polarized perpendicular to the second filter, and thus be blocked and the pixel will appear black. By controlling the voltage applied across the liquid crystal layer in each pixel, light can be allowed to pass through in varying amounts thus constituting different levels of gray.

Energy efficiency

Among newer TV models, LCDs require less energy on average than their plasma counterparts. A 42-inch LCD consumes 203 watts on average compared to 271 watts consumed by a 42-inch plasma display. *(This information is outdated - current plasma tv's such as the panasonic TH-42 X10 consume between 80-200W. When measuring the average powerconsumption, it is usually between 120W and 150W.)* Energy use per inch is another metric for comparing different display technologies. CRT technology is more efficient per square inch of display area, using 0.23 watts/square inch, while LCDs require 0.27 watts/square inch. Plasma displays are on the high end at 0.36 watts/square inch and DLP/rear projection TVs represent the low end at 0.14 watts/square inch.

Bistable displays do not consume any power when displaying a fixed image, but require a notable amount of power for changing displayed image.