ELE110 - Semiconductor Devices (Part 2)

Content: Solid state electronics study mainly deals with electrical behavior of solids and especially concentrates on semiconductors, their crystal structure, crystal growth technology and applications in electronic devices.

Textbook: "Solid state electronic devices", Ben G. Streetman, Prentice Hall.

Reference Books: "

1.) "Semiconductor Electronics", J.F. Gibbons, Mc Graw Hill.
2.) "Physical and Solid State Electronics", K.A. Kano, A. Wesley
3.) and other available books on physical electronics.

Course Outline:

- Atom theory (classical & modern)
- Quantum theory
- Bond theory of solids
- Basic semiconductor definitions and measurement techniques.
- Metal-semiconductor, semiconductor-semiconductor contacts
- Diodes (pn junction)
- Transistors (BJT and FET's)

Allison, "Electronic Engineering Materials and Devices'
TK 7871 A44 1971

Structure of Matters:

Elements are all composed of atoms carrying all properties of matter. Elements are classified according to their chemical and physical properties in periodic table by Mendelief.

Crystal Structure:

In a crystal of an element, atoms are placed obeying some determined geometry which is repeated along the solid material. However all solids are not crystals having periodic atomic seperation. Such non-periodic solid structures are called as amorphous. If a solid has locally different crystal structures, then it is called as polycrytals. The order of the crystal structure of the solid defines its lattice and its smallest part is called as unit cell.

There are 7 crystal types as : cubic, triclinic, monoclinic, orthorhombic, tetragonal, trigonal and hexagonal.

* Note : See the cubic crystals from your lecture notes.

Bonding Types in Solids:

Very important function of holding the crystal together, is provided by interaction of electrons in neighboring atoms of a solid. The resulting bonding can be grouped as :

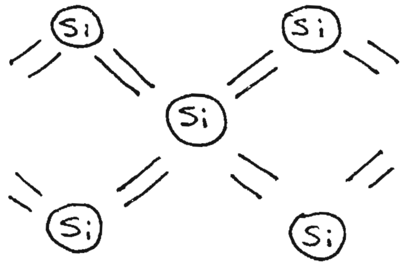
Ionic Bonding : is encountered in the compounds of very electropositive elements (alkali metals) and very electronegative elements (hallogens). NaCl is the typical example of this type of bonding. There is one electron in the last circle (shell) of the Sodium and there is one missin in the last circle of Klor which share one electron from Na to form two bonded ions. Thus such compounds are good insulators having well defined crystal shape such as cubic type.

Metallic Bonding :

Metals are characterized by the number of their valance electrons (electrons in outermost orbitals) being less than the number of its valance orbitals (the free positions in the outermost orbitals). Thus outer electron of an atom can move easily to neighboring orbital. These electrons hold any positive ion to be seperated from the structure. This structure is called as metallic bonding. Since the electron and ions are free to move, metals are good conductors and they can be easily shaped.

Covalent Bonding :

This bond is formed by the shared electrons of two atoms. Some atoms like Ge and Si makes such bonding by sharing their 4 valance electrons in order to provide chemical stability of the atom.



In such a lattice free electrons are not available and material is a good insulater. But pure crystal structure is not possible and some impurities do exist, making electrical conduction possible. Thus material behaves as semiconductor.

Electron Properties:

Electron is one of the most important part of the atom having very small radius of $R = 1.9 \times 10^{-15}$ m and negative charge of $e = 1.602 \times 10^{-19}$ Coulomb. The mass of the electron is also very small ($m_0 = 9.11 \times 10^{-31}$ kg). This means that electron carries high charge with respect to its mass.

According to relativity theory, if electron rotates in high speeds, then its mass is the function of its speed v as:

$$m_v = \frac{m}{\sqrt{1 - \frac{v}{c}}}$$

v : e^- speed
m : e^- mass $(= m_0)$
c : speed of light

Electron Movement in Static Electric Field:

Static electric field gives energy to the electron but magnetic field only changes its orbit.

If an electron is placed in electric field $\bar{E}$, field exerts force on it as:

$$\bar{f}_e = -e \cdot \bar{E}$$

Thus, force has the parallel direction with $\bar{E}$ field and in opposite direction due to negative charge of an electron.

It is known that static $\bar{E}$ field can be expressed as the gradient (directional derivative) of a scaler field V as;

$$\bar{E} = -\nabla V \qquad \left(\nabla = \frac{\partial}{\partial x} \hat{a}_x + \frac{\partial}{\partial y} \hat{a}_y + \frac{\partial}{\partial z} \hat{a}_z \right)$$

and thus

$$\bar{E}_x = -\frac{\partial V_x}{\partial x}\hat{a}_x \quad , \quad \bar{E}_y = \frac{-\partial V_y}{\partial y}\hat{a}_y \quad \text{and} \quad \bar{E}_z = \frac{-\partial V_z}{\partial z}\hat{a}_z$$

and

$$f_e = -e\nabla V = m a$$

$$f_e = m\frac{dV}{dt} = m\frac{\partial^2 r}{\partial t^2}$$

$$\Rightarrow \boxed{\frac{\partial^2 r}{\partial t^2} = \frac{e}{m}\nabla V}$$ defines electron movement in electrostatic field.

Electron Movement in Magnetic Field:

It is well known that time varying charge flow results with a current i as:

$$i = \frac{\partial Q}{\partial t} = \frac{-e}{\partial t}$$

Force exerted by magnetic field onto a charge q is expressed as:

$$\bar{f}_m = q\,\bar{V}\times\bar{B} = q\,V\,\bar{B}\,\mathrm{Sin}\,\theta$$

q: charge amount
V: speed of charges
$\bar{B}$: magnetic flux density around.

As a result of accelerating electron movement;

$$\bar{f}_m = m a$$

$$= m\frac{\partial^2 r}{\partial t^2} = -e\,\bar{V}\times\bar{B} = -e\,V\,B\,\mathrm{Sin}\,\theta\,\hat{a}$$

$$\Rightarrow \boxed{\frac{\partial^2 r}{\partial t^2} = \frac{-e}{m}\,\bar{V}\times\bar{B}}$$ defines electron movement in magnetostatic field.

Thus force fm is always perpendicular to the $\bar{B}$ field which means no work is done by magnetostatic field.

If both $\bar{E}$ and $\bar{B}$ field are present, then total force on electron can be found from the superposition as :

$$f_{total} = f_m + f_e$$

The above given theory gives the position and velocity of an electron in classical dynamics but not sufficient to explain behavior of particles in atomic size having high velocity. In such a case quantum theory provides more precise model.

Quantum Theory :

Before the quantum theory was proposed, Max Planck (1911) showed that mechanical system cannot have any arbitrary energy, but only special energy values related to natural frequency of the system; (classical quantum theory) as:

$$\boxed{E = nh\gamma}$$

where n is any positive integer, h is the Planck's constant and γ is the oscillator frequency ($h = 6.63 \times 10^{-34}$ Joule/sec).

Then Niels Bohr tried to develop better model using Planck's ideas, and postulated the following :

1-) The electron in an atom has only certain definite states of motion allowed to it, each of these stationary states has a definite fixed energy.

$$\frac{e^2}{4\pi\varepsilon_0 r^2} = \frac{mv^2}{r}$$

$v = \omega r \Rightarrow$

$$\frac{mv^2}{r} = mr\omega^2 \equiv \frac{e^2}{4\pi\varepsilon_0 r^2}$$

4th post : $mvr = nh/2\pi \quad n = 1, 2, \ldots$

$\Rightarrow$ ~~mvr~~ $v = \frac{nh}{2\pi m r}$

$$\Rightarrow \frac{mv^2}{\cancel{r}} = \boxed{\frac{e^2}{4\pi\varepsilon_0 r\cancel{^2}} = m \cdot \left(\frac{nh}{2\pi m r}\right)^2}$$

$= mv^2$

$r = \frac{\cancel{e^2}}{m 4\pi\varepsilon_0 v^2}$

$$\Rightarrow r = \frac{n^2 h^2}{(2\pi)^2 m} \left(\frac{4\pi\varepsilon_0}{1}\right)$$

$$= r = \frac{\varepsilon_0 h^2 n^2}{e^2 m \pi}$$

Only cert. orbits s having radius r is possible.

$= 0.529 \times 10^{-10} n^2$

2-) When an atom is in one of these states it does not radiate but when changing from a high energy state to a lower energy the atom emits a quantum of radiation whose energy "$h\nu$" is equal to the difference in the energy of the two states, as:

$$\Delta E = E_1 - E_2 = h\nu$$

3-) In any of these states electron moves in a circular orbit around the nucleus.

4-) The states of allowed electronic motion are those in which angular momentum of the electron is an integral multiple of $h/2\pi$ as $L = nh/2\pi$.

of these four postulates first two are correct and are retained in <u>modern</u> quantum theory. The fourth is partially correct; angular momentum is fixed but not in the way Bohr proposed. The third is completely incorrect and does not appear in modern quantum theory.

Bohr's theory is mainly proposed for Hydrogen atom and does not provide good results for large atoms. For small atom, for balance Coulomb and centrifugal forces must be equal to each other as:

Coulomb force = centrifugal force

$$\frac{e^2}{4\pi\varepsilon_0 r^2} = \frac{mv^2}{r}$$

$e = 4.8 \times 10^{-18}$ esu
$m = 6.68 \times 10^{-24}$ gm

$$= Z\frac{e^2}{r^{\cancel{2}}} = \frac{mv^2}{\cancel{r}}$$

$V = wr$; $mVr = nh/2\pi$; 4th postulate $n = 1, 2, \ldots$

$v = \frac{nh}{2\pi mr}$

and $\frac{mv^2}{r} = mrw^2 \Rightarrow mrw^2 = \frac{e^2}{4\pi\varepsilon_0 r^2}$

thus
$$\boxed{r = \frac{\varepsilon_0 h^2}{e^2 m\pi} n^2 = 0.529 \times 10^{-19} n^2}$$

: only certain orbits having radius r are possible !

Then energy of these possible orbits can be found.

It is assumed that no force acts on an electron at infinity due to zero potential and force :

$$\text{force} = f(r) = \frac{e^2}{4\pi\varepsilon_0 r^2} \qquad r \to \infty \; ; \; \text{force} \to 0$$

If this electron is moved towards a proton then work is required :

$$\text{work} = \int_{x=\infty}^{x=r} f(r)\,dr = -\frac{e^2}{4\pi\varepsilon_0 r} = \text{potential energy required !}$$

$$\Rightarrow E_p = -\frac{e^2}{4\pi\varepsilon_0 r}$$

We know that kinetic energy of an electron is equal to $E_k = \frac{1}{2} m v^2$. We have $mrw^2 = \frac{e^2}{4\pi\varepsilon_0 r^2}$ previously.

Thus by using $V = wr$, E_k can be obtained in the form

$$E_k = \frac{1}{2} . m (r^2 w^2) = \frac{e^2}{4\pi\varepsilon_0 r} . \frac{1}{2} = \boxed{\frac{e^2}{8\pi\varepsilon_0 r} = E_k}$$

So total energy $E_t = E_p + E_k$;

$$E_t = -\frac{e^2}{8\pi\varepsilon_0 r}$$ is obtained. (-) due to $\overset{(-)}{\text{charge}}$ of an electron!

We have previously determined possible r (distance between electron and nucleus; radius of the orbit);

$$r = \frac{\varepsilon_0 h^2 n^2}{e^2 m \pi}$$ ⇒ substitude in E_t ;

$$\boxed{E_t = -\frac{me^4}{8\varepsilon_0^2 h^2} \cdot \frac{1}{n^2} = -\frac{13.6}{n^2}}$$

: only certain energies are allowed to the atom

(eV : electron.volt)

Ex: For n=1 , E=-13.6 means that -13.6 eV energy is required to move electron from r to infinity.

Modern Quantum Theory : is statistical in nature. For an electron both particle and wave characterististics can be observed. The behavior of an electron is governed by a wave equation called Schrodinger's equation. Solution of this equation gives atomic orbitals corresponding certain energy levels having allowed certain number of electrons.

Quantum Numbers: In order to determine position of an electron in solid, four quantum number are defined as

1-) Main quantum number: (integer) n : indicates the energy of an electron and big diameter of the elliptic orbit of it in Bohr's model.

2-) Orbital quantum number: (integer) ℓ : indicates the small diameter of the elliptic orbit.

3-) Orbital magnetic quantum number : (integer) m : indicates the slope of the orbital plane with respect to the applied magnetic field direction.

4-) Spin quantum number : s : In addition to the orbital motion, electron also rotates around its axis called as spin. This rotation creates electric and magnetic fields and determines magnetic properties of materials. $s = \pm 1/2$ due to turning of electron either left or right.

Conduction Process in Semiconductors:

Intrinsic Semiconductor: Semiconductor material which is very pure is called as intrinsic semiconductor. (For Ge 1 part impurity / 100 million parts, for Si 1 part impurity / 10 trillion parts will cause the material to be impure).

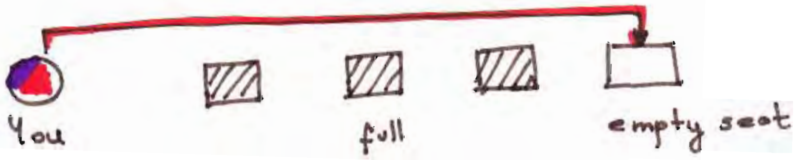
At 0°K (= -273.15 °C) semiconductor electrons are tightly bound to each other and material is an insulator. An increase in temperature, increases the kinetic energy of the atoms which is transferred to an electron and makes it break away from its parent atom as a free electron.

As the temperature of a semiconductor is raised from 0K, some electrons in the valence band receive enough thermal energy to be excited across the band gap to the conduction band. The result is a material with some electrons in an otherwise empty conduction band and some unoccupied states in an otherwise filled valence band.
An empty state in the " " is called as hole. If the conduction band electron and the hole are created by the excitation of a valence band electron to the cond. band, they are called an electron-hole pair (EHP.

A free electron lefts behind a vacancy called as hol
This process of formation of free electrons and holes is called as thermal generation.

It is interesting that hole behaves as current carrier. If a hole is created, an electron from a neighboring atom can easily jump into this hole, thus hole seemed to be moving. This process requires less energy for the electron to become free.

Think a place in a stadium:



Case I: (You = free e-)

Directly go to empty seat

: free e^- movement

Case II: After each persion is moved to right sight, you may sit to the first place, empty seats seems to be moving twds left : hole movement.

Therefore, we consider a hole as a positive charge. It is very important to understand that even though hole current is actually due to the motion of electrons, it is not the same type of electron motion as that of a free electron. The electrons which cause the hole current jump from hole to hole (if any) and do not have enough energy to become really free. Thus we can conclude that free electrons are able to move faster through the crystal than the holes thus have higher mobility (moving ability).

If an electron (from a battery) filles a hole, then filling electron will dissapper. This process is called as recombination.

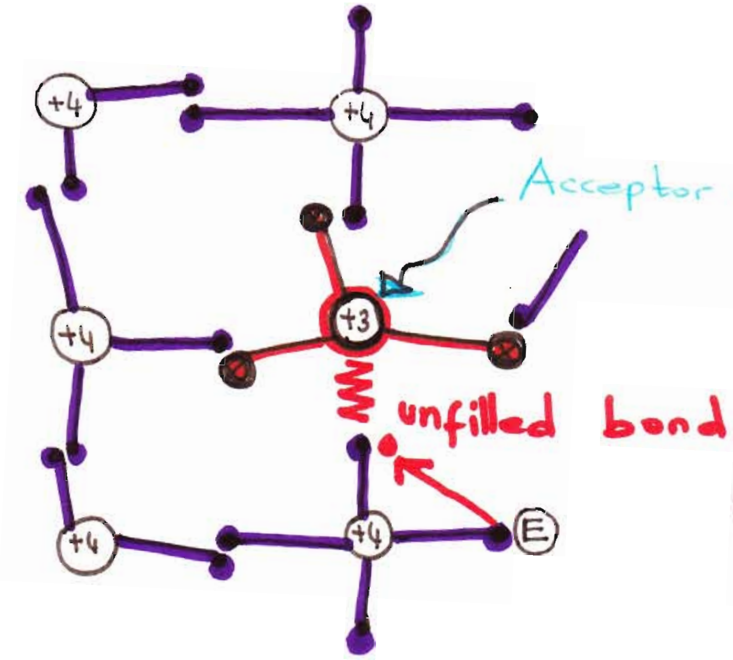
In an intrinsic semiconductor with no applied voltage, the total number of free electrons equals the total number of holes. In the case of temperature increase, more electrons can become free due to decrease in the resistance of the material.

Extrinsic Semiconductor:

Semiconductor materials having certain impurities are called as extrinsic semiconductors and have found various applications. The concentration of added impurity is small ≈ 1 part / 1 million parts of pure semiconductor). Impurities are usually one of two types : an element which has three valance electrons or one which has five valance electrons, the former is called as p-type impurity, the latter n-type.

P_type impurity:

An added impurity will take the place of an semiconductor atom. Since it has three valance electrons, one of the four covalent bonds will not be formed. This makes it very easy for an electron in a bond between two neighboring semiconductor atoms to fill this vacant bond, leaving a hole behind it. In this way, addition of p-type impurities leads to the formation of excess holes. Common p-type impurities are Boron, Gallium and Indium (B, Ga, In).

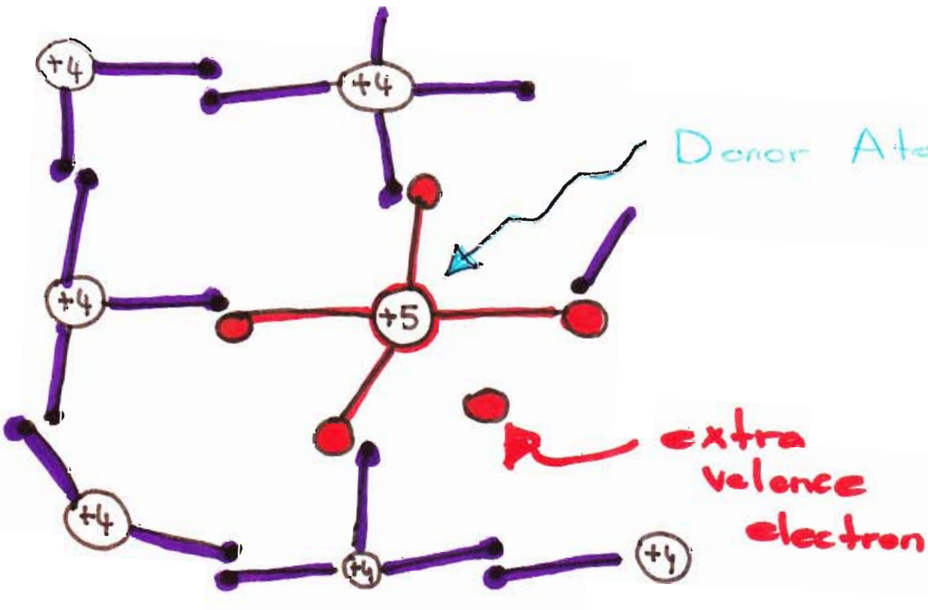


Acceptor Atom = p-type impurity = called as acceptor impurity

⇒ : Ⓔ can jump into the unfilled bond.

n-type impurity :

In this type only four of the valance electrons of the added impurity can form covalent bonds. The fifth electron can be easily removed from the parent atom, thus becoming a free electron. n-type impurities results in excess free electrons. Typical n-type impurities are Sb : Antimony, P : Phosphorus and As : Arsenic. n-type impurities are called as donor impurities. (p-type as acceptor). The intentional process of adding impurities is called as doping.



Donor Atom = n-type impurity

= can easily become a free electron

After doping, the resistance of the semiconductor greatly changes (typically multiplied by a factor of ~100). A semiconductor doped with p-type (n-type) impurity is called as p-type (n-type) semiconductor. The number of current carriers due to doping (excess carriers) are much larger than the number which would be present in the intrinsic (undoped) semiconductor.

For some materials both doping types may be done by adjusting the number of carriers called as compensation.

Energy Levels (Bands) for Crystals:

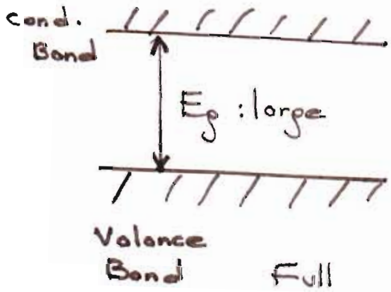
It was previously shown that;

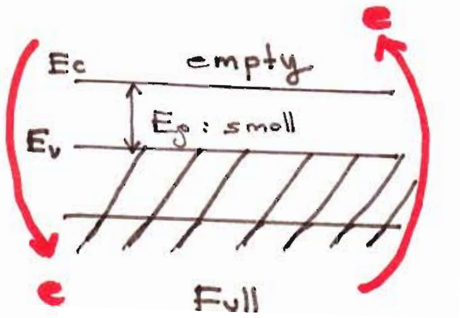
- Any system tends to move in the direction to decrease its total energy.
- Pauli-Exclusion principle: no two electrons in an interacting system can have the same set of quantum numbers, which means same energy levels.
- In an atom, the electrons occupy the lowest possible energy states. Thus atom is said to be in its ground state.

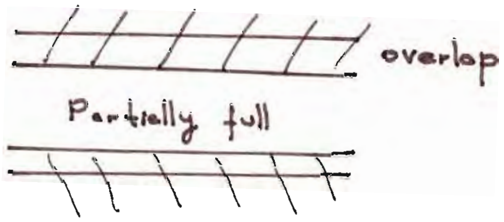
When two atoms are completely isolated from each other they are both in their ground state and have identical electronic structures. As the spacing between two atoms becomes smaller, due to Exclusion principle, the discrete energy levels of the isolated atoms must split into new levels. The upper band is called as conduction band, the lower band as valence band. These two bands are seperated by an energy gap E_g which is also called as forbidden band.

Metals, Semiconductors, and Insulators:

Energy band structure of solids is responsible for the electrical characteristics of them. The electron in a solid can move only to allowed energy levels to create an electric current.







⇓

Insulator:

E_g is large

e^- can not jump into E_c.

⇓

Semiconductor

E_g is small

↳ enough e's can be found in cond. band at room temp.

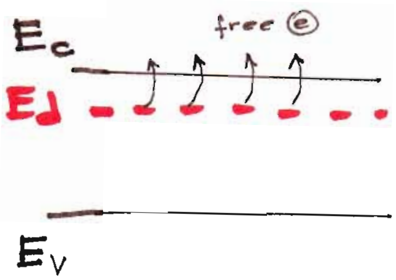
⇓

Metal

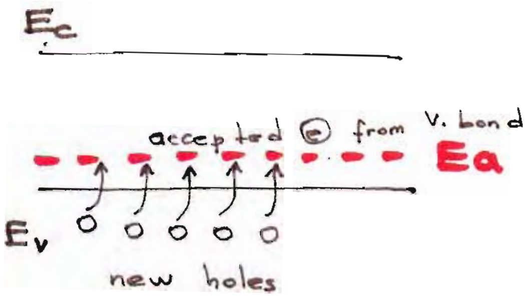
Cond. and valence bands are partially full.

Energy Band Diagrams for Extrinsic S. Material:

When impurities are introduced into semiconductor crystal, additional energy levels are created in the energy band structure. Addition of n-type impurity introduces an energy level very near to the conduction band in Ge or Si, whereas addition of p-type impurity introduces an energy level very near to the valence band.



-n-type

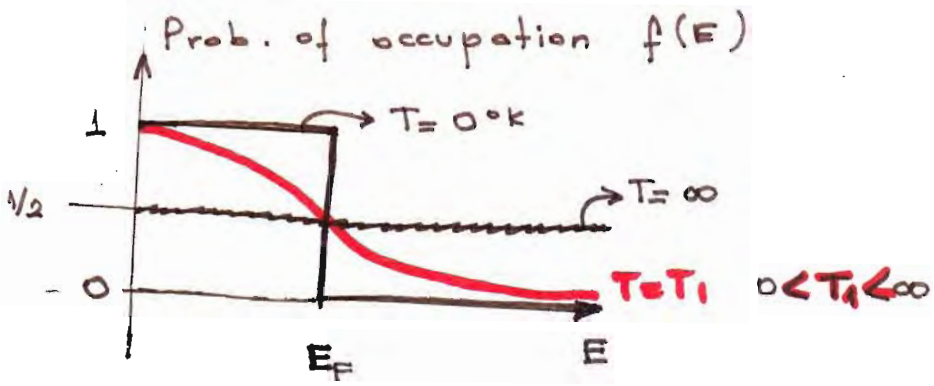


p-type

Fermi-Dirac Distribution, Fermi Level:

We have pointed out that <u>some</u> electrons are found in the conduction band in semiconductors as the temperature is increased. We want to know <u>exact</u> number of electrons found at each energy state at a given temperature. This is a statistical problem. Thus we need a function that gives the probability of being occupied for a state at a given temperature

Consider a solid at 0°K. The lowest possible energy state for the electrons in the solid is the one where all quantum states are occupied with just one electron (Exclusion principle) ranging from the lowest energy state up to some energy E_f. The level E_f is determined by the number of electrons and the allowed energy states of the material. One can say that for all energy states below E_F, the probability of being occupied at T=0°K is unity, while for all others above E_f, the probability of occupation is zero.



As the temperature is increased we expect to have some probability above E_F. At $T=\infty$, since all the electrons have enough energy to occupy any state, the probability of occupation must 1/2 for any E.

The exact probabilities are given by the Fermi-Dirac distribution :

$$\boxed{f(E) = \frac{1}{1+e^{(E-E_F)/kT}}}$$

$f(E)$: is the prob. of a state with energy E.

: probability of occupation

k : Boltzman const.

$= 1.38 \times 10^{-23}$ Joule / °K

E and E_F are in Joules and T in °K .

The value of E_F is called as the Fermi level of the material. Notice that :

$$f(E_F) = \frac{1}{1+e^{0}} = \frac{1}{1+1} = \frac{1}{2}$$

and the function satisfies the requirements for T=0 °k and T=∞

An important factor about Fermi-Dirac distribution is that it gives only the probability of being occupie for a state whose energy is E, but the existance of such a state is out of interest. If a state is not available, then of course it cannot be occupied.

For a semiconductor Fermi level E_F must lie in the energy gap as:

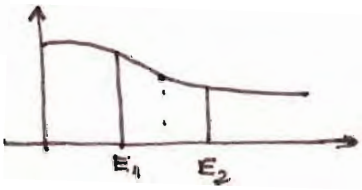
$$E_v < E_F < E_c$$

Probability of being not occupied for a state of energy E can be determined as $1 - F(E)$. Fermi energy E_F of a system can be thought as the average energy of the electrons in a determined temperature. Fermi functions for donor and acceptors are different from each other, (See notes) due to the introduced new energy states.

Carrier Densities, Density of States:

We showed that Fermi function only gives the probability of occupation of a state with energy E. Between any energies $E + dE$, there will be some number of quantum states. The number of states is proportional to dE.

→ The density of a state at energy level E, $N(E)$ can be shown as:



$$N(E) = \lim_{\Delta E \to 0} \frac{\#\text{ of states}}{\Delta E} \quad \leftarrow$$

n = Density of (concentration of) electrons in cond. band;

$$n = N_c \, e^{-(E_c - E_F)/kT}$$

$$N_c = 2\left(\frac{2\pi m_n^* kT}{h^2}\right)^{3/2}$$ m_n: effective mass for e

p = Concentration of holes in valance band

$$p = N_v \, e^{-(E_F - E_v)/kT}$$

$$N_v = 2\left(\frac{2\pi m_p^* kT}{h^2}\right)^{3/2}$$ m_p: effective mass for hole.

$$n = n_i \, e^{(E_F - E_i)/kT}$$

$$P = p_i \, e^{(E_i - E_F)/kT}$$

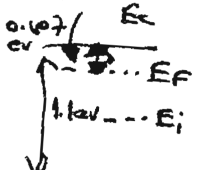
$$n_i = \sqrt{N_c N_v} \, e^{-E_g/kT.2} = p_i$$

$$n_i p_i = N_c N_v \, e^{-E_g/kT}$$

Ex: A Si sample is doped with 10^{17} As atoms/cm^3. Find equilibrium hole concentration at 300 k? Where E_F relative to E_i?

$N_d \gg n_i \Rightarrow n \cong N_d$ $\quad p = \frac{n_i^2}{N_D = n} = 2.25 \times 10^3 \ cm^{-3}$

$E_F - E_i = kt \ ln \frac{n_0}{n_i} = 0.0259 \ ln \frac{10^{17}}{1.5 \times 10^{10}} = 0.407 \ ev$



For example $N(E) = 0$ for energy gap.

Let us consider energy levels above conduction band having energies $E_c + E$. The density of conduction electro… n, in the material (free e / per unit volume), in c. band:

$$n = \int_{E_c}^{\infty} \underbrace{f(E)}_{\substack{\text{Possibility of} \\ \text{occupation of} \\ \text{a state}}} \underbrace{N(E)}_{\substack{\text{number of} \\ \text{available energy states}}} dE$$

$n = \sum \underline{f(E)}\, \underline{N(E)d}$…

Similarly hole density at valance band

$$p = \int_{-\infty}^{E_v} f(E)\, N(E)\, dE$$

$$N(E) = \lim_{\Delta E \to 0} \frac{\#\text{ of states}}{\Delta E}$$

OR

$$n = N_c\, e^{-(E_c - E_F)/kT}$$

N_c: effective density of ~~of~~ states in cond. band (depends on temp.) ✓

$$p = N_v\, e^{-(E_F - E_v)\, kT}$$

$$np = N_c N_v\, e^{-(E_c - E_v)/kT} = N_c N_v\, e^{-E_g/kT} = n_i p_i$$

N_v = effective density of ~~the~~ states in valance band. (depends on temp.) ✓

For an intrinsic (pure) semiconductor, let us call

$n_i = n$ and $p_i = p$ then

$$\boxed{n_i = p_i}$$:

$$n = n_i\, e^{(E_F - E_i)/kT}$$

$$p = p_i\, e^{(E_i - E_F)/kT}$$

** $kT = 0.026\,\text{eV}$ at 300 K

Concentration of free electrons = Concentration of holes

(e / m³ = p / m³)

$$n_i = N_c\, e^{-(E_c - E_i)/kT}$$

$$p_i = N_v\, e^{-(E_i - E_v)/kT}$$

$$n_i = N_c N_v\, e^{-E_g/2kT}$$

The product of n and p is constant at equilibrium for a particular material at a cnst temperature even if the doping is varied

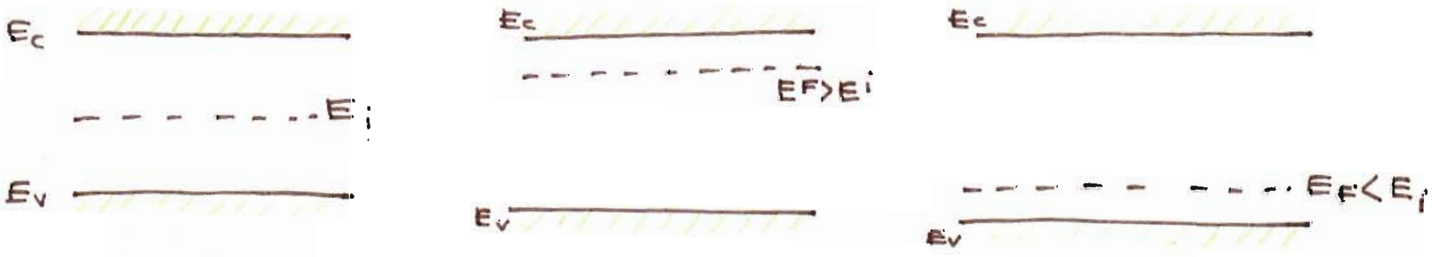
$$\boxed{\begin{array}{c} np = n_i^2 \\ \text{or} \quad np = n_i p_i \end{array}}$$

→ for all semiconductors at cnst T.

n_i is called as <u>intrinsic concentration</u>. ($\propto$ temperature , p_i also)

For a p-type semiconductor, the hole density ($p > n_i$) is larger than the intrinsic concentration. This also means that E_F^i is increased to a higher value; E_F^p. ($E_F^p < E_F^i$)

For an n-type semiconductor, electron density $n > n_i$ and intrinsic fermi level E_F^i is decreased to a lower value say, E_F^N.



intrinsic (pure) case

$E_i = \dfrac{E_c + E_v}{2}$: intrinsic level

n-type

$n > n_i$

p-type

$p > n_i$

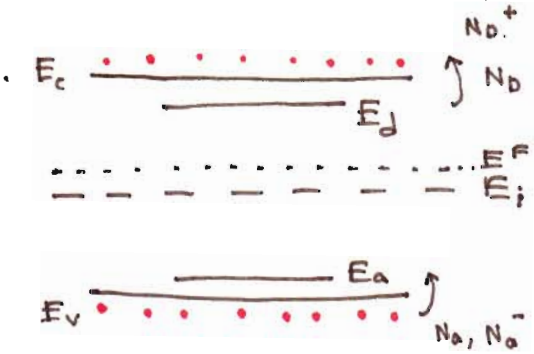
if $E_F = E_i \Rightarrow n = n_i$

if $E_F > E_i$ n increases exponentially

if $E_F < E_i$ p increases exponentially (**)

$$\Rightarrow \quad E_F^p < E_F^i < E_F^n \leftarrow \text{type of semiconductor}$$

In general, any doped material contains both donor atoms with density N_d and acceptor atoms with density N_a. The type of the material depends on doping conditions which is the result of adjustment of number of carriers called <u>compensation</u>.



If $N_D > N_a$ then, n-types dominate and E^F moves toward conduction band. For p-type $N_A > N_D$ and E^F moves toward valence band.

Space-charge neutrality principle indicates that net charge is zero over whole material and total positive charge equal to total negative charge as:

$$p + N_D^+ = n + N_A^-$$

N_D, N_A: excess carrier density as a result of doping

or

$$n = p + N_D^+ - N_A^-$$

For n-type $n \gg p$, thus $n \cong N_D - N_A \cong N_D$; $n \cong N_D$

" p- " $p \gg n$, thus $p = N_A - N_D \cong N_A$; $p \cong N_A$

We know that;

$$np = n_i^2$$

$$\Rightarrow \quad \boxed{n \cong \frac{n_i^2}{N_D} \quad ; \quad p \cong \frac{n_i^2}{N_A}}$$

Forming of Charge Carriers and Recombination:

Charge means an electron or a hole. If an electron breaks off from valence band (to c.b), a hole is formed. This is called as electron-hole pair - EHP. Electrons provide conduction either due to freely moving in the crystal or due to jumping from one hole to another.

At temperature $T = 0\,°K$, semiconductors behaves as insulators. In order to provide electron movement from valance band to conduction band, the required amount of energy must be provided to the electrons due to three different ways:

1.) Photon energy generation
2.) Thermal energy 〃
3.) Doping: adding impurities

1.) Photoelectric generation and photoconductivity

Light is assumed to be consisting of photons. Photons are energy packets having no mass but with frequency. When light is incident on a crystal, all photon energy is transferred to the electrons to make them jump to conduction band or leave from the crystal to the environment. If photon energy is less than E_g, (energy of bandgap), then this energy is absorbed by the crystal. When $E_{ph} = h\nu > E_g$, EHP occurs. This process is called as photoconductivity.

2.) Thermal generation

Electron-hole pairs (EHP) can also be formed due to heating of the crystal, as a function of applied temperature and crystal structure. Let us define thermal forming ratio $g(t)$ as:

$$g(T) = AT^a e^{-E_g/kT}$$

which gives the number of EHP in unit time and unit volume, at temperature T. In this equation

$e^{-E_g/kT}$ is called as Boltzman distribution function giving the probability of a valance electron to absorb E from the crystal atoms.

3. Doping : as explained above is the addition of impuritie to the crystal structure in order to change the original covalent bonded structure of the semiconductor resulting with free e^- and hole movement in the lattice.

Recombination :

For a given material the rate of generation of electron-hole pairs is a constant at a given temperature. This means that certain number of electrons are produced in unit time, and we denote this thermal generation rate by $g(T)$

$g(T)$: # of free e generated / unit volume / unit time

An e^- in conduction band can give up energy due to lattice and electron interactions and fall back into the valence band, and recombines with an available hole. The rate of recombination must depend on the temperature since electron energy is the function of T. This rate also depends on the number of conduction electrons available to recombine and number of available holes in the Valance band. Thus

$$R = r(T)\,pn$$

where $r(T)$ reflects (back) the temperature dependence of R, n & p are the density of free electrons and holes in valance band, respectively.

$$R = g(t) \quad \text{in equilibrium} \quad g(t) = r(t)\,pn \Rightarrow np = \frac{g(T)}{\underset{\substack{\downarrow \\ \text{depends only on} \\ \text{temperature !}}}{r(T)}}$$

At a given T, R increases to a value which is equal to $g(T)$ and an equilibrium is achieved:

$$g(T) = r(T)\, np$$

for an intrinsic material $n = n_i$ $p = p_i$ $\Rightarrow$

$$n_i p_i = n_i^2 = \frac{g(T)}{r(T)}$$

For a doped s.c., due to small no of impurities $g(T)$ is not affected by doping. Thus:

$$np = \frac{g(T)}{r(T)}$$

where $p = n_i^2 / N_D$ and $n = n_i^2 / N_A$.

Conduction Process in Semiconductors

There two main conduction process in semiconductors:

1.) Drift current : due to an application of electric field
2.) Diffusion " : " " unequal carrier density distribution

If an electric field $\bar{E}$ is applied to a charge q, the force acting on q can be found from classical methods as:

$$\boxed{\bar{F} = q\bar{E} = m\bar{a}}$$

m : mass of the charge
$\bar{a}$: acceleration "

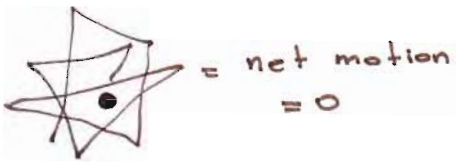
However, motion of free electron or hole in a semiconductor lattice can not be explained with classical terms directly due to many interactions between the carrier q and the lattice.

However, replacement of real mass of the carrier with effective mass solves this problem (see table at p. 33 of lecture notes). Effective mass of an electron for some materials is greater than the real mass, and for some others lesser.

Drift Conduction by Electrons:

In metals and n-type semiconductors, there are many free electrons that can move due to the application of external field. Current is defined as the total charge that passes through a specified crossection, in unit time.

In the absence and presence of external field, the motion of a carrier can be shown as:



a-) without ext. field

b-) with external field

In the case of applied field, electrons try to accelerate but this is not achieved due to lattice collisions and it is useful to characterize them with average velocity. The measure of moving easines of a particle is its mobility; μ, and defined as the drift velocity/applied field;

$$\boxed{\mu = \frac{|\bar{V}|}{|\bar{E}|}}$$

: Mobility is the function of temperature.

$= cm^2/volt\_sec$

μ_n : e^- mobility
μ_p : hole "

If n is the electron concentration and $\bar{v}_n$ is the average drift velocity, the time rate of charge flow across a plane perpendicular to the electron flow is the current density and given as:

$$\boxed{\bar{J} = -n q \bar{v}_n}$$

$\downarrow$ due to charge of an e^-

$$\bar{J} = \frac{\bar{I}}{A}$$

where

$$\bar{v}_n = -\mu_n \bar{E}$$

Thus $$\boxed{\text{conductivity } \sigma = \frac{|\bar{J}|}{|\bar{E}|} = q n \mu_n} \quad : \ \text{mho} : \mho/\text{cm}$$

q : Coulomb
n : number/cm^3
μ_n : cm^2/volt-sec

$$\Rightarrow \ \sigma : \text{mho}/\text{cm} = \mho/\text{cm}$$

Note that conductivity is the function of n, and thus can be controlled by doping.

$$\sigma = \frac{1}{\rho} \rightarrow \text{resistivity} \qquad \left(R = \rho \frac{l}{A} \ : \text{resistance} \right)$$

Drift Conduction by Holes :

A very useful representation for a hole is a +q charged particle. Due to more restricted motion of hole with respect to free electron, mobility of a hole is expected to be lower than that of the free e^-. Ex:

	Si	Ge	
μ_p :	480	1800	cm^2/volt.s
μ_n :	1300	3800	

at room temp. 300°K

The conductivity of p-type material, σ_p, can be similarly obtained as:

$$\boxed{\sigma_p = q p \mu_p}$$

p: Concentration of holes

Thus, if one of the charge carriers is not dominant, then current is carried by both of them and

$$\sigma = \sigma_n + \sigma_p = q n \mu_n + q p \mu_p$$

for an intrinsic material;

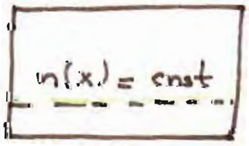
$$\sigma = q n_i (\mu_n + \mu_p)$$

for n-type $n \gg p \Rightarrow \sigma \cong \sigma_n$
" p- " $p \gg n \Rightarrow \sigma \cong \sigma_p$

Current Flow due to Diffusion:

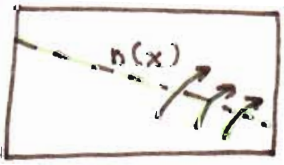
The second conduction mechanism in semiconductors is due to diffusion as a result of nonuniform carrier distribution.

For a uniform material with no external field, no net electron flow occurs, but each individual electron continiously moves in a random manner. But if non-uniform distribution is present, electrons re-distribute to achieve uniform distribution due to thermal energy, and diffusion current flows.



a.) Uniform charge distr.
Stable

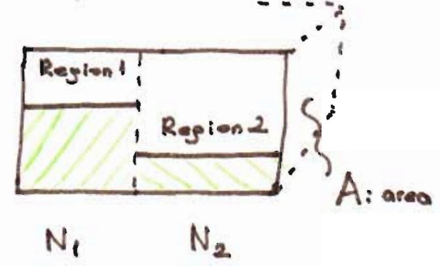
;



b.) Non-uniform
Not-stable

: diffusion current is observed!

Assume a distribution having N_1 and N_2 total number of particles in each region;

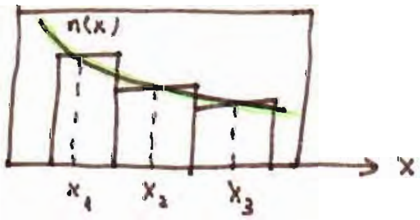


Flow rate from ① to ②;

$$F_{12} = KN_1 - KN_2$$

K: proportionality cnst

for a general distribution;



$$F_{12} = -DA\frac{dn(x)}{dx}$$

D is called as diffusion cnst.

(−) sign indicates particle flow from higher concentration.

The rate of charge motion is multiplied by the charge of a carrier, current is obtained. Thus;

Electron diffusion current
〃 〃 〃 dens.

$$\boxed{\begin{aligned} I_n &= q \cdot D_n \frac{dn}{dx} A \\ J_n &= q D_n \frac{dn}{dx} \end{aligned}}$$

Similarly for holes;

$$\boxed{\begin{aligned} I_p &= -q D_p \frac{dp}{dx} A \\ J_p &= -q D_p \frac{dp}{dx} \end{aligned}}$$

Diffusion constant(s) (D_n and D_p), D, depends on carrier type, doping levels, mobility of the carriers and the temperature.

Typical values :

	Si	Ge	
D_n :	34	98	cm^2/sec
D_p :	13	46	

$D_n > D_p$ because $\mu_n > \mu_p$

It is also known that

$$\boxed{\frac{D_n}{\mu_n} = \frac{D_p}{\mu_p} = \frac{kT}{q}}$$: Einstein relation

Therefore, for a given semiconductor, if an electric field exist in addition to a non-uniform charge distribution, total current density has both drift and diffusion components as:

$$\boxed{\begin{aligned} J_n(x) &= q\mu_n n(x) E(x) + q D_n \frac{dn(x)}{dx} \\ J_p(x) &= q\mu_p p(x) E(x) - q D_p \frac{dp(x)}{dx} \end{aligned}}$$: drift + diff.

then total current can be found as:

$$I(x) = J_x \cdot A$$

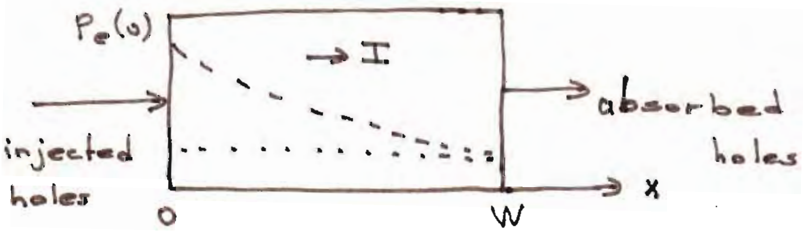
where A is the crossectional area into which current flows.

Recombination of charge carriers (excess carriers):

Remember that under equilibrium conditions recombination and generation rates are equal for both intrinsic and extrinsic semiconductors. For intrinsic case $n = p = p_i$ even though recombination and generation are continuing. The difference of extrinsic case is that $n \gg p$ or $p \gg n$. The net change in free carrier concentration is zero.

In various applications of semiconductors, excess carriers are injected into them, upsetting equilibrium conditions. The concentration of excess carriers are much smaller and does not affect the generation rate.

Let us consider a n-type material into which holes are injected from left edge resulting with a current. Due to recombination fewer holes reach to the right edge and current is observed to be exponentially decreasing with a time constant τ.



hole density is a function of position

In order to make total current flow constant, (current entering = current leaving), equal amount of electrons must also enter from the right edge in order to account for the difference between I and the hole diffusion current. These electrons are provided by the external source.

Majority and minority carriers:

Electrons are called as majority carriers and holes as minority carriers for n-type $(n \gg p)$ and for p-type reverse is true. If external field is present electron drift current $\cong 10^{10} \times$ hole drift current and thus hole drift current is ignored for n-type. $n_i = 1.4 \times 10^{10} / cm^3$, $n = 1 \times 10^{15} / cm^3 = N_D$ and $p = n_i^2 / N_D$ $\Rightarrow p = 1.96 \times 10^5 / cm^3$.

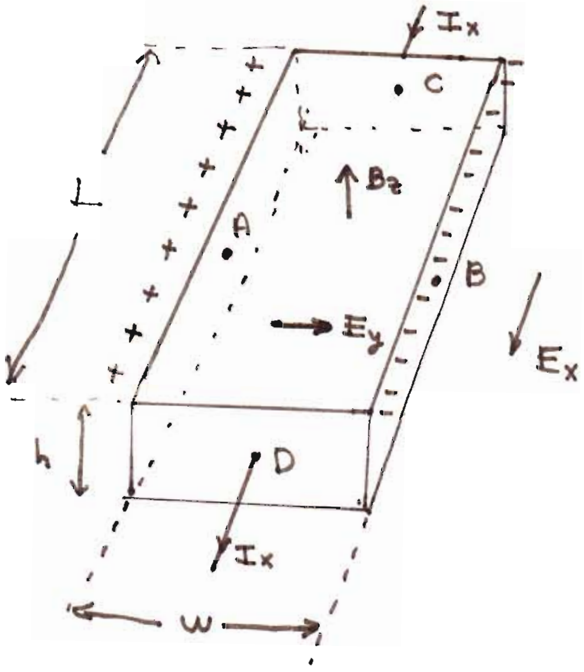
* Minority carriers are important for diode and BJT FET operations.

Hall Effect : The total force on a charge Q in the presence of both electric and magnetic fields is given as:

$$\bar{F} = Q(\bar{E} + \bar{v} \times \bar{B})$$

where $\bar{v}$ is the velocity of the charge, $\bar{E}$ and $\bar{B}$ are the electric and magnetic fields. Assume a piece of semiconductor material of p-type is connected to a battery and carries a current I_x, and uniform magnetic field is applied in z direction. The y. component of force on the majority charg carriers, i.e. holes is;

$$F_y = q(E_y - v_x B_z)$$



$$E_x = \frac{V_{CD}}{L} \qquad \bar{v} = v_x \hat{a}_x$$

$$\bar{B} = B_z \hat{a}_z$$

$$q(\bar{v} \times \bar{B}) = -q v B_z \hat{a}_y$$

$$= -q E_y \hat{a}_y$$

under equilibrium $F_y = 0$ and;

$$-q V_x B_z = -q E_y \Rightarrow V_x B_z = E_y$$

$$\Rightarrow V_{AB} = E_y \cdot w = V B_z w$$

$$\Rightarrow V_x = \frac{V_{AB}}{w B_z} = \text{drift velocity}$$

Note that initially E_y is zero but since the holes are moving in +x direction, they will experience a force in y direction and will move in that direction.

This non-uniform charge distribution in y direction will create an electric field E_y and equilibrium is achieved when $E_y = V_x B_z$ for $F_y = 0$. The establishment of E_y is known as Hall Effect. The electric field E_y causes potential difference between the sides A & B; $V_{AB} = E_y . w$ which is called as Hall voltage.

$$\mu_p = \frac{v_x}{E_x} \quad ; \quad J_x = \sigma_p E_x = q p \mu_p E_x = q p v_x \Rightarrow v_x = \frac{J_x}{qp}$$

Thus ; $$E_y = \frac{J_x}{qp} B_z = R_H J_x B_z \quad ; \quad R_H : \text{Hall coefficient}$$

If we know I_x , then $J_x = \frac{I_x}{wh}$ and

$$p = \frac{1}{q R_H} = \frac{J_x B_z}{q E_y} = \frac{(I_x / wh) B_z}{q . (V_{AB} / w)} = \frac{I_x B_z}{q h V_{AB}}$$

* A measurement of Hall voltage for a known current and magnetic field yields a value for the hole density p.

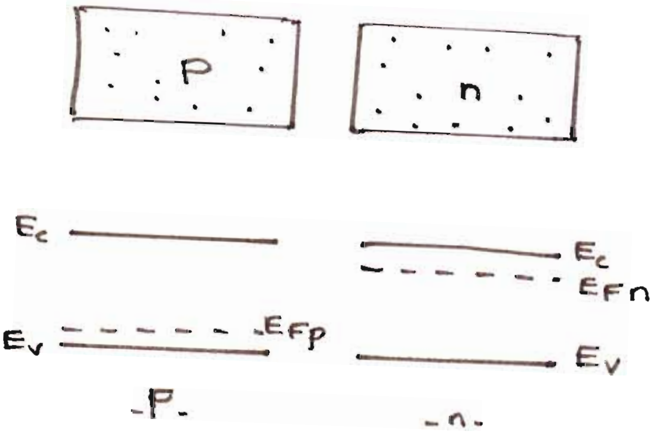
$$\mu_p = \frac{\sigma_p}{qp} = \frac{1/\rho \ \text{: resistivity}}{1/R_H} = \frac{R_H}{\rho} \quad \text{: mobility of holes} \checkmark$$

$$\rho = \frac{R wh}{L} = \frac{V_{CD}\, wh}{L I_x}$$

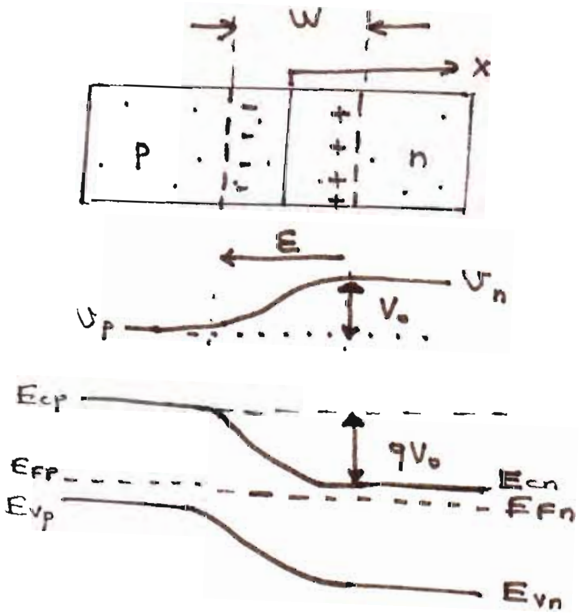
and $\sigma = 1/\rho$.

P-N Junction :

Let us consider seperate blocks of p- and n-type semiconductors brought together to form a pn junction. Before they are joined n-type has large concentration of electrons and few holes and reverse is true for p-side. After joining, holes diffuse from the p-side to n-side and electrons from n- to p-side. The resulting diffusion current can not build up continiously due to opposing electric field created at the junction. If we consider that electrons diffusing from n to p leave behind uncompensated (due to junction) donor ions (N_d^+), and holes leaving p-region leave behind uncompensated acceptors (N_a^-), it is easy to visualize the development of a region of positive space charge near the n-side of the junction and negative charge near the p-side. The resulting electric field $^{\varepsilon}$ is directed from the positive charge toward the negative charge. Thus $\mathcal{E}$ is in the direction opposite to that of the diffusion current for each type of carrier, (remember that electron current has opposite direction w.r.to electron flow). Therefore, the field creates a drift component of current from n to p, opposing the diffusion current.



a.) Before joining



b.) After joining W: transition region, space charge region

Particle flow direction		Current type, direction	
hole	⟶	⟶	hole diffusion curr.
"	⟵	⟵	" drift curr.
electron	← - - -	- - - →	electron diffusion curr.
"	- - - →	← - - -	" drift curr.

c.) Directions of four components of particle flow within the transition region W and resulting (corresponding) current flow directions.

We know that at equilibrium net current is zero. Thus the current due to drift of carriers in the $\mathcal{E}$ field must exactly cancel the diffusion current : $J_{\substack{p\\n}(\text{drift})} + J_{\substack{p\\n}(\text{diff})} = 0$ (due to also no net build up of electrons and holes on either side as a function of time)

Therefore electric field $\mathcal{E}$ builds up to point where net current is zero at equilibrium. Electric field appears only in the region W about the junction and equilibrium potential difference across W is V_o. $\mathcal{E}$ field in W can be expressed as $\mathcal{E}(x) = -\partial V(x)/\partial x$ where it is zero outside of the W. Thus, there is a constant potential V_n in the neutral n region and contant V_n in the neutral p region and a potential difference $V_o = V_n - V_p$ between the two. The region W is called as transition or depletion region (space charge region also).

In energy diagrams it is observed that valence and conduction energy bands are higher on the p-side of the junction than on n side by an amount of qV_o called as work function. At equilibrium Fermi levels E_F must be equal at both sides otherwise this requires current. Thus $E_F = F_n = F_p$ at equilibrium.

In order to obtain relation between V_o and the doping concentrations on each side of the junction we can use equilibrium condition for diffusion and drift components of hole current as:

$$J_p(x) = q\left[\mu_p\, p(x)\, E(x) - D_p \frac{dp(x)}{dx}\right] = 0$$

where p(x) : density of holes as a function of x

At equilibrium

$$J_p(x) = q\left[\mu_p\, p(x)\, E(x) - D_p \frac{dp(x)}{dx}\right] = 0$$

$$\Rightarrow \quad \frac{\mu_p}{D_p} E(x) = \frac{1}{p(x)} \frac{dp(x)}{dx}$$

$$\Rightarrow \quad -\frac{q}{kT} \frac{dV(x)}{dx} = \frac{1}{p(x)} \frac{dp(x)}{dx}$$

Since $E(x) = -dV(x)/dx$ and $D_p/\mu_p = q/kT$ (Einstein's relation). Now we integrate from $x = -W/2$ to $+W/2$ to obtain:

$$-\frac{q}{kT}\int_{x_1}^{x_2} \frac{dV(x)}{dx}\, dx = \int_{x_1}^{x_2} \frac{1}{p(x)} \frac{dp(x)}{dx}\, dx$$

$$\Rightarrow \quad -\frac{q}{kT}\int_{V_p}^{V_n} dV = \int_{p_p}^{p_n} \frac{1}{p}\, dp$$

p_p : density of holes at p-side

p_n : density of p's at n-side

$$-\frac{q}{kT}\left(V_n - V_p\right) = \ln p_n - \ln p_p = +\ln \frac{p_n}{p_p}$$

$$\Rightarrow \quad * \; \boxed{V_o = -\frac{kT}{q} \ln \frac{p_n}{p_p}}$$

diode eqn. $= \frac{kT}{q} \ln \frac{p_p}{p_n}$

Thus V_o can be found in terms of the doping levels in the p and n regions. This eqn can be put into more useful form as:

N_a acceptors on p side and N_D donors on n side $\Rightarrow$

$p_p = N_a$; $p_n = n_i^2 / N_D$; majority carrier concentr. $\cong$ doping carrie concentr.

$P = N_a^-$
$n = N_D^+$

$$\Rightarrow \boxed{V_o = \frac{kT}{q} \ln \frac{N_a}{n_i^2/N_D} = \frac{kT}{q} \ln \frac{N_a N_D}{n_i^2} = V_o} \quad *$$

$$\Rightarrow q V_o = kT \ln \frac{N_a}{n_i}$$

Another useful form

$$\boxed{\frac{p_p}{p_n} = e^{qV_o/kT} = \frac{n_n}{n_p}}$$

; $p_p n_p = n_i^2$
$n_n p_n = n_i^2$ $\Big\} \Rightarrow$

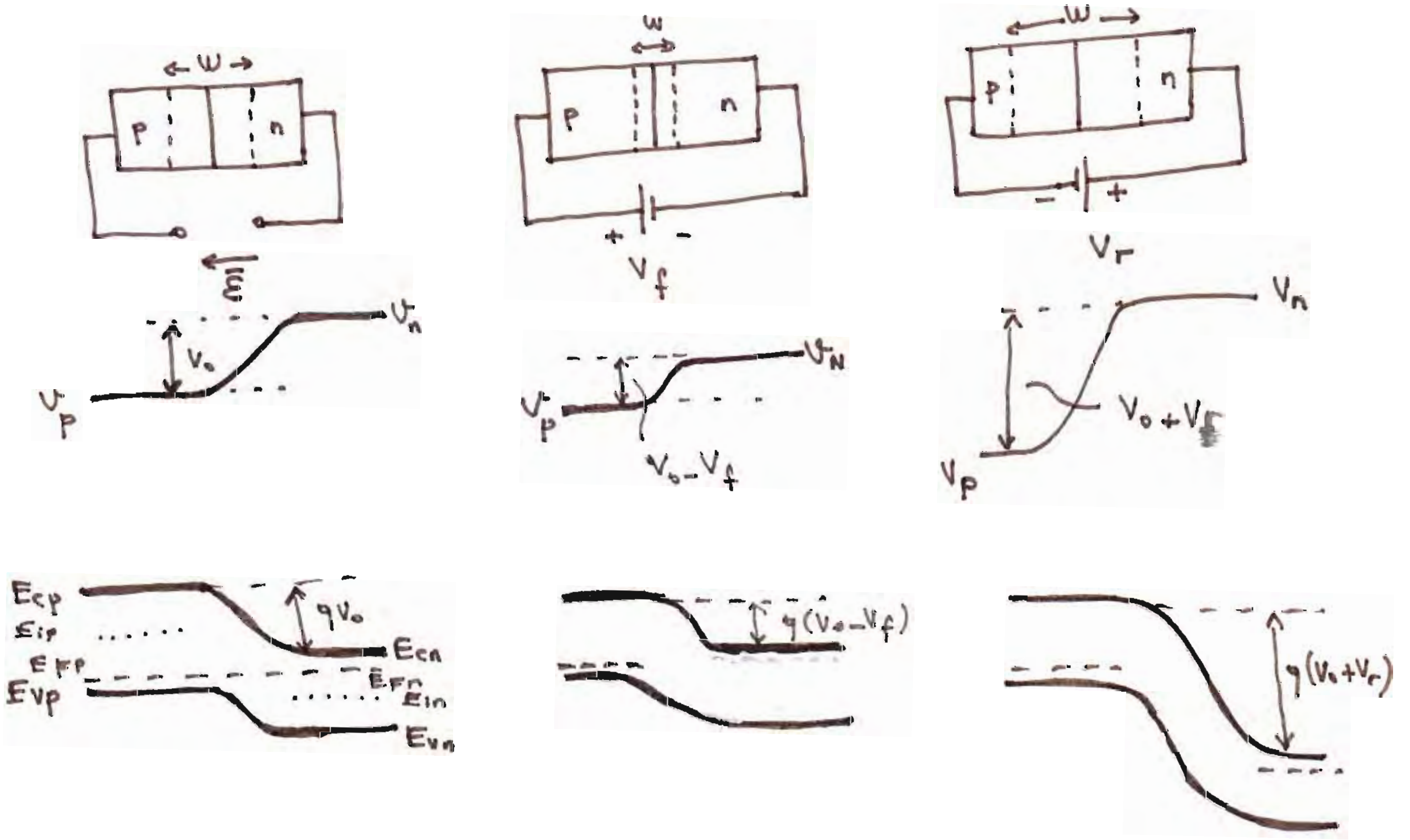
$$\frac{p_p}{p_n} = \frac{n_n}{n_p} = \frac{\bullet^{qV_o}}{e^{kT}}$$

V_o is called as built in potential and gives potential across the (depletion) transition region W in thermal equilibrium. Since thermal equilibrium implies that the Fermi energy is constant throught the pn junction (or pn diode), the built in potential is equal to the difference in the Fermi energies E_{Fn} and E_{FP} divided by the electronic charge. $\boxed{(E_{F_n} - E_{F_p}) = q V_o}$

p-type region of the pn junction or diode is named as anode and n-type region as cathode. If positive voltage is applied to p side and negative to n side, this type of biasing of the diod is called as forward biasing and the opposite case is called as reverse biasing. In forward biasing case net current is obtained over diode and in reverse biasing current flow is negligible as we will see in the next section.

Forward and Reverse Biasing of a Diode

If positive voltage is applied to the p-side and (-)ve voltage to n-side this is called as forward biasing and opposite as reverse biasing. We assume that an applied voltage V appears across the transition region of the junction rather than the neutral p and n regions due to their smaller length with respect to their area (small resistance) and due to moderate or heavy doping.



a.) equilibrium b.) forward biasing c.) reverse biasing

$$qV_o = kT \ln \frac{p_p}{p_n} = E_{cp} - E_{cn}$$

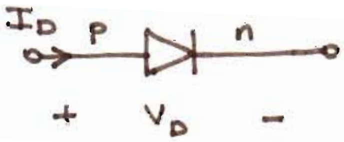
For forward biasing case V_o (equilibrium) ⟶ $V_o - V_f$ (lowered) and width W of depletion region decreases, applied field decreases the built in field (reverse is true for reverse biasing) and thus also potential barrier reduces. This results with current flow. For reverse biasing case barrier potential is high resulting with no current flow.

Total diode current eqn. is in the form :

$$\boxed{I_D = I_s \left(e^{qV_D/kT} - 1 \right) = I_s \left(e^{V/\gamma} - 1 \right)} \qquad \gamma = \frac{kT}{q}$$

where

V_D : Voltage through diode
I_D : Current on the "
I_s : Reverse saturation current



Note that if V_D is (+)ve exponential term is in effect corresponding to forward biasing. For (-)ve V_D, $I_D \approx I_s \approx 0$, corresponding to reverse biasing.

Electric field distribution in transition region, the width of the transition region and junction capacitance can be obtained by using Gaus's Law :

$$\nabla . \bar{E} = \frac{\rho}{\varepsilon} \qquad \text{where } \rho \text{ : charge density in trans. region} \quad \bar{E} \text{ : } \bar{E} \text{ field}$$

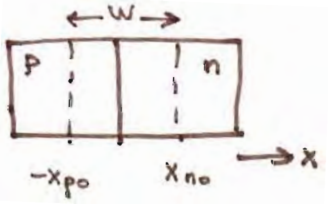
For solid state devices uni-directional variation only is assumed such as x-direction in this case :

$$\nabla . \bar{E} \rightarrow \frac{\partial E(x)}{\partial x}$$

$$\frac{\partial E(x)}{\partial x} = \frac{q}{\varepsilon} \left(p - n + N_d^+ - N_a^- \right)$$

In transition region carriers (p-n) can be neglected with respect to space charge carriers N_d^+ and N_a^-. Thus, we have two regions of constant space charge :

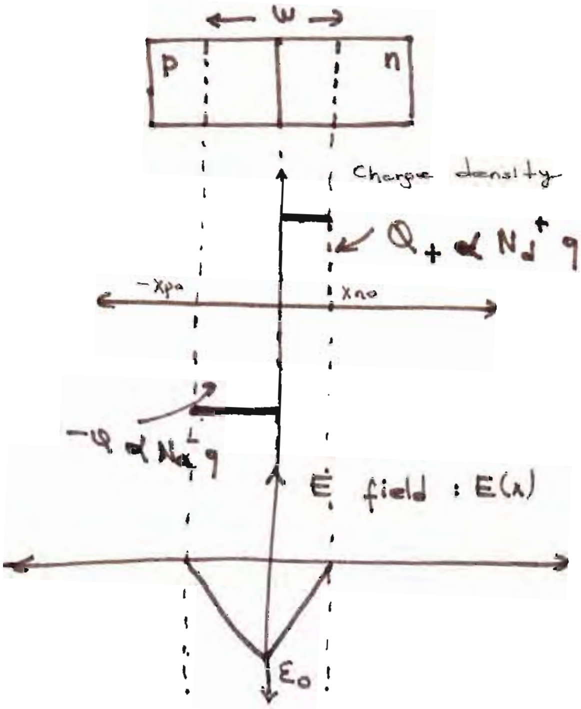
$$\frac{\partial E}{\partial x} = \frac{q}{\varepsilon} N_d \quad ; \quad 0 < x < x_{no}$$



$$\frac{\partial E}{\partial x} = -\frac{q}{\varepsilon} N_a \quad ; \quad -x_{po} < x < 0$$

assuming complete ionization of the impurities as $N_d^+ = N_d$, $N_a^- = N_a$

E field is assumed to be ≈ zero ($\to 0$) at the edges of transition region, and it has max. value at the junction :



x_{po} may be different from x_{no} due to different doping levels.

$$\Rightarrow \int_{\mathcal{E}_0}^{0} \partial E = \frac{q}{\varepsilon} N_d \int_{0}^{x_{no}} dx \qquad 0 < x < x_{no}$$

$$\int_{0}^{\mathcal{E}_0} \partial E = -\frac{q}{\varepsilon} N_a \int_{-x_{po}}^{0} dx \qquad -x_{po} < x < 0$$

ε : perm. of the semicond.

$$\Rightarrow E_{max} = \mathcal{E}_0 = -\frac{q}{\varepsilon} N_d x_{no} = -\frac{q}{\varepsilon} N_a x_{po}$$

$$\Rightarrow \quad N_d x_{no} = N_a x_{po}$$

$$W = x_{no} + x_{po}$$

And we know that $\bar{E} = -\nabla V$;

$$\Rightarrow \quad \mathcal{E}(x) = -\frac{dV(x)}{dx} \quad \text{or} \quad -V_0 = \int_{-x_{p0}}^{x_{n0}} \mathcal{E}(x)\,dx$$: contact potential

$$\Rightarrow \quad V_0 = -\frac{1}{2}\,\mathcal{E}_0\,W = \frac{1}{2}\,\frac{q}{\varepsilon}\,N_d\,x_{n0}\,W$$

from the previous eqns :

$$\boxed{\begin{aligned} x_{n0} &= W\,\frac{N_a}{N_a+N_d} \\ x_{p0} &= W\,\frac{N_d}{N_a+N_d} \end{aligned}}$$

; the widths of transition region in both sides are determined by doping levels

$$\Rightarrow \quad V_0 = \frac{1}{2}\,\frac{q}{\varepsilon}\cdot\frac{N_a N_d}{N_a+N_d}\cdot W^2$$

or $$\boxed{W = \left[\frac{2\varepsilon V_0}{q}\,\frac{(N_a+N_d)}{N_a N_d}\right]^{1/2}} \qquad (*)$$

Then junction capacitance or varactor capacitance is obtained as :

$$\boxed{C = \varepsilon\frac{A}{W} = \varepsilon A\,\frac{1}{\sqrt{\frac{2\varepsilon V_0}{q}\left(\frac{1}{N_a}+\frac{1}{N_d}\right)}}}$$

See ex 5.2 from Streetman !

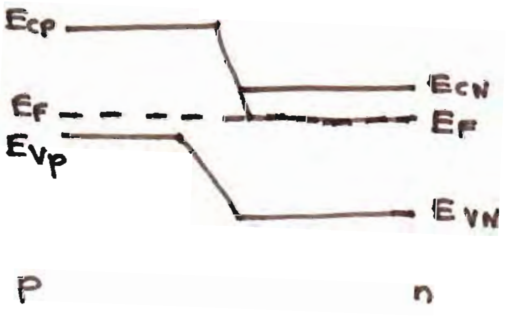
In the biased case of the junction, for example W changes due to change in the barrier voltage. For forward biasing case $V_0 \rightarrow V_0 - V_f$. Thus by replacing V_0 with $V_0 - V_f$ (bias volt.), W value can be obtained which is smaller wrto the equilibrium (unbiased) value.

Reverse Bias Breakdown :

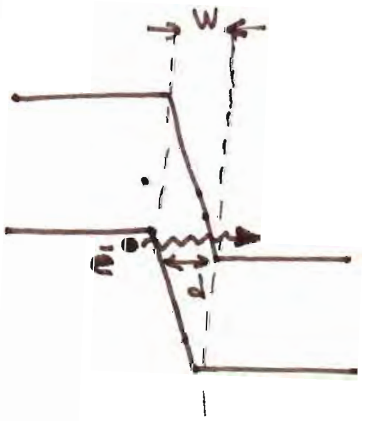
P-n junction biased in reverse direction exhibits a small, ≈0 current independent of voltage. This is true until a critical reverse bias is achieved (reached) for which reverse breakdown occurs. At this voltage V_{br}, the reverse current through the diode increases sharply and relatively large currents can flow with little further increase in the voltage. Reverse breakdown can occur by two mechanisms :

a-) Zener Breakdown

When a heavily doped junction is reverse biased the potential barrier increases to a level so that the n-side conduction band appears opposite the p-side valence band. When such a crossing occurs, valence band electrons in p-side can move into the conduction band of the n-side. This is possible due to wave nature of the electron and called as tunnelling. This gives rise to a reverse current.



a.) heavily doped unbiased junction



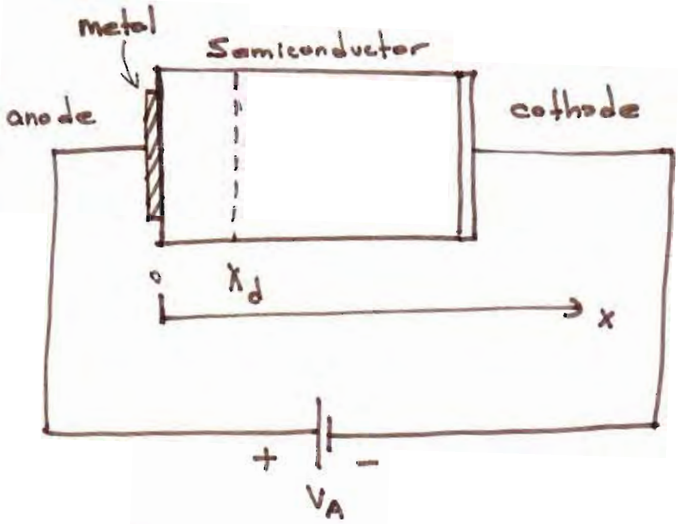
b.) heavily doped junction with rev. bias; electron tunnelling occurs from p→n.

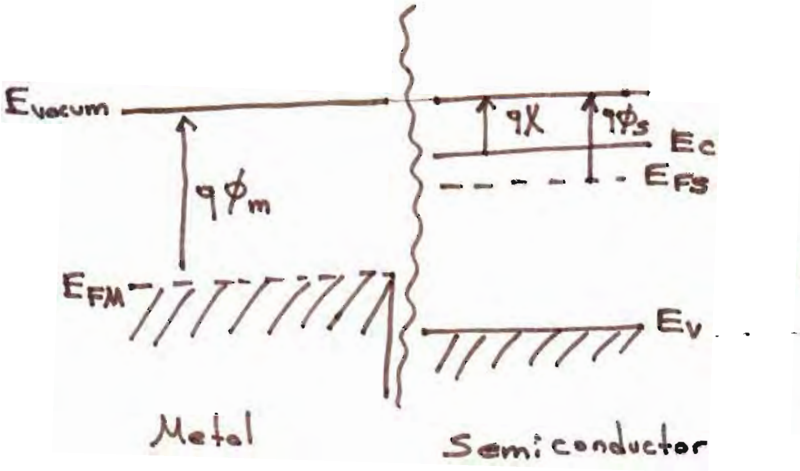
: d; depends on doping levels, bias field and material properties.

b-) Avalanche Breakdown: For lightly doped junctions, tunnelling is negligible but drifting electrons can accelerate to high enough kinetic energies. Each such energetic electron can create other carriers thus multiplying the number of carriers. This is an avalanche process and results with reverse current.

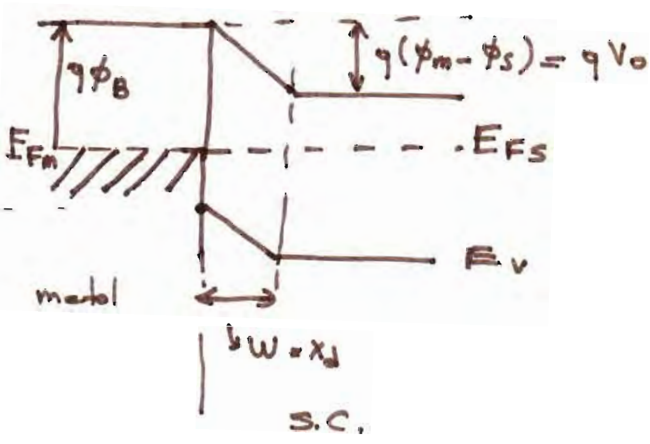
Metal-Semiconductor Junctions

Many useful properties of p-n junction can be achieved from metal-semiconductor junction within high speed.





a.) Band diagrams of metal & s.c. before joining $\phi_m > \phi_s$



b.) After joining, Fermi levels aligned due to electrostatic potential increase in S.C.

Workfunction $q\phi_m$ of a metal in a vacuum means energy required to remove an electron at the Fermi level to vacuum : outside the metal. Typical values of Φ_m depend on material (see notes), for ex. : 4.3 V for Al and 4.8 V for Au.

If forward biasing (V) is applied, contact potential reduces from $V_0 \rightarrow V_0 - V$, electrons in the semiconductor conduction band can diffuse across the depletion region to the metal. This results with a forward current from metal to S.C. Conversely if reverse biasing is applied, contact potential increases from V_0 to $V_0 + V_r$ and electron flow from s.c. to metal becomes negligible. The current equation has the same form with the p-n junction case.

P-N Junctions as Optoelectronic Devices

Some optoelectronic devices are constructed from p-n junctions. Some examples of such devices can be given as:

a. Photodiodes : for such a pn junction, current is the result of the exposed light onto the surface due to the photogeneration of electrons and holes as a function of incident light energy.

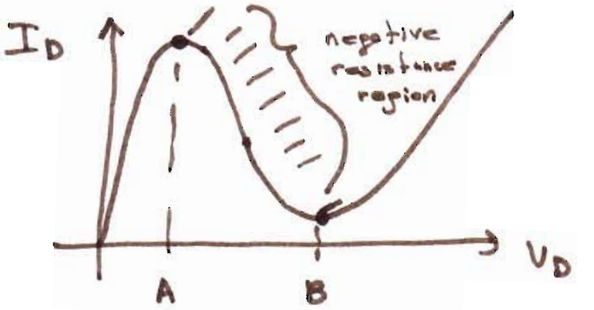
b. Solar cells : illuminated with sunlight having photons with energy larger than the energy bandgap of the S.C. to generate electron-hole pairs.

: sign convention of p-n solar cell.

c.) LED's (light emitting diodes): are p-n junctions produced from compound semiconductors such as GaAs and I
The type of the S.C defines the output light color.

d.) Tunnel Diodes: Operation mainly depends on "tunnelling" due to higher impurity concentration as explained before.

The main property of a tunnel diode is its negative resistance region. In this region, when diode voltage is increased, current decreases as the opposite case of the conventional diode



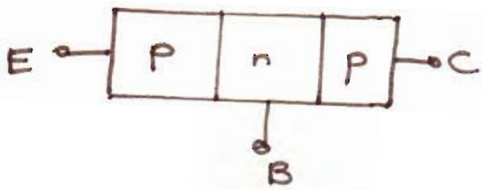
e.) Varactor Diode: Under reverse biasing, it is used as variable junction capacitor by varying reverse voltage. Nonlinear increase is observed in capacitance as the reverse voltage is decreased (see notes for details!)

Bipolar Junction Transistor - BJT :

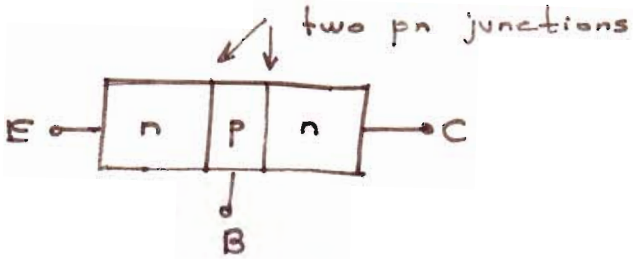
BJT was firstly investigated in 1948 by Bardeen and Brattain at the Bell Laboratories. The name of the transistor comes from transfer + resistor. Transistor basically transfers the current from small resistance to high resistance region and thus provides amplification. Basic structure of a transistor is composed of two pn junctions as npn or pnp, using doped semiconductors. Bipolar name comes from usage of both p and n type materials in the structure.

Basic Structure

BJT is a three terminal device having npn and/or pnp forms as:



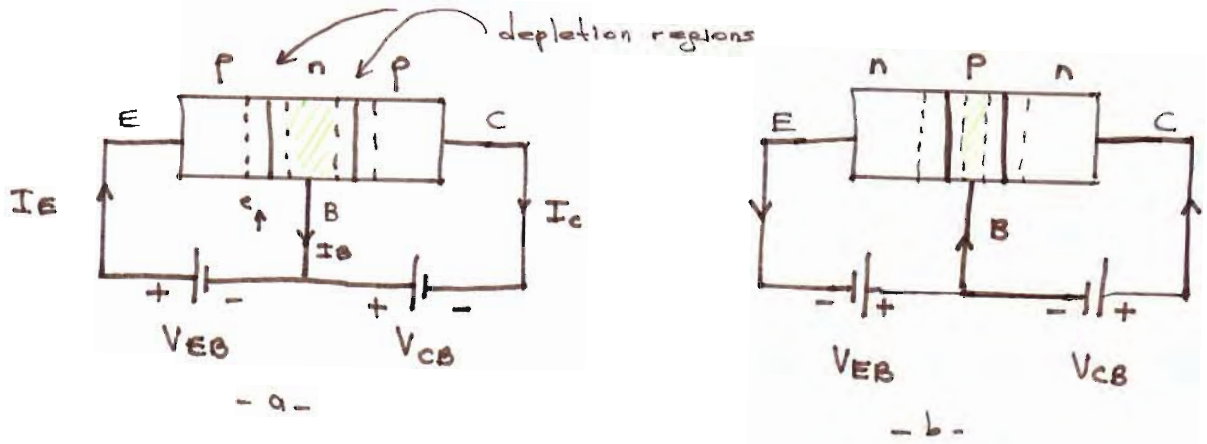
pnp type



npn type

E: Emitter C: Collector B: Base terminals

Base region is very narrow wrt emitter and collector regions (0.0001 to 0.001 cm). The main transistor operation take place in base region. For the proper operation of the transistor, external voltage supplies must be correctly connected to EB and BC terminals. For the normal operation (forward active mode) of the BJT, emitter-base junction is forward biased and collector-base junction is reverse-biased.



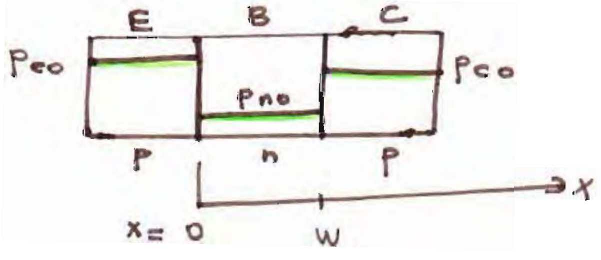
- a -

- b -

BJT in normal mode, a-) pnp type, b-) npn type.

The effective base region of the BJT is the region between the two depletion layers and depends on applied voltages. Base is actually very thin and lightly n-doped (p-doped) while the emitter is highly p-doped (n-doped) for pnp (for npn) type BJT. Collector is lightly p-doped (n-doped) wrt emitter in pnp (npn) type BJT. Minority carrier distribution in base region is very important for the BJT operation because when a voltage is applied to a pn junction (diode), majority carrier distribution is negligibly changed, but the minority carrier distribution depends strongly on the applied voltage.

Now, let us discuss pnp type transistor operation:



$p_{eo} > p_{co} \gg p_{no}$

: p_{eo}, p_{co}, p_{no} : hole distribution in each region for no bias case

In the above figure hole concentration is shown because holes are minority carries for n-type base region of pnp transistor.

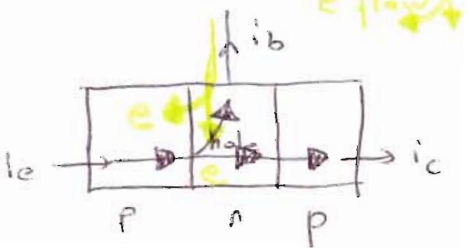
$p_{eo} = N_{AE}$: acceptor concentration in emitter region
$p_{co} = N_{AC}$: " " " collector "

$p_{eo} > p_{co}$; $N_{AE} > N_{AC}$

$n_{bo} = N_{DB}$: donor (e) concentratin in base region
$n_{eo} = N_{DE}$: " " " emit. "

For a proper operation $N_{AE} \gg N_{DB}$ $\Rightarrow$

$$\underbrace{n_{eo} = \frac{n_i^2}{N_{AE}}}_{\text{minority carrier concentration in E region}} \ll \underbrace{p_{bo} = \frac{n_i^2}{N_{DB}}}_{\text{minority carrier concentration in B region}} ;$$



Concentration of minority carriers at junction edge

$$= \text{equilibrium minority concentr.} \times \exp\left[\frac{q}{kT} \times (\text{applied voltage})\right]$$

We know that CB - junction is reverse biased, thus:

$$\boxed{p(w) = p_{bo}\, e^{-V_{CB}/0.026} \to 0}$$

means that concentration of holes at $x = w$ ≈ 0, for a reasonable bias voltage and at $x = 0$;

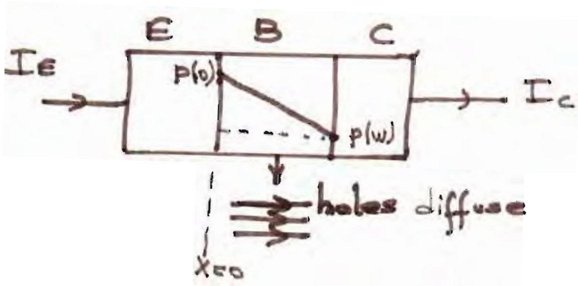
$$\boxed{p(0) = p_{bo}\, e^{V_{EB}/0.026}}$$

These equations indicate that base-collector junction absorbs holes while emitter-base junction is the source for the holes.

This is achieved due to proper connection of V_{EB} and V_{CB} potentials (one forward, one reverse biasing requirement for normal mode operation). Therefore diffusion curren is expected through the base region as a function of V_{EB} potential.

Base Diffusion Current:

The average distance for a hole to travel before recombining is given by L_p which is ≈ 0.005 cm and this is a wider length wrt base width. For, this reason little recombination of holes is expected in base region and thus diffusion current is constant at all points in this region. This also requires constant slope of the carriers as shown:



$$I_p = -qAD_p \frac{dp}{dx}$$

$$= qAD_p \frac{p(0) - p(w)}{W}$$

$$= \frac{qAD_p}{W} p_{bo} e^{V_{EB}/0.026}$$

⇒ This means that current from E to C; at room temp. depends mainly on V_{EB} potential. Due to Kirchoff's Law:

$$I_E = I_B + I_C$$

Thus, current through transistor can be controlled by adjusting the V_{EB} potential.

We have neglected recombination in the base region. But actually small fraction of the injected holes diffusing through base region recombine with electrons. In order to replace this electrons to conserve equilibrium, electrons flow from base lead resulting with outward current flow from base region. The base current is only about $\frac{1}{50}$ or $\frac{1}{100}$ of the emitter current, but it is very important for transistor operation.

$$I_E = I_B + I_c \quad ; \quad I_B \ll I_E \Rightarrow I_c \approx I_E$$

Generally parameter α is used to relate I_c and I_E as:

$$\alpha_{dc} = \left.\frac{I_c}{I_E}\right|_{V_{EB} = cnst} \lesssim 1$$

α = Current transfer ratio

for ac $\alpha = \frac{\Delta I_c}{\Delta I_E}$, but generally $\alpha = \alpha_{dc}$ is taken.

Typical range for α : 0.97 to 0.999.

Another important parameter for transistor operation is the forward transfer ratio of the collector current to base current defined as:

$$\beta = \frac{i_c}{i_B} = \frac{\alpha . i_E}{\underset{\substack{\downarrow \\ i_E - i_c}}{i_B}} = \frac{\alpha\, i_E}{i_E - \alpha\, i_E} = \frac{\alpha}{1-\alpha}$$

Thus, small base current can control large collector current. Typical range for small signal BJT's β: 100-300 and for power transistor β: 5-10.

Summary:

Under some approximations, transistor operation can be simply explained. If we ignore saturation current at the collector, collector current is assumed to be made up of holes injected at the emitter due to V_{EB} potential, which are not lost to recombination in the base. Base current is due to recombination of some of the holes in base region with the electrons of n-doped base material. Collector current i_c is proportional to the hole component of the emitter current i_{Ep} :

$$i_c = B\, i_{Ep}$$

Constant B is the fraction of injected holes from E to C through base region and called as base transport factor.

The total emitter current i_E is made up of the hole component (of) I_{Ep} and an electron component i_{En}, due to electrons injected from base-n to emitter-p. I_{En} is small wrt I_{Ep}. Then emitter efficiency is defined as :

$$\gamma = \frac{i_{Ep}}{i_{Ep} + i_{En}} \Rightarrow \frac{i_c}{i_E} = \frac{B\, i_{Ep}}{i_{En} + i_{Ep}} = B\gamma = \alpha$$

$B\gamma$ is called as factor α and named as current transfer ratio. For efficient transistor B and $\gamma \approx 1$.

We know that lost electrons in base region are re-supplied by I_B current which is in the form :

$$i_B = i_{En} + (1-B)\, i_{Ep}$$

$$\Rightarrow \quad \frac{i_c}{i_B} = \frac{B\, i_{Ep}}{i_{En} + (1-B)\, i_{Ep}} = \frac{B\gamma}{1-B\gamma} = \frac{\alpha}{1-\alpha} \equiv \beta$$

β is called as base to collector current amplification factor. i_c / i_B can also be determined from:

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

τ_p : hole lifetime before recombination

τ_t : hole transit time across base junction

Ex: $\tau_p = 10\ \mu sec$
$\tau_t = 0.1\ \mu sec$ $\Rightarrow \beta = \frac{\tau_p}{\tau_t} = 100$

Ex . (5.1) –