Sl. No:	Contents	P. No
1.	**Semiconductors:**	
2.	Bonding forces in solids	2
3.	Energy bands	
4.	Metals, Semiconductors and Insulators	3-8
5.	Direct and Indirect semiconductors	9
6.	Electrons and Holes	10
7.	Intrinsic and Extrinsic materials	13-16
8.	Conductivity and Mobility	17
9.	Drift and Resistance	21
10.	Effects of temperature and doping on mobility	23
11.	Hall Effect	26
12.	**P-N Junctions:**	
13.	Forward and Reverse biased junctions	32
14.	- Qualitative description of Current flow at a junction	34-47
15.	reverse bias	48
16.	Reverse bias breakdown- Zener breakdown	49
17.	Avalanche breakdown	50
18.	Rectifiers	53
19.	Optoelectronic Devices Photodiodes: Current and Voltage in an Illuminated Junction,	58
20.	Solar Cells,	62
21.	Photodetectors: Light Emitting Diode	65
22.	Light Emitting materials	69
23.	**Bipolar Junction Transistor:** Fundamentals of BJT operation	72
24.	Amplification with BJTS	77
25.	BJT Fabrication	83
26.	The coupled Diode model (Ebers-Moll Model)	88
27.	Switching operation of a transistor, Cutoff, saturation	98
28.	switching cycle, specifications	104
29.	Drift in the base region	106
30.	Base narrowing	108
31.	Avalanche breakdown	110
32.	Fabrication of PN Junction: Thermal Oxidation	113
33.	Diffusion	115
34.	Rapid Thermal Processing	118
35.	Ion implantation	119
36.	Chemical vapour deposition.	121
37.	photolithography	122
38.	Etching, metallization	127

periodic Table

S. No.	Date	Title	Page No.	Teacher's Sign / Remarks
(1)		Solid State Electronic Devices Ben. G. Streetman Sanjay Banerjee		

The definition of ionic bond is when a positively charged ion forms a bond with a negatively charged ion and one atom transfers electrons to another Eg. NaCl

* Ionic bonding is the complete transfer

Shells : K, L, M, N, { The general formula is that the nth shell can in principle hold up to $2(n^2)$ electrons.

Electron configuration Table : Describes how many electrons are in each energy level of an atom & how electrons are arranged within each energy level.

Eg. Gold

$1s^2$
$2s^2$ $2p^6$
$3s^2$ $3p^6$ $3d^{10}$
$4s^2$ $4p^6$ $4d^{10}$ $4f^{14}$
$5s^2$ $5p^6$ $5d^{10}$
$6s^1$

Video: How to draw atom
https://education.jlab.org/qa/electron_config.html

3.1.1, 3.1.2

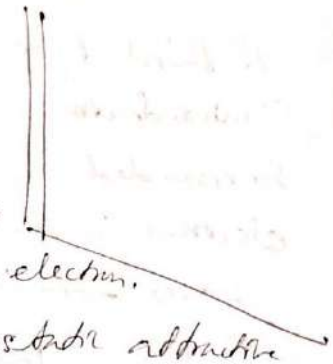
Bonding forces in Solids

Electrons interact with neighbouring atoms hence held the crystal together.

Eg:- Alkali halides Eg. Sodium chloride (typical & characterised) NaCl are typified by ionic bonding.

In NaCl lattice each Na atom is surrounded by 6 Cl atoms & vice versa. The two dimensional figure shown in Fig. shows 4 nearest Cl atoms.

The electronic structure of Na (Z=11) is [Ne] $3s^1$, and Cl (Z=17) has the structure [Ne] $3s^2 3p^5$.



In the lattice each Na atom gives up its outer 3s electron to a Cl atom. The Na^+ ion has a net positive charge, having lost an electron & Cl^- ion has a net negative charge after gaining an electron.

That is each Na^+ ion exerts an electrostatic attractive force upon six Cl^- neighbors, & vice versa. These coulombic forces pull the lattice together until a balance is reached with repulsive forces.

Note: In NaCl structure, all the electrons are tightly bound to atoms. That is after electron exchange, between Na & Cl the outer orbits of all atoms are completely filled, (The ions (Na^+, Cl^-) have configurations of the inert atoms Ne and Ar) there is no loose bound electrons to participate in current flow. As a result NaCl is a good insulator.

Metals: In metals the outer electronic shell is partially filled, usually not more than 3-electrons. Alkali metals (highly reactive) Eg. Na have only one electron in its outer orbit, & is loosely bound. Can give up easily in ion formation.

In metal the outer one electron of each alkali atom is contributed to the crystal as a whole, so that solid is made up of ions with closed shells immersed in a sea of free electrons.

The forces holding the lattice together arise from an interaction between +ve ion cores & the surrounding free electrons. This is one type of metallic bonding.

In general, metals have the sea of ~~electrons~~ electrons in common & these electrons are free to move about the crystal under the influence of electric field.

Semiconductor:

A third type of bonding is exhibited by diamond lattice semiconductors. Atoms in Si, Ge and C diamond lattice is surrounded by four nearest neighbours, each with four electrons in the outer orbit. In this case each atom shares its valence electrons with its four neighbors (Fig 15). The bonding forces arise from a quantum mechanical interaction between shared electrons. This is known as covalent bonding. Each electron pair constitutes a covalent bond.

~~As~~ similar to ionic crystals, no free electrons are available to the lattice in the covalent diamond structure. Hence we may think Ge & Si should be insulators. However an ~~atom~~ electron can be thermally (above 0°k) or optically excited out of covalent bond & free to participate in conduction. This is an important concept in semiconductors.

Compound semiconductors (Ex: GaAs) have mixed bonding in which both ionic and covalent bonding forces participate

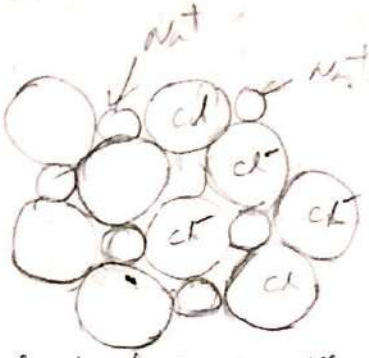


Fig.(1a) Example of ionic bonding in NaCl.

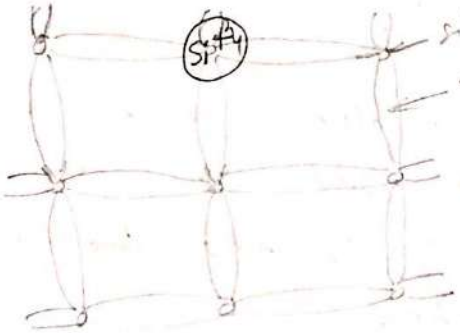


Fig (1b) covalent bonding in the silicon crystal.

Energy Bands: Atomic distance/lattice constant of Si = 5.43 A°

As isolated atoms are brought closer, various interactions occur between neighbouring atoms. The forces of attraction & repulsion between atoms find a balance at the proper interatomic spacing for the eq. crystal. In the process, important changes occur in the electron energy level configurations, which is responsible for various electrical properties of solids.

Qualitatively we say that as atoms are brought together, the application of the Pauli exclusion principle becomes important. When two atoms are completely isolated from each other, then there is no interaction of electron wavefunctions, they can have identical electronic structures. As the distance between atoms decreases, the wave functions begins to overlap. The exclusion principle dictates that no two electrons in a given interacting system can occupy the same quantum state.

P.T.O.

There the discrete energy levels of the isolated atoms must split into new levels belonging to the pair rather than to the individual atoms.

In a solid, many atoms are brought together, so that the split energy levels form essentially continuous bands of energies. (called Energy bands)

Eg:

Let us consider the imaginary ~~structure of a silicon~~ formation of a silicon crystal formed by isolated silicon atoms. Each isolated silicon atom has an electronic structure ~~$1s^2 2s^2 2p^6 3s^2 3p^2$~~ $1s^2 2s^2 2p^6 3s^2 3p^2$ in the ground state. Each atom has available two 1s state, two 2s state, six 2p states, two 3s states, six 3p states and higher states.

For N atoms, there will be 2N, 2N, 6N, 2N and 6N states of type 1s, 2s, 2p, 3s and 3p respectively. When the interatomic spacing is large, the energy levels of these subshells are isolated

As the interatomic spacing decreases, these energy levels split into bands, beginning with the outermost ($n=3$). As the "3s" and "3p" bands grow they merge into a single band — composed of a mixture of energy levels. This band of "3s – 3p" levels contains 8N available states.

As the distance between atoms approaches the equilibrium interatomic spacing of silicon, this band splits into two bands separated by an energy gap E_g.

The upper band (conduction band) contains 4pN states and (all are empty at 0°K)
the lower band (valence band) also contains —11—, all are occupied at 0°K.

The energy gap contains no available states for the electrons to occupy & hence is called forbidden gap.

Note: Valence electrons contain all the valence electrons

Conduction band is primarily responsible for conduction of excited electrons into it.

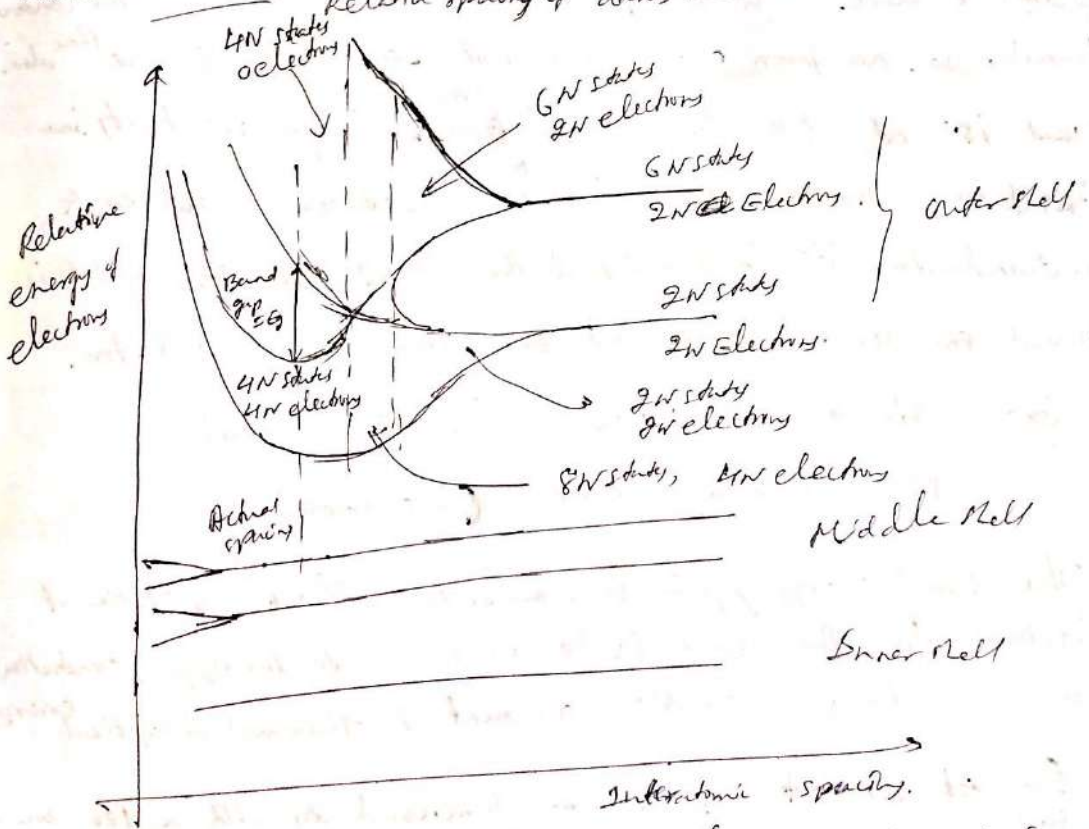
note: The low lying levels are tightly bound.

At this stage, the lower "1S" band is filled with 2N electrons which originally resided in the collective 1s states of the isolated atoms. The 2s band and 2p bands will have 2N and 6N electrons in them respectively.

There were 4N electrons in the originally isolated n=3 shells (2N in 3s states and 2N in 3p states). These 4N electrons will occupy lowest energy state at 0°k.

In Silicon, there are exactly 4N states in the valence band available to the 4N electrons, thus at 0°k, every state in valence band is filled, while conduction band is completely empty of electrons.

This ~~will affect~~ arrangement of empty & filled energy bands has an important effect on the electrical conductivity of solids.



Fig(2) Energy bands of Si atom as a function of inter atomic spacing

Metals, Semiconductors & insulators:

Every solid has its own characteristic energy band struct. This variation in band structure is responsible for different electrical characteristics of different materials.

Prior to discussing ~~current~~ mechanisms of current flow we see that for electrons to experience acceleration in an applied electric field, they must be able to move into new energy states. That means, there must be empty states available to the electrons.

Ex: If few electrons reside in an empty band, ample unoccupied states are available, into which the electrons can move.

OR

In silicon band structure, the valence band is completely filled with electrons at 0°K and conduction band is completely empty. There is no charge transport within the valence band, since no empty states are available for electrons to move. Since there is no electrons in the conduction band, so no charge transport can take place there also. That is at 0°K, the silicon behaves as a high resistivity insulator. However the difference between insulator & semiconductor lies in the size of the band gap Eg, which is much smaller in semiconductors than in insulators.

Ex- Si → Eg = 1.1 eV. (Semiconductor)

Diamond → Eg = 5 eV (Insulator)

The small energy gap in semiconductors allows excitation of electrons from the lower (valence) band to the upper (conduction) band with a reasonable amount of thermal or optical energy.

Ex: At room temp. for a semiconductor with a 1ev band gap will have significant number of electrons excited

thermally across the energy gap into the conduction band, whereas an insulator with $E_g = 10$ eV will have negligible number of such excitations.

i.e. no of electrons available for conduction can be increased greatly in semiconductors by thermal or optical energy.

In metals the bands either overlap or are partially filled. Thus electrons & empty energy states are intermixed within the bands so that electrons move freely under the influence of an electric field. [Fig. (3)].
The metals have high electrical conductivity as expected.

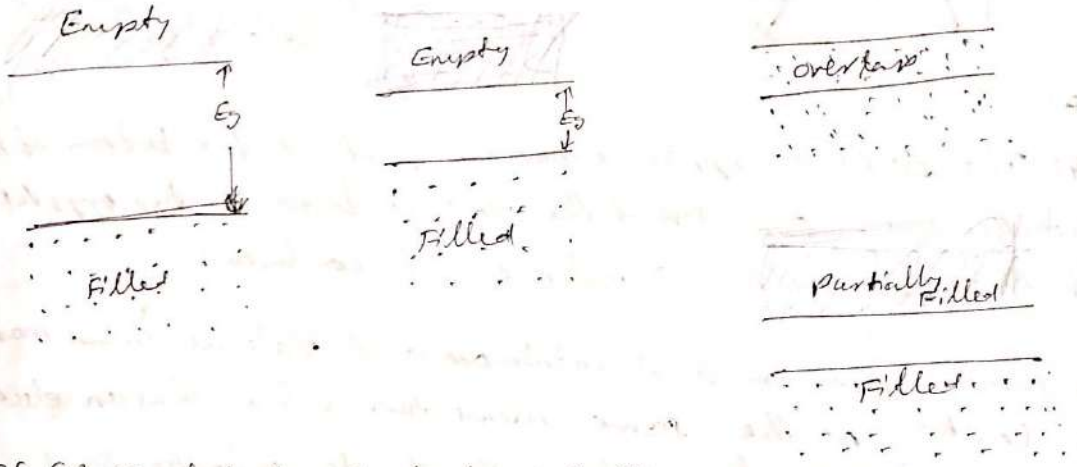
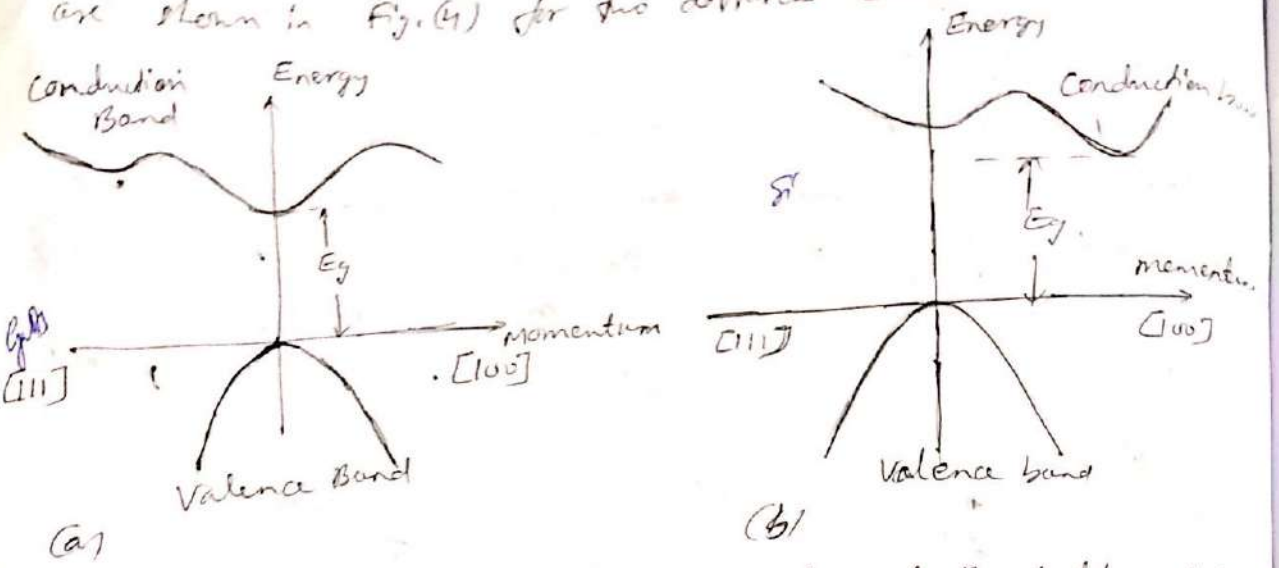


Fig. (3) Typical band structure at 0°K.
(Simplified representation of a complex band structure).

PT.

A little complicated energy band diagram showing the electron energy Vs momentum in two crystal directions is as shown in Fig.(4) for the different cases.



Fig(4) The electron energy Vs momentum plots at the bottom of the conduction band and top of the valence band in two crystal directions (a) direct (b) indirect semiconductor.

In case (a) conduction band minima and valence band maxima are located at the same momentum value. Thus an electron can make transition from Valence Band to Conduction Band without any change in momentum (Ex GaAs). Such materials are called direct semiconductors.

In case (b) The conduction band (CB) minima and valence band (VB) maxima are not located at the same momentum values. Thus an excitation of an electron from the valence band to the conduction band needs a change in momentum in addition to the extra energy input. (Ex Si) Such semiconductors are called indirect semiconductors.

In direct semiconductor (SC), a photon of energy $h\nu = E_g$ can excite an electron from VB to CB (direct transition).

However in indirect semiconductors, this type of direct transition is not possible. Since photons have very small momentum while the electron has to undergo a large change in momentum.

In this case, electron transition from valence band to conduction band can occur by involving a lattice phonon (thermal energy), which can support the required momentum change.

3.2.1 (1.2.1) Electrons & Holes:

As the temperature of semiconductor raises above 0°K, some electrons in VB receive enough thermal energy to excite into the conduction band. This creates some electrons in the conduction band (some occupied states in CB) & some unoccupied states in the VB. [Fig (5)]

[The energy needed to break the bond is actually the band gap energy in the energy band diagram]

For convenience, an empty state in the VB is called a hole. That means this process created an electron (in CB) – hole (in VB) pair (EHP).

After excitation an electron is surrounded by a large number of unoccupied energy states. Hence few electrons in the conduction band are free to move through the available empty states.

[At room temp. intrinsic carrier concentration of Si $n_i(Si) \approx 10^{10}$ EHP/cm³. compared to the silicon atom density of 5×10^{22} atoms/cm³].

In the following discussion we shall concentrate on electrons in the CB & holes in the VB. The movements of these two types of charge-carriers accounts for the current flow in semiconductor (SC).

PTO

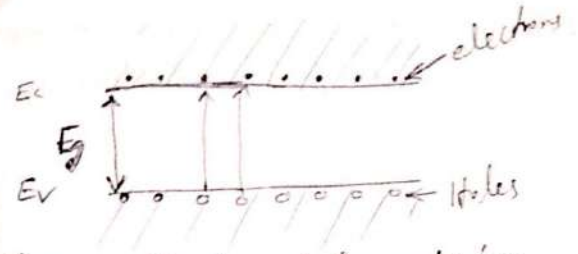


Fig.(5) Electron-hole pairs in a Semiconductor.

Note(1) In subsequent discussions we consider the bottom of the conduction band as E_c and the top of Valence band as E_v.

Note(2) To discuss hole movement opposite to that of electron movement & both will contribute for current can be explained by empty & filled bottles analogy on page 64-65 of text.

The energy band diagram actually reflects the electronic energy, ie. if an electron gains energy it moves up in the energy band diagram. On the otherhand, when a hole gains energy, it moves down due to its opposite charge. Thus holes seeking the lowest energy state are available at the top of the VB (E_v) while electrons are available at the bottom of the conduction band (E_c). The unit of energy generally used is electron volt (eV), which is equal to 1.6×10^{-19} J.

The unit of electron volt is so named due to the fact that if an electron falls through a potential of 1V (the kinetic energy gained), which is equal to decrease in its potential energy.

ie $E = qV = 1.6 \times 10^{-19} C \times 1V = 1.6 \times 10^{-19} J$

(Electron × Volt = Joule)

So in the energy band diagram if electron energy E goes up by 1eV it means equivalently it has fallen through a potential V of 1V. Note that although V and E are dimensionally different they have the same magnitude. So we say that energy band diagram reflects the electronic potential, which is -ve of the electrostatic potential. [The electrostatic potential is defined

For an electron in the conduction band having an energy E, the energy corresponding to the bottom of the conduction band (E_c) is its potential energy, while ($E - E_c$) is its kinetic energy, as in Fig. (a)

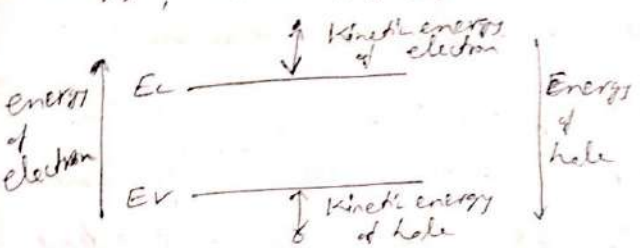
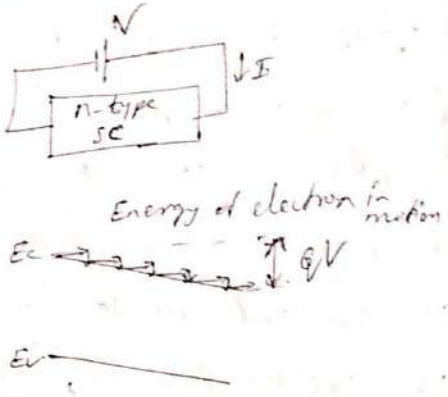


Fig (a)



(b)

Fig (a) & (b) Energy band diagram of a semiconductor under (a) Zero bias and (b) under biasing conditions.

If we apply a voltage V across SC, the energy band diagram looks as shown in Fig (1b). As electron moves under the influence of the applied potential V, its loss in potential energy (qV) is equal to the gain in kinetic energy. However as electron moves, it suffers a number of collisions, giving its kinetic energy to the lattice (generates heat) & falls to the bottom of the conduction band.

Analogy: consider a mass 'm' at a height 'h', which has a potential energy mgh, where 'g' is the acceleration due to gravity. when it falls & reaches the ground level under the influence of gravitational field, its entire potential energy is converted into kinetic energy.

However as it 'collides' with earth, it ultimately loses all kinetic energy & comes to rest.

3.2.3 (1.7 + 1.0 = 2.7)

Intrinsic Material:

A perfect semiconductor crystal with no impurities or defects is called an intrinsic semiconductor. In such materials, the valence band is filled and the conduction band is empty at 0°K & hence there is no charge carriers. At higher temperatures electron-hole pairs (EHPs) are generated as covalent bonds break & e⁻ valence band electrons are excited to E_c [Refer Theory of Pdf]. These EHPs are the only charge carriers in intrinsic material.

Since electrons & holes are created in pairs, the electron concentration, n (electrons/cm³) in the conduction band is equal to the hole concentration (holes/cm³), p in valence band.

i.e $n = p = n_i$ ——— (1)

where n_i is called the intrinsic carrier concentration. As the temp^re increases, lattice vibrations also increases and more EHP's are created. Thus n_i increases with temp^re.

That is there is a certain concentration of electron-hole pairs n_i at a given temp^re. That is to maintain steady state carrier concentration there must be a recombination of EHPs at the same rate at which they are generated.

Recombination occurs when an electron in the conduction band makes a transition to the an empty state (hole) in the valence band, thus annihilating the pair.

If generation of EHPs as g_i (EHP/cm³-s) and recombination rate as r_i, equilibrium requires

$r_i = g_i$ ——— (2)

These rates are temp^re dependent.

As T increases, $g_i(T)$ increases and new carrier concentration n_i is established & hence higher recombination rate $r_i(T)$. for balancing.

At any temp^{re} we can predict that rate of recombination of electrons & holes r_i is proportional to the equilibrium concentration of electrons n_0 & concentration of holes.

$$r_i = \alpha_r n_0 p_0 = \alpha_r n_i^2 = g_i \quad \text{---(3)}$$

where α_r is a constant depends recombination mechanism.

NOTE: n_i is a function of both temp^{re} & energy gap

Eg. For Si $n_i \simeq 1.5 \times 10^{10}$ / cm^3 at 300 k

($E_g = 1.12$ ev) $\simeq 3 \times 10^{14}$ / cm^3 at 500 k

Ge has which has larger energy gap than Si has.

($E_g = 1.42$ ev) $n_i \simeq 1.75 \times 10^6$ / cm3 at 300 k.

3.2.4. Extrinsic material:

Additional carriers can be generated in semiconductors by introducing impurities into the crystal. This process is called doping. This process is commonly used for varying the conductivity of semiconductors.

By doping material can be made as n-type or p-type depending on dopant.

If dopant is pentavalent → ~~material~~ material becomes n-type

" trivalent → " p-type.

* When the equilibrium carrier concentration n_0 and p_0 are different from intrinsic carrier concentration n_i, the material is said to be extrinsic.

Explanation with energy band diagram

When impurities are introduced in a perfect crystal, additional energy levels are created in the energy band structure.

Eg: a group V element (P, As and Sb) (Phosphorus, Antimony) generates an energy level E_D which is very close to E_c in silicon or germanium.

PTO

This level is filled with electrons at 0K and with a little thermal energy these electrons can be excited into the conduction band. Such an impurity is called a donor as it donates electrons.

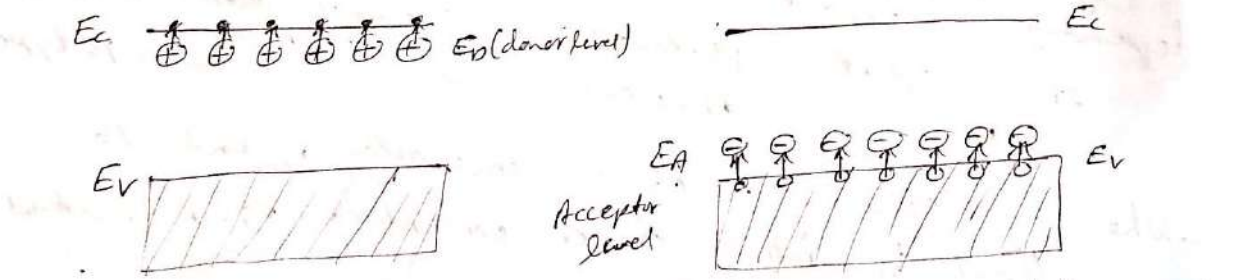
That is even when the temperature is low, the number of thermally generated carriers is small, there is a large concentration of electrons in the conduction band of this the type of doped semiconductor ie. $n >> n_i, p$. Such a sc is called n-type semiconductor.

p-type sc:

If an element from group III is used to dope silicon or germanium, an energy level E_A is created very close to E_V. This level is empty at 0K but with a little thermal energy, electrons from the valence band can be excited into this level, creating a large number of holes in the valence band.

As this level accepts electrons, this type of dopants are termed acceptors & material is called p-type where $p >> n_i, n$.

Energy band diagrams of n-type and p-type sc with the donor and acceptor levels are shown in Fig(7)



(a) n-type sc with donor level E_D

(b) p-type sc with acceptor level E_A.

Fig(7)

note: The process of donating or accepting electrons ionizes the impurity atoms, the donors becoming +vely charged, while the acceptors become -vely charged.

When a semiconductor is n- or p-doped, only one type of carriers dominate. These are called majority carriers.

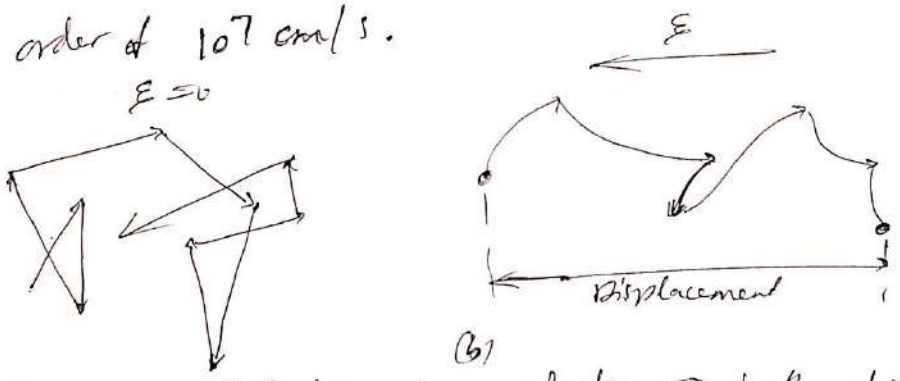
Thus, electrons in the n-type sc &
holes in the p-type sc are majority carriers.

Conductivity & Mobility:

Note(1) Knowledge of carrier concentration in solids is necessary for calculating current flow in the presence of electric & magnetic fields. In addition to the values of n & p. we must also need to account collisions of the charge carriers. in the lattice.

Note(2) The scattering processes depend on temp, which affects the thermal motion of the lattice atoms & the velocity of carriers.

Even at thermal equilibrium, the charge carriers in a solid are in constant motion. The thermal motion of individual electron may be visualized as random scattering from lattice atoms, impurities, other electrons & defects, Fig(8). The random motion of electron leads to a net zero displacement over a sufficiently long period of time, for a group of n electrons/cm^3. Fig(8) depicts the trajectory of an electron consisting of a series of straight lines between collisions. The average time between collisions is called the mean free time and the average distance travelled between collisions is the mean free path. The root mean square (rms) value of the random thermal velocity v_{th} is of the order of 10^7 cm/s.



Fig(8) The possible trajectory of an electron (a) in the absence of electric field and (b) in the presence of electric field.

When an electric field E is applied to the semiconductor, the induction charges. Each electron experiences a force $-qE$ due to the field and is consequently accelerated between

each collision in a direction opposite to the field, thereby a net displacement of the electron, as in Fig.(5b). That is an additional velocity component is superposed on the random thermal velocity, & is called drift velocity, v_d, which is defined as the net displacement per unit time.

That is in the presence of electric field, electron acquires a drift velocity & its resultant velocity is therefore the sum of v_x and v_d. Since the mean free path does not alter due to the application of field, this increases in electron velocity reduces the mean free time. Consequently the electron suffers more frequent collisions and therefore loses more energy to the lattice. Hence a resistive force develops which does not allow a continuous increase of the carrier velocity. At equilibrium, an electron acquires an average drift velocity. Assuming that in steady state, all momentum gained between collisions is lost to the lattice, hence we can write.

$$m_e^* v_d = -q\mathcal{E}\tau_c \qquad \text{——— (1)}$$

where m_e^* = effective mass of the electron & collisions.

τ_c = mean free time between two successive

$$\therefore v_d = -\frac{q\tau_c}{m_e^*}\mathcal{E} = -\mu_n \mathcal{E} \qquad \text{——— (2)}$$

Thus the average drift velocity of the electrons is proportional to the applied electric field & this constant of proportionality $\left(\frac{q\tau_c}{m_e^*}\right)$ is called the mobility of electrons.

The -ve sign in eq(2) indicates that the direction of v_d is opposite to that of the applied field.

Similarly for holes

$$v_d = \frac{q\tau_c}{m_h^*}\cdot\mathcal{E} = \mu_p\mathcal{E} \qquad \text{——— (3)}$$

When m_h^* = effective mass of holes.

note: 1) Since the holes carry +ve charge, their movement is in the same direction as the electric field & hence there is no negative sign.

2) Usually $m_e^* < m_h^*$ & hence the mobility of electrons is larger than that of the holes for any SC.

Table: Electron & hole mobilities in some common SC

SC	Electron Mobility (cm^2/Vs)	Hole Mobility (cm^2/Vs)
Silicon	1350	450
Germanium	3900	1900
GaAs	8500	400
Indium Phosphide (InP)	4600	150
Indium Arsenide	33000	460

The current density resulting from net drift is just the number of electrons crossing a unit area/unit time (or flux) multiplied by the charge ~~on the electron~~ (eq)

$$\text{i.e. } J_x = -q n \langle v_d \rangle \quad \text{---} (4)$$

$$\frac{\text{ampere}}{cm^2} = \frac{\text{coulomb}}{\text{electron}} \cdot \frac{\text{electrons}}{cm^3} \cdot \frac{cm}{s} \quad \text{---} (4)$$

From eq (2)

$$J_x = \frac{q^2 \tau_c n}{m_e^*} \mathcal{E}_x \quad \text{---} (5)$$

Thus current density is proportional to the electric field as expected from ohms law

$$J_x = \sigma \mathcal{E}_x \text{ . then } \sigma = \frac{n q^2 \tau_c}{m_e^*} \quad (6)$$

The conductivity σ ($\Omega cm)^{-1}$ can be written

$$\sigma = q n \mu_n, \text{ where } \mu_n = \frac{q \tau_c}{m_e^*} \quad (7)$$

where μ_n → electron mobility

Here m_c^* is the conductivity effective mass for electrons, different from the density of - states effective mass (m_{dn}^* or m_{dp}^*)

The density of effective mass is used to count the number of carriers in bands.

Conductivity effective mass is used for charge transport problems.

The mobility defined in eq (1) can be expressed as the average particle drift velocity / unit electric field. Comparing eq (1) & (2) we have

$$\mu_n = -\frac{\langle v_d \rangle}{\mathcal{E}_x} \quad \text{——— (5)}$$

The units of mobility are $(cm/s)/(V/cm) = cm^2/V\text{-}s$.

The $-$ ve sign in eq (5) results in a positive value of mobility, since electrons drift opposite to the field.

The current density in terms of mobility is

$$J_n = q n \mu_n \mathcal{E}_x \quad \text{——— (9)}$$

Eq (9) is valid if the current is carried primarily by electrons.

For hole conduction, we change n to p, $-q$ to $+q$ and μ_n to μ_p.

$$\text{where } \mu_p = +\frac{\langle v_d \rangle}{\mathcal{E}_x} \quad \text{——— (10)}$$ is the mobility of holes.

If both electrons & holes participate, we must modify eq (9) to

$$J_n = q(n\mu_n + p\mu_p)\mathcal{E}_x = \sigma \mathcal{E}_x \quad \text{——— (11)}$$

PT

Drift and Resistance:

Let us discuss the drift of electrons & holes. If the semiconductor bar of Fig(6) contains both types of carriers eq.(11) gives the conductivity of the material. Therefore the resistance of the bar is

$$R = \frac{\rho L}{wt} = \frac{L}{wt} \cdot \frac{1}{\sigma} \qquad \text{---(12)}$$

Where ρ → resistivity (Ω-cm)
L → Length of the bar,
w → width of the bar
t → Thickness of the bar.

The mechanism of carrier drift requires that the holes in the bar move as a group in the direction of the electric field & that the electrons move as a group in the opposite direction. Both e & h component of current are in the same direction of the E field.

Conventional current is positive in the direction of hole flow & opposite to the direction of electron flow. The drift current described by eq (11) is constant throughout the bar.

Note: The contacts at the end of the bar are ohmic, meaning they are perfect sources & sinks of both carrier types & they will not inject or collect either electrons or holes.

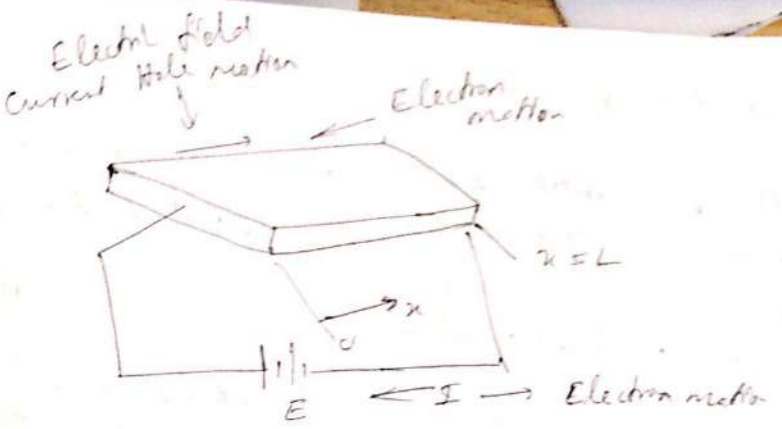


Fig (9) Drift of electrons & holes in a SC bar

If we consider current is carried around the external circuit by electrons, we can visualize electrons flowing into the bar at one end & out at the other (always opposite to I). That is every electron leaving left end ($x=0$) [Fig (9)], there is a corresponding electron entering at $x=L$, so that the electron concentration remains constant 'n' in the bar.

If the holes reaches ohmic contact at $x=L$, it recombines with an electron. As the hole disappears, a corresponding hole must appear at $x=0$ to maintain space charge neutrality.

note: consider the source of this hole as the generation of the EHP at $x=0$, with the hole flowing into the bar, & the electron flowing into external ckt.

Ex: Find the resistivity of intrinsic Si at 300K

Given $\mu_n = 1350$ & $\mu_p = 480 \text{ cm}^2/\text{V-s}$ for intrinsic Si.

Soln: Since $n_0 = p_0 = n_i$

$$\sigma_i = q(\mu_n + \mu_p) n_i = 1.6\times10^{-19}(1830)\times(1.5\times10^{10})$$

$$= 4.39\times10^{-6} (\Omega\text{-cm})^{-1}$$

$$\rho_i = \sigma_i^{-1} = 2.28\times10^{5}\ \Omega\text{-cm}.$$

Note: Charge neutrality occurs when all the charge in a volume adds to zero, it is neutral (neither +ve or -ve).

If $\rho = q(p_0 - n_0 + N_D - N_A) = 0$, then sum of the charges associated with the carriers must equal zero.

i.e $p_0 - n_0 + N_D - N_A = 0$ → Acceptor Concentration, Donor Concentration

We have $p_0 n_0 = n_i^2$

PTO

Effects of Temperature & doping on mobility

Drift velocity of carriers stabilizes to an average value (instead of increasing continuously) proportional to the applied electric field, as the carriers have lose their energy gained from the field by collision (scattering).

The basic scattering mechanisms that influence electron & hole mobility are lattice scattering & impurity scattering.

In lattice scattering, a carrier moving through the crystal is scattered by a vibration of the lattice (phonon scattering, vibrations of atom is called phonons). The freq of such scattering events increases as temper increases, given by the relation

$$\mu_L \propto T^{-3/2} \quad \text{——— (xi)}$$

(since thermal agitation of the lattice becomes increases with temper)

That is mobility decreases as temper increases as in Fig. (10)

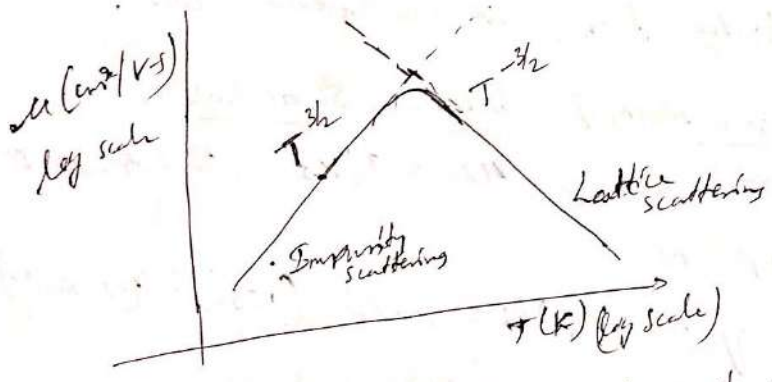


Fig. (10) Approximate temper dependence of mobility with both lattice & impurity scattering.

Scattering from crystal defects such as ionized impurities becomes the dominant mechanism at low temps.

Impurity Scattering:- is the dominant factor at low temperatures. When a charge carrier travels past an ionized impurity, its path gets deflected due to Coulomb force interaction. The probability of impurity scattering depends on the total concentration of ionized impurity & therefore, the carrier mobility is reduced for highly doped sc.

However the impurity scattering is insignificant at higher temperature. At higher temperatures, carriers move faster & therefore spend less time in the vicinity of an ionized impurity. Thus they are less effectively scattered by the ionized dopants.

The dependence of carrier mobility due to impurity scattering $\mu_i \propto \frac{T^{3/2}}{}$.

~~The In general the~~

In general, the carrier mobility in a sc can be expressed as $\frac{1}{\mu} = \frac{1}{\mu_L} + \frac{1}{\mu_i}$

The electron mobility in silicon for various doping concentrations of the sample is plotted as a function of temperature in Fig. (11). It is seen that for lightly doped samples, mobility decreases with increase in temp^{re} clearly showing the dominance of lattice scattering.

However for heavily doped samples, mobility is low at low temp^{re}atures, where the impurity scattering dominates. It increases ~~with~~ with temp^{re}, reaches a peak & then at higher temperatures lattice scattering takes over and mobility decreases again. Also at any temperature, mobility is lower for samples with higher doping concentration.

The variation of electron & hole concentration as a function of doping is shown in Fig. (12).

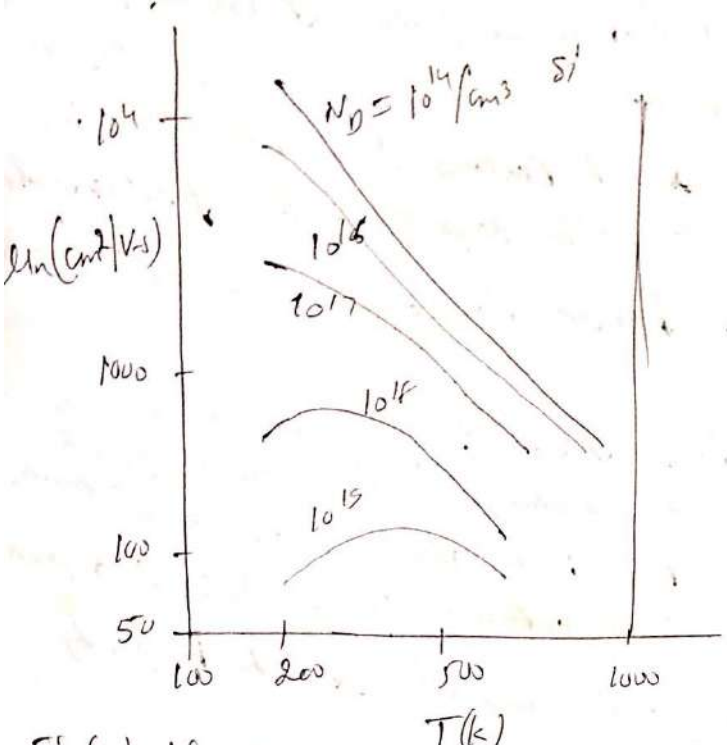


Fig (11). Variation of electron mobility in Si as Vs doping concentration & Temperature

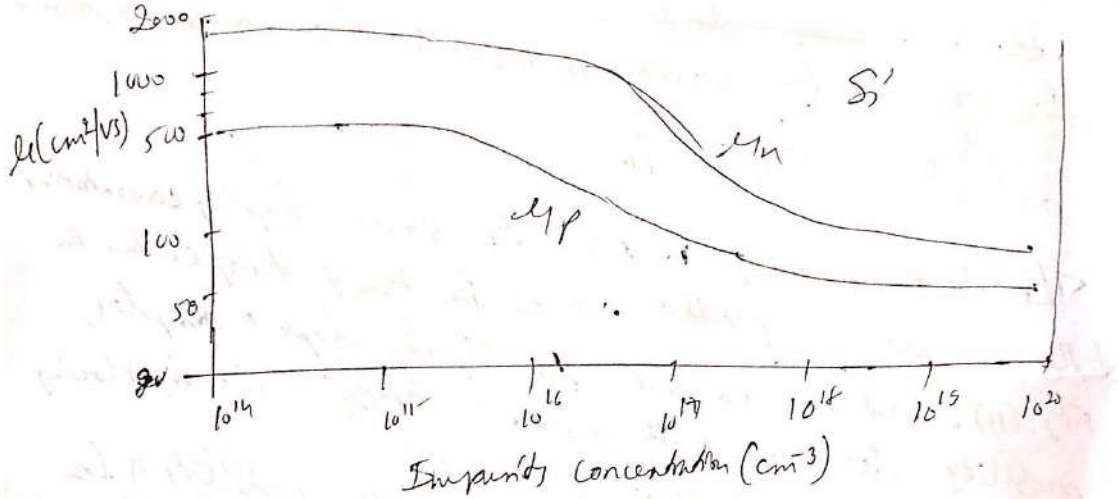
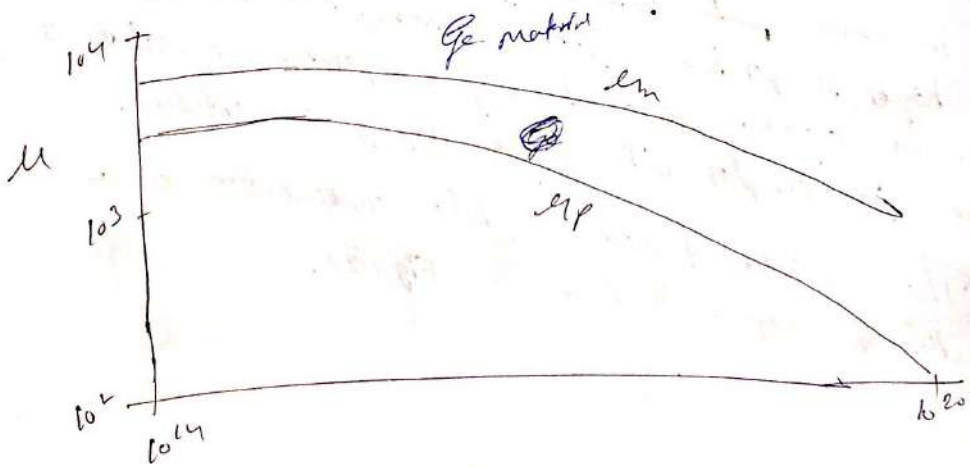


Fig.(12) n & h mobilities in Si at room temperature as a function of impurity concentration.



The Hall effect.

Is used to measure carrier concentration.

Principle: When a magnetic field is applied in a direction perpendicular to the flow of charge carriers, the path of the carriers gets deflected.

Fig(13) shows a p-type semiconductor bar of length 'L', width 'W', & thickness d. Electric & magnetic fields are applied to this bar along the x-axis and z-axis respectively. Due to the magnetic field B_z, an upward Lorentz force given by qV_xB_z is exerted on the carriers moving along x-direction with a drift velocity V_x.

The upward force results in accumulation of holes at the top of the sample which gives rise to a corresponding downward electric field E_y. Since there is no net current flow in the y-direction in the steady state, the forces due to electric field must exactly balance the Lorentz force

i.e. $qE_y = qV_xB_z$

$E_y = V_xB_z$

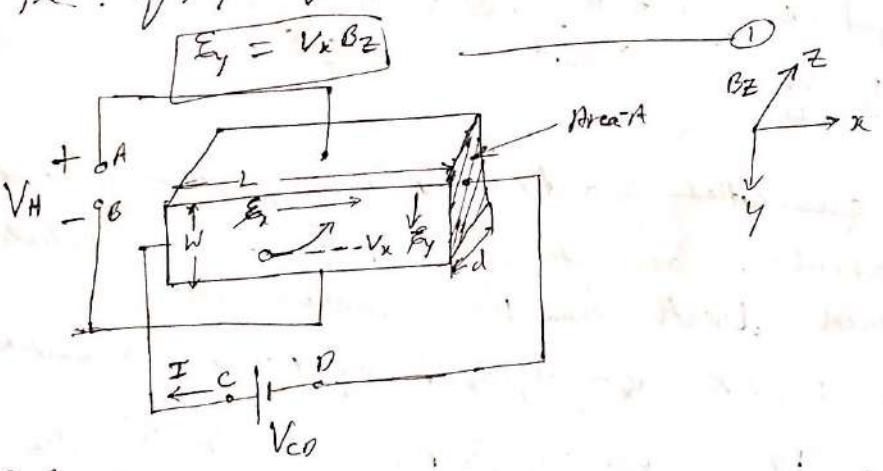


Fig. (13)

This effect is because the magnetic field results in an increase in the hole concentration at the top surface of the sample, as in Fig. (13), till the electric field E_y becomes as large as V_xB_z. Once this electric field is established, no net lateral force is experienced by the holes as they drift along the bar in the x-direction. Till $E_y = V_xB_z$

The establishment of this electric field is known as the Hall effect, & the field itself is called the 'Hall field'. The voltage across A and B is called the Hall voltage V_H.

P.T.O

From fig.(13), we have $V_H = E_y W$

From eq ① and hole drift current density $J_p = q p v_x$

$v_x = J_p / q p$

We can express Hall field as

$$E_y = v_x B_z = \frac{J_p B_z}{q p} = J_p B_z R_H \quad \text{——②}$$

i.e. the Hall field is proportional to the product of current density and applied magnetic field with a proportionality constant $R_H = 1/q p$ call Hall co-efficient. ③

Measurement of Hall voltage for a known current (I) and magnetic field (B_z) is carried out to obtain the hole concentration 'p' in the sample.

From eq ③ & ②

$$p = \frac{1}{q R_H} = \frac{J_p B_z}{q E_y}$$

$$= \frac{\left(\frac{I}{wd}\right) \cdot B_z}{q \left(\frac{V_H}{w}\right)} = \frac{I B_z}{q V_H \cdot d} \quad \text{——④}$$

$$E_y = \frac{J_p B_z}{q p} = J_p B_z R_H$$

$$\frac{1}{q R_H} = p$$

$$\frac{J_p B_z}{q E_y} = p$$

Since all quantities quantities on the right hand side of eq ④ are measurable, accurate values of the hole concentration can be obtained directly from Hall measurements.

now substituting $v_x = \mu_p E_x$ in eq ②, we can write

$$E_y = \frac{V_H}{w} = B_z \cdot \mu_p E_x B_z = \mu_p \frac{V_{CD}}{L} \cdot B_z \quad \text{——⑤}$$

$$\therefore \mu_p = \frac{L \cdot V_H}{w V_{CD} B_z}$$

If mobility can also be obtained by direct measurement as

$$\mu_p = \frac{L V_H}{w V_{CD} B_z} \quad \text{——⑥}$$

Hall measurement carried out over a wide range of temperature yields plot of majority concentration & mobility Vs temperature.

Similarly hall measurement can be done for n-type material also.

If a measurement of resistance R is made, the resistivity can be calculated as

$$\rho\,(\Omega\text{-cm}) = \frac{R\,dw}{L} = \frac{V_{CD}/I_x}{(L/dw)} \quad \text{—(1)}$$

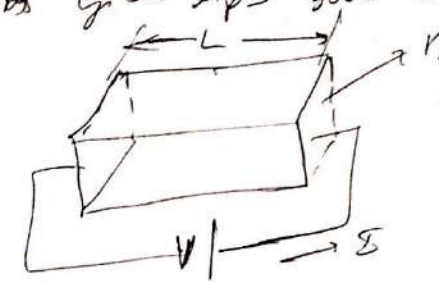
& conductivity $\sigma = 1/\rho$ is given by $q\mu_p p_o$, the mobility is simply the ratio of Hall co-efficient & the resistivity.

$$\mu_p = \frac{\sigma}{qp} = \frac{1/\rho}{q(1/qR_H)} = \frac{R_H}{\rho} \quad \text{—(2)}$$

Problem

A bar of p-type silicon, as in Fig. (below) has a cross sectional area of $A = 10^{-6}\ cm^2$ & a length $L = 1.2 \times 10^{-3}$ cm. For the applied voltage of 5V, a current of 2mA is required. What is the required (a) resistance (b) resistivity of the silicon, and (c) impurity doping concentration.

Ans Given $\mu_p = 300\ cm^2/Vs$



Given
$A = 10^{-6}\ cm^2$ $I = 2mA$.
$L = 1.2 \times 10^{-3}$ cm
$V = 5V$, $I = 2mA$

(a) Resistance = ? (b) Resistivity = ? (c) doping conc = ?

$$R = \frac{\rho L}{A} = \frac{V}{I} = \frac{5}{2 \times 10^{-3}} = 2.5\ k\Omega$$

$$\rho = \frac{RA}{L} = \frac{2.5 \times 10^3 \times 10^{-6}}{1.2 \times 10^{-3}} = 2.08\ \Omega cm$$

$$\sigma = \frac{1}{\rho} = \frac{1}{2.08\ \Omega cm} = 0.48\ (\Omega cm)^{-1}$$

$$\sigma = q \mu_p \cdot p = q \mu_p \cdot N_a$$

For intrinsic material
$\sigma = q(\mu_n + \mu_p) n_i$

for p-type material $\sigma = q(\mu_n n + \mu_p p)$
$\simeq q \mu_p \cdot p$ (for p-type material)

$$\therefore N_a = \frac{\sigma}{q \mu_p} = \frac{0.48}{1.6 \times 10^{-19} \times 300}$$

$$= 1 \times 10^{16}$$

3.20 half problem

Test-1

3.13:

A 2-cm long piece of Si with cross sectional area of $0.1\ cm^2$ is doped with donors at $10^{15}\ cm^{-3}$, & has a resistance of 90 ohms. The saturation velocity of electrons in Si is 10^7 cm/s for field above 10^5 V/cm. calculate the electron drift velocity, if we apply a voltage of 100V across the piece

What is the current through the piece if we apply a voltage of 10^6 V across it?

Ans Given

$L = 2\text{cm}$ $A = 0.1\ \text{cm}^2$, $N_d = 10^{17}/\text{cm}^3$ $R = ?\ \Omega$

$v_s = 10^7\ \text{cm/s}$ for $\mathcal{E} > 10^4\ \text{V/cm}$.

drift velocity $v_d = ?$ if ~~$I = ?$~~ $V = 100\text{V}$ ~~$I = ?$~~

$I = ?$ if $V = 10^6$ V.

$I = q \cdot A \cdot N_d v_s = 1.6 \times 10^{-19} \times 0.1 \times 10^{17} \times 10^7$

$=$

$\mathcal{E} = \frac{100\text{V}}{2} = 50\ \text{V/cm}$

The velocity is not saturated.

low field $v_d = \mu_n \cdot \mathcal{E}$

If $\mu = 1350\ \text{cm}^2/\text{V-s}$

$v_d = 1350 \times 50 = 67500$ $6.75 \times 10^4\ \text{cm/s}$.

If $V = 10^6$V, $\mathcal{E} = \frac{10^6}{2} = \frac{10,00000}{2} = 500,000$

$= 5 \times 10^5\ \text{V/cm}$

$\therefore I = q A \cdot N_d \cdot v_s$

$= 1.6 \times 10^{-15} \times 0.1 \times 10^{15} \times 10^7$

$= 160\text{A}$

Hall effect Ex:

A sample of Si is doped with 10^{17} phosphorous atoms/cm³. What would you expect to measure for its resistivity? What Hall voltage would you expect in a sample of 100 μm thick if $I_x = 1\text{mA}$ & $B_z = 1\text{kg} = 10^5\ \text{Ws/cm}^2$?

Given $\mu_n = 700\ \text{cm}^2/(\text{V-s})$

Ans. $\sigma = q \mu_n n_0 = (1.6 \times 10^{-19})(700) \cdot (10^{17}) = 11.26\ (\Omega\text{-cm})^{-1}$

since p_0 is negligible. The resistivity is

$\rho = \sigma^{-1} = 0.0889\ \Omega\text{-cm}$.

PTO

The Hall co-efficient is

$$R_H = -(q n_0)^{-1} = \left(1.6 \times 15^{19} \times 10^{17}\right)^{-1}$$

$$= -62.5 \text{ cm}^3/\text{C}$$

we have.

$$\frac{1}{q} R_H = \frac{I_x B_z}{q V_{AB}}$$

$$\Rightarrow V_{AB} = \frac{I_x B_z \cdot R_H}{t}$$

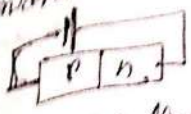
$$= \frac{(10^{-3} A)(10^{-5} \text{Wb/cm}^2)(-62.5 \text{ cm}^3/\text{C})}{10^{-2} \text{cm}}$$

$$= -62.5 \text{ mV}$$

Module-1

3/4/1 Forward - and Reverse - Biased Junctions: Steady state conditions.

current flows freely in the p to n direction when p has +ve external voltage relative to n (forward bias & forward current),



Virtually no current flows when p is made +ve relative to n (reverse bias and reverse current)

- This asymmetry of the current flow makes p-n junction diode very useful as a rectifier.

* Rectification is an application of diode (pn junction)

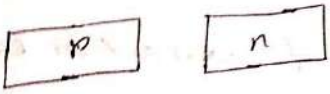
Other applications: Biased p-n junctions can be used as a voltage variable capacitor, photocells, light emitters etc

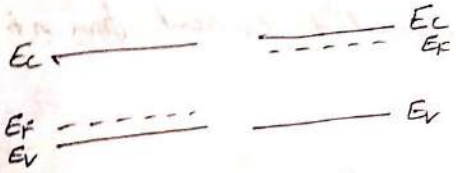
Qualitative Description of Current flow at a Junction:

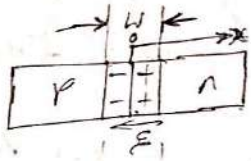
Apparent

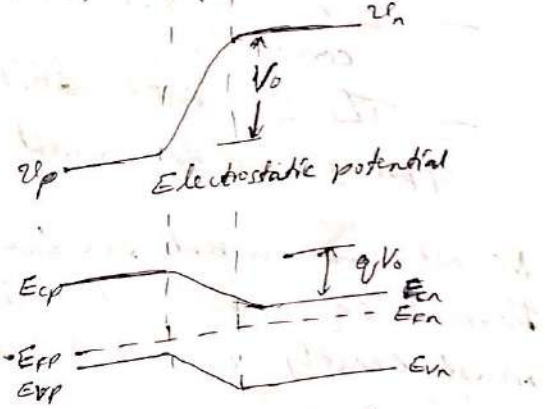
(4b)

Node: (5.2)

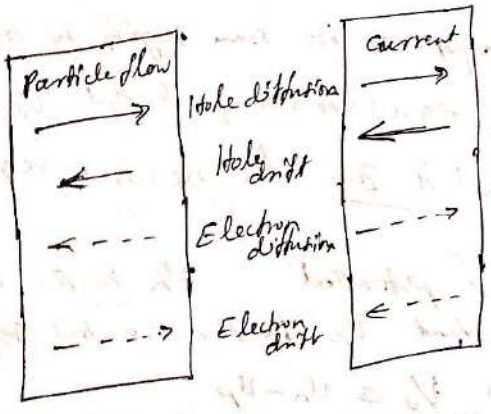








(a) (b)



(c)

Fig:) properties of an equilibrium pn Junction

(a) Isolated neutral regions of p-type & n-type material & energy bands for isolated regions.

(b) Junction showing space charge in the transition region W, the resulting electric field E, & constant potential Vo & separation of energy bands

Before joining:

* The n-material has large concentration of electrons & few holes, conversely true for p-material.
* Upon joining two regions, diffusion of carriers take place, because of large carrier concentration gradient at the junction.
  - Then holes diffuse from the p-side into the n-side and electrons diffuse from n to p.
* Because of opposing electric field (E) created at the junction [Fig.a] due to space charge, the diffusion process stops after some time.
  - Consider electrons diffusing from n to p leave behind uncompensated donor ions (N_d^+) in the n-material and holes leaving the p-region leave behind uncompensated acceptors (N_a^-).
  - This creates +ve space charge near n-side junction and -ve charge near the p-side.
  - Electric field (E) directed from positive charge towards the -ve charge.
  - E is opposite to the direction of current for each carrier
  - The E creates drift component of current from n to p, opposing diffusion current.

- No net current across the junction at thermal equilibrium. The drift current due to E field must exactly cancel the diffusion current.

$$J_{p(drift)} + J_{p(diff)} = 0 \quad \text{(1a)}$$

$$J_{n(drift)} + J_{n(diff)} = 0 \quad \text{(1b)}$$

The electric field E appears in some region W about the junction, & there is an equilibrium potential V_0 across W.

- Assume electric field is Zero in neutral regions outside W.
- i.e. there is a constant potential U_n in the neutral n-material & a constant U_p in the neutral p-region & potential difference $V_0 = U_n - U_p$

∴ The region W is called the transition region & V_0 is called the contact potential. (is called built in potential).

Qualitative Description of Current flow at a Junction:

Assume applied bias voltage V appears across the transition region of the junction, and in ~~neutral~~ neutral n and p regions.

* There will be some voltage drop in neutral region if a current flows through it.

since in neutral region length is small & doping is moderate to heavy. The resistance is small in neutral region. Hence only small voltage drop can be maintained outside space charge region.

Practical purpose: Assume applied voltage appears entirely across transition region.

Let 'V' is +ve when external voltage is +ve on the 'p' side −ve on the n-side

~~Applied vo~~ Applied Voltage: changes electrostatic potential barrier Hence the electric field in the transition region. This changes various components of current, changes in separation of energy bands along with width of the depletion region.

Reward Bias: (V_f)

Electrostatic potential barrier at the junction is lowered by forward Bias (V_f) from the equilibrium potential V_0 to ~~small~~ $V_0 - V_f$.

P.T.O

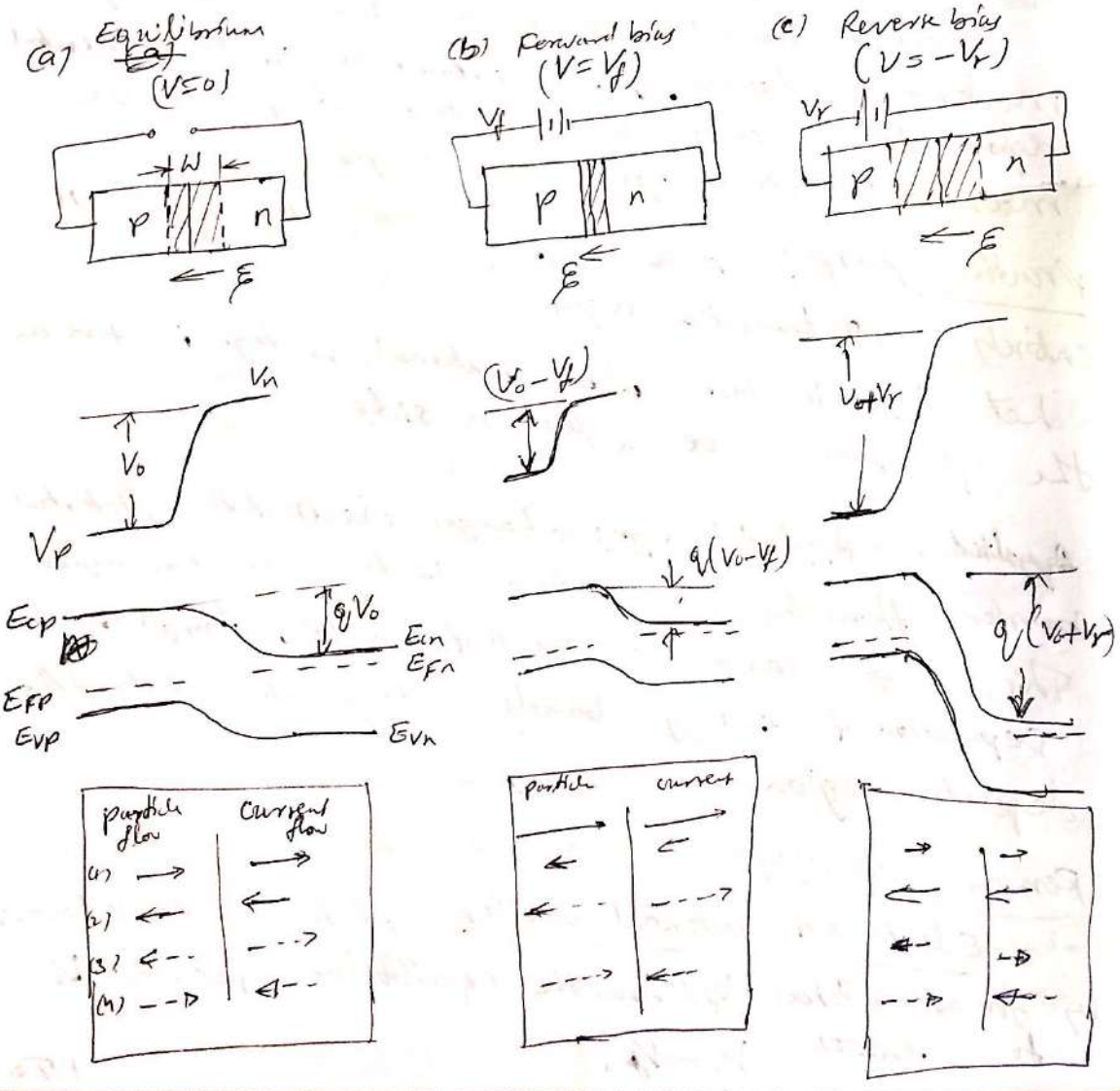
For reverse bias ($V = -V_r$), the opposite occurs. The electrostatic potential is depressed on p-side relative to n-side. The potential barrier at the junction becomes $(V_0 + V_r)$.

The electric field within the transition region can be obtained from the potential barrier.

Note: The field decreases with forward bias since the applied electric field opposes the build-in field.

With reverse bias the field at the junction is increased by the applied field.

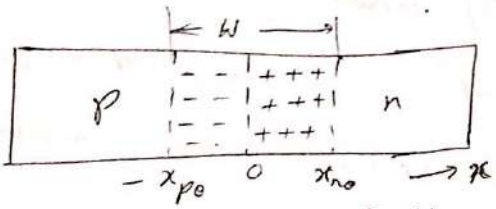
The change in electric field (E) at the junction changes the transition region width 'W'. The 'W' decreases under forward bias (smaller E, & fewer uncompensated charges) & to increase under reverse bias



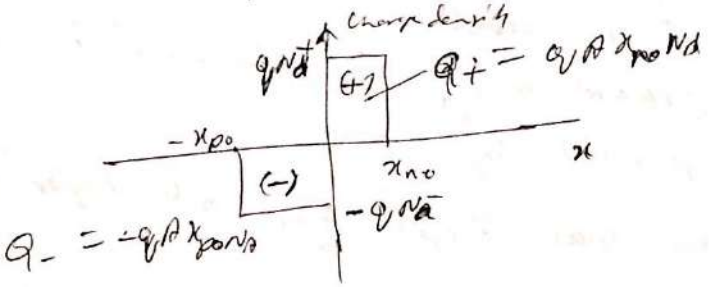
(1) Hole diff, (2) Electron diffusion
(2) Hole drift (2) Electron drift.

Fig (1)

Node



(a)



(b)

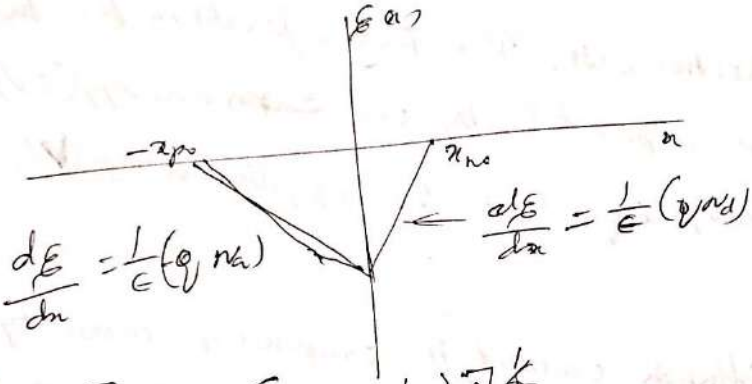


Fig (2)

can be calculated

$$W = \left[\frac{2\epsilon V_0}{q}\left(\frac{1}{N_a} + \frac{1}{N_d}\right)\right]^{1/2} \quad -(1)$$

$$x_{po} = \left\{\frac{2\epsilon V_0}{q}\left[\frac{N_d}{N_a(N_a+N_d)}\right]\right\}^{1/2} \quad -(2)$$

$$x_{no} = \left\{\frac{2\epsilon V_0}{q}\left[\frac{N_a}{N_d(N_a+N_d)}\right]\right\}^{1/2} \quad -(3)$$

Where V_0 is replaced by the new barrier height $V_0 - V$.

— The separation of the energy bands is a direct function of the electrostatic potential barrier at the junction

* The height of the electron energy barrier is simply the electronic charge (q) times the height of the electrostatic barrier.

- The bands are separated less $[q(V_0 - V_F)]$ under forward bias than at equilibrium,

  ↳ more $[q(V_0 + V_F)]$ under reverse bias.

* We assume Fermi level is at equilibrium value.

* Shifting of energy bands under bias implies a separation of the Fermi levels

5.2.2

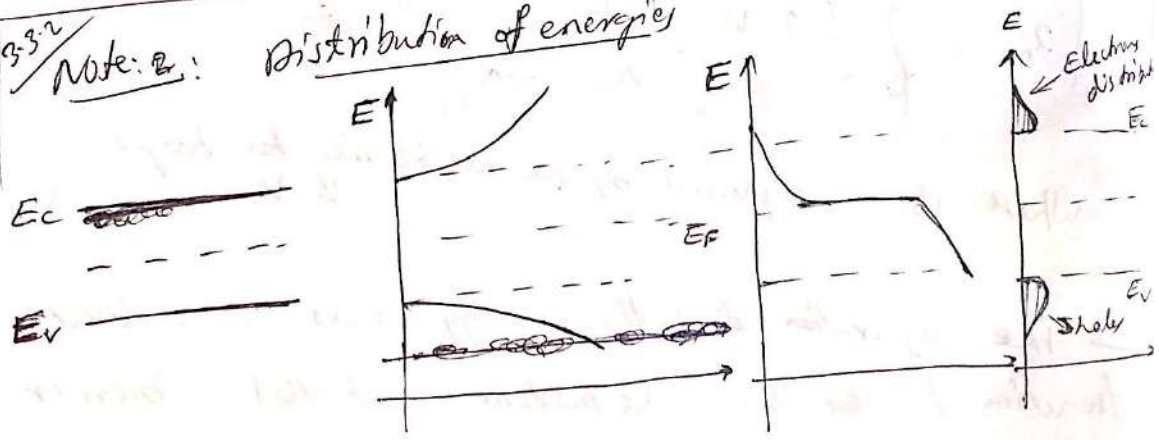
- Under forward bias (V_F) E_{Fn} ~~(for n-side)~~ is above E_{Fp} (p-side) by qV_F
- For reverse bias E_{Fp} is qV_F joules higher than E_{Fn}.

* In ~~eV~~ electron volts, The Fermi levels in the two neutral regions are separated by an ~~energy~~ energy (eV) numerically equal to the applied voltage (V).

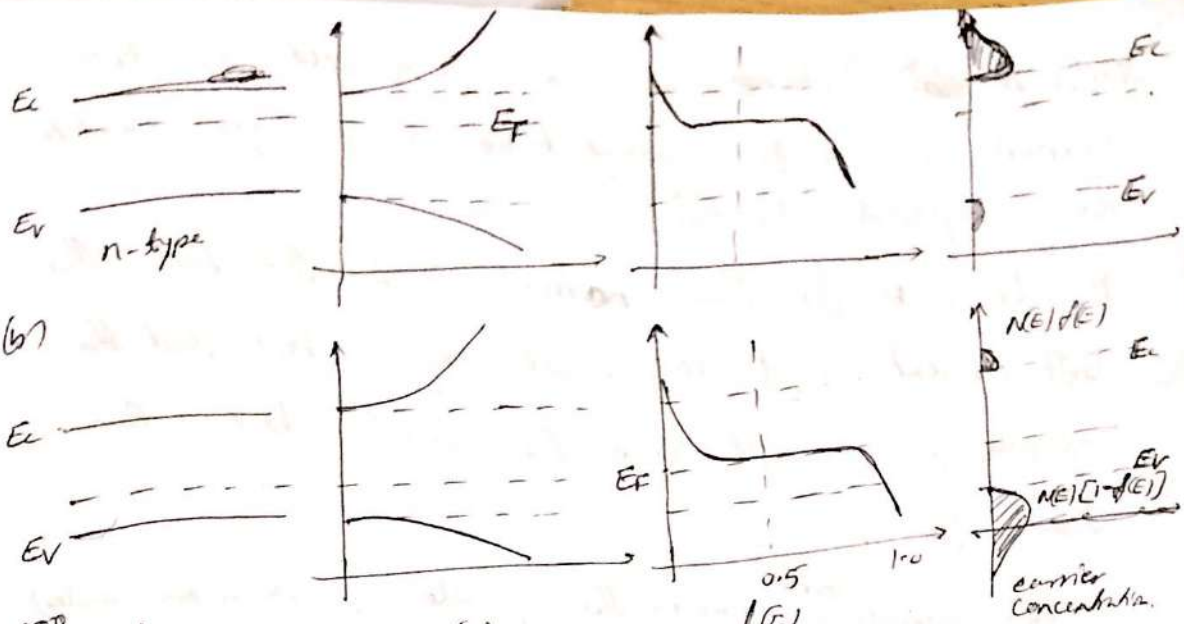
Currents:-

Diffusion The diffusion current is composed of majority carriers electrons on the n-side, surmounting the potential energy barrier to diffuse to p.-side. & holes surmounting the barrier from p to n.

5.3.2 Note: 2: Distribution of energies



(a) Intrinsic



(a) Schematic band diagram, (b) density of states, (c) Fermi-Dirac distribution, (d) carrier concentration at thermal equilibrium

Fig (3)

In fig (c), some electrons in the high-energy "tail" of the distribution have enough energy to diffuse from n to p at equilibrium in spite of ~~barrier~~ barrier.

* With forward bias, the barrier is lowered to $(V_0 - V_f)$ & many more electrons have sufficient energy to diffuse from n to p.

* Electron diffusion current is large with forward bias.

* Illy more holes can diffuse from p to n under forward bias.

— For reverse bias, the barrier becomes so large $(V_0 + V_r)$ that virtually no electrons in the n-side conduction band or holes in the p-side valence band have energy to surmount it.

∴ diffusion current is negligible in reverse bias.

Drift current:

Is insensitive to the height of the potential barrier.

This is strange, because we expect ample carriers & drift current to be simply proportional to the applied field E.

* The reason for this anomaly is the fact that the drift current is limited not by how fast the carriers are swept down the barrier, but rather how often.

Eg: Minority carrier electrons on the p-side which move (wander) into the transition region will be swept down the barrier by the E field, giving rise to electron component of drift current. This current is small not due to size of barrier, but due to very few minority carriers in the p-side to participate.

- Every electron on the p-side which diffuses to the transition region will be swept down the potential energy hill. (Weather hill is small or large).

— i.e drift current does not depend on how fast an individual electron is swept from p to n but rather how many electrons are swept down the barrier/sec
Ditto for holes from n-side to p-side

* The good approximation that the electron & hole drift currents at the junction are independent of applied voltage.

The minority carriers on each side of the junction which participate in the drift component is generated by thermal excitation of electron-hole pairs.

Eg: EHP created near the junction of the p-side provi minority electron in the p-material.

* If EHP is generated within diffusion length L_n of the transition region, this electron can diffuse to the junction & be swept away to n-side.

This drift current is called the generation current, since its magnitude depends on EHP's.

* EHP generation can be increased greatly by optical excitation of EHP's near Tn. [Gn: photocurrent]

— The total current crossing the junction is the sum of diffusion & drift components.

From Fig(1), the electron & hole diffusion currents are both directed from p to n & drift currents are from n to p.

The net current crossing the junction is zero at equilibrium, since drift & diffusion components cancel each other.

Reverse Bias

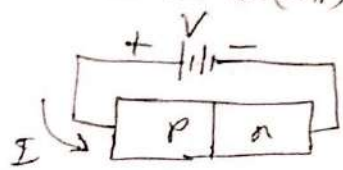
Under reverse bias, both diffusion components are negligible because of large barrier at the junction. The only current is very small (& is voltage independent) (generation current from n to p) and is shown in Fig. (2) for a pn junction. It is also I-V plot.

At +ve current (I) is from p to n & V is positive, when +ve terminal is connected to p & -ve terminal to n.

The only current flowing in this pn junction for -V is the small current I(gen), due negative V to carriers generated in the transition region or minority carriers diffused into the junction.

The current at $V=0$ (equilibrium) is zero since generation & diffusion currents cancel

$$I = I(diff) - |I(gen)| = 0 \text{ for } V=0 \quad \text{—— (8)}$$



$$I = |I(gen)| \left(e^{qV/KT} - 1\right)$$

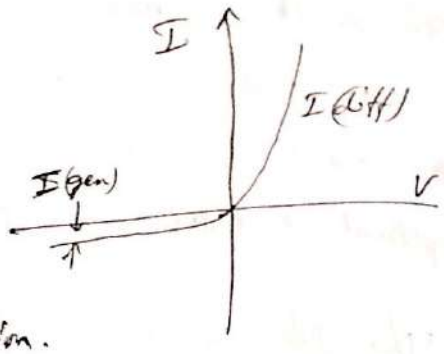


Fig.(4) I-V c/c's of a p-n junction.

~~Note~~ An applied voltage $V = V_f$ increases the probability that a carrier can diffuse across the junction by $e^{(qV_f/KT)}$

∴ The diffusion current under forward bias is given by its equilibrium value multiplied by $e^{(qV/KT)}$ [i.e. $I(gen) \cdot e^{(qV/KT)}$]

Similarly for reverse bias, the diffusion current is the equilibrium value reduced by same factor, with $V = -V_r$.

∴ The equilibrium diff current equal to $|I(gen)| \left(e^{qV/KT}\right)$. since

∴ The total current I is the diff current minus the absolute of generation current (now call it as I_0)

$$\therefore I = I_0 \left(e^{qV/KT} - 1\right) \quad \text{—— (3)}$$

In eq(3), V can be +ve (V_f) or negative ($-V_r$)

When 'V' is +ve & greater than few $\frac{KT}{q}$ (Thermal equivalent Volts) ($\frac{KT}{q} = 0.0259$ V @ at room temp), the exponential term is much greater than unity.

With forward bias current increases exponentially.

When V is -ve, exponential term approaches Zero & $I = -I_0$ is called reverse Sat current.

For pn junction, current flows freely in forward bias of diode, & no current flows in the reverse direction.

Reverse Bias

Note: Relating the electric field to the constant potential V_0, since the E field at any 'x' is the negative of the potential gradient at that point

$$E(x) = -\frac{dV(x)}{dx} \quad \text{or} \quad -V_0 = \int_{-x_{po}}^{x_{no}} E(x)\,dx \quad \text{---} (1)$$

If $V = -V_r$ (p is negative biased w.r.to n), we can approximate

$$\Delta p_n = p(x_{no}) - p_n = p_n\left(e^{qV/kT} - 1\right) \text{ by}$$

$$\Delta p_n = p_n\left(e^{q(-V_r)/kT} - 1\right) \simeq -p_n \text{ for } V_r \gg \frac{kT}{q} \quad \text{---} (1)$$

III ly $\Delta n_p \simeq -n_p$

— For a reverse bias of more than a few tenths of a volt, the minority carrier concentration at each edge of transition region becomes Zero.

& As excess concentration approaches -ve of equilibrium conc.

— The excess minority carrier concentration in the neutral regions are

$$\delta_n(x_p) = n_p\left(e^{qV/kT} - 1\right) e^{-x_p/L_n}$$

$$\delta_p(x_n) = p_n\left(e^{qV/kT} - 1\right) e^{-x_n/L_p}$$

Where $L_n \rightarrow$ diffusion length of electrons

$L_p \rightarrow$ —"— of holes

So that depletion of carriers below the equilibrium values extends approximately a diffusion length beyond each side of transition region.

The reverse-bias depletion of minority carriers can be thought of as minority carrier extraction.

— Extraction occurs because minority carriers at the junction of the depletion region are swept down the barrier at the junction to the other side so are not replaced by an opposing diffusion of carriers.

Ex: Holes at x_n are swept across the junction to the p-side by the E field, a gradient in the hole distribution in the n-material exists, & holes in n-region diffuse toward the junction.

* Hole distribution in the n-region has the exponential shape Fig (5a)

* Note: Although the reverse saturation current occurs at the junction by drift of carriers, this current is fed from each side by diffusion towards the junction of minority carriers in the neutral regions.

The rate of carrier drift across the junction (the reverse sat. current) depends on the rate at which holes arrive at x_n (electrons at x_p) by diffusion from the neutral material.

* These minority carriers are supplied by thermal generation.

We can show that

$$p_n = n_i^2 e^{(qV/kT)} e^{(F_n - F_p)/kT} = n_i^2 e^{qV/kT}$$

represents the rate at which carriers are generated thermally within a diffusion length of each side of the transition region

In reverse bias, Fermi levels split in the opposite sense than in forward bias [Fig (5b)] Fn moves further away from Ec (close to Ev) & Fp moves further away from Ev, reflecting the fact that there are fewer carriers

than in equilibrium, (In forward bias we have an excess carriers)

In reverse bias, in depl region, we have

$$p_n = n_i^2 \, e^{(F_n - F_p)/kT} \simeq 0$$

Note that ~~Fermi~~ the quasi-Fermi levels in the reverse bias can go inside bands.

Eg- F_p goes inside the conduction band on the n-side of the depletion region. However ~~we must remember~~ F_p is a measure of hole concentration, & should be correlated with the Valence band edge E_v not with E_c.

i.e. the band diagram simply reflects that we have few holes in this region, even fewer than the already small equilibrium minority carrier hole concentration (Fig 5a)

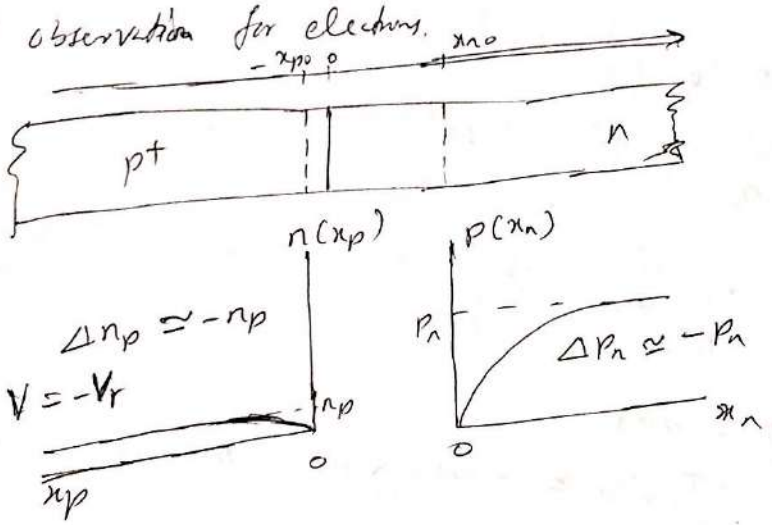


Fig (5a)

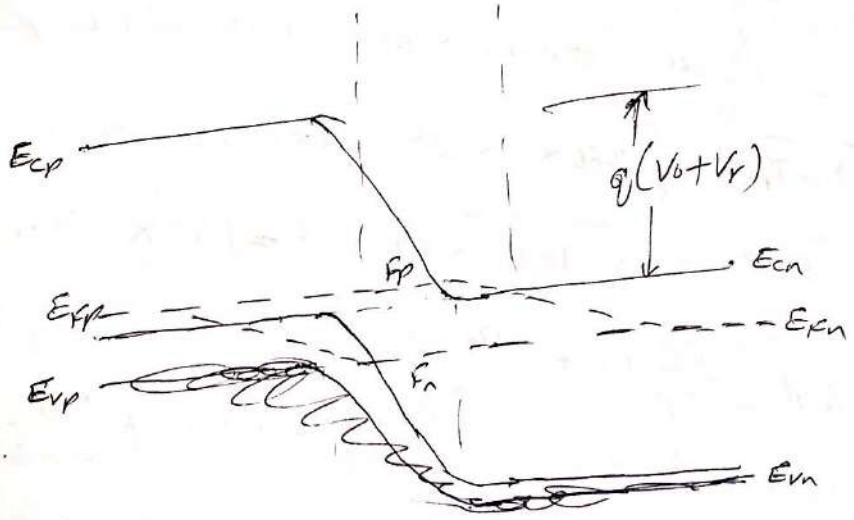


Fig.(5b)
Reverse bias pn-junction ① minority carrier distribution near the reverse-biased junction ② Variation of quasi Fermi levels.

Example problems

Ex: An abrupt Si p-n junction ($A = 10^{-4}\ cm^2$) has the following properties at 300k.

p-side	n-side
$N_A = 10^{17}/cm^3$	$N_D = 10^{15}/cm^3$
$\tau_n = 0.1\ \mu s$	$\tau_p = 1\ \mu s$
~~$\mu_p = 200\ cm^2/V\cdot s$~~	~~$\mu_n = 360\ cm^2/V\cdot s$~~
$\mu_n = 700\ cm^2/Vs$	$\mu_p = 450\ cm^2/Vs$

$T = 300k$
$K = 1.38 \times 10^{-23}\ J/k$
$q = 1.60 \times 10^{-19}\ C$
$\frac{KT}{q} = 0.025\ V$
or $KT = 0.025\ eV$

The junction is forward biased by 0.5V, what is the forward current? What is the current at a reverse bias of $-0.5V$?

Solution:

$$I = qA\left(\frac{D_p}{L_p}p_n + \frac{D_n}{L_n}n_p\right)\left(e^{qV/KT} - 1\right)$$

$$= I_0\left(e^{qV/KT} - 1\right)$$

$$p_n = \frac{n_i^2}{n_n} = \frac{(1.5 \times 10^{10})^2}{10^{15}} = 2.25 \times 10^5/cm^3$$

$$n_p = \frac{n_i^2}{p_p} = \frac{(1.5 \times 10^{10})^2}{10^{17}} = 2.25 \times 10^3/cm^3$$

For minority carriers

$$D_p = \frac{KT}{q}\mu_p = 0.0259 \times 450 = 11.66\ cm^2/s \text{ on n-side}$$

$$D_n = \frac{KT}{q}\mu_n = 0.0259 \times 700 = 18.13\ cm^2/s \text{ on p-side}$$

$$L_p = \sqrt{D_p \tau_p} = \sqrt{11.66 \times 10 \times 10^{-6}} = 1.08 \times 10^{-2}\ cm$$

$$L_n = \sqrt{D_n \tau_n} = \sqrt{18.13 \times 0.1 \times 10^{-6}} = 1.35 \times 10^{-3}\ cm$$

$$I_0 = qA\left(\frac{D_p}{L_p}p_n + \frac{D_n}{L_n}n_p\right)$$

$$= 1.6 \times 10^{-19} \times 0.0001\left(\frac{11.66}{0.0108} \times 2.25 \times 10^5 + \frac{18.13}{0.00135} \times 2.25 \times 10^3\right)$$

$$= 4.320 \times 10^{-15}\ A.$$

$I = I_0 \left(e^{0.5/0.025} - 1\right) \simeq 1.058 \times 10^{-6}$ A in forward bias

$I = -I_0 = -4.37 \times 10^{-17}$ A in reverse bias

Prob: A Si pn junction with cross sectional area $A = 0.001\ cm^2$ is formed with $N_a = 10^{15}/cm^3$ and $N_d = 10^{20}/cm^3$. Calculate:

(a) contact potential, V_0,

(b) space charge width at equilibrium (zero bias)

(c) current with a forward bias of 0.7V. Assume that the current is diffusion dominated. Assume $\mu_n = 1500\ cm^2/V\text{-}s$, $\mu_p = 450\ cm^2/V\text{-}s$, and $\tau_n = \tau_p = 2.5$ ns. which carries most of the current, electrons or holes, & why? If you wanted to double the electron current, what should you do?

Ref. Text-1 p. 163, p. 201

Assume $n_i = 1.5 \times 10^{10}$.

$$V_0 = \frac{kT}{q} \ln\left(\frac{N_a N_d}{n_i^2}\right) = \frac{1.38 \times 10^{-23} \times 300}{1.6 \times 10^{-19}} \ln\left(\frac{10^{17} \times 10^{20}}{(1.5 \times 10^{10})^2}\right)$$

$$= 0.025875 \ln(4.44 \times 10^{14}) = 0.87\ V$$

$$W = \left[\frac{2\epsilon V_0}{q}\left(\frac{N_a + N_d}{N_a N_d}\right)\right]^{1/2} \quad \text{(equilibrium)}$$

$$W = \left[\frac{2\epsilon (V_0 - V)}{q}\left(\frac{N_a + N_d}{N_a N_d}\right)\right]^{1/2} \quad \text{(with bias)}$$

$\epsilon_0 = 8.85 \times 10^{-12}$ F/m

$= 8.85 \times 10^{-14}$ F/cm

$\epsilon_r \simeq 11.9$ (Si)

$\epsilon_r (SiO_2) \simeq 4$

$$= \left[\frac{2 \times 8.85 \times 10^{-12} \times 11.9 \times 0.87}{1.6 \times 10^{-19}}\left(\frac{10^{17} + 10^{20}}{10^{17} \times 10^{20}}\right)\right]^{1/2}$$

$$= \left[\frac{2 \times 8.85 \times 10^{-14} \times 11.9 \times 0.87}{1.6 \times 10^{-19}}\left(\frac{1.001 \times 10^{17}}{10^{17} \times 10^{20}}\right)\right]^{1/2} = 11,479,335 \cdot (1.0001 \times 10^{-17})$$

$$= \left[\frac{11479335}{1.6 \times 10^{-19}} \times 1\right] = 1.0719 \times 10^{-4}\ cm$$

$$I_0 = qA\left(\frac{D_p}{L_p} p_n + \frac{D_n}{L_n} n_p\right)$$

$$p_n = \frac{n_i^2}{n_a} = \frac{(1.5 \times 10^{10})^2}{10^{20}} = \frac{2.25 \times 10^{20}}{10^{20}} = 2.25\ /cm^3$$

$$n_p = \frac{n_i^2}{p_p} = \frac{(1.5 \times 10^{10})^2}{10^{17}} = \frac{2.25 \times 10^{20}}{10^{17}} = 2.25 \times 10^{3}/cm^3$$

$$D_p = \frac{kT}{q} \cdot \mu_p = \frac{1.38 \times 10^{-23} \times 300}{1.6 \times 10^{-19}} \times 450 = 11.64\ cm^2/s$$

$$D_n = \frac{kT}{q}\mu_n = \frac{1.38\times10^{-23}\times300}{1.6\times10^{-19}}\times1500 = 38.8\ cm^2/s$$

$$L_p = \sqrt{D_p\tau_p} = \sqrt{11.64\times2.1\times10^{-3}} = 0.166\ cm$$

$$L_n = \sqrt{D_n\tau_n} = \sqrt{38.8\times2.1\times10^{-3}} = 9.7\times10^{-2}\ cm$$

$$I_0 = qA\left(\frac{D_p}{L_p}p_n + \frac{D_n}{L_n}n_p\right)$$

$$= 1.6\times10^{-19}\times0.001\left(\frac{11.64}{0.166}\times2.1\times10^{3} + \frac{38.8}{9.7\times10^{-2}}\cdot2.1\times10^{5}\right)$$

$$= 1.6\times10^{-22}\left(157.771 + 9\times10^{7}\right)$$

$$= 1.44\times10^{-14}\ A$$

$$I = I_0\left(e^{V/V_T} - 1\right)$$

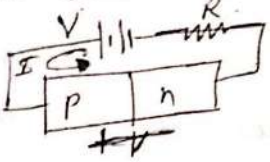
$$= 1.44\times10^{-14}\left(e^{0.7/0.259} - 1\right)$$

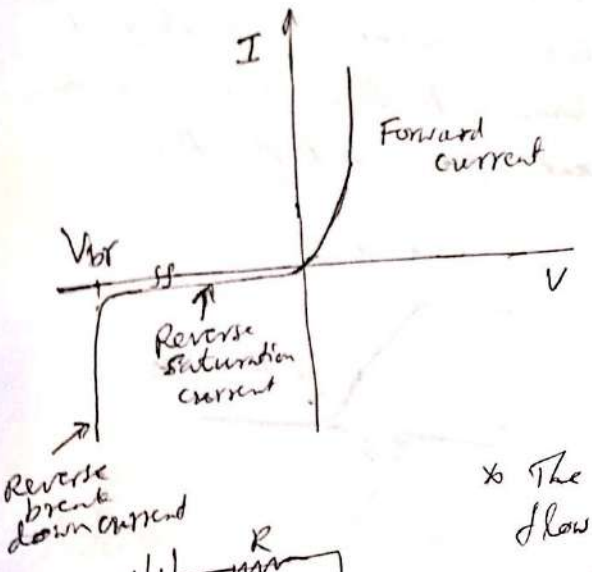
$$= 1.44\times10^{-14}\left(e^{27.02} - 1\right) = 7.8716\times10^{-3}$$

Reverse-Bias Breakdown:

We have seen that, in reverse biased p-n junction there is voltage independent reverse saturation current. This is true until critical reverse bias (V_{BR}) at which reverse breakdown occurs, & reverse current increases sharply. Even a small increase in voltage, causes large current.

If the current is limited to a reasonable value by the external circuit, pn junction can be operated safely. in reverse breakdown, like forward bias condition.





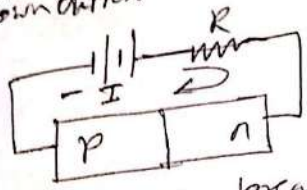


Fig (1) Reverse break down in a pn junction

* The maximum reverse current which can flow in the device of fig(1). is $(V-V_{BR})/R$,
  - R can be chosen to limit the current to a safe level.
  - If current is not limited externally the junction can be damaged by excessive reverse current.

  Due to excessive power dissipation like situation can occur in forward bias also.

* Useful devices called breakdown diodes are designed to operate in the reverse breakdown region of the ckt's.
* Reverse breakdown can occur by two mechanisms.
  (i) Zener effect is operative at low voltages (upto a few volts reverse bias)
  (ii) Avalanche breakdown (breakdown occurs at few volts to thousands of volts).

Zener Breakdown:

For heavily doped reverse biased junctions, the energy bands become crossed at even at low voltages, as Fig (1b). (The n-side conduction band appears opposite to p-side Valence band). The crossing of bands aligns the larger number of empty states in the n-side conduction band opposite to the filled states of the p-side Valence band.

* If the barrier separating these two bands is narrow, tunneling of electrons can occur. Tunneling of electrons from p-side Valence band to the n-side conduction band causes reverse current. from n to p, this is called Zener effect.

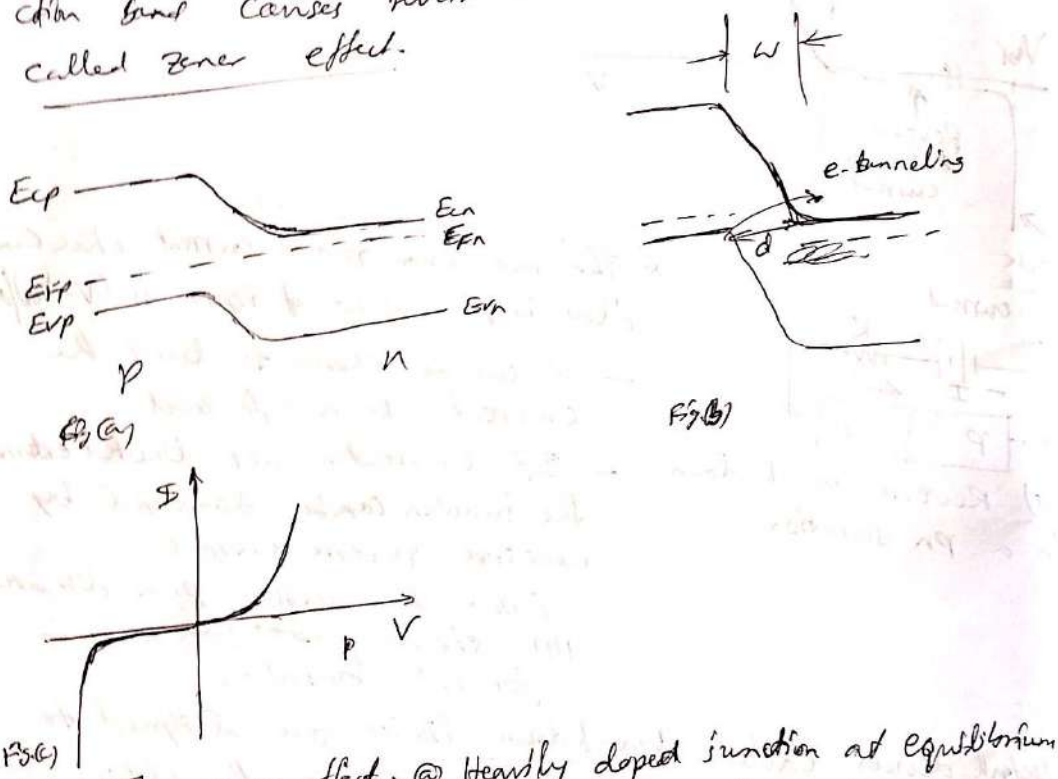


Fig. (1) The Zener effect. @ Heavily doped junction at equilibrium (b) reverse bias with electron tunneling from p to n. © I-V chic's

Requirement for tunneling:

(1) Large number of electrons separated from a large number of empty states by a narrow barrier of finite height.

(2) Metallurgical junction be sharp & doping is high. (Since tunneling probability depends on width of barrier (d in Fig (b)

* So that transition region W extends only short distance from each side.
* If Junction is not ~~sharp~~ abrupt, or if either side of doping at Junction is lightly doped, the 'W' is too wide for tunneling
— For small reverse bias, tunneling distance 'd' may be too large, for appreciable tunneling.
* However 'd' becomes smaller as reverse bias is increased (since higher electric field results in steeper slope for the band edges)
— For heavily doping on each side & for low reverse bias, 'W' does not increase appreciably is a good assumption

Avalanche Breakdown

. For lightly doped junctions electron tunnelling is ~~negl~~ negligible. However breakdown mechanism involves the impact ionization of host atoms by energetic carriers.
- Normal lattice scattering events can result EHP's if the carrier being scattered has sufficient energy.

Eg. If electric field $\mathcal{E}$ in the transition is large, an electron entering from the p-side may cause an ionization collision with the lattice. This results in carrier multiplication, the original electron & the generated electron are ~~swept~~ swept into n-side. ~~of the~~ Generated hole is swept to p-side (Fig. 2b)
* ~~Multiplication~~ multiplication can become very high if carriers generated within transition region also have ionization collision with the lattice.

& create EHP, each of these carriers may create a new EHP, each of these can also create an EHP, & so forth (Fig (2c)) This is avalanche process, i.e. an incoming carrier can initiate the creation of a large no new carriers.

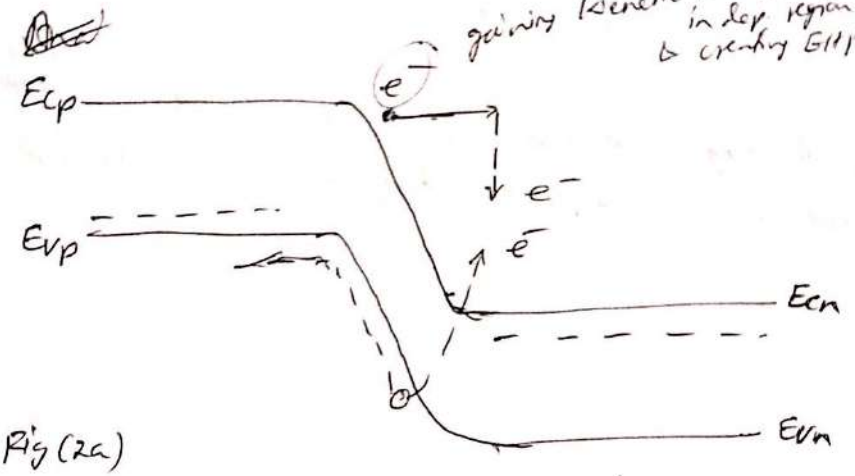
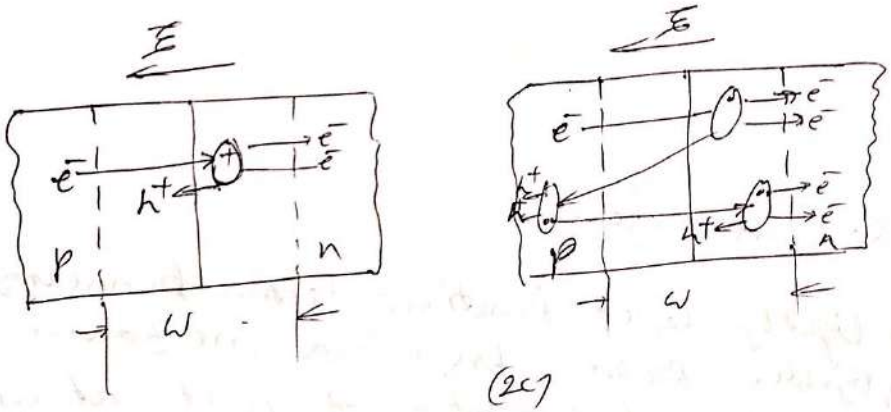


Fig (2a)



(2b) (2c)

Fig. (2) EHP created by impact ionisation (a) Band diagram of p-n$^+$ in reverse bias showing electron gaining kinetic energy in dep. region & creating secondary EHP.

(b) a single ionizing collision by an incoming electron in the dep region. (c) primary, secondary & tertiary collisions.

Analysis.

Assume that a carrier of either type has a probability 'p' of having ionization collision with lattice, while it is ~~accelerated~~ transmitted through transition region 'W'.

∴ For N$_{in}$ electrons entering from p-side there will be pN$_{in}$ ionizing collisions & an EHP (secondary)

(secondary carriers) from each collision.

- After the n_{in} collisions by primary electrons, we have $n_{in}(1+P)$ ~~secondary~~ primary plus secondary electrons $n_{in}(1+P)$. After a collision each EHP moves effectively a distance $\frac{W}{2}$ within transition region.

Eg: If an EHP is created at the centre of the region, the electron drifts a distance $W/2$ to n and the hole $W/2$ to P.

- Probability that an ionizing collision will occur due to secondary pairs will be $(n_{in}P)\cdot P$ ionizing collisions.

& hence $n_{in}P^2$ tertiary pairs

∴ The total no. of electrons out of region at 'n' after many collisions is (assuming no-recombination)

$$n_{out} = n_{in}(1 + P + P^2 + P^3 + \cdots) \quad \text{—①}$$

Including recombination, & assuming different probabilities for e & holes

From simple theory electron multiplication

$$M_n = \frac{n_{out}}{n_{in}} = 1 + P + P^2 + P^3 + \cdots = \frac{1}{1-P} \quad \text{—②}$$

as P approaches unity, (1) the carrier multiplication increases without limit. The limit on the current is decided by the external ckt.

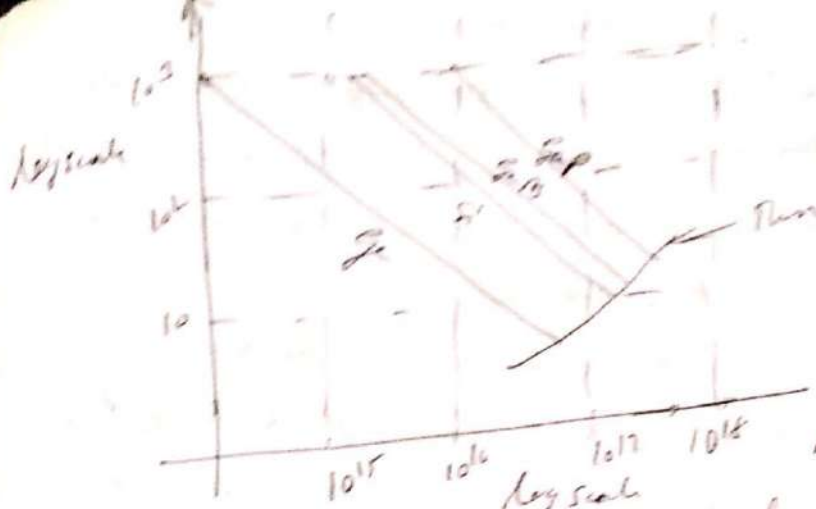
Relation between P & parameters of junction

In general ionization probability increases with increasing electric field, & hence depend on reverse bias. Measurement of carrier multiplication M in junctions near breakdown tends to empirical relation

$$M = \frac{1}{1-(V/V_{BR})^n} \quad \text{——③}$$

Where n → varies from 3 to 6 depending on the type of material used for the junction.

* V_{BR} decreases as doping increases
* The critical reverse voltage for breakdown increases in increase in band gap of the material Fig. (8)



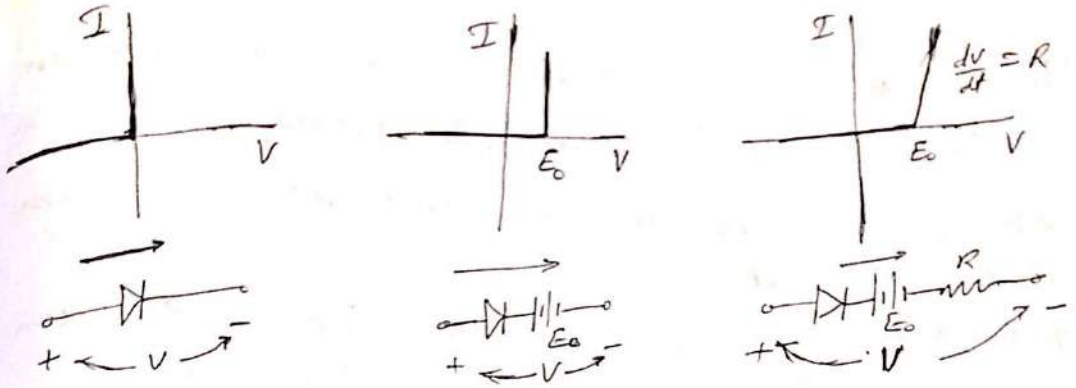
Fig(3) Variation of avalanche breakdown voltage with doping & energy gap

Rectifiers:

The property of a pn junction is its unilateral nature. i.e. it conducts current in one direction & blocks current in other direction.

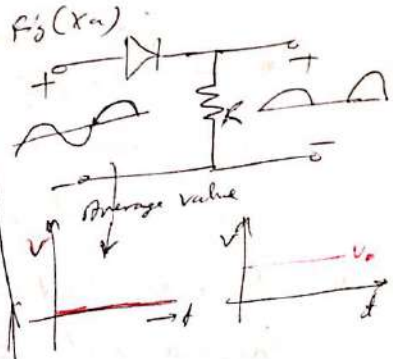
We can think of an ideal diode has a short circuit when forward biased and as an open circuit when reverse biased. (Fig.-1a).

However most forward-biased diodes exhibit an offset voltage E_0, & can be approximated in a circuit model by a battery in series with ideal diode (Fig-1b). This keeps the ideal diode turned off for applied voltages less than E_0. The $E_0 \approx$ contact potential of the junction. This approximation can be improved by adding series resistance R to the circuit (Fig-1c). The equivalent circuit approximations illustrated in Fig(1) are called piecewise-linear equivalents.



Ideal diode connected in series with ac signal provides rectification signal, since current can flow only in forward direction. Only +ve half cycle of the input sine wave is passed.
i.e. o/p is half-rectified sine wave.
The i/p sinusoid has zero average value, the o/p has positive average value, contains dc component.
* By appropriate filtering, this dc level can be extracted from filtering.

Fig(x)



Fig(x5).

App: Unilateral nature of diode is ~~useful~~ used in effective clipping. This involves alteration of ac-signals by passing only certain portions of signal.

Diodes designed to use as rectifiers should have applies I-V chc's. The reverse current should be negligible, & forward current (E_f) should have little voltage dependence (i.e. negligible forward ~~R~~ resistance R).
The reverse breakdown voltage should be large.
offset voltage E_0 in forward direction should be small.
However meeting all meeting all the above requirement in a single device is difficult, hence

PTO

The various requirements for good rectifier functions are

(i) Band gap of the material for rectifier diodes.

* As E_g ↑, n_i decreases hence decrease in reverse saturation current.

* Rectifier made of wide band gap material can be operated at higher temperatures, since thermal excitation of EHP's is reduced by ↑ E_g.

~~(Ei impt in rectifier diodes)~~

* – Temp effects are impt in rectifiers, since they carry large current in forward direction, leads to large heating.

(ii) However contact potential & offset voltage E_0 generally increase with E_g (This draw back outweighted by the low n_i)

Eg:- Si is preferred over Ge for power rectifiers because of wider band gap, lower leakage current, & higher breakdown voltage, & more convenient fabrication properties

(iii) Doping concentration on each side, ~~of the junction~~ influences avalanche breakdown voltage, the contact potential & series resistance of diode.

* If the junction has one heavily doped side (Eg p^+n or n^+p), the lightly doped region determines many of the properties

– High-resistivity region should be used for at least one side to increase break down V_{BR}.

– This increases forward resistance R_f & contributes to thermal effects due to I^2R heating.

– To reduce R, increase area & reduce L of lightly doped side.

(vi) Physical dimensions of the diode is another important design variable.

* Localized flaws in junction uniformity can cause premature reverse breakdown.

* Lightly doped region cannot be arbitrarily made small, which may cause Punch Through: since transition region width 'W' increases with reverse bias & extends into lightly doped side mode.

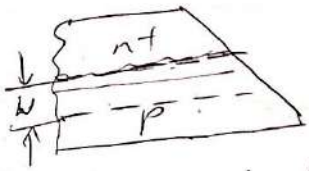
Generally Punch through is a breakdown occurs below V_{BR}.

- Devices designed to use at high reverse bias

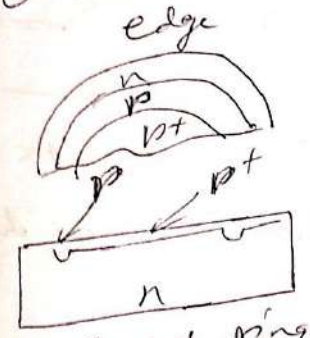
Care must be taken to avoid premature breakdown across edges.

* This effect can be reduced by beveling the edges or by diffusing guard ring to isolate the junction from the edges of the sample.

(a) Diode with beveled edge



(b) Close up view of edge, showing reduction of depletion region near bevel.



(c) Guard ring.

* The electric field is lower at the beveled edge of sample as in Fig (b).

* Hence the junction at the lightly doped p- guard ring (Fig (c)) breaks down at higher voltages than p+ to n.
(The depletion region is wider in the p- ring than in p+ region & average electric field is smaller in ring)

Fabrication.

In fabricating p^+-n or p-n^+ junction, it is common to terminate, lightly doped side is terminated with heavily doped layer of the same type (P^+ (or) n^+) to make ohmic contact to the device.

This results in either p^+-n-n^+ or p^+-p-n^+ device.

* The lightly doped region determines the avalanche breakdown voltage.
* If lightly doped region is smaller than diffusion length of minority carriers, the excess carrier injection for large forward currents can increase conductivity of the region.

This current modulation, reduces forward resistance, it is useful for high-current devices.

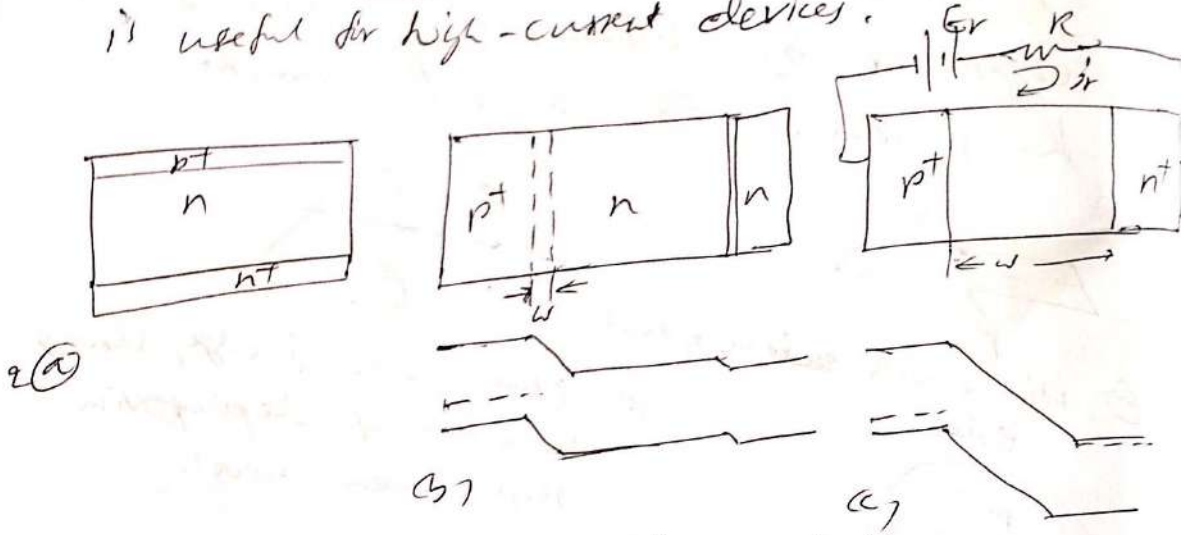


Fig (2) A p^+-n n^+ Jn diode (a) device configuration (b) Zero bias condition (c) Reverse bias to punch thru.

— Mounting of device is also important

* For low power circuits glass, or plastic encapsulation or a simple header mounting is adequate.
* For high current devices requires special mounting to transfer thermal heat away from Jn.

Optoelectronic Device

These devices provide optical sources & detectors, & provide ~~the optic~~ broadband telecommunication & data transmission over optical fibers. This area of application is called optoelectronics. These are devices that detect photons & also emit photons.

* These are devices convert optical energy into electrical energy (Eg. photodiodes, & solar cells).

photon emitters < Light emitting diodes (LED's) → incoherent source
Lasers → coherent sources.

Photodiodes:

Devices designed to absorb photon energy are called photodiodes. Some devices photodiodes have extremely high sensitivity & response speed. Modern electronics involves optical as well as electrical signals. ~~Photodiodes~~ photodiodes serve important functions as electronic devices.

Let us study the response of p-n junctions to optical generation of electron-hole pairs (EHP's) and discuss typical photodiode detector structures.

Current and Voltage in an illuminated Junction:

We know that drift of minority carriers ~~acro~~ across the junction is called as generation current. The carriers generated with in the ~~depletion~~ depletion region W are separated by the junction field, electrons are collected in the n-region and holes in the p-region. Also carriers generated ~~with~~ thermally within the diffusion length of each side of the junction diffuse to depletion region & are swept to the other side by electric field.

P.T.O

If the junction is illuminated by photons with $h\nu > E_g$, an added generation rate g_{op} (EHP/cm³-s) participate in the current. (Fig. (1))

- No of holes created/s within a diffusion length of the transition region on the n-side is

$A L_p g_{op}$

illuminated $A L_n g_{op}$ electrons generated/s within L_n of x_{po}

and $AW g_{op}$ carriers are generated within W.

The resulting current due to collection of these optically generated carriers in the junction is

$$I_{op} = q A g_{op} [L_p + L_n + W] \quad \text{——①}$$

Call thermally generated current (I_o) as I_{th} & adding eq ① (1) to get total reverse saturation current with illumination.

Since current is directed from n to p, the diode current

$$I = I_{th} \left(e^{qV/KT} - 1 \right) - I_{op}$$

$$= \underbrace{q A \left(\frac{L_p}{\tau_p} p_n + \frac{L_n}{\tau_n} n_p \right) \left(e^{qV/KT} - 1 \right)}_{\text{Term 1}} - \underbrace{q A g_{op} (L_p + L_n + W)}_{\text{Term 2.}} \quad \text{——②}$$

The I-V curves are shown in Fig. (1c) depending on generation rate.

There are two parts in eq ② (i) current described by usual diode eq & current due to optical generation.

If device is short circuited ($V = 0$), the first term of eq ② is Zero. However there is short-circuit current from n to p = I_{op}.

The I-V characteristics cross I-axis at +ve values proportional to g_{op}.

If there is an open circuit across device, $I = 0$, & $V = V_{oc}$ is

$$V_{oc} = \frac{kT}{q} \ln\left[I_{op}/I_{th} + 1\right]$$

$$= \frac{kT}{q} \ln\left[\frac{L_p + L_n + W}{(L_p/\tau_p)p_n + (L_n/\tau_n)n_p} g_{op} + 1\right] \quad \text{— (3a)}$$

For symmetrical junction, $p_n = n_p$, & $\tau_p = \tau_n$ eq (3a) can be written in terms of thermal generation rate $\frac{p_n}{\tau_n} = g_{th}$ & optical generation rate g_{op}. i.e neglecting generation within W:

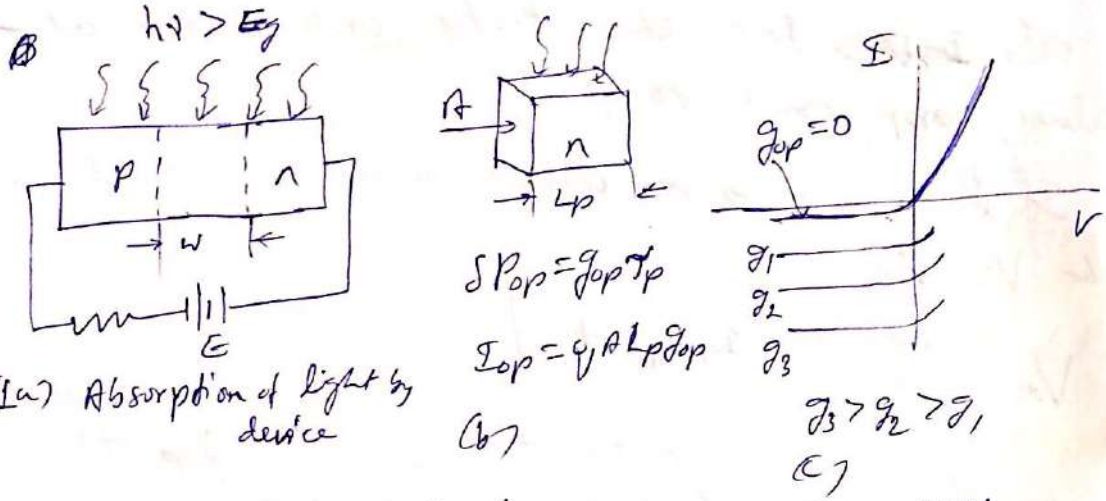
$$V_{oc} \simeq \frac{kT}{q} \ln \frac{g_{op}}{g_{th}} \quad \text{for } g_{op} \gg g_{th}. \quad \text{— (3b)}$$

Where $g_{th} = p_n/\tau_n$ represents the equilibrium thermal generation-recombination rate. As minority carrier concentration is increased by optical generation of EHP's → lifetime τ_n becomes shorter & p_n/τ_n rises. (However p_n is fixed, for given N_A & T)
∴ V_{oc} cannot increase indefinitely, & limit on V_{oc} is the equilibrium contact potential V_0 (Fig. (a))
V_0 is the maximum forward bias that can appear across a junction.

* The appearance of forward voltage across illuminated junction is called photovoltaic effect.

Depending on the application, the photodiode of Fig.(1) can be operated in either 3rd or 4th quadrant of I-V chr.c's.

P.T.O

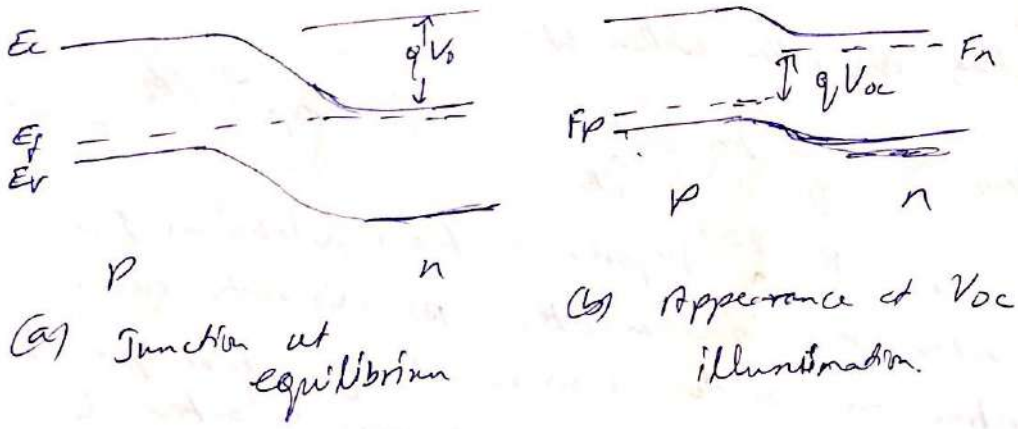


(a) Absorption of light by device (b) (c)

(b) current Iop resulting from EHP generation within a diffusion length of the junction on the n-side

(c) I-V chcl's of illuminated junction.

Fig.(1) Optical generation of carriers in a pn junction.

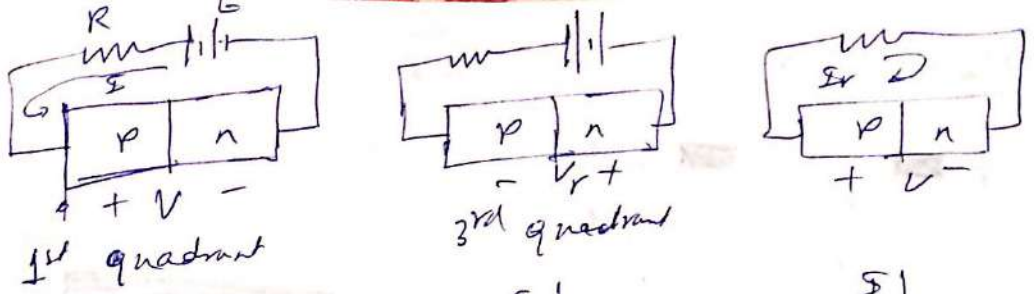


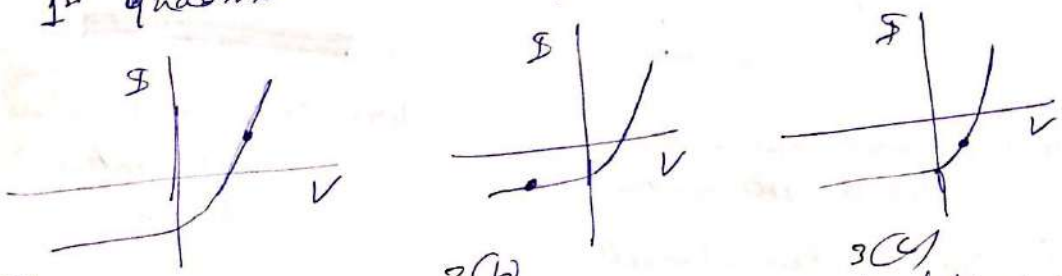
(a) Junction at equilibrium (b) Appearance of Voc with illumination.

In Fig.(3) power is delivered to the device from the external circuit when current & junction voltage are both +ve or both -ve (First or third quadrant)

In IV-quadrant, junction voltage is +ve and current is -ve. In this case power is delivered from the the junction to the external circuit. (If power is to be extracted from device, IV- quadrant is used).

In photodiode photodetector application the junction is reverse biased & operated in III-quadrant.





3(a) 3(b) 3(c) Device delivers power to load

Power is delivered to the device by external circuit.

Solar cells: (SC)

SC's can deliver power to an external circuit by an illuminated function, hence possible to convert solar energy into electrical energy. If we consider IV quadrant of Fig.(3c), only, (since only one device) much power may not be delivered. In this device, the voltage is restricted to less than contact potential, & less than bandgap voltage E_g/q. For Si $V_{oc} \simeq$ 1V. The current generation depends on illuminated area, typically I_{sc} is in the range 10–100mA, for an area of 1 cm^2.

* If we use many such devices significant power can be generated.

* Arrays of p-n junction solar cells are used to supply electrical power to many space satellites, & can supply power over long period, have advantages over batteries. (life, weight etc)

To utilize maximum amount of available optical energy, solar cells are designed with larger area junction near surface of device (Fig.(4)) p-to

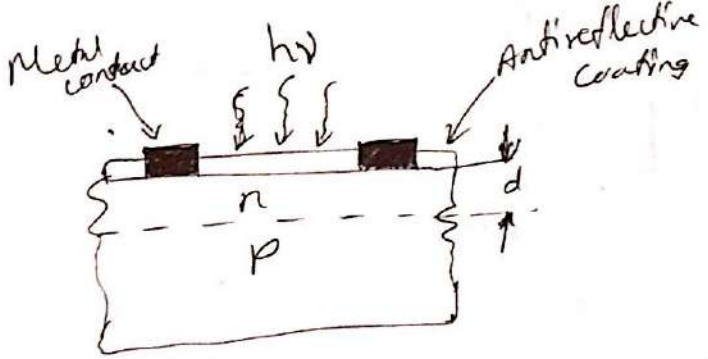
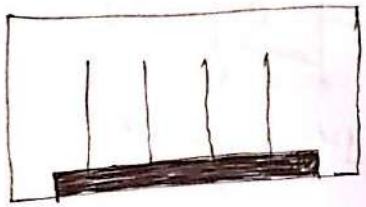


Fig. (4a) Enlarged view of the planar ~~(illegible)~~ junction



(4b) Top view showing metal contact "fingers"

Fig (4) configuration solar cell

The planar junction is formed by diffusion or ion implantation, & surface is coated with anti reflection coating ~~used~~ to reduce reflection & to decrease surface recombination.

Compromises in solar cell.

(i) The junction depth 'd' must be less than L_p in the 'n' material to allow holes generated near the surface to diffuse to the junction before they recombine.

(ii) Thickness of p region must be such that, electron can diffuse to the junction before recombination take place. Requires matching between diffusion length (L_n) & thickness of 'p' region; & mean optical penetration depth $1/\alpha$. ($\alpha \rightarrow$ absorption co-efficient in cm^{-1})

(iii) It is desired to have a large contact potential V_0, to obtain large photovoltage & hence heavy doping is required.

(iii) In general long lifetimes are desirable, however gets reduced by doping too heavily.

(iv) Series resistance of the device need to be very small to avoid ohmic losses in the device itself. Even few ohms will reduce o/p of solar power greatly.

* Since area made large, resistance of the p-type body of the device can be made small.
* To avoid contact resistance, contacts are distributed over n-region as in Fig (4b). (Reduces series resistance)

Figure (5) shows, fourth-quadrant portion of solar cell characteristics. The open circuit voltage V_{oc} & short-circuit current I_{sc} are determined for given light level by the cell properties.

Maximum power delivered to load by the solar cell occurs when the product VI_r is maximum. Then voltage & current are marked as V_m & I_m (Fig. 5), is less than $I_{sc} V_{oc}$ product.

The ratio $V_m I_m / I_{sc} V_{oc}$ is called fill factor, is a figure of merit for solar cell design.

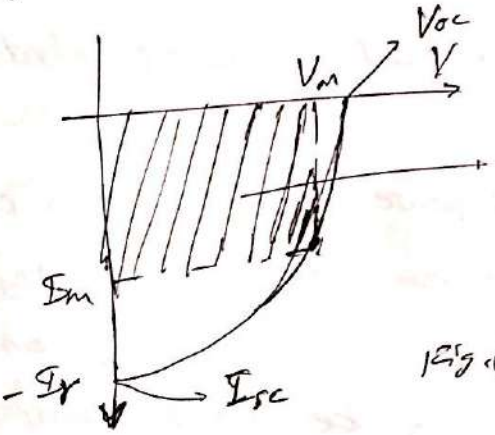


Fig. (5) I-V. chic's of illuminated solar cell. The maximum power rectangle is shaded

Application:

(i) outer space, (in satellites)
(ii) In the world ~ 15 TW is generated (old data)
(iii) nearly 1 kw/m² power is available in a partially sunny location. (i.e mean 600 TW potentially available world wide)

Efficiency of solar cells ~ 25%. provides ~ 250 W/m². Under full illumination

Ex: Si solar cell has short circuit current (I_{sc}) of 100 mA & an open circuit voltage (V_{oc}) of 0.8V under full solar illumination. The fill factor is 0.7. What is the maximum power delivered to a load by this cell

$$P_{max} = (ff)\, I_{sc} V_{oc} = 0.8 \times 100 \times 0.7 = 56\ mW.$$

Photodetectors:-

Photodiode operated in III-quadrant of its I-V ch/c's (Fig. 3b), the current is independent of voltage but proportional to optical generation rate. Such devices can be used to measure illumination levels or converting time-varying optical signals to electrical signals.

* Detectors speed of response or bandwidth is critical.

Ex: If photodiode is to respond to series of 1ns apart light pulses, photogenerated minority carriers must diffuse to the junction & swept across to the other side in time < 1ns. The carrier diffusion step in this process is time consuming & should be eliminated if possible.

* Desired to have depletion width W large enough - so that most photons absorbed within W.

* EHP's created in the depletion region, the electric field sweeps electrons to n-side & holes to p-side, since carrier drift occurs in short time, the response of photodiode can be fast.

When carriers are generated within depletion layer 'W' the detector is called depletion layer photodiode. Hence it is necessary dope at least one side of the junction lightly, so we can have larger W.

* ~~In general~~ W is chosen as a compromise between sensitivity & speed of response.
* If W is wide most of incident photons will be absorbed in the depletion region, leading to high sensitivity,
  - wide W results in small junction capacitance $\left[C_j = \frac{\epsilon A}{W}\right]$, hence reducing RC time constant.

disadv
- If W is too wide, then time required for drift of photogenerated carriers out of depletion region is large, hence low bandwidth.

Building p-i-n photodetector (Fig(6)), is one method of controlling depletion region width. The 'i' region need not be intrinsic as long as the resistivity is high. i-region is grown epitaxially on the n-type substrate, & p-region is ~~grown~~ obtained by implantation. When device is reverse biased, the applied voltage appears across 'i' region.

- Carrier ~~life~~ time within 'i' region is long compared to drift time, most of photogenerated carriers will be collected by n & p regions.
- Figure of merit of photodetector is the external quantum efficiency η_q.

Defined as number of carriers that are collected for every photon impinging on the detector.

- For a photocurrent density J_{op}, we collect $\frac{J_{op}}{q}$ carriers per unit area per sec.

p.T.O.

For an incident optical power density P_{op}, number of photons striking per unit area per sec $= P_{op}/h\nu$

$$\therefore \eta_q = \frac{(J_{op}/q)}{(P_{op}/h\nu)} \quad \text{——— (4)}$$

– For a no current gain photodiode maximum $\eta_q = 1$.

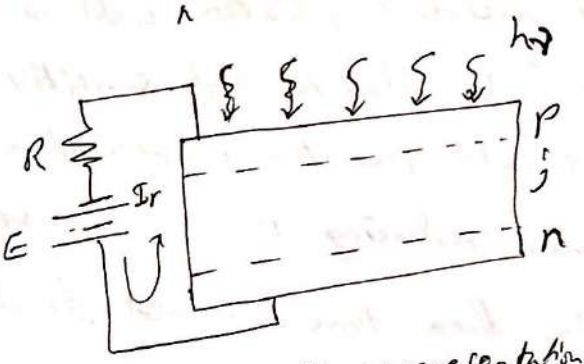


Fig (6) Schematic representation of p-i-n photodiode.

– Operate the photodiode in the avalanche region of its characteristics to detect low-level signals

* In avalanche mode each photogenerated carrier results in large change in current due to avalanche multiplication, leading to gain & external quantum efficiency > 100%.

* Avalanche photodiodes (APDs) are useful as detectors in fiber-optic systems.

* If $h\nu < E_g$ photons will not be absorbed.

or If $h\nu \gg E_g$, they will be absorbed very near the surface (where recombination rate is high)

* Hence choose photodiode material with bandgap corresponding to particular region of the spectrum.

——— Using lattice matched multilayers of compound semiconductors, bandgap of the absorbing region can be tailored

Eg:
InGaAs with an In mole fraction of 53%. can be grown epitaxially on InP with excellent lattice-matching, having Bang gap of 0.75eV useful in long term fibre optic systems.

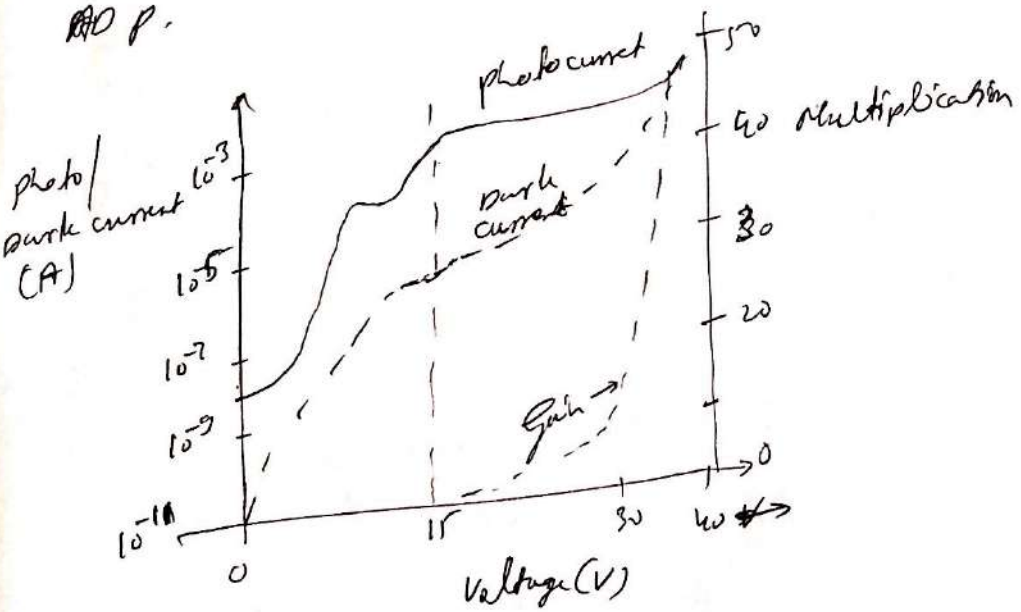
— In APD's, having narrow band ~~gap~~ gap material, it is advantageous absorb light in narrow-gap semiconductor (Ex. InGaAs) & transport the resulting carriers to a junction made in wider band gap material (Ex. InAlAs) where avalanche ~~multiplication~~ multiplication takes place at high fields.

* Avoids excessive leakage currents. In reverse bias narrow-gap materials.

* The photocurrent & dark current both increase with bias because of avalanche multiplication (Fig. 7)

* One need to maximize the difference between ΔI between the photocurrent I_p & the dark current I_d

* The ratio of ΔI at different voltages to that at a low ~~the~~ reference voltage is defined as gain of APD.



Light Emitting Diodes (LED's)

When carriers are injected across a forward-biased junction, the current is accounted for recombination in the transition & neutral regions near junction.

— In indirect band gap (Ge or Si) the recombination releases heat to the lattice.

— In direct band gap materials the recombination, light is given off from the junction (under forward bias).

This effect is called injection electroluminescence.

Application: diodes as generator of light.

∴ LED's in digital displays, traffic signals

— Another important device make use of radiative recombination in a forward biased p-n junction is the semiconductor laser.

* Lasers emit coherent light narrow wavelength band than LED's.
  → Used in fiber optic communication.

Light emitting Materials

The bandgaps of various binary compound semiconductors are different & there is wide variation in band gaps [Fig (X1)] Hence available photon energies, extending from ultraviolet (GaGaN, 3.4ev) to infrared (InSb, 0.18ev)

However by utilizing ternary and quaternary compounds the no of available energies can be increased significantly

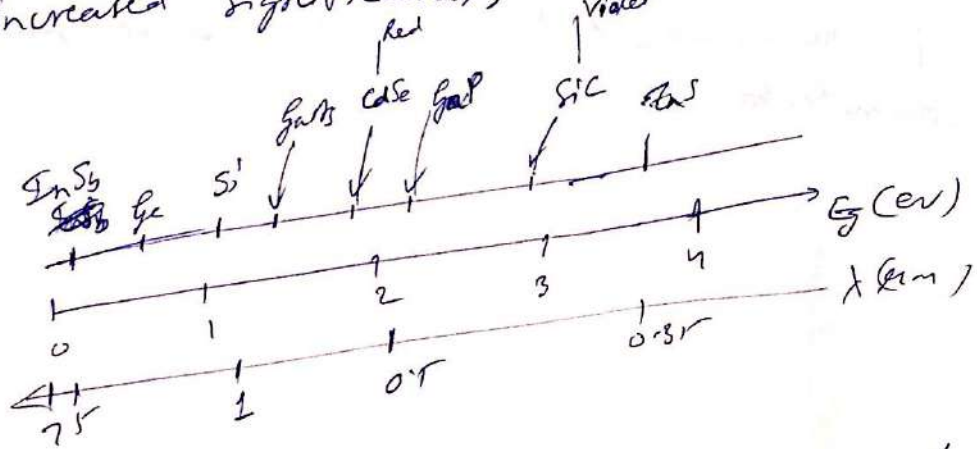
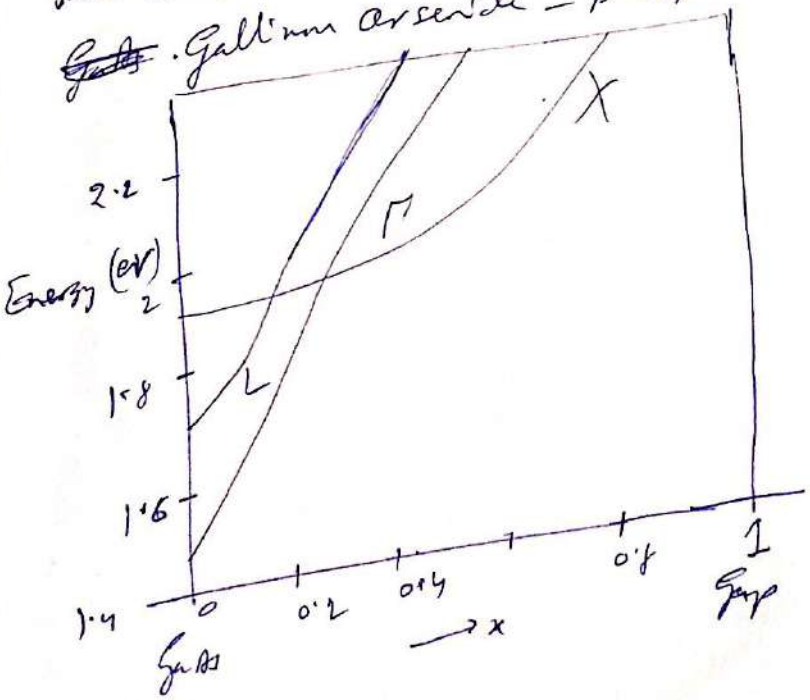


Fig (X1)

Ex: of variation of photon energy obtainable from the compound semiconductors is the ternary alloy Gallium arsenide – phosphide ($GaAs_{1-x}P_x$) [Fig (X2)]



Conduction band energies as a function of alloy composition for $GaAs_{1-x}P_x$.

Fig. (X2)

P.T.O.

When the percentage of As is reduced & P is increased in AlAs material, the band gap varies from direct 1.43ev gap of GaAs (infrared) to the indirect 2.26ev gap of GaP (green)

The band gap varies almost linearly with x until the 0.45 composition is reached, & e-h recombination is direct over this range.

The ~~co~~ common alloy composition used in LED display is $x \simeq 0.4$., for this composition band gap is direct. Γ minimum (at k=0) is the lowest part of ϕ conduction band

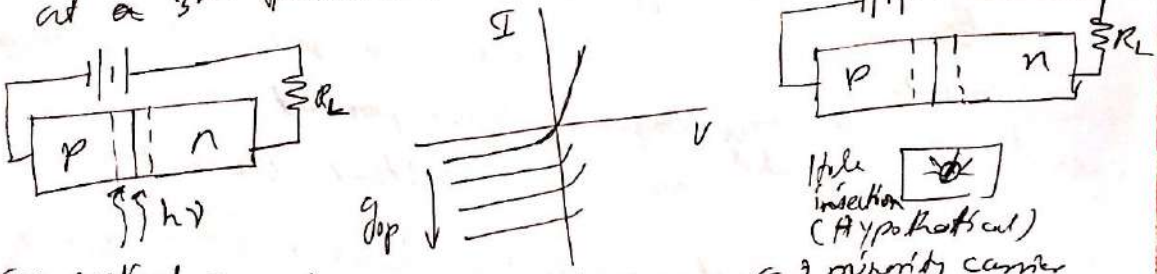
Bipolar Junction Transistors:

- First we we shall do qualitative discussion of charge transport in a Bipolar Junction Transistor (BJT)@
- Investigate the charge distribution in the transistor and relate three terminal currents to the physical ch(c)s of device.
- Aim is to gain solid understanding of current flow & control of transistor & understand secondary effects.
- Discuss the properties of the transistor with proper biasing for amplification & switching.
- we shall use p-n-p transistor for mod illustration.

Adv. hole flow & current flow are in the same direction. & hence visualize the charge transport easier.

Fundamentals of BJT operation:

- First we shall define some terms & gain physical understanding of how carriers are transported through the the device.
- Next discuss how the current through two terminals can be controlled by small changes in the current at a 3rd terminal.



(a) Optical generation (b) I-V ch(c)s as a function of E/HP generation. (c) minority carrier injection

Fig. (1) External control of the current in reverse biased p-n Jn.

PTO

Let us discuss BJT by considering reverse biased p-n junction as in Fig(1a). The reverse saturation current through diode depends on the rate at which minority carriers are injected in the neighborhood of the junction.

* we know that reverse saturation current (I_0) due to holes swept from n to p is independent of the size of the junction & hence independent of reverse bias.

* - Since hole current depends on how often minority holes are generated by EHP, with in diffusion length of the jn.

*: It is possible to increase the reverse current through diode by increasing rate of EHP generation (Fig (1b)).

- one method is by optical excitation of EHP with ($h\nu > E_g$).

- If it is possible to increase reverse current by by injecting minority carriers as in Fig.(1c). There are several advantages,

Eg: The current through reverse-biased junction would vary, very little if R_L were changed, since magnitude of junction voltage is unimportant.

* would act as constant current source

- The convenient method of hole injection device is a forward biased p+-n junction.

The current in such a junction is primarily due to holes injected from p+ region into n- material.

- If we make n-side of the forward-biased junction the same as the n-side of the reverse biased junction, the p^+-n-p structure of Fig.(a) results.

In this configuration injection of holes from p+n junction into the center of n-region supplies the minority carrier holes to participate in the reverse current through the p-n junction.

* However it is required that injected holes do not recombine in the n-region before they can diffuse to the depletion layer of reverse-biased junction. Hence make 'n' region narrow compared to hole diffusion length.

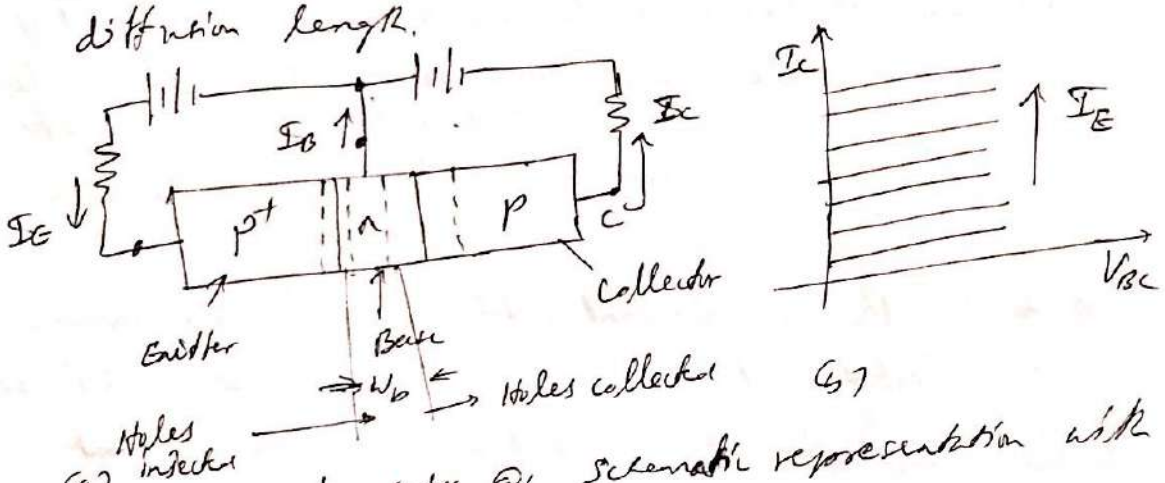


Fig.(2) pnp transistor (a) schematic representation with Je forward biased & Jc reverse biased.
(b) I-V characteristics of reverse biased p-n - Jn.

The forward biased junction in Fig (a) injects holes into the center of n-region is called emitter junction & reverse biased junction that collects the injected holes is called collector junction.

- p^+ region acts as a source of injected holes is called emitter,
- p - region into which holes are swept by the reverse biased junction is called collector.
- n → centre n-region is called base.

P70

The biasing arrangement shown in Fig (2a) is called Common base configuration, since base electrode 'B' is common to emitter & collector circuits.

To have good p-n-p transistor, (i) assume almost all the holes injected from emitter is collected in the base.

n-type base region is narrow (W_b)

hole life time (τ_p) should be long.

ie $W_b \ll L_p$

Where W_b → length of the neutral 'n' material of the base (measured between the depletion regions of emitter & collector Jn)

L_p → diffusion length for holes in the base $(D_p \tau_p)^{1/2}$.

* This requirement satisfied, an average hole injected at the emitter junction will diffuse to the depletion region of the collector. (without recombining in base)

and requirement

(ii) Current I_E crossing emitter junction should be composed of entirely of holes injected into the base,

* Rather than electrons crossing from base to emitter

* The above is satisfied by doping base region lightly compared to emitter, results in p^+-n emitter Jn. (Fig (2a))

If I_E flows into the cable emitter of properly biased p-n-p transistor, I_C flows out of collector.

Regarding base current I_B.

- In a good transistor, the base current will be very small, since I_E is essentially hole current.

And collected hole current is $I_C \approx I_E$.

* There is some base current I_B due to requirements of electron flow into n-type base region (Fig.(3))

Base current (I_B) is physically accounted by three dominant mechanisms.

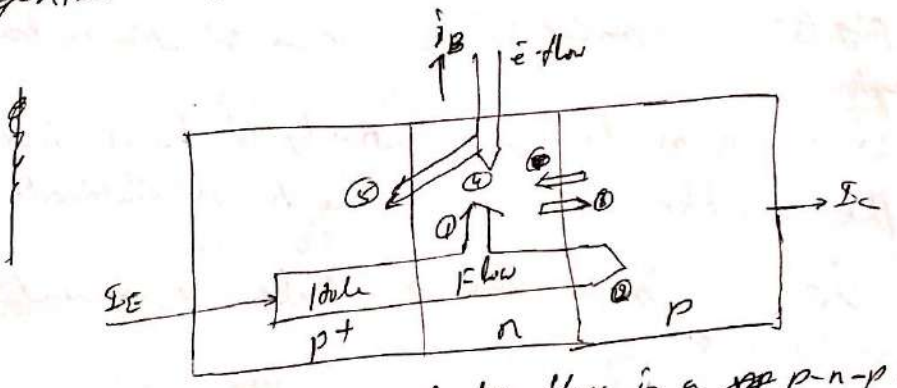
(i) There is some ~~recombination of~~ recombination of injected holes with electrons in the base (even with $W_b \ll L_b$)
These lost electrons is resupplied ~~with~~ through base contact.

(ii) Some electrons are injected from n to p in the forward-biased emitter junction. These electrons must also be supplied by I_B.

(iii) Some electrons are swept into the base from ——— reverse-biased collector junction due to thermal generation in the collector.
This small current reduces I_B by supplying electrons to base.

The dominant sources of base current are (i) recombination in base & (ii) injection into emitter region.
* These effects can be greatly reduced by desire design.

In general $I_B \ll I_E$.



Fig(3) Summary of hole & electron flow in a ~~pnp~~ p-n-p transistor with proper biasing.

① Injected holes lost to recombination in base ② Holes reaching reverse-biased collector junction ③ Thermally generated ē & h^+ making up reverse saturation current at collector I_C. ④ Electrons supplied by the base contact for recombination with holes.
⑤ electrons injected across - forward biased emitter Jn.

In n-p-n transistor, the three current directions are reversed, since electrons flow from emitter to collector, & holes must be supplied to the base.

7.2 Amplification with BJT

- we shall discuss the various factors involved in transistor amplification.
- Transistor is used in amplifiers, since currents at the emitter & collector are controlled by relatively small base current.

* neglect secondary effect, to understand essential mechanisms.
* we shall use both dc & ac current in this discussion.
  Assume simple analysis applies only to dc and to small signal ac at low frequencies.
- we can relate terminal currents of the transistor, i_E, i_B and i_C by several important factors.

* Let us neglect saturation current at the collector (Fig. (3) component 3.) & recombination in transition region.

(1) $\therefore$ I_C is made up entirely of holes injected at the emitter. (Most lost due to recombination in base)

That is i_C is proportional to hole component of emitter current.

$$i_C = B\, i_{Ep} \quad \text{—— (1)}$$

where $B \rightarrow$ is fraction of holes injected which make it across base to the collector.

$B \rightarrow$ is called Base transport factor.

(1) Total emitter current I_E is made up of hole component i_{EP} and electron component i_{EN} (due to electrons injected from base to emitter [components in Fig. (3)]

The emitter injection efficiency

$$\gamma = \frac{i_{EP}}{i_{EN} + i_{EP}} \quad \text{——— (2)}$$

For efficient transistor, β & γ must be very near unity. i.e. emitter current should be mostly due to holes ($\gamma \approx 1$) & most of the injected holes should eventually participate in the collector current. ($\beta \approx 1$).

∴ The relation between collector & emitter current is

$$\frac{i_C}{i_E} = \frac{\beta i_{EP}}{i_{EN} + i_{EP}} = \beta\gamma = \alpha \quad \text{——— (3)}$$

where $\alpha \to$ is called current transfer ratio, represents the emitter-to-collector current amplification.

* There is no real amplification between these currents since $\alpha < 1$.

* However relation between i_C and i_B is more promising for amplification.

$\left(\beta = \frac{i_C}{i_B}\right) \cdot \frac{i_{EP}}{i_{EN} + i_{EP}}$

$\beta\cdot\gamma = \frac{i_C}{i_E} \cdot \frac{i_{EP}}{i_{EN} + i_{EP}}$

$\beta\gamma = \frac{i_C}{i_E} = \alpha$

Note: α is called Common base current gain.

——— To account base to current, include the rates at which electrons are lost from base by injection across the emitter function (i_{EN}) & the rate of electrons recombine with holes in the base.

In both cases lost electrons must be resupplied through base i_B.

If the fraction of injected holes making it across base without recombination is β, it follows that $(1-\beta)$ is the fraction recombining in the base.

∴ Base current $i_B = i_{EN} + (1-\beta) i_{EP}$ ——— (4)

Neglecting collector saturation current. The relation between collector & base current is obtained from eq ① to ④

From ① & ④

$$\frac{i_c}{i_B} = \frac{B i_{EP}}{i_{EA} + (1-B) i_{EP}} = \frac{B i_{EP}}{i_{EA} + i_{EP} - B i_{EP}}$$

$$= \frac{\frac{B i_{EP}}{i_{EA} + i_{EP}}}{1 - B\left[\frac{i_{EP}}{i_{EA} + i_{EP}}\right]} \quad \text{——— (5)}$$

$$= \frac{BT}{1 - BT} \quad \text{from eq (3)}$$

$$= \frac{\alpha}{1-\alpha} = \beta \quad \text{——— (6)}$$

The factor β relating collector current to base current is the base-to-collector current amplification factor (it also called common base current gain α) (is called common emitter current gain)

auto: $\alpha \simeq 1$, & β can be large for good transistor, & collector current is large compared with the base current.

Controlling of i_c by small i_B:

We have discussed control of i_c by the emitter current i_E with base current characterized as a small side effect.

However it can be shown from space charge neutrality concept that i_B can be used to determine the magnitude of i_c.

Consider Fig.(4) in which i_B is determined by a biasing circuit.

Assume: (i) unity emitter injection efficiency and negligible collector saturation current

Since n-base is electrostatically neutral between two transition regions, presence of excess holes in transit from emitter to collector calls for compensating excess electrons from the base contact.

* However there is a differences in times the electrons & holes spend in the base.

* The average excess hole spends a time τ_t defined as transit time from emitter to collector

* Since base width W_b is small compared with L_p

$\therefore L_p \ll \tau_p$ (average hole time) in the base.

— Average excess electron supplied from the base contact spends τ_p seconds in the base supplying space charge neutrality during the life time of an excess holes

* while average electron weight τ_p seconds for recombination.

- many individual holes can enter & leave the base region each with an average transit time τ_t.

In particular, for each electron entering from base contact, τ_p / τ_t holes can pass from emitter to collector while maintaing space charge neutrality.

$\therefore$ The ratio of collector current to base current is simply

$$\frac{i_c}{i_B} = \beta = \frac{\tau_p}{\tau_t} \qquad \text{——— (1)}$$

for $\gamma = 1$ & negligible collector sat current.

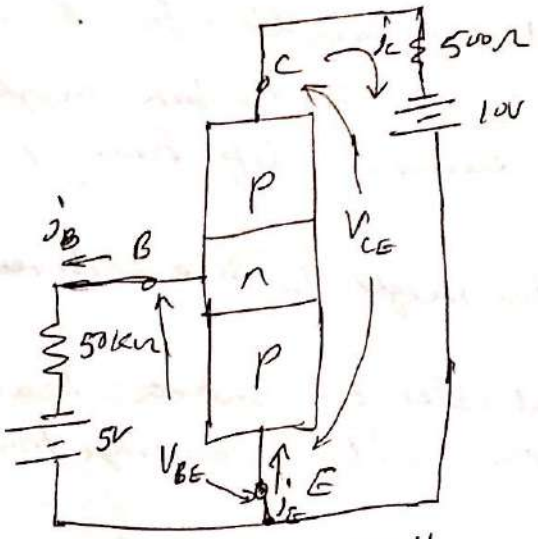
— If electron supply to the base (i_B) is restricted, then holes from emitter to base is correspondingly reduced.

P.T.O

~~Correct~~ Counter argument:

Suppose the hole injection does continue despite the restriction on electrons from the base contact. The result would be net buildup of positive charge in the base and a loss of forward bias (∴ loss of hole injection) at the emitter junction

∴ clearly the supply of electrons through i_B can be used to fix or lower the hole flow from emitter to collector.



Biasing circuit

$\tau_p = 10\ \mu s$

$\tau_t = 0.1\ \mu s$

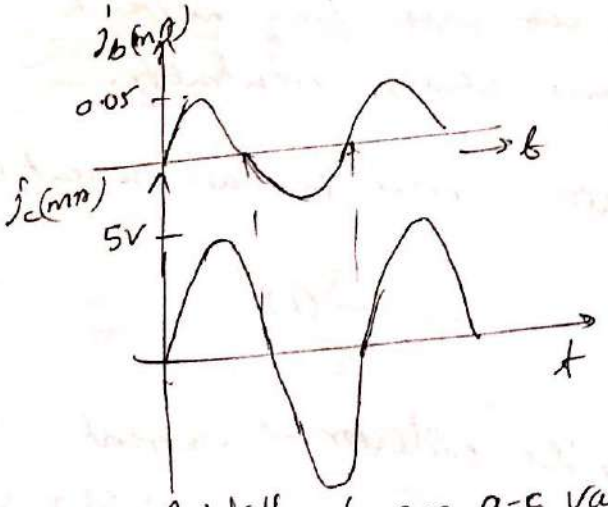
$\frac{i_c}{i_B} = \beta = \frac{\tau_p}{\tau_t} = 100$

Neglecting V_{BE}

$I_B = \frac{5V}{50k\Omega} = 0.1\ mA$

$I_c = \beta I_B = 10\ mA$

Fig.(4)(a)



Fig(4b) Addition of an a-c variation of base current i_b to the dc value of I_B resulting in an ac component i_c

In Fig.(4) base current is controlled independently, & is called common-emitter circuit. Since — emitter terminal is common to both base & collector ckt's.

The emitter junction is forward biased by battery, & drop across it is small. Almost all the voltage from collector to emitter appears across the reverse biased collector junction.

Since V_{BE} is small (for forward bias), neglecting it, the base current

$$i_b = \frac{5V}{50k\Omega} = 0.1mA$$

If $\tau_p = 10\mu s$, $\tau_t = 0.1\mu s$, ∴

$$\therefore \beta = \tau_p / \tau_t = 100. \leftarrow$$

collector current $i_c = \beta i_b = 100 \times 0.1mA$

$= 10mA$

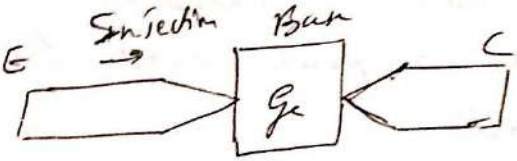
note: i_c is determined by β and base current rather than by the battery & resistor in the collector circuit.

If a small a-c current i_b is superimposed on the steady state base current (Fig (4a)), corresponding ac current i_c appears in the collector circuit.

The time varying portion of the collector current will be i_b multiplied by the factor β, termed the current gain.

BJT Fabrication

History
In 1947 Bardeen & Brattain invented first point contact transistor. Here two sharp metal wires were simply pressed (formed an emitter & a collector of carriers) onto a slab of Ge (provided base or mechanical support). Through which injected carriers flowed.



* This basic invention lead to the BJT, in which charge collection injection & collection was achieved using two p-n junctions in proximity to each other.

* p-n junctions in BJT's can be formed using thermal diffusion, but modern devices use ion implantation.

Double polysilicon self-aligned n-p-n Si BJT.

- Use of n-p-n transistor is more popular than p-n-p ~~transistor~~ devices ~~by~~ due to higher mobility of electrons compared to holes. process steps are shown in Fig.(5).

(1) A p-type Si substrate is oxidized, windows are defined using photolithography & etched in oxide. Using photoresist and oxide as an implant mask, dopants such as As or Sb is implanted into the open window to form high conductivity n+ layer Fig(5a).

* Subsequently, the photoresist & the oxide are removed & a lightly doped n-type epitaxial layer is grown.

- During high temp. epitaxial growth, the implanted n+ layer diffuses slightly toward the surface & ~~bar~~ becomes ~~co~~

conductive buried collector (also called n^+ sub-collector).

The n^+ subcollector layer guarantees a low-collector series resistance to collector ohmic contact, made through collector contact region (Fig 5c).

* Lightly doped n-type collector above n^+ sub-collector in the part of BJT where base & emitter are formed ensures high base-collector reverse breakdown voltage.

2. For integrated circuits involving many transistors, electrical isolation of BJT's are important to ensure no electrical cross-talk between them.

Such an isolation can be achieved by LOCOS to form filled or isolation oxides after a B-channel implant [Fig. 5b].

Another isolation scheme well suited for high-density bipolar circuits involves the formation of shallow — trenches by reactive ion etching (RIE) backfilled with oxide & polysilicon. Using RIE, a narrow trench about 1 μm deep can be formed with straight sidewalls. Oxidation inside the trench forms an insulating layer, & the trench is then filled with oxide using LPCVD.

— polysilicon layer is then deposited by LPCVD & doped heavily p^+ with boron (B) during deposition. An oxide layer is next deposited by LPCVD. Using photolithography with base/emitter mask, a window is etched in the polysilicon/oxide stacks by RIE [Fig. 5c].

Heavily doped "extrinsic" p^+ base is formed by diffusion of B from the doped polysilicon layer into the substrate in order to provide a low-resistance, high speed base ohmic contact.

PTO

An oxide layer is then deposited by LPCVD, which has the effect of closing up the base window that was etched previously, & 'B' is implanted into this window [Fig. (5d)]

This base implant forms a more lightly p doped "intrinsic" base through which most of the current flows from emitter to collector

* Heavily doped extrinsic base forms a collar around the intrinsic base, and serves to reduce the base series resistance.

— The base is enclosed well within the collector, else it would be shorted to the p^- substrate.

— Finally another LPCVD oxide layer is deposited to close up the base window further, & oxide is etched all the way to the Si substrate by RIE, leaving oxide spacers on the sidewalls.

Heavily n^+ doped polysilicon is then deposited on the substrate, patterned & etched, forming polysilicon emitter & collector contacts [Fig. (5e)].

—— Self alignment is referred to the ~~process~~ as referred, since separate lithography step is not required to form the n^+ emitter-region

—— Finally an oxide layer is deposited by CVD, windows are etched to make emitter (E), base (B) and collector (C) contacts & suitable contact metal, such as Al is sputter deposited to form ohmic contacts.

Al is patterned using photolithography & etched with RIE. The IC's on the wafer are then separated

into individual dies by sawing.

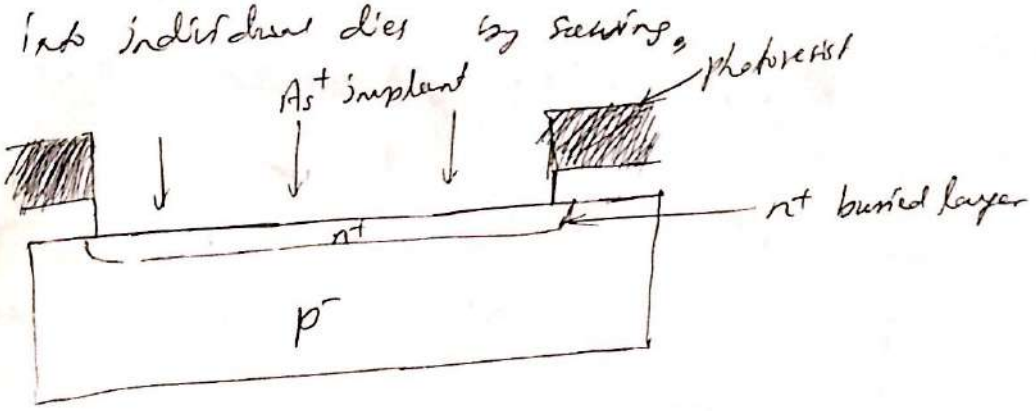


Fig (5a)

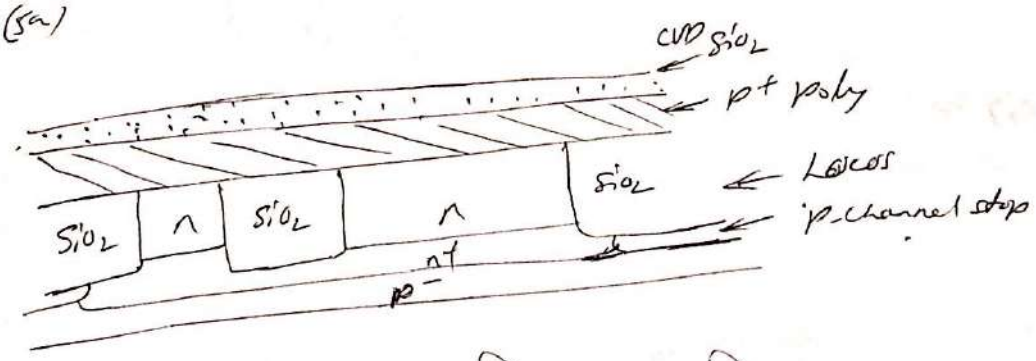
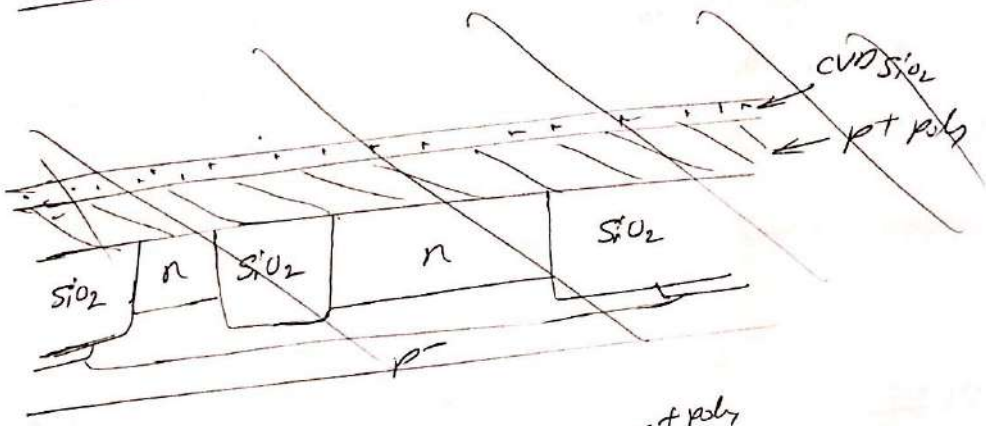


Fig (5b)



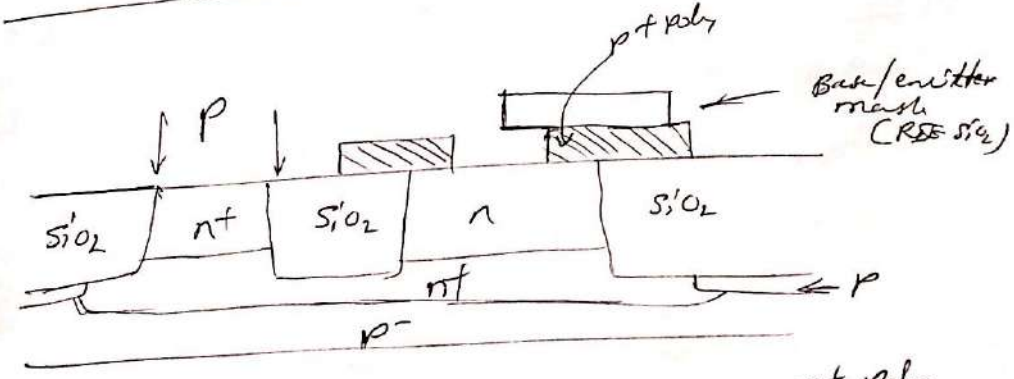


Fig. (5c)

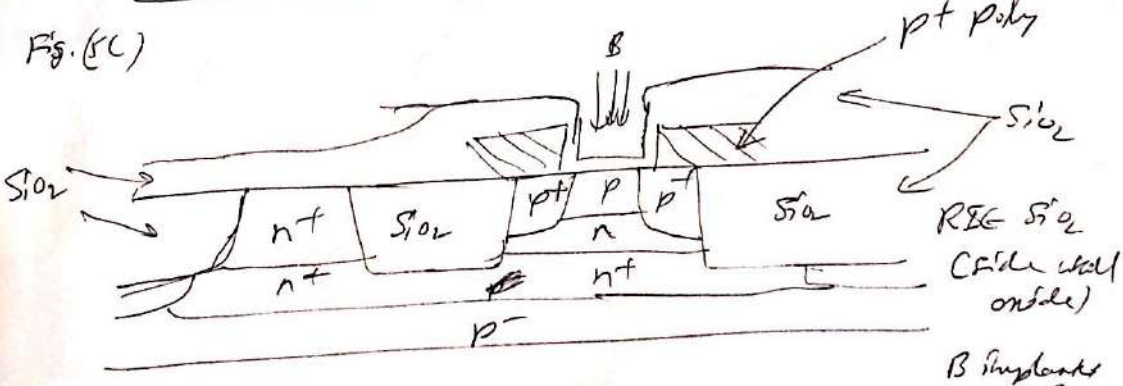


Fig (5d)

P-10

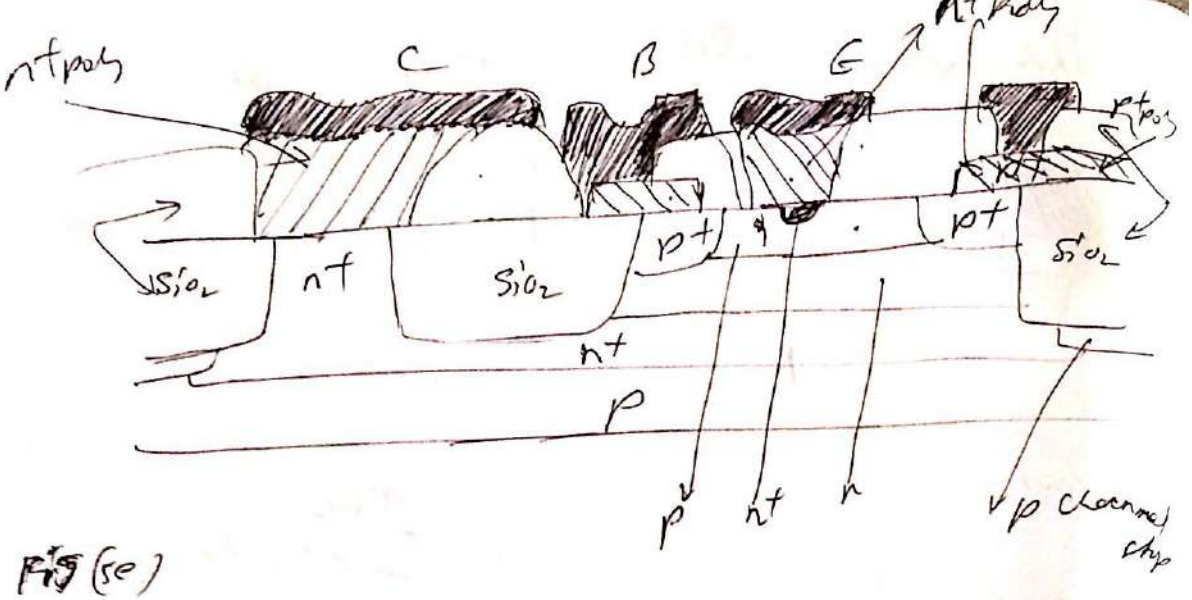


Fig (5e)

Generalized Biasing:

In the actual BJT, collector & emitter junctions may differ in area, saturation currents and other parameters. That means the BJT is not symmetrical between emitter & collector

Note: (1) In the normal active mode emitter junction is forward biased & the collector is reverse biased.

(2) In switching operations normal biasing rule is violated, in these cases it is important to account differences in injection & collection properties of the two junctions.

- we shall discuss generalized approach which accounts for transistor operation in terms of a coupled diode model valid for all combinations of emitter and collector bias.
- 4 measurable parameters are involved in this model, that can related to geometry & material properties of the device.
- This model in conjunction with the Large control approach, we can describe physical operation of a transistor in switching ckts & other applications

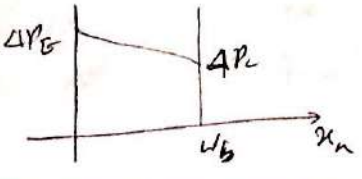
The coupled – diode model

ΔP_E → excess hole concentration at the edge of emitter depletion region. (cm^{-3})

$$\Delta P_E = P_n \left(e^{q V_{EB}/kT} - 1\right) \quad \text{—(xa)}$$

ΔP_C = at the collector side of base (cm^{-3})

$$= P_n \left(e^{q V_{CB}/kT} - 1\right) \quad \text{—(xb)}$$



D_p → diffusion co-efficient of holes (cm^2/s)

L_p → Diffusion length of holes in cm = $(D_p \tau_p)^{1/2}$

L_n → " " of electrons in cm

A → Area of cross section in cm^2

W_b → Length of the neutral n material of the base in cm (measured between depletion region of emitter & collector regions)

I_{EN} = ~~Electron component of emitter current~~ normal mode components injected from emitter to collector.

I_{CN} = ~~Collector~~

~~AB~~

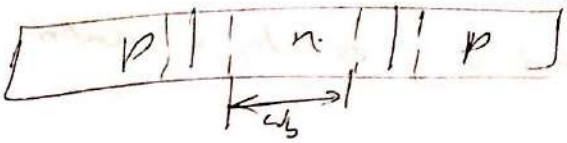
I_{EI} } inverted mode of injection from collector to emitter.

I_{CI} }

τ_p → hole life time (Recombination life time of holes in s)

τ_n → Recombination life time of electrons in s.

~~Wb~~ W_b → ~~Be~~ Length of the neutral n material of the base



7.2

~~Base~~

β → base-to-collector current amplification factor

γ → emitter injection efficiency $\frac{I_{Ep}}{I_{En} + I_{Ep}}$

α → $= \beta\gamma$ → current transfer ratio, represents collector to emitter current amplification

B → base transport factor (Fraction of holes which make it across base to the collector)

The coupled Diode Model

If collector junction is forward biased, we cannot neglect ΔP_c, hence use more general hole distribution in base. Fig. (6a) shows situation in which both emitter & collector junctions are forward biased, & hence ΔP_E & ΔP_c are positive numbers.

The situation can be handled using following equations

$$I_{EP} = qA \frac{D_p}{L_p} \left(\Delta P_E \coth \frac{W_b}{L_p} - \Delta P_c \operatorname{csch} \frac{W_b}{L_p} \right) \quad -(1a)$$

$$I_C = qA \frac{D_p}{L_p} \left(\Delta P_E \operatorname{csch} \frac{W_b}{L_p} - \Delta P_c \coth \frac{W_b}{L_p} \right) \quad -(1b)$$

$$I_B = I_E - I_C = qA \frac{D_p}{L_p} \left[(\Delta P_E + \Delta P_c) \tanh \frac{W_b}{2L_p} \right] \quad -(2)$$

for symmetrical transistors.

These equations [(1) & (2)] can be considered as linear superpositions of injection by each junction. Hence Fig (6a) can be broken into two components of Fig. (6b) & Fig (6c).

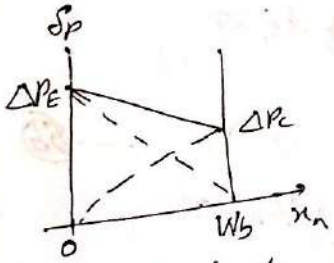


Fig (6a) Approximate hole distribution in base with I_E & I_C forward biased.

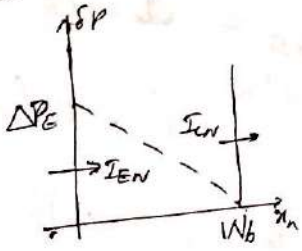


Fig. (6b). Component due to injection & collection in normal mode

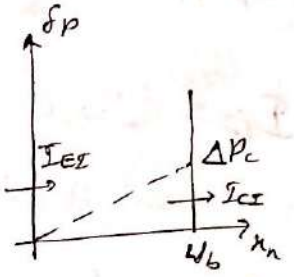


Fig. (6c) Component due to inverted mode

Fig. (6) Hole distribution evaluation in terms of normal & inverted modes.

One component in Fig (6b) accounts for holes injected by the emitter & collected by the collector. And call the resulting currents (I_{EN} & I_{CN}) the normal mode components.

P.T.O.

The other component of current [Fig.(6b)] results in currents I_{EI} and I_{CI} due to injection from collector to emitter called inverted mode of injection.

* These currents (I_{EI} & I_{CI}) oppose (negate) the I_{EN} & I_{CN} since holes flow opposite to the original definition of I_E & I_C.

For symmetrical transistor, various components are described by Eq(1). Defining

$$a \equiv (qA D_p/L_p)\coth(W_b/L_p) \quad \text{and}$$
$$b \equiv \left(\frac{qA D_p}{L_p}\right)\operatorname{cosech}(W_b/L_p) \quad \text{we have} \qquad \text{—(2)}$$

$I_{EN} = a\,\Delta P_E$ and $I_{CN} = b\,\Delta P_E$ with $\Delta P_C = 0$ —(4a)

$I_{EI} = -b\,\Delta P_C$ and $I_{CI} = -a\,\Delta P_C$ with $\Delta P_E = 0$ —(4b)

These four components are combined by linear superposition in eq (1) we have

$$I_E = I_{EN} + I_{EI} = a\cdot\Delta P_E - b\,\Delta P_C \qquad \text{—(5a)}$$
$$= A\left(e^{qV_{EB}/KT} - 1\right) - B\left(e^{qV_{CB}/KT} - 1\right)$$

$$I_C = I_{CN} + I_{CI} = b\,\Delta P_E - a\,\Delta P_C \qquad \text{—(5b)}$$
$$= B\left(e^{qV_{EB}/KT} - 1\right) - A\left(e^{qV_{CB}/KT} - 1\right)$$

Where $A = a P_n$ and $B = b P_n$

The above equations show that the superposition of the normal and inverted components gives the results for symmetrical transistor.

For a more general case, we must relate four components of current by factors which allow for asymmetry in two junctions.

Eq. emitter current in normal mode can be written

$$I_{EN} = I_{ES}\left(e^{qV_{EB}/KT} - 1\right), \quad \Delta P_c = 0 \qquad \text{—(6a)}$$

Where I_{ES} is the magnitude of the emitter saturation current in the normal mode. Since we assume $\Delta P_c = 0$ that means, $V_{CB} = 0$ (from eq $\Delta P_c = P_n\left(e^{qV_{CB}/KT} - 1\right)$.

∴ we consider I_{ES} is the magnitude of the emitter saturation current with collector junction shorted circuited.

With the collector current in the inverted mode is

$$I_{CI} = -I_{CS}\left(e^{qV_{CB}/KT} - 1\right), \quad \Delta P_c = 0 \qquad \text{—(7)}$$

Where I_{CS} is the magnitude of the collector saturation current with $V_{EB} = 0$, The − sign with I_{CI}, means, inverted mode holes are injected opposite to the defined direction of I_C.

— The collector currents for each mode of operation can be written by defining a new α for each case,

$$I_{CN} = \alpha_N I_{EN} = \alpha_N I_{ES}\left(e^{qV_{EB}/KT} - 1\right) \qquad \text{—(8a)}$$

$$I_{EI} = -\alpha_I I_{CI} = -\alpha_I I_{CS}\left(e^{qV_{EB}/KT} - 1\right) \qquad \text{—(8b)}$$

Where α_N & α_I are the ratios of collected current to injected current in each mode

Note: In inverted mode injected current is I_{CI} & collected current is I_{EI}.

The total current can also again be obtained by superposition of components.

$$\left.\begin{aligned} I_E &= I_{EN} + I_{EI} = I_{ES}\left(e^{qV_{EB}/KT} - 1\right) - \alpha_I I_{CS}\left(e^{qV_{CB}/KT} - 1\right) \qquad \text{—(9a)} \\ I_C &= I_{CN} + I_{CI} = \alpha_N I_{ES}\left(e^{qV_{EB}/KT} - 1\right) - I_{CS}\left(e^{qV_{CB}/KT} - 1\right) \qquad \text{—(9b)} \end{aligned}\right\}$$

derived by J.J. Ebers & J.L. Moll hence referred to as Ebers – Moll equations.

The general form of eq (9a) & (9b) is similar to (5a) & (5b) of symmetrical transistor.

But (9a) & (9b) allow for variations in D_{ES}, D_{CS}, α_I & α_N due to asymmetry between junctions.

It is possible to show reciprocity argument

that $\alpha_N D_{ES} = \alpha_I D_{CS}$ ——— (10)

even for non-symmetrical transistor.

Note: In Ebers-Moll equations, the I_E & I_C are described, in terms of resembling diode relations (I_{EN} & I_{CI}) & the terms that provide coupling between the properties of emitter & collector (I_{ES} & I_{CN}).

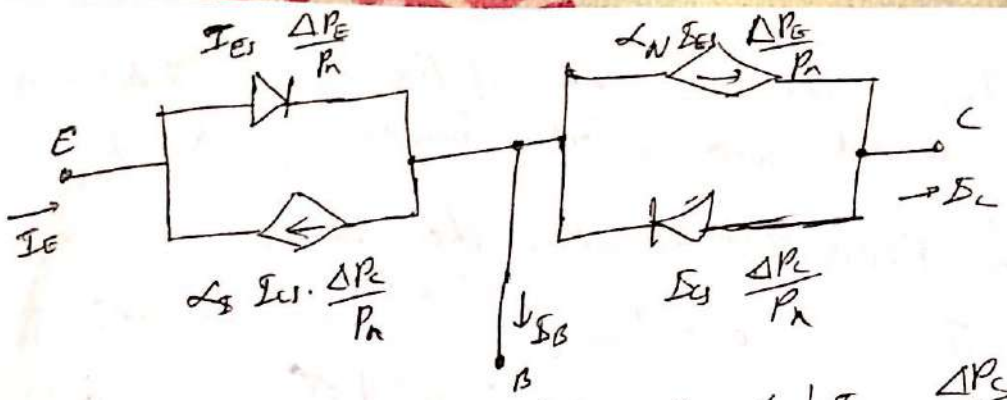
The coupled-diode property is illustrated using equivalent circuit [Fig. (7)]. In this figure, eq (10a) & (10b) are used to write the Ebers-Moll equations in the form.

$$I_E = I_{ES} \cdot \frac{\Delta P_E}{P_n} - \alpha_I I_{CS} \cdot \frac{\Delta P_C}{P_n}$$

$$= \frac{I_{ES}}{P_n} \cdot (\Delta P_E - \alpha_N \Delta P_C) \quad \text{——— (12a)}$$

$$I_C = \alpha_N I_{ES} \frac{\Delta P_E}{P_n} - I_{CS} \frac{\Delta P_C}{P_n}$$

$$= \frac{I_{CS}}{P_n} (\alpha_I \cdot \Delta P_E - \Delta P_C) \quad \text{——— (11b)}$$



$$I_B = (1-\alpha_N) I_{ES} \frac{\Delta P_E}{P_n} + (1-\alpha_I) I_{CS} \cdot \frac{\Delta P_C}{P_n}$$

Fig (7) Equivalent Circuit representing the Ebers-Moll equations

It is useful to relate the terminal currents to each other as well as to the saturation currents.

- we can eliminate the saturation currents from the coupling term in eq (9)

Eq. multiplying eq (10) by α_N

$$\alpha_N I_E = \alpha_N (I_{EN} + I_{EI}) = \alpha_N I_{ES} \left(e^{q V_{EB}/KT} - 1\right) - \alpha_N \alpha_I I_{CS} \left(e^{q V_{CB}/KT} - 1\right)$$

i.e $$\alpha_N I_{ES} \left(e^{q V_{EB}/KT} - 1\right) = \alpha_N I_E + \alpha_N \alpha_I I_{CS} \left(e^{q V_{CB}/KT} - 1\right)$$

Substituting in eq (9) we get

$$I_C = \alpha_N I_E + \alpha_N \alpha_I I_{CS} \left(e^{q V_{CB}/KT} - 1\right) - I_{CS} \left(e^{q V_{CB}/KT} - 1\right)$$

$$I_C = \alpha_N I_E - (1 - \alpha_N \alpha_I) I_{CS} \left(e^{q V_{CB}/KT} - 1\right) \quad \text{---(12)}$$

Similarly the emitter current can be written in terms of collector current

$$I_E = \alpha_I I_C + (1 - \alpha_N \alpha_I) I_{ES} \left(e^{q V_{EB}/KT} - 1\right) \quad \text{---(13)}$$

The term $(1-\alpha_N \alpha_I) I_{CS}$ and $(1-\alpha_N \alpha_I) I_{ES}$ can be abbreviated as I_{CO} and I_{EO} respectively.

where I_{CO} → is the magnitude of the collector saturation current with emitter junction open ($I_E = 0$).

P.T.O

– I_{EO} is the magnitude of the emitter saturation current with collector junction open ($I_C = 0$)

The Ebers-Moll equations then become

$$I_E = \alpha_I I_C + I_{EO}\left(e^{qV_{EB}/KT} - 1\right) \quad \text{—(14a)}$$

$$I_C = \alpha_N I_E - I_{CO}\left(e^{qV_{CB}/KT} - 1\right) \quad \text{—(14b)}$$

& equivalent circuit is shown in Fig. (8a).

Here the equations describe both emitter and collector currents in terms of simple diode chars plus a current generator proportional to the other current.

Eg. Under normal biasing the equivalent circuit reduces to the form shown in Fig (8b).

The collector current is α_N times the emitter current plus collector saturation current.

The resulting collector characteristics of the transistor appear as a series of reverse-biased curves as in Fig (8c).

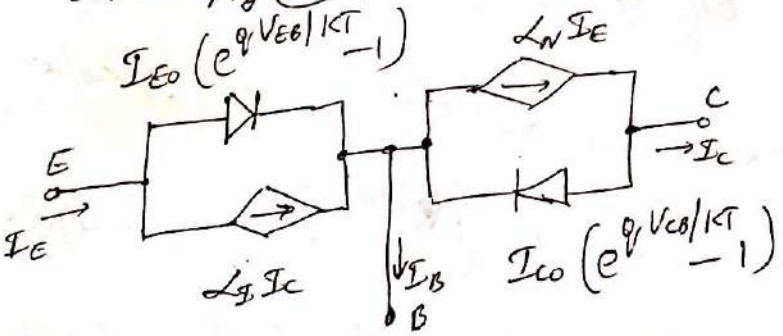


Fig. (8a)

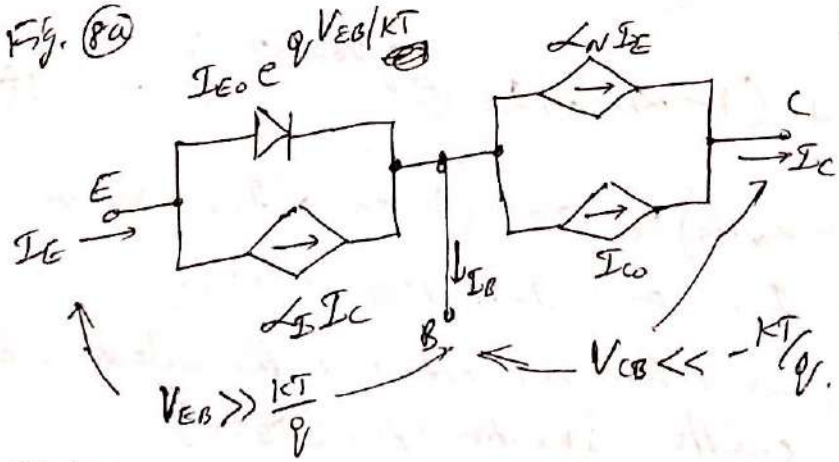


Fig (8b)

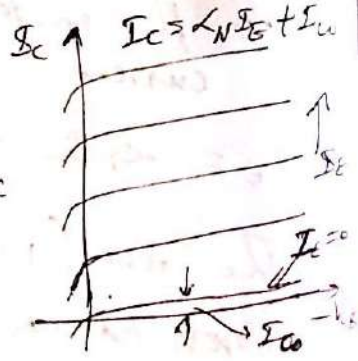


Fig. (8c)

Example problems:

A symmetrical $p^+ n p^+$ bipolar transistor has the following properties.

	Emitter	Base
$A = 10^{-4}\ cm^2$	$N_a = 10^{17}$	$N_d = 10^{15}\ cm^{-3}$
$W_b = 1\mu m$	$\tau_n = 0.1\mu s$	$\tau_p = 1\mu s$
	$\mu_p = 200$	$\mu_n = 1300\ cm^2/V\text{-}s$
	$\mu_n = 700$	$\mu_p = 450\ cm^2/V\text{-}s$

(a) Calculate the saturation current $I_{ES} = I_{CS}$.

(b) with $V_{EB} = 0.3V$ and $V_{CB} = -40V$, calculate base current I_B, assuming perfect emitter injection efficiency.

(c) Calculate the base transport factor β, emitter injection efficiency γ, and amplification factor β, assuming that the emitter region is long compared with L_n.

Soln. In the base

In the base $p_n = n_i^2/n_n = (1.5 \times 10^{10})^2 / 10^{15} = 2.25 \times 10^5$

$D_p = KT\mu_p \cdot \left(\frac{KT}{q}\right) = (450)(0.025) = 11.66,$

$L_p = \sqrt{D_p \cdot \tau_p} = (11.66 \times 10 \times 10^{-6})^{1/2} = 1.08 \times 10^{-2}$

$\frac{W_b}{L_p} = \frac{10^{-4}\ cm}{1.08 \times 10^{-2}\ cm} = 9.26 \times 10^{-3}$

$I_{ES} = I_{CS} = qA\left(D_p/L_p\right) p_n \cdot \coth\left(\frac{W_b}{L_p}\right)$

$= (1.6 \times 10^{-15}) \times (10^{-4}) \cdot \left(\frac{11.66}{1.08 \times 10^{-2}}\right) \times (2.25 \times 10^5)$

$\times \coth 9.26 \times 10^{-3}$

$= 4.2 \times 10^{-13}\ A$

$\Delta P_E = P_n e^{qV_{EB}/KT}, \quad \Delta P_C \simeq 0$

$= 2.25 \times 10^5 \times e^{(0.3/0.025)} = 2.4 \times 10^{10}$

P.T.O.

$$I_B = qA\left(\frac{D_p}{L_p}\right)\Delta P_E \tanh\left(W_b/2L_p\right)$$

or

$$I_B = \frac{Q_b}{T_p} = \frac{qA W_b \Delta P_E}{2T_p}$$

$$= \frac{1.6\times10^{-15}\times10^{-4}\times10^{-4}\ 2.4\times10^{10}}{2.\times10\times10^{-6}} = 1.9\times10^{-12}\ A.$$

In the emitter

$$D_n = \frac{kT}{q}(\mu_n) = (0.025)(700) = 18.13$$

$$L_n = \sqrt{D_n\tau_n} = \left(18.13\times10^{-7}\right)^{1/2} = 1.35\times10^{-3}$$

$$I_{En} = \frac{qA D_n^E}{L_n^E} n_p^E e^{qV_{EB}/kT}$$

$$I_{Ep} = \frac{qA D_p^B}{L_p^B} p_n^B \coth\frac{W_b}{L_p^B} e^{qV_{EB}/kT}$$

$$\gamma = \frac{I_{Ep}}{I_{En}+I_{Ep}} = \left[1+\frac{I_{En}}{I_{Ep}}\right]^{-1}$$

$$\gamma = \left[1+\frac{D_n^E/L_n^E n_p^E}{D_p^B/L_p^B\cdot n_n^B}\tanh\frac{W_b}{L_p}\right]^{-1}$$

$$\left(\text{Use } \frac{n_p^E}{p_n^B} = \frac{n_n^B}{p_p^E}\right)$$

$$= \left[1+\frac{18.13\times1.08\times10^{-2}\times10^{15}}{11.66\times1.35\times10^{-3}\times10^{17}}\tanh 9.26\times10^{-3}\right]^{-1}$$

$$= 0.9988$$

$$\beta = \operatorname{sech}\frac{W_b}{L_p} = \operatorname{sech} 9.26\times10^{-3} = 0.99996$$

$$\alpha = \beta\gamma = (0.9988)(0.99996) = 0.9988$$

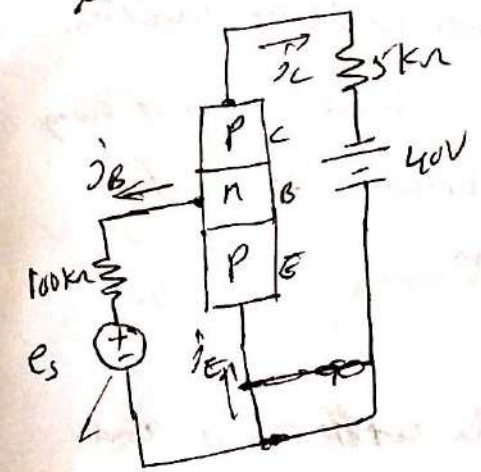
$$\beta = \frac{\alpha}{1-\alpha} = \frac{0.9988}{0.0012} = 832$$

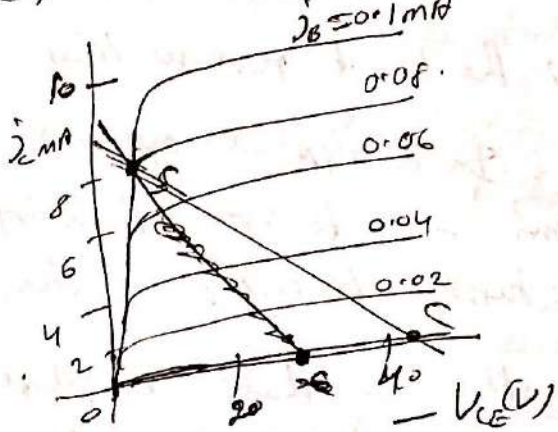
Switching:

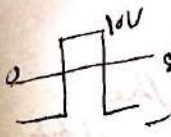
In switching operation transistor operates in two states, namely on-state and off-state. Ideally a switch should appear as short circuit when turned on and an open circuit when turned off. Further ~~the switch~~ ~~ideal switch must be able be~~. It is desirable to move from one state to another with no lost time in between.

* However transistor switches can be approximated to practical electronic circuits. (They cannot ~~correct~~ fit as ideal switches)

The two states of the transistor in switching can be seen in common emitter configuration as in Fig.(1a) In this figure i_C is controlled by the base current i_B. The load line specifies the locus of allowable (i_C, V_{CE}) points for the circuit



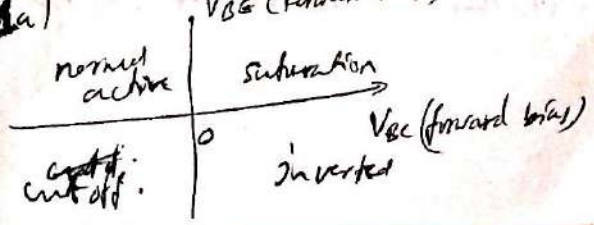




(1a)

(1b)

Fig(1) simple switching circuit in the CE configuration
(a) Biasing ckt
(b) Collector ckt's & load line with sat & cutoff indicated
(c) operating regions of BJT.



(1c)

If i_B is such that the operating point lies between two end points of the load line (Fig (1b)), the transistor operates in the normal active mode. i.e emitter junction is forward biased & collector is reverse biased, with reasonable i_B flowing out of the base.

If base current is zero or -ve, the point 'C' is reached at the bottom end of the load line & collector current is negligible. This is "off-state" of the transistor, & device is operating in the cutoff regime.

If base current is +ve & sufficiently large, and the device is driven to saturation region, marked 'S'. This is "on-state" of the transistor, in which a large value of i_C flows with only very small voltage drop V_{CE}.

Note: The beginning of saturation regime corresponds to the loss of reverse bias across collector junction.

In a typical operation the base current swings from +ve to -ve thereby driving device from saturation to cutoff & vice versa.

Next: we shall explore

(i) nature of conduction in the cutoff and saturation region.

(ii) Factors affecting the speed with which the transistor can be switched between the two states.

Different regions of operation of BJT are illustrated in Fig. (1c)

normal active mode: Emitter junction is forward biased & collector junction is reverse biased

Inverted mode: Emitter junction is reverse biased and collector junction is forward biased.

Cut off region: Both junctions are reverse biased leading to very high impedance state of BJT.

Saturation region: Both junctions are forward biased & low impedance state of BJT.

Cutoff:

If J_E is reverse biased in the cutoff regions ($-ve\ i_B$), we can approximate the excess hole concentration at the edge of the reverse-biased collector & emitter junctions as

$$\frac{\Delta P_E}{P_n} \approx \frac{\Delta P_C}{P_n} \approx -1 \qquad \text{(15)}$$

which implies $P(x_n) = 0$. The excess hole distribution in the base is approximated constant at $-P_n$.

The base current i_B can be approximated for a symmetrical transistor as a charge storage basis as $-q A P_n W_b / \tau_p$. The -ve excess hole concentration corresponds to generation in the same way that a +ve distribution indicates recombination.

physically, a small saturation current flows from n to p in each reverse biased junction & is supplied by the base current i_B (which is -ve when flowing into the base as per our definition)

pr 0

- General evaluation of currents can be obtained from Ebers-Moll equations.

By applying eq (15) to eq (11)

$$i_E = \frac{D_E}{P_n}\left(\Delta P_E - \alpha_N \Delta P_C\right)$$

$$\frac{\Delta P_E}{P_n} \simeq \frac{\Delta P_C}{P_n} \simeq -1$$

$$i_E = -I_{ES} + \alpha_I$$

$$I_E = D_{ES} \cdot \frac{\Delta P_E}{P_n} - \alpha_I D_{ES} \cdot \frac{\Delta P_C}{P_n}$$

$$I_E = -D_{ES} + \alpha_I D_{CS} = -(1-\alpha_N) D_{ES} \quad -(16a)$$

$$I_C = -D_{ES} + \alpha_I D_{CS}$$

$$D_C = \alpha_N D_{ES} \frac{\Delta P_E}{P_n} - D_{CS} \frac{\Delta P_C}{P_n}$$

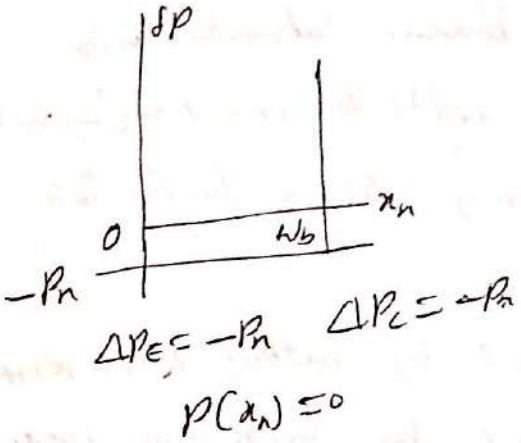
$$i_C = -\alpha_N D_{ES} + D_{CS} = (1-\alpha_I) D_{CS} \quad -(16b)$$

$$i_B = I_E - i_C = -(1-\alpha_N) D_{ES} - (1-\alpha_I) D_{CS} \quad -(16c)$$

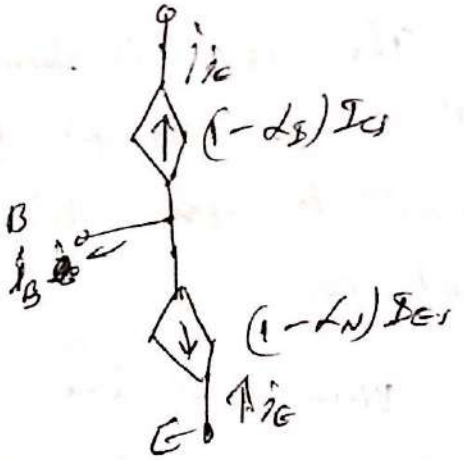
In the steady-state circuit saturation currents D_{ES} & D_{CS} are small and α_N & α_I are both near unity. Now currents will be negligible & cutoff region will closely approximate the off condition of the ideal switch.

The equivalent circuit corresponding to eq (10) is shown in (2a) (2b)

PTO



(2a)



(2b)

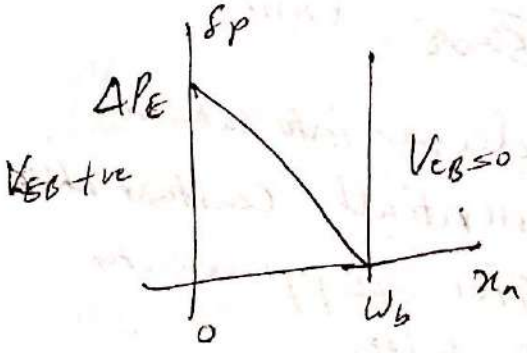
Fig.(2) The active regime of pnp transistor

(a) Excess hole distribution with both J_E & J_C reverse biased

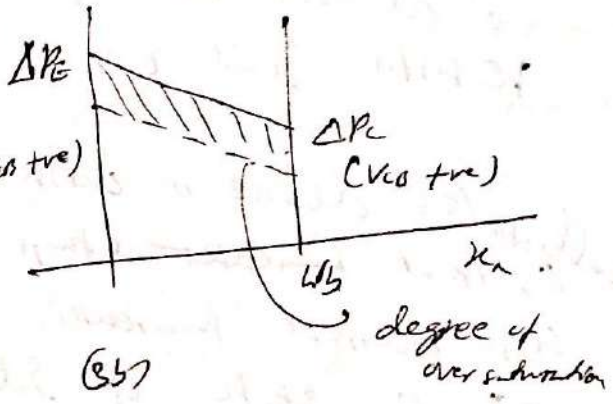
(b) Equivalent circuit corresponding to eq (10)

Saturation:

This region begins when reverse bias across collector junction is reduced to zero, & continues as collector becomes forward biased. The excess hole distribution is shown in Fig. (3).



(3a)



(3b)

Fig.(3) Excess hole distribution in base of saturated transistor

(a) Begining of saturation (b) oversaturation.

P.T.O

The device is ~~forward biased~~ saturated when $\Delta P_c = 0$ & forward bias of collector junction (Fig(3)) leads to +ve ΔP_c, driving device further into saturation.

When load line is fixed by battery & $5k$ resistor in Fig(1) saturation is reached by increasing base current i_B. With increase in i_B (since the stored charge required to accommodate a given i_B (& vice versa)) calls for an increase in the area under the $\Delta P(x_n)$ distribution.

In Fig(3a) The device has just reached saturation, & collector junction has zero bias, & very little voltage drop appears across the device from collector to emitter. The magnitude of V_{CE} is only a fraction of a volt.

$\therefore$ almost all battery voltage drop across resistor and $I_c = \frac{10V}{5k\Omega} = 8mA$.

Collector current ~~increases~~ stays essentially constant while base current increases. This approximates "On" state of ideal switch.

The Switching cycle:

Mechanism of switching cycle are shown in Fig. (4). If the device is in cutoff, a step increase of base current to I_B causes the hole distribution to ~~increase~~ increase — approximately as in Fig. (4b). For simplicity assume that distribution maintains a simple form in each time interval of transient.

At time t_0 the device enters saturation & the hole distribution reaches its final state at t_2. As stored charge in the base Q_b increases, there is increase in the collector current i_c. (The ~~Here~~ However, the collector current does not increase beyond its value at the beginning of saturation t_0).

Approximate the saturated collector current as $I_C \simeq E_{CC}/R_L$ where E_{CC} is the value of the collector circuit battery & R_L is the load resistor

There is an exponential increase in the collector current while Q_b rises to its value Q_s at t_s; this time serves as ~~a~~ ~~of the~~ limitations of the transistor in a switching application. Similarly when base current is switched negative (say $-I_B$ in Fig (4c)), the stored charge must be withdrawn from the base before cutoff is reached.

While Q_b is $\geq Q_s$, the collector current remains at the value I_C, (and is decided by battery & RC resistor). There is storage delay time t_{sd} after the base current is switched & before i_C begins to fall toward zero.

After the stored charge is reduced below Q_s, i_C drops exponentially with chip's fall time.

After stored charge is withdrawn, the base current decays to small cutoff value.

Specifications for switching Transistor:

It is possible to determine t_s and t_{sd} by solving for the time-dependent base current $i_B(t)$.

writing expression in terms of

$$i_B(t) = \frac{Q_p(t)}{\tau_p} + \frac{d\,Q_p(t)}{dt}$$

Do not neglect the charging time of the emitter junction capacitance in going from cutoff to saturation.

* Since emitter junction is reverse biased in cutoff, it is necessary for the emitter space charge layer to be changed to forward bias condition, before i_c can flow.

Hence include delay time t_d as in Fig(5).

Rise time t_r defined as the time required for the collector current to rise from 10 to 90 % of its final value.

- Another specification is the fall time t_f required for i_c to fall through ~~a similar~~ from 90 % of its maximum value to 10 % of its maximum value.

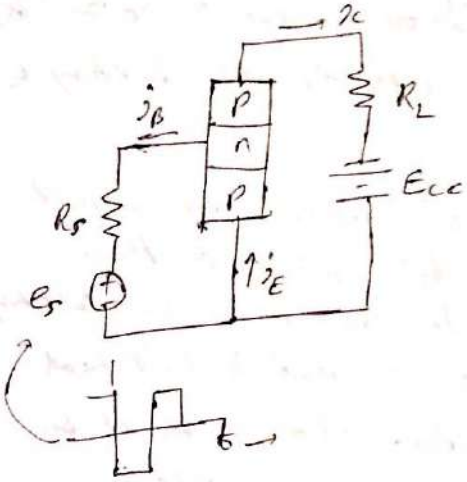


Fig. (4a)

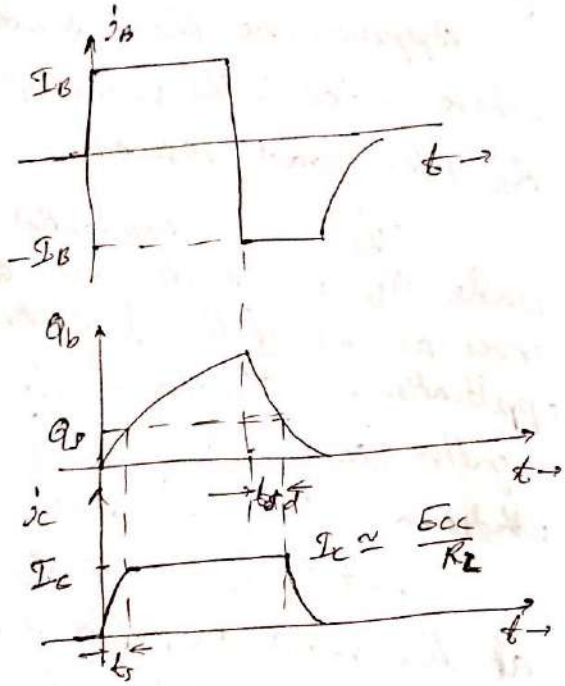


Fig. (4c)

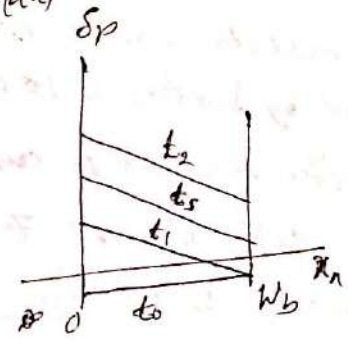
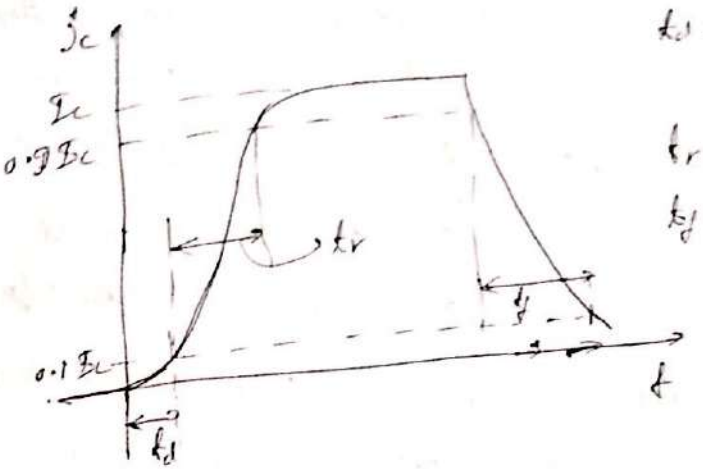


Fig. (4b)

Fig. (4) Switching effects in CE transistor. (a) Circuit diagram, (b) Approximate hole distribution in the base during cutoff to saturation. (c) base current, stored charge, & collector current during turn-on & turn-off transients.



t_d → Delay time while junction capacitance is charging

t_r → Rise time from 0.1 to 0.9 I_c

t_f → Fall time from 0.9 to 0.1 I_c

Fig (1)

Other important Effects:

In the previous analysis of transistor, a number of simplifying assumptions are made. Some of these assumptions need to modify in dealing with practical devices.

In this section we shall discuss non-uniformities of (i) non-uniform doping in base region

(ii) Effect of large reverse bias in collector junction.

(iii) Avalanche multiplication

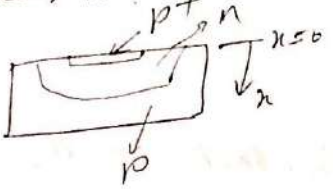
Drift in Base Region:

The assumption of uniform doping breaks down for implanted junction transistors, that involve impurity grading. In general the implanted transistor has a doping similar to the one shown in Fig. (6). There is fairly sharp discontinuity in the doping profile, when the donor concentration in the base region becomes smaller than the constant p-type background doping in the collector.

P.T.O

Similarly, the emitter is assumed to be heavily doped (p^+) providing a second sharp boundary for the base. Within base, the net doping concentration $(N_d - N_a) = N$ varies along a profile that decreases from emitter edge to the collector edge.

We see from Fig (b) the effect of impurity gradient, $N(x_n)$ varies exponentially with in base region.



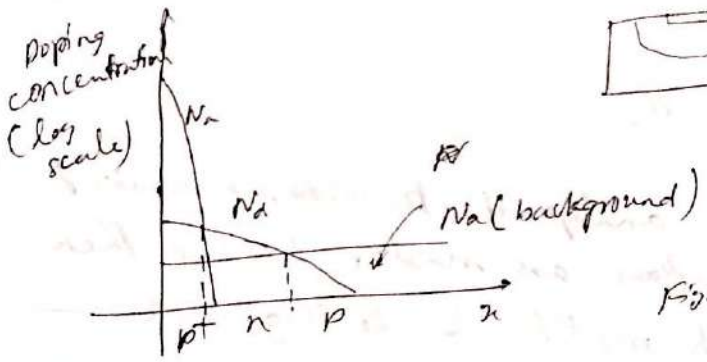


Fig. (6a)

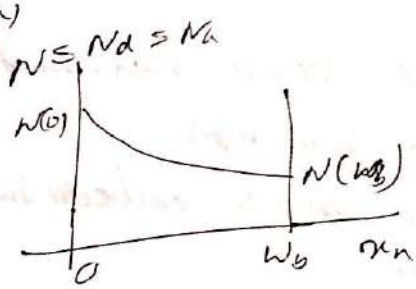


Fig. (6b)

Fig. (6) Graded doping in the base region of a p-n-p transistor
(a) Typical doping profile
(b) Approximate exponential distribution of net donor in base.

The result of a graded base region is that a built-in electric field exists from emitter to collector (for a p-n-p) ~~Here by adds~~. This adds drift component to the transport of holes across the base.

We shall demonstrate this effect simply by considering balance of drift & diffusion in the base at equilibrium.

If net doping of the base is large enough to allow the approximation $n(x_n) \simeq N(x_n)$, the balance of electron drift and diffusion currents at equilibrium requires

$$J_n(x_n) = q\mu_n N(x_n) E(x_n) + qD_n \frac{dN(x_n)}{dx_n} = 0 \quad \text{---(17)}$$

∴ The built in electric field is

$$E(x_n) = -\frac{D_n}{\mu_n} \frac{1}{N(x_n)} \frac{dN(x_n)}{dx_n}$$

$$= \frac{-kT}{q} \cdot \frac{1}{N(x_n)} \cdot \frac{dN(x_n)}{dx_n} \quad \text{---(18)}$$

For a doping profile $N(x_n)$, that decreases in +ve x_n direction, this field is +ve, directed from ~~the~~ emitter to collector

* For exponential doping profile, the electric field $E(x_n)$ turns out to be constant with position in the base. we can represent exponential distribution as

$$N(x_n) = N(0)\, e^{-ax_n/W_b} \quad \text{where } a \equiv \ln \frac{N(0)}{N(W_b)} \quad \text{---(19)}$$

Taking derivative of this distribution and substituting in eq (18) we obtain constant field

$$E(x_n) = \frac{kT}{q} \frac{a}{W_b} \quad \text{---(20)}$$

This field aids the transport of holes across the base region from emitter to collector, thus transit time τ_f is reduced below that of a comparable uniform base ~~transistor~~ transistor.

Is important for high freq devices.

Base Narrowing:

So far we have assumed base width W_b is independent of bias voltages applied to J_E & J_C junctions. And is not always valid.

Eg: pt-n-pt transistor of Fig (7) is affected by the reverse bias applied to the collector.

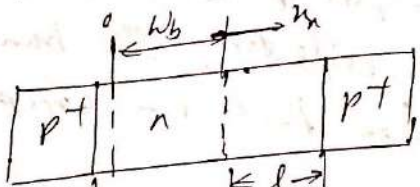
P.T.O

If the base is lightly doped, the depletion region at the reverse biased collector junction can extend significantly into the n-type base region, which effectively reduces the base width W_b. This effect is called "base narrowing" / "base-width modulation" or Early effect (after J.M. Early).

* The effect of base narrowing is visible in the collector characteristics for CE configuration. [Fig. (25).]

The decrease of W_b causes β to decrease, as a result collector current I_c increases with collector voltage. (Rather than staying as constant).

* The slope introduced by the Early effect is almost linear, with I_c, and common-emitter characteristics extrapolated to an intersection V_B on the voltage axis, called Early voltage.



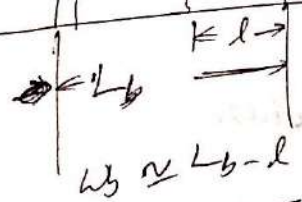
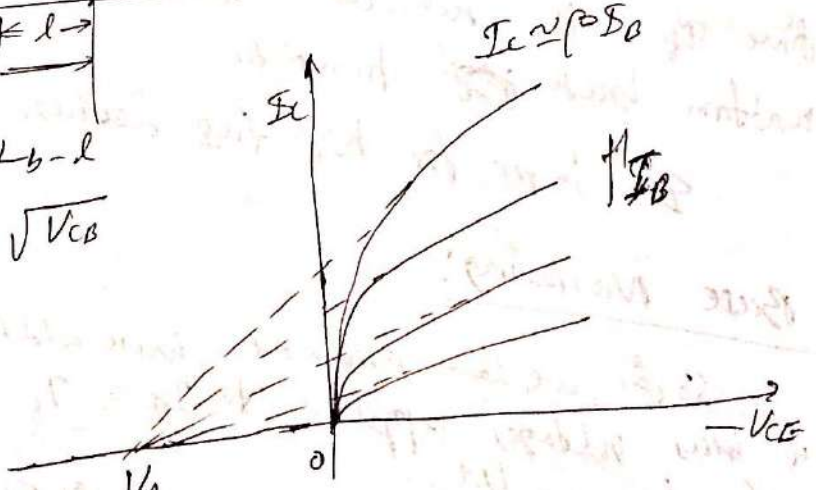


Fig. (24)



(25)

For p+-n-p+ device of Fig.(7) we can approximate the length 'l' of the collector junction depletion region in the n-material with V_0 replaced by $V_0 - V_{CB}$.

$$\therefore l = \left(\frac{2\epsilon V_{BC}}{q N_d}\right)^{1/2} \quad \text{---(21)}$$

If reverse bias of collector junction is increased, it decreases W_b, and is possible that collector depletion region fills entire base. This is called punch through condition. And holes are swept directly from emitter to collector & transistor action is lost.

punch through is a breakdown effect generally avoided in circuit design.

In most cases avalanche breakdown of collector junction occurs before punch through is reached.

Avalanche Breakdown:

The avalanche multiplication at the collector junction becomes important & collector current increases sharply at a well-defined breakdown voltage BV_{CBO} for common base configuration.

For CE configuration, there is a strong influence of carrier multiplication over fairly broad range of collector voltage. The breakdown voltage in CE configuration is BV_{CEO} & is much smaller than BV_{CBO}.

To understand these effects, consider breakdown for the condition $I_E = 0$ in the ~~common~~ CB base case & $I_B = 0$ in CE case. The are implied by '0' in BV_{CEO} & BV_{CBO}.

* In each case, terminal current I_c is the current entering the collector depletion region multiplied by the factor M, is given by

$$I_c = (\alpha_N I_E + I_{CO}) M \quad \text{——(22)}$$

$$= (\alpha_N I_E + I_{CO}) \frac{1}{(1 - V_{BC}/BV_{CBO})^n}$$

(Includes multiplication due to impact ionization)

it an empirical expression

— For limiting case of $I_E = 0$. The lowest curve in Fig. (b), I_c is simply $M I_{CO}$, & breakdown voltage is well defined.

BV_{CBO} → signifies the collector junction breakdown voltage in CB with emitter open.

In CE case setting $I_B = 0$ & therefore

$I_c = I_E$ in eq (21)

we have

$$I_c = \frac{M I_{CO}}{1 - M \alpha_N} \quad \text{——(23)}$$

ie. collector current increases indefinitely when $M\alpha_N$ approaches unity.

On the other hand M must approach infinity in (CB) common base case before BV_{CBO} is reached.

Since α_N is close to unity in most transistors, M need to be only a slightly larger than unity, for eq (23) to approach breakdown.

∴ Avalanche multiplication dominates the current in CE transistor well below the breakdown voltage of isolated collector junction.

∴ BV_{CEO} is smaller than BV_{CBO}.

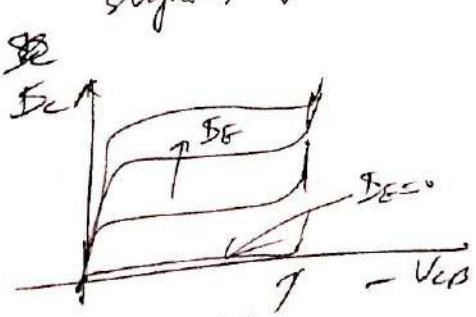
Physical Understanding: Importance of multiplication in CE configuration by considering effect of M on base current.

When ionizing collision occurs in the collector junction depletion region, secondary hole & electrons are created. The primary & secondary holes are swept into collector & in a p-n-p, & electron is swept into the base by the junction field

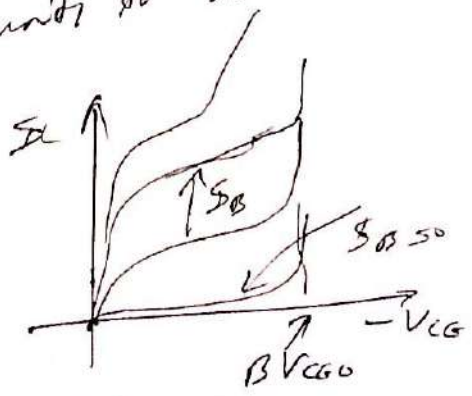
∴ The supply of electrons to the base increases & from charge control analysis we conclude that hole charge in the section at the emitter must increase to maintain charge neutrality. This is regenerative process, in which, increased injection of holes from emitter causes an increased multiplication current at the collector junction.

This increases the rate of secondary electrons are swept into the base, calling for more hole injection.

That is why multiplication factor M are need to be slightly greater than unity to start avalanche process.



8(a)



8 (b)

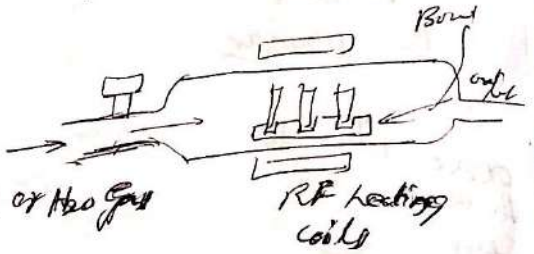
Fabrication of P-n Junctions:

So far we have learnt little about how the devices work. Now we shall see how the devices are made. (specifically p-n junctions which are fundamental to the performance of functions such as rectification, amplification, switching & other operations)

The basic processing required for fabricating devices are (i) ~~How~~ doping ~~can be varied as a function of depth to laterally across the surface~~ (ii) Thermal oxidation, (iii) diffusion (iv) Rapid thermal oxidation, (v) Ion implantation (vi) chemical vapour deposition (CVD) (vii) photolithography (viii) Etching (ix) Metallization

Thermal oxidation:

Many Fabrication steps involve heating up the wafer to enhance chemical process. Eg: Thermal oxidation of Si to SiO_2. This involves placing a batch of wafers in a quartz tube which can be heated to (800 – 1000°C) using heating coils. An oxygen containing gas such as dry O_2 or H_2O is flowed into the tube at atmospheric pressure, & flowed at other end.



normally horizontal furnaces are used. However vertical surfaces are employed recently.

The ~~the~~ Si wafers are placed in ~~silicat~~ silica (quartz) holder called boat. & Then they are moved into the furnace at the specified location. (where required temp is maintained) The gas flow is into the furnace from one end & flow out at the other end.

P.T.O

The following reactions occur during oxidation are

$$Si + O_2 \longrightarrow SiO_2 \text{ (dry oxidation)}$$
$$Si + 2H_2O \longrightarrow SiO_2 + 2H_2 \text{ (wet oxidation)}$$

In both cases Si is consumed from the surface of substrate. For every micron of SiO_2 grown, 0.44 μm of Si is consumed. leading to volume expansion upon oxidation.

The oxidation proceeds by having the oxidant (O_2 or H_2O) molecules diffuse through the already grown oxide to the Si-SiO_2 surface interface, where the above reactions take place.

* One of the important reason why Si IC's exists is because of stable thermal oxide can be grown on Si with excellent electrical properties.

* In other semiconductors it is not possible to grow native oxide

oxide thickness is a function of both time & temperature

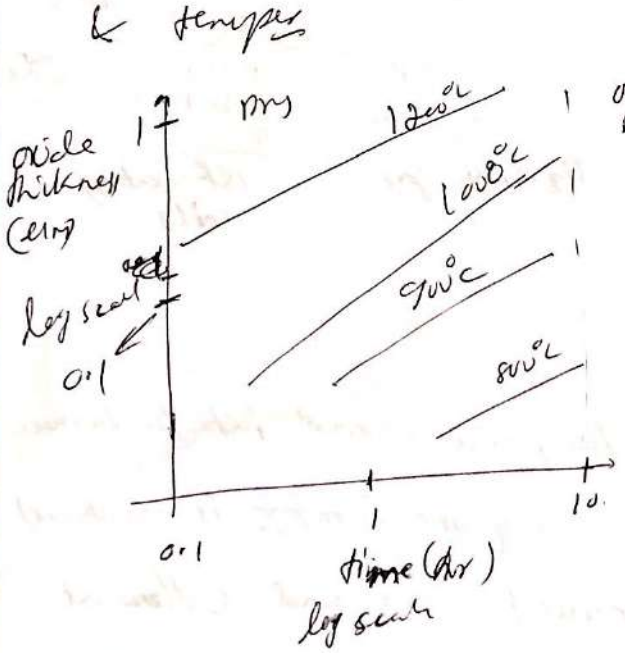


Fig (1a)

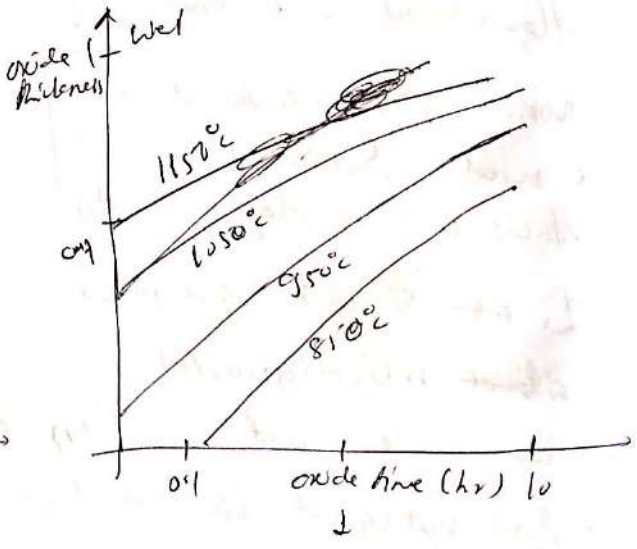


Fig (1-b)

Diffusion.

This is another thermal process used for diffusion of dopants in furnaces. Here wafers are first oxidized and windows are opened in the oxide using photolithography & etching

Dopants (Eg: B, P or As) is introduced into the patterned wafers at high temp^{re} (800 – 1100°C)

The dopants are diffused generally using a gas or vapour source.

The dopants are transported from high concentration region near surface into the substrate through diffusion

(The maximum impurities that can be dissolved is limited by the solid solubility of material).

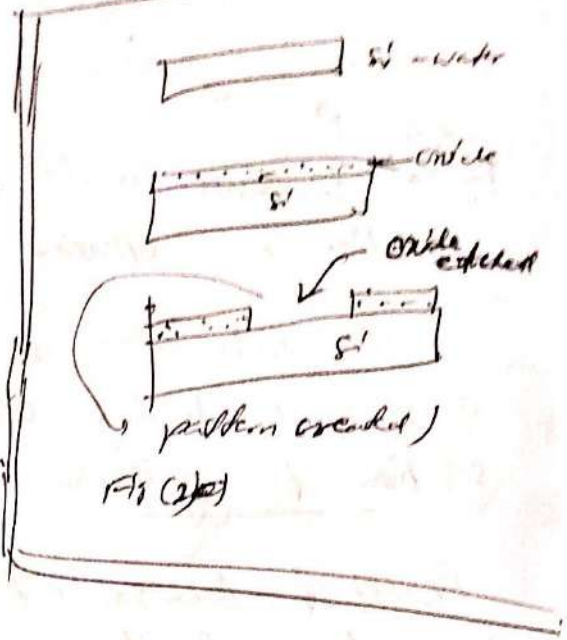


Fig (2)(a)

- The diffusivity of dopants in solids (D) is strongly dependent on temp^{re} (T). & is given by

$$D = D_0 e^{(E_A/kT)},$$

where D_0 is a doping constant dependent on material & dopant.

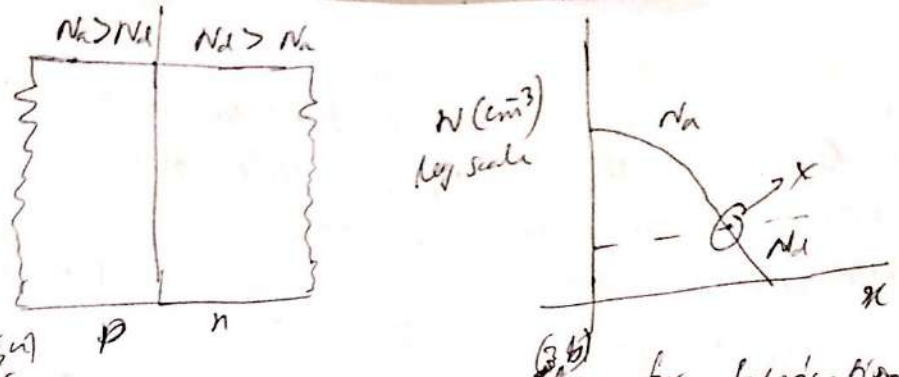
E_A → Activation energy.

- The Average distance the dopants diffuse is related to diffusion length. In this case diffusion length is $\sqrt{Dt}$

where 't' is processing time

product Dt → is called thermal budget

- The diffusivity varies exponentially with T, (increases with T), hence controlling precisely the furnace temp^{re} (within several degrees), in order to have control over diffusion profile is important.
- The dopants are blocked by oxide (acts as mask layer) since diffusivity is very low).



(3a) (3b)

Fig (3a) Impurity concentration profile for fabrication of p-n junction by diffusion.

— The distribution of impurities in the sample at any time during diffusion can be calculated from a solution of diffusion eq. with boundary conditions.

Eg: (1) If the source of dopant atoms at the surface of sample is limited (given no of atoms deposited on the surface before diffusion), a gaussian distribution is used.

$$\delta_{pp}(x,t) = \left[\frac{\Delta p}{2\sqrt{\pi D_p t}}\right] e^{-x^2/4D_p t} \quad \text{—— (1)}$$

when $\Delta p \rightarrow$ no of holes/unit area created over a negligibly small distance at $t = 0$

$t \rightarrow$ time

$D_p \rightarrow$ ~~diffusivity of hole~~ hole diffusion co-efficient

$x \rightarrow$ distance from surface

(2) If the the dopant atoms are supplied continuously, concentration at the surface is maintained at a constant value, then distribution follows complementary error function. $N(x,t) = N_0\ erfc\left(\frac{x}{2\sqrt{Dt}}\right)$

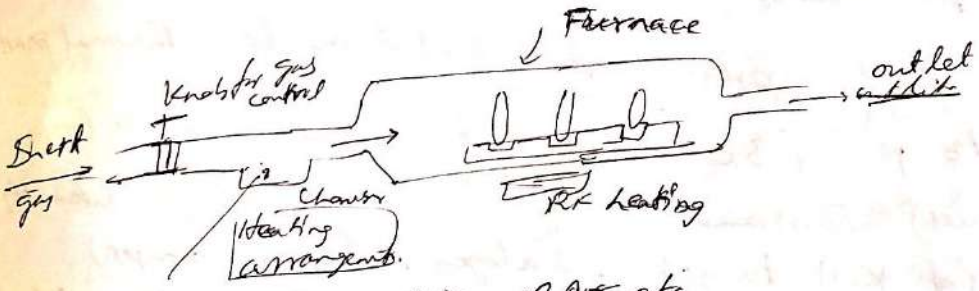
In Fig. (3b), at point 'x' the introduced acceptor concentration ~~in the original~~ is equal to, donar concentration of the original n-type sample.

This point is the location of p-n junction. To the left of this junction acceptors are predominate & material is p-type. whereas to the right of the junction, donor atoms are predominate & material is n-type.

- The depth of junction beneath surface can be controlled by time & temper of diffusion.

The common diffusion materials used for diffusion in Si are B_2O_3, BBr_3, & BCl_3 for boron, Phosphorous source include PH_3, P_2O_5 & $POCl_3$.

Solid sources are placed in a separate heating zone of the furnace. Gases sources (Inert gas) are passed through liquid dopants (After heating solids become liquid, & then passed through the furnace tube. The Si wafers are pushed into the furnace & removed by silica rod.



Dopants such as B_2O_3, P_2O_5 etc

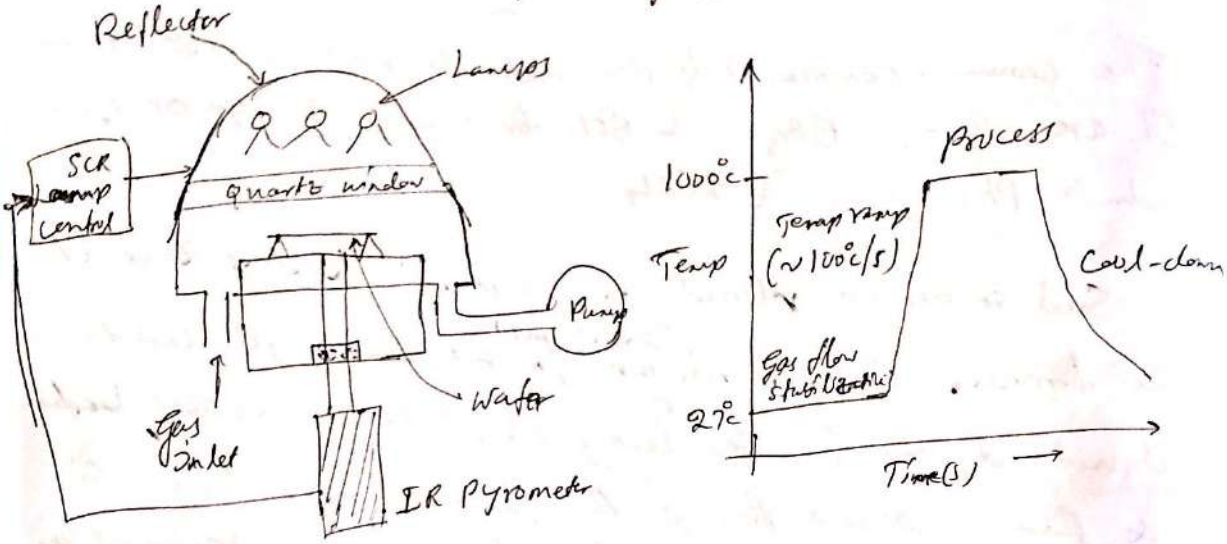
note: All these processes must take place in a clean room, since doping concentrations present one part per million or less.

* Even the purity of material & gas it is must be extremely pure
* Silica tube, sample holder & pushing rod must be cleaned & etched. in hydrofluoric acid (HF) before use
* Si wafers also must undergo various cleaning procedure before diffusion.

Rapid Thermal Processing: (RTP)

Many thermal steps (Heating procedure) done using furnace are now done using rapid thermal processing (RTP). This include, rapid thermal oxidation, annealing of ion implantation & chemical vapour deposition.

Fig.(4) shows the simple ~~system for~~ RTP system.



Fig(4) Schematic diagram of a rapid thermal processor, and time ~~temp~~ temp graph.

Here single ~~pcs~~ wafer is held on low-thermal mass quartz pins (Instead of batch of wafers, in like in furnaces). Surrounded by bank of high-intensity (Tens of ~~Kw~~) tungsten-halogen infrared lamps) ~~tungsten~~ tungsten-halogen infrared lamps, with gold plated reflectors around them.

By turning on lamps high-intensity infrared ~~radiation~~ radiation shines through quartz chamber. ~~It~~ This is absorbed by wafer, to rise its temp rapidly (~ 50-100°c/s). The processing temp reaches quickly, after gas flow is stabilized in chamber. At the end lamps are turned off, allowing wafer temp to drop rapidly.

Thermocouples or pyrometers are used to accurately measure temp. It is imp to maintain temp uniformly across the wafer.

- The key parameter in all thermal processing steps is the thermal budget Dt.

  where D is diffusion co-efficient
  t is diffusion time.

Generally Dt is to be maintained small to avoid loss of control over doping profile.

Ion Implantation.

In this process energetic ions are implanted into the semiconductor (SC). A beam of impurity ions is accelerated to kinetic energies from few KeV to several MeV & is directed on to the surface of SC. The impurity atoms give up their energy to the lattice in collisions and finally come to rest at average penetration depth, called projected range (Rp). This range may be varied depending on impurity & implantation energy. This range may vary from a few hundred angstroms to about 1 μm.

The ions distribute ~~symmetrically~~ evenly about the projected range as in Fig (5a). The distribution is ~~gaussian~~ for a dose of ϕ ions/cm² is gaussian, & is given by

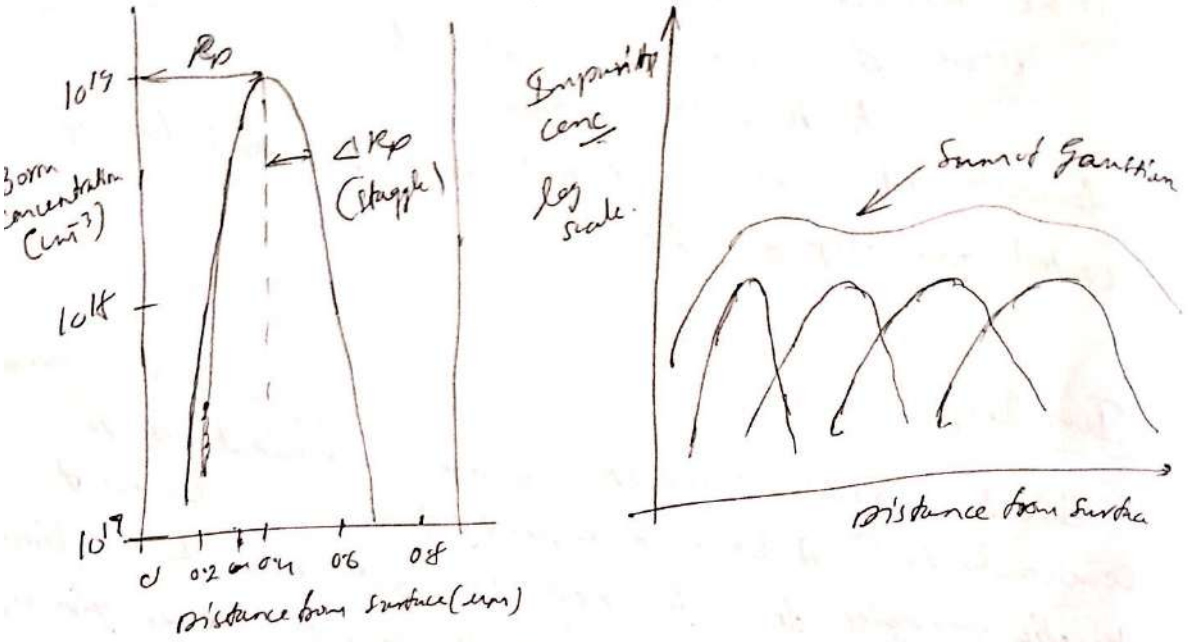
$$N(x) = \frac{\phi}{\sqrt{2\pi}\,\Delta Rp} \exp\left[-\frac{1}{2}\left(\frac{x - Rp}{\Delta Rp}\right)^2\right] \quad \text{— (1a)}$$

where ΔRp is called straggle measure half-width of the distribution at $e^{-1/2}$ of peak (Fig (5a))

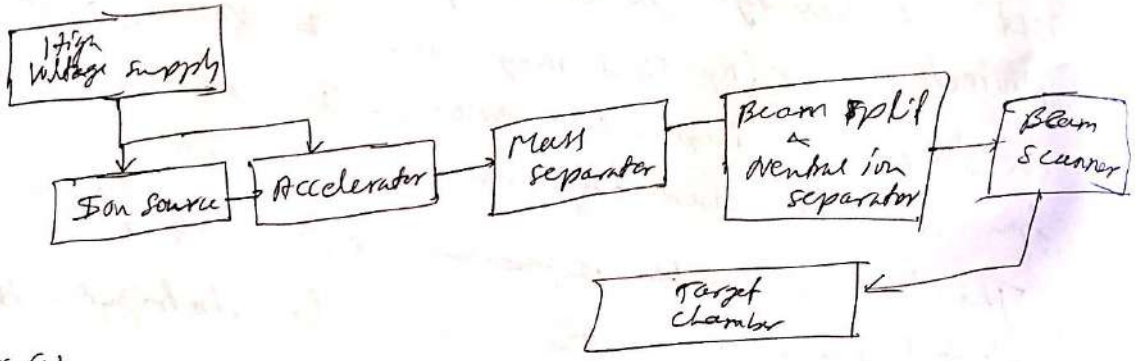
Both Rp & ΔRp increase with increasing implantation energy.

PTO

By performing several implantations at different energies, it is possible to obtain desired impurity profile as in Fig. (5b).



5(a) Gaussian distribution



Fig(b)

Block diagram of ion implanter is shown in Fig (6a). A gas containing desired impurity is ionized with in the source, & extracted into the acceleration tube, The ions accelerated ions are passed through mass separator, to ensure desired ion species to pass further. The ion beam is then focussed & scanned electrostatically over the surface of wafer. The target chamber includes automatic wafer handling facilities.

Adv. (1) can be done at relatively low temp (so previously diffused atoms will not get disturbed)

(2) Very shallow & well defined doping profiles can be achieved

(3) Thin doping layers can be achieved.

(4) precise control of doping conc is possible

Disadv

(1) Lattice gets damaged due to collisions between the ions & the lattice atoms.

* However can be removed by heating the crystal after implantation. This is called annealing.

* Annealing leads to unintended diffusion of implanted dopants. It can be minimized by controlling annealing temp & time.

The profile after annealing is

$$N(x) = \frac{\phi}{\sqrt{2\pi}\left(\Delta R_p^2 + 2Dt\right)^{1/2}} \cdot \exp\left[-\frac{1}{2}\left(\frac{(x-R_p)^2}{\Delta R_p^2 + 2Dt}\right)\right]$$

Chemical Vapor Deposition: (CVD)

At different stages of fabrication, thin films of dielectrics, Semiconductors & metals have to be formed on the wafer & then patterned and etched.

Eg: Thermal oxidation of Si.

SiO_2 films can also be formed by low pressure (~100 m Torr) chemical vapour deposition (LPCVD) or plasma enhanced CVD (PECVD).

Adv: The thermal oxidation consume Si from substrate & require high temp, whereas CVD of SiO_2 does not consume Si from substrate & can be done at lower temp

12-70

In CVD process Si-containing gas such as SiH_4 reacts with an oxygen-containing precursor, causing a chemical reaction, leads to deposition of SiO_2 on the substrate.

The technique of depositing SiO_2 is very important in many applications.

Ex: (i) In building complicated device structure, Si substrate may not be available for reaction.

(ii) There may be metallization on the wafer that cannot withstand high temp.

In such cases CVD is an alternative.

LPCVD is also used to deposit other dielectrics such as silicon nitride (Si_3N_4), polycrystalline or amorphous Si.

Vapor-phase Epitaxy (VPE) & metal-organic vapor-phase epitaxy (MOVPE) are other challenging examples of CVD, where single-crystal growth is maintained.

Photolithography:

Patterns corresponding to circuitry are formed on a wafer using photolithography.

1. The first step of photolithography involves generating a reticle (mask) which is transparent silica (Quartz) plate containing the pattern.

The opaque region on the mask made of ultraviolet-light absorbing layer (Cr, Iron oxide)

* Reticle typically contains patterns corresponding to a single chip or die
* Per mask contains patterns corresponding to a entire wafer.

mask / reticle are created by computer controlled electron beams, driven by circuit layout data.

* Thin layer of electron beam sensitive material (called electron beam resist) is placed on the iron-oxide-covered quartz plate. The resist is exposed by electron beam & undergoes chemical changes

* Resist is exposed selectively corresponding to the pattern required.

* After exposure, the resist is developed in a chemical solution.

Two types of resist

- positive resist : Exposed portion can be removed after exposure
- negative resist : Unexposed portion can be removed

* The iron oxide layer is then selectively etched off in a plasma to generate required pattern.

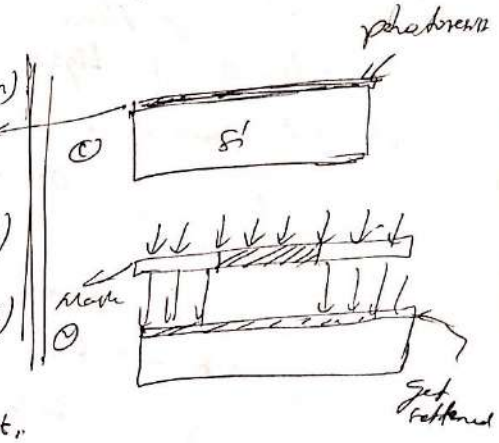
- These patterns can be used repeatedly to pattern Si wafer.

* To make IC's dozen or more reticles are required.

The Si wafers are first covered with photosensitive material called photo resist

Few drops are placed & spun to achieve uniform thickness (~0.1 μm)

photoresist: positive photoresist (PPR), negative photoresist (NPR)



Creates opposite polarity image on wafer compared to the mask.

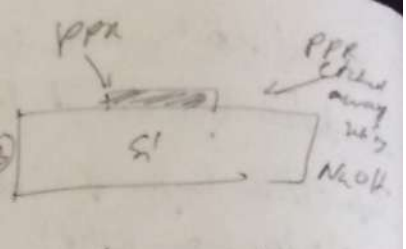
~~Compared to the positive photoresist.~~

* PPR can achieve better resolution, down to ~0.25 μm using UV light.

* In PPR exposed portion get softened after passing UV light & can be etched away by NaOH soln; the pattern on the reticle/mask is transferred to the die on the wafer.

PTO

; Then the remaining PPR can be baked at ~125°c to harden it, then further processing steps can be performed. (Eg. Implanting dopants, or plasma etching etc)



— The ~~exposure~~ exposure of the wafer is achieved die-by-die in a step & repeat system is called a stepper. (Fig (x)). Here UV light shines selectively through the reticle on to a single die. After the exposure, ~~the wafer mechanically — then translates on a precisely controlled~~ it moves on to the next die location & is exposed again. These tools are also called mask aligners.

The photolithography along with etching, determines how closely packed devices (Eg. transistors) can be made.

note:
1) Smaller the device → higher the speed of operation & lower power dissipation.

Challenges of LIthography.
(1) The pattern dimensions are comparable to the wavelength of light of ~~that~~ used. The diffraction-limited minimum geometry is given by

$$D_{min} = 0.8\lambda / NA \quad \text{——— (1)}$$

where λ → wavelength of light &
NA → (~0.5) is the numerical aperature or size of ~~the~~ the lens used in aligner.

note: (1) Smaller geometries require shorter wavelengths.
∴ UV lamp source (0.365 μm) can be replaced with argon fluoride (ArF) lasers (λ = 0.193 μm)

(2) novel exposure techniques, employing (i) phase-shift masks, (ii) optical proximity correction, & off-axis illumination can be used to allow resolution near or below the

dimension of the wavelength being used.

* Extreme ultraviolet source (13 nm), uses plasma to generate such short wavelengths for next generation lithography.

- Other important parameter in lithography is depth-of-focus (DOF), given by

$$DOF = \frac{\lambda}{2(NA)^2} \quad \text{—②}$$

Gives the range of distance around the focal plane where the image quality is sharp.

Pg. Stepper

Etching:

photoresist can be used as a mask to etch the material underneath.

* Earlier etching was done using wet chemicals.

Eg. dilute HF can be used to etch SiO_2 layer grown on Silicon substrate with excellent selectivity.

Selectivity here means HF attacks SiO_2 & does not affect the Si substrate underneath or photoresist mask.

- Most etchents are selective, however isotropic. That is they etch as fast laterally as they etch vertically.

* This is not acceptable for ultra-small features.

∴ Wet etching is supplimented by dry, plasma-based etching which can be made both selective & anisotropic. (etches vertically & not laterally)

* In modern IC processing wet etching is mainly used for cleaning of ∅ wafers.

Most popularly used plasma-based etching is Reactive Ion Etching (RIE) F_2(X1). Here appropriate etch gases such as chlorofluorocarbons (CFCs) flow into the chamber at reduced pressure (~1 – 100 mTorr) & plasma is struck by applying an RF voltage across cathode & anode. The RF voltage accelerate the light electrons in the system to higher kinetic energies (~10eV) than heavier ions. High energy elergy electrons collide with neutral atoms & molecules to create ions & molecular fragments (called radicals).

The wafers are held on the rf powered cathode, while the grounded chamber acts as anode.

Generally plasma is highly conducting, less conducting sheath region form next to two electrodes. The sheath voltage next to cathode can be increased by making cathode area smaller. A high dc voltage (~100 – 1000V) developed across the sheath next to cathode,

PTO

Hence +ve ions gain kinetic energy due to accelerate in RIS region, & bombard the wafer normal to the substrate. This bombardment contributes to etch physical component anisotropically. [This physical etching is unselective, The highly reactive radicals in the system give rise to a chemical etch component; i.e. very selective & real anisotropic. The result of RIE achieves a good compromise between anisotrophy & selectivity.]

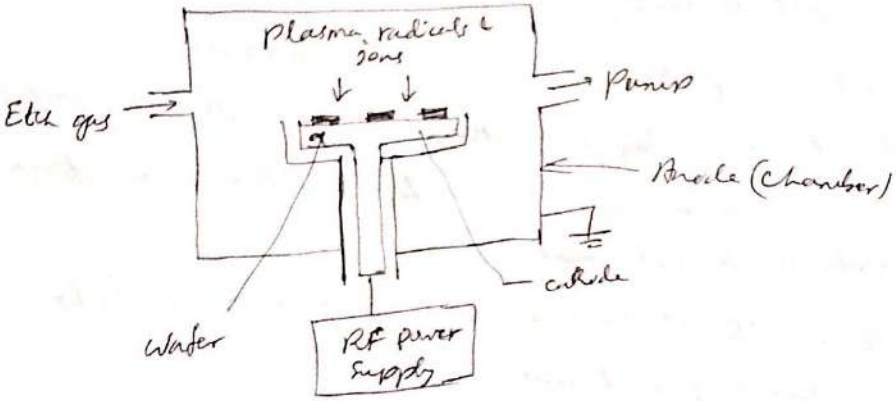


Fig (X1) Reactive ion etcher. (Blade etcher)
[RF freq = 13.56 MHz].

Metallization:

After the semiconductor devices are made in an IC, they need to be connected to each other, by metallization. Metals films are deposited by a physical vapor deposition technique such as evaporation (Eg. Au on GaAs) or sputtering (Eg. Al on Si).

Sputtering of Al is achieved by immersing an Al target (Al alloys with 1% Si, & 4% Cu for good electrical & metallurgical properties). in an Ar plasma. Argon ions bombard the Al & physically dislodge Al atoms by momentum transfer Fig (X2).

Many Al atoms eject from target & deposit on Si wafer. The Al is then patterned using metallization reticle & subsequently etched by RIE, & sintered at 450°C for 30 minutes, to form good electrical, ohmic contact to the Si.

Recently Si integrated started using low resistivity Cu metallization instead of Al. And it deposited using electroplating procedure.

Drawback: (1) Cu diffuses fast in Si and acts as a deep trap.

(2) To avoid this, a barrier metal such as Ti is first sputtered deposited to block Cu diffusion into Si.

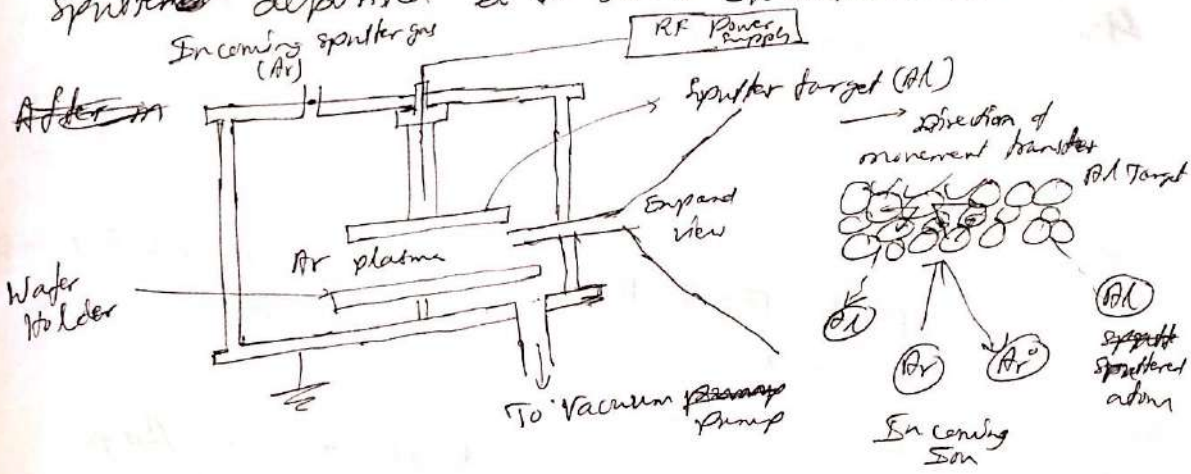
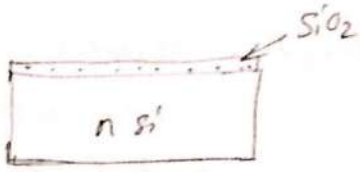


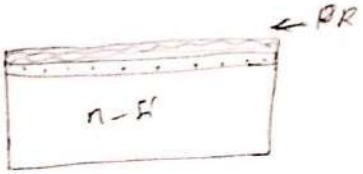
Fig. (X2)

After completing metallization, protective silicon nitride layer is deposited using PECVD. Then individual integrated circuits can be separated by sawing or by scribing & breaking the wafer. & finally packaged.

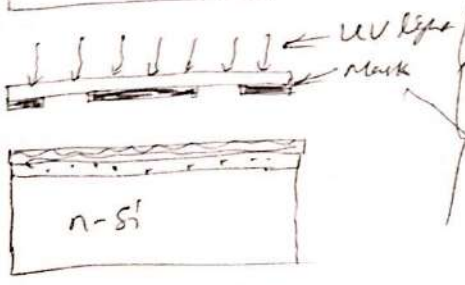
The main steps necessary to make p-n junction is illustrated in Fig. (X3).

1. 

oxidize the Si sample

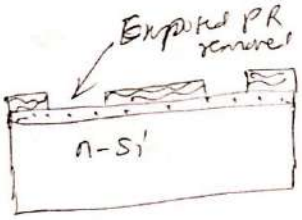
2. 

Apply a layer of the photo resist (PR)

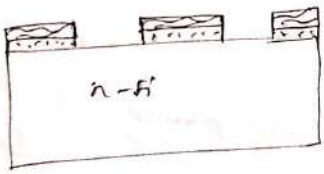
3. 

Expose PR through mask A.

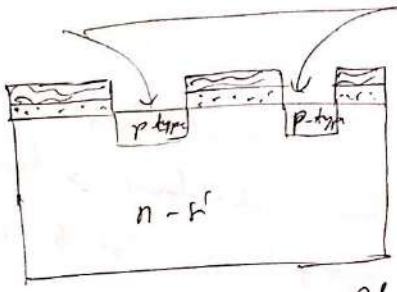

Mask-A (~~dopant~~ doping)

4. 

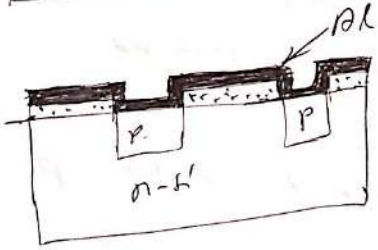
Remove exposed PR.

5. 

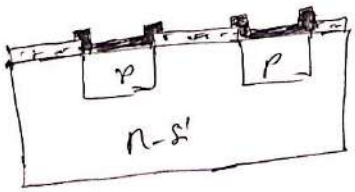
Use RIE to remove SiO_2 in windows.

6. 

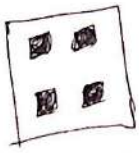
Implanted boron through windows in PR & SiO_2 layer

7. 

Remove PR and sputter Al onto the surface.

8. 

Using PR and mask B repeat steps 2-4, etch away Al except in p contact areas.


Mask B

Prob: Phosphorus is diffused into Si from an infinite source at 1000°c for 30 minutes. Assuming constant diffusivity.
(a) Find out the total amount of phosphorus that has gone into silica (b) After step (a), the source is shut off and the sample is subjected to drive in at 1200°c. If the final surface concentration is to be maintained at $5 \times 10^{19}/cm^3$, what should be the duration of drive-in?
(c) If the original substrate doping was $10^{15}/cm^3$, what is the junction depth after step (b).
(d) Assuming constant mobility $\mu_n = 600\ cm^2/V\text{-}sec$, find the sheet ~~resistance~~ resistivity of the diffused layer after step (b).

Given $D_{1000} = 3 \times 10^{-14}\ cm^2/sec$, $D_{1200} = 2.5 \times 10^{-12}\ cm^2/sec$
Solid solubility of Ph at 1000°c $= 10^{21}/cm^3$.

Any Q (a) $= 2N_s \sqrt{\frac{Dt}{\pi}}$

$= 2 \times 10^{21} \sqrt{\frac{3 \times 10^{-14} \times 0.5 \times 60 \times 60}{\pi}}$

$= 8.28 \times 10^{15}/cm^2$

$N_s \rightarrow /cm^3$
$Dt = cm^2/sec \times sec$

(b) We have

$N(x,t) = \frac{Q}{\sqrt{\pi D_2 t_2}} \times \exp\left(\frac{x^2}{4 D_2 t_2}\right)$ — N_B

At surface $x=0$, $N \rightarrow N_s$

ie $N_s(0,t) + N_B = \frac{Q}{\sqrt{\pi D_2 t_2}}$

N_s

$\Rightarrow \sqrt{\pi D_2 t_2} = \frac{Q}{N_s + N_B}$

$t_2 = \frac{1}{\pi D_2} \times \left[\frac{Q}{N_s + N_B}\right]^2 = \frac{1}{\pi \cdot 2.5 \times 10^{-12}} \left[\frac{8.28 \times 10^{15}}{5 \times 10^{19} + 10^{15}}\right]$

$= 3490$ sec ⟶ verify

(c) $$\frac{Q}{\sqrt{\pi D_2 t_2}} \exp\left(\frac{-x_j}{4 D_2 t_2}\right) = N_B$$

i.e $$-\frac{x_j^2}{4 D_2 t_2} = \ln\left(\frac{N_B \cdot \sqrt{\pi D_2 t_2}}{Q}\right)$$

or $$\frac{x_j^2}{4 D_2 t_2} = \ln\left[\frac{Q}{N_B \sqrt{\pi D_2 t_2}}\right]$$

i.e. $$x_j = \sqrt{4 D_2 t_2 \cdot \ln\left[\frac{Q}{N_B \sqrt{\pi D_2 t_2}}\right]}$$

$$= \sqrt{4 \times 2.5 \times 10^{-12} \times 3490 \ln\left[\frac{8.28 \times 10^{15}}{10^{15}}\right]}$$

$$x_j = \sqrt{4 \times 2.5 \times 10^{-12} \times 3490 \ln\left[\frac{8.28 \times 10^{11}}{10^{11} \sqrt{\pi \times 2.5 \times 10^{-12} \times 3490}}\right]}$$

$$= 0.614 \times 10^{-3} \text{ cm}$$ verify

(2) Find the time required to form a PN junction from the surface by performing a constant source solid solubility limited boron diffusion into a background phosphorus concentration of $1 \times 10^{16}/\text{cm}^3$ at 900°C & 1100°C. Given $D_0 = 10.5 \text{ cm}^2/\text{sec}$, $E = 3.69 \text{ eV}$ & solid solubility is 10^{20} (at 900°C) & 2×10^{20} (at 1100°C)

Any diffusion co-efficient $D = D_0 e^{-E/kT}$ $\quad e^1 (2.71) = 0.9999$

$T = 900 + 273 = 1173^\circ k$ (at 900°C)

$\therefore D = 10.5 \times e^{\left(\frac{-3.69 \text{ ev}}{8.62 \times 10^{-5} \times 1173}\right)}$

$= 1.47 \times 10^{-11} \text{ cm}^2/\text{sec}$

$K = 1.38 \times 10^{-23} \text{ J/k}$
or $= 8.62 \times 10^{-5} \text{ ev/k}$

For a constant source

Given $N_B = 10^{16}/\text{cm}^3$, & $x_j = 2 \mu\text{m}$

$$N(x,t) = N_0 \operatorname{erfc}\left(\frac{x_j}{2\sqrt{Dt}}\right) - N_B$$

At the junction $N(x,t) = 0$

$$\operatorname{erfc}\left(\frac{x_j}{2\sqrt{Dt}}\right) = \frac{N_B}{N_0} = \frac{10^{16}}{10^{20}} = 0.0001$$

$$\operatorname{erf}(x) \doteq 1 - \operatorname{erfc}(0.0001) = 0.9999$$

we have $\operatorname{erf}(2.75) = 0.9999$

$$\therefore \frac{x_j}{2\sqrt{Dt}} = 2.75 \Rightarrow \sqrt{Dt} = \frac{x_j}{2 \times 2.75}$$

$$= \frac{2 \times 10^{-4}}{2 \times 2.75} \leftarrow 1\mu m \ (2 \text{ erm})$$

$$\sqrt{Dt} = 3.64 \times 10^{-5}$$

$$\therefore t = \frac{(3.64 \times 10^{-5})^2}{D}$$

$$\therefore t = \frac{(3.64 \times 10^{-5})^2}{1.47 \times 10^{-11}} = 9009.52 \text{ sec}$$

$$= 2.49 \text{ hr}$$

(3) Boron is diffused into an n-type Si with background concentration of $1 \times 10^{16}/cm^3$ at 1100°C. Assuming infinite source solid solubility ~~diffusion~~ diffusion, find out

(a) time necessary to form a pn junction 1 μm away from the surface if the doping profile is erfc type.

(b) The total amount of boron incorporated into silicon during above process

(c) If the boron profile in Si is approximated to be $N(x,t) = N_S(1 - y^{2/3})$ where N_S is the surface doping concentration and $y = (x^2/6D_S t)^{3/2}$, then assuming D_S to be the diffusion co-efficient as calculated in part (a) find out junction depth for the same duration of diffusion as in part (a)