

**KWAME NKRUMAH UNIVERSITY OF SCIENCE AND
TECHNOLOGY, KUMASI**

FACULTY OF DISTANCE LEARNING

**EE 271: SEMICONDUCTOR DEVICES
[Credits: 2]**

By
K. DIAWUO
LECTURER
KNUST
July 2007

Contact Address

Dean

Faculty of Distance Learning

New Library Building

Kwame Nkrumah University of Science and Technology

Kumasi, Ghana

Phone: +233-51-60013

+233-51-60014

Fax: +233-51-60014

E-mail: cdce@fdlknust.edu.gh

kvcit@fdlknust.edu.gh

kvcit@avu.org

Website: www.fdlknust.edu.gh

Course Author

DR. KWESI DIAWUO is a senior lecturer at the Department of Electrical/Electronic Engineering and for the past 24years been teaching Semiconductor Devices, Digital Systems, Linear Electronic Circuits, Computer Architecture, Introduction to VLSI and Fault Diagnosis and Failure Tolerance at the Faculty of Electrical/Electronic and Computer Engineering. He holds a first degree, Masters and PhD degrees in Electrical/Electronic Engineering.

His areas of research interests include Microprocessor and VLSI applications, and Semiconductor devices.

He is also a member of the following professional institutions:

- Ghana Institute of Engineers (GhIE),
- Institute of Electrical and Electronic Engineers (IEEE).

Course Introduction

Course Content

PN Junction: Junction diodes and their static and dynamic properties at high frequencies, switching.

Metal-Semiconductor junctions: Energy band diagram of the Schottky barrier, MIS Schottky diode, ohmic contact, applications of Schottky-Barrier diodes, heterojunctions. Solar Cell and light-emitting diode: optical absorption in a Semiconductor, Photovoltaic effect, Schottky-barrier and MIS solar cells, LEDs, eye sensitivity and brightness, Quantum efficiency.

MOS Devices MOS capacitor, MOS transistor (DC characteristics, depletion - MOST and JFET). Polysilicon gate technology, Metal gate technology, comparison. Control of threshold voltage through ion implantation. Frequency and small signal properties of MOST, special MOS devices (CMOS, VMOS, DMOS, CCD). Examples of CCD in electro-optical applications and signal processing.

Bipolar Junction Devices: The BJT transistor action, Ebers-Moll's model, Hybrid- π and its equivalent circuit, the BJT as a switch, breakdown voltages, lateral PNP transistor, PNP devices, Noise source in BJT.

Course Objective

The objective is to study and understand the properties of electronic devices and their relation to circuit applications. It emphasizes on basic properties of semiconductors and also how these properties influence the behavior of semiconductor devices. Thus the course will deal with the inside of devices and it is highly conceptual.

References

1. Michael Shur, Introduction to Electronic Devices
2. Ben G. Streetman and Sanjay Banerjee, Solid State Electronic Devices, Fifth Edition.

Table of Contents

Contact Address	2
Course Author	3
Course Introduction	4
Table of Contents	5
List of Figures	7
Unit 1	9
FUNDAMENTALS OF SEMICONDUCTORS	9
SESSION 1-1: CRYSTAL STRUCTURE AND ENERGY BANDS, AND ELECTRONIC STRUCTURES	9
1-1.1 Introduction: Crystal Structure And Energy Bands	9
1-1.2 Electronic Structure	10
1-1.2.1 Number of states in a band	10
1-1.2.2 Insulators, metals and semiconductors	12
1-1.2.3 Conduction	13
1-1.2.4 Electrons and holes	15
1-1.2.5 Occupancy of available energy levels and filling the bands	16
1-1.2.6 Calculating n	17
1-1.2.7 Conduction band density of states	18
SESSION 2-1: INTRINSIC SEMICONDUCTORS AND EXTRINSIC SEMICONDUCTORS	19
2-1.1 Intrinsic Semiconductors	19
2-1.2 Extrinsic Semiconductors	20
2-1.2.1 Temperature dependence	22
Unit 2	25
CARRIER FLOW	25
SESSION 1-2: CONDUCTIVITY OF SEMICONDUCTOR AND DRIFT CURRENT	25
1-2.1 Conductivity Of Semiconductor	25
1-2.2 Drift Current	26
SESSION 2-2: DIFFUSION OF ELECTRONS AND HOLES AND GENERATION/RECOMBINATION	28
2-2.1 Diffusion Of Electrons And Holes	28
2-2.2 Generation/Recombination	29
2-2.2.1 Generation of electrons and holes	29
2-2.2.2 Recombination of electrons and holes	31
Unit 3	33
PN- DIODE DEVICES	33
SESSION 1-3: JUNCTION FORMATION, EQUILIBRIUM PROCESS AND BIASING	33
1-3.1 Junction Formation	33
1-3.2 Equilibrium Process	34

1-3.2.1 After equilibrium.....	35
1-3.2.2 Built-in-potential.....	36
1-3.2.3 Fields.....	36
1-3.3 Biasing.....	39
1-3.3.1 Reverse bias	39
1-3.3.2 Forward bias.....	40
SESSION 2-3: DIODE CURRENT, DC MODEL, METAL-SEMICONDUCTOR JUNCTION AND “OHMIC” CONTACTS TO SEMICONDUCTORS.....	42
2-3.1 Diode Current	42
2-3.2 DC Model	46
2-3.2.1 Small signal analysis.....	46
2-3.2.2 Small-signal model	47
2-3.3 Metal-Semiconductor Junction.....	48
2-3.3.1 Schottkey barriers (metal-semiconductor junctions)	48
2-3.3.1.1 metal band structure (half filled band).	48
2-3.3.1.2 reverse bias ($V < 0$)	50
2-3.3.1.3 forward bias	50
2-3.4 “Ohmic” Contacts To Semiconductors.....	51
Unit 4	54
TRANSISTORS AND PHOTODETECTORS.....	54
SESSION 1-4: TRANSISTORS.....	54
1-4.1 The Junction FET	54
1-4.1.1 Linear equations.....	55
1-4.1.2 Saturation case	55
1-4.1.3 AC response	55
1-4.2 Bipolar Transistors.....	56
1-4.3 MOSFET Transistors.....	58
1-4.3.1 The MOSFET current – voltage characteristics.....	61
1-4.3.2 Computation of MOS currents.....	61
1-4.3.3 Small signal analysis.....	62
SESSION 2-4: PHOTODETECTORS	68
2-4.1 Structure Of pn Photodiode	69
2-4.2 Avalanche Detectors	70
2-4.3 Photodetector Materials.....	70

List of Figures

Figure 1.1 “free” Bloch electrons
Figure 1.2 Region of k-space
Figure 1.3 States in Bands
Figure 1.4 Energy band of carbon
Figure 1.5 Energy band of silicon
Figure 1.6 Energy band of an alkali metal
Figure 1.7 Electrons moving under influence of crystal potential ϕ , and externally applied electric field $\vec{\mathcal{E}}$
Figure 1.8 E-k curve
Figure 1.9 E-k curves
Figure 1.10 Band structure of a semiconductor
Figure 1.11 Probability ($f(E)$) of occupation of the energy level at $E = E_F$
Figure 1.12 Semiconductor for which the bottom of the conduction band has spherical constant energy
Figure 1.14 Silicon crystal doped with phosphorus
Figure 1.15 Energy levels of a doped silicon crystal.
Figure 1.16 Energy band structure of a phosphorus doped silicon crystal.
Figure 1.17 Silicon crystal doped with boron and Energy levels of a doped silicon crystal.
Figure 1.18 Energy band structure of a boron doped silicon crystal.
Figure 1.19 Plot of electron density against temperature.
Figure 2.1 Model for electron motion
Figure 2.2 Movement of electrons between scattering events
Figure 2.3 v is limited by v_{sat} the “scatter limited velocity”
Figure 2.4 An electron with an area A
Figure 2.5 A piece of conductor
Figure 2.6 Random thermal motion of electrons
Figure 2.7 Linear drop in concentration
Figure 2.8 E-k relationship for real solids.
Figure 2.9 Direct gap material
Figure 2.10 Model for light absorption
Figure 2.11 Indirect gap material
Figure 2.12 Direct radiative recombination
Figure 2.13 Recombination through Midgap Energy levels (traps)
Figure 2.14 Thermal Generation (trap aided)
Figure 3.1 n and p type semiconductor samples placed together.
Figure 3.2 Electrons and holes recombine, producing a depletion region
Figure 3.3 Equilibrium is achieved (no net current flow).
Figure 3.4 The same Fermi energy throughout the system at equilibrium
Figure 3.5 Charge Density, Field, Potential $V(x)$, carrier Density and Electron Energy Diagrams
Figure 3.6 Reverse bias
Figure 3.7 Depletion area becomes wider and the potential barrier grows higher.
Figure 3.8 Forward bias
Figure 3.9 Forward bias lower the height of the potential barrier ($\phi - V$)
Figure 3.10 A typical I-V curve for a diode
Figure 3.11 Diode in breakdown region
Figure 3.12 DC model
Figure 3.13 Circuit and Loadline

Figure 3.14 DC and AC loadlines
Figure 3.15 Small signal model
Figure 3.16 Metal band structure
Figure 3.17 Semiconductor Band Structure
Figure 3.18 Metal and semiconductor in close proximity
Figure 3.19 Metal and Semiconductor in contact (equilibrium condition).
Figure 3.20 Reverse bias
Figure 3.21 Forward bias
Figure 3.22 A metal-semiconductor junction
Figure 3.23 Bands of the metal and semiconductor placed together
Figure 3.24 Bending of the bands
Figure 3.25 I-V characteristics of a Schottkey barrier and an ohmic contact
Figure 4.1 The Junction FET (JFET)
Figure 4.2 AC model
Figure 4.3 Basic npn structure
Figure 4.4 NPN band structure in equilibrium with no applied biases.
Figure 4.5 npn biased in forward active mode
Figure 4.6 Electron transport current in a npn transistor
Figure 4.7 Flow of electrons
Figure 4.8 The NPN MOSFET
Figure 4.9 Flow of electrons to the positive gate potential
Figure 4.10 No inversion layer
Figure 4.11 Small signal model of MOSFET
Figure 5.1 Separation of an electron and hole
Figure 5.2 External current flows between the p and n regions
Figure 5.3 The current-voltage characteristics of a p - n junction under various levels of illumination
Figure 5.4 Simple pn photodiode structure
Figure 5.5 A p - i - n diode structure
Figure 5.6 The principle of operation of an avalanche photodiode

FUNDAMENTALS OF SEMICONDUCTORS

SESSION 1-1: CRYSTAL STRUCTURE AND ENERGY BANDS, AND ELECTRONIC STRUCTURES

1-1.1 Introduction: Crystal Structure And Energy Bands

A great majority of semiconductor materials finding practical application in electronics are in the form of crystalline solids like that of diamond and zincblende. For example, the crystalline structure of silicon (Si) is that of diamond whilst that of the gallium arsenide (GaAs) is of the form of zincblende. In the crystalline solid, atoms forming the primitive cell (lattice) are arranged in an ordering manner, which by repeating itself forms the crystal. It is often necessary to refer to systems of planes within the crystal that run in certain directions, intersecting certain sets of atoms. It is the orientation of such systems of planes that determine the properties of the crystal. These sets of planes are specified by sets of three numbers called Miller indices.

The bond formation of crystalline silicon material is of the covalent –electron pair bond in which an atom shares its four valence electrons with its four nearest neighbours. Thus, no free electrons are available for conduction at absolute zero temperature. The conductivity is however, very sensitive to tiny amounts of impurity atoms and increases with rising temperature. In the semiconductor, the valence electrons group together to occupy a band of energy levels, called the valence band. The next higher band of allowed energy levels called the conduction band is separated from the former by a forbidden gap of energy, e.g., it is this energy gap E_g that classifies a material as a metal, insulator or semiconductor. In the case of metals, E_g is zero whilst E_g is large in insulators for an electron to be successfully ejected from the valence band to the conduction band. In a semiconductor the forbidden energy gap is not too large, for example, $E_g = 1.12$ eV in silicon and 1.43 eV in GaAs. This allows electrons in the valence band to be promoted to the conduction band by thermal agitation where they contribute to current conduction under the influence of an external electric field. For each successful transition of valence electron to the conduction band, a hole is left behind in the valence band. Thus, current conduction in semiconductors is due to both electrons (negatively charged) and holes (positively charged). In terms of quantum mechanics, the electron in a crystalline lattice is considered as a particle in constant motion with momentum k and energy E . The energy-momentum (E-K) relationship is usually obtained by solving the Schrodinger equation.

$$\left[-\frac{h^2}{8\pi^2 m} \nabla^2 + V(r) \right] \Phi_k(r) = E_k \Phi_k(r) \quad (1.1)$$

Where $V(r)$ is the periodic potential energy of the lattice, h is the Planck's constant and m the mass of electron. $\phi_k(r)$ is the solution of equation (1.1) which has the form:

$$\phi_k(r) = e^{jk \cdot r} U_n(k, r) = \text{Bloch function} \quad (1.2)$$

Where $U_n(k, r)$ is periodic in r with the periodicity of the direct lattice and n is the band index. The band structure resulting from equation (1.1) will be discussed later in the course for Si and GaAs. In either crystal the energy gap E_g is the difference between the lowest conduction-band minimum and the valence band maximum. The mass of the carriers (which is different from the electron mass m_o in free space) depends on the energy band they occupy, the momentum of the carriers and the direction of the applied field.

1-1.2 Electronic Structure

1-1.2.1 Number of states in a band

The $E(k)$ relationship tells us where allowable electronic states are, we also need to know how many there are:

A crystal of size L is a finite potential well or “Box” for a “free” (we find that electrons in a crystal could move between the walls like free electrons) electron, Fig 1.1

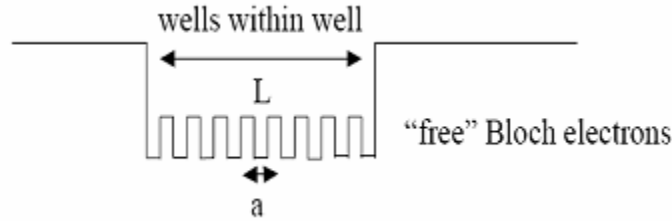


Figure 1.1: “free” Bloch electrons.

We know that $L = Ma$ where M = the number of atoms in crystal and a is the atomic spacing. To account for the fact that the crystal has finite size, we need to impose the following boundary condition on the wave function Ψ .

$$\Psi(x) = \Psi(x + L) \quad (1.3)$$

This results in only certain discrete values of k being allowed. It can easily be verified that the previous B.C. leads to k values of the form (apply this B.C. to the free electron solution):

$$k = n \frac{2\pi}{L} \quad (1.4)$$

in other words, Δk between states in k -space is $\frac{2\pi}{L}$. So we have Fig. 1.2

Therefore the spacing between allowed k values = $\frac{2\pi}{L}$ due to finite crystal size L . A region of k -

space of volume Ω will contain $\frac{\Omega}{(2\pi/L)^3} = \frac{\Omega V}{8\pi^3}$ allowed values of k . Or the number of k values per unit volume of k -space (density of k states) is $\frac{\Omega}{8\pi^3}$ where V is the volume of the crystal (L^3).

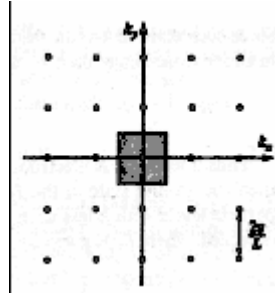


Figure 1.2: Region of k -space

In one dimension, the density of k states in k -space is just $L / 2\pi$ as we have one state every $2\pi / L$. We also find that for a crystal electron K ranges from $-\frac{\pi}{a}$ to $\frac{\pi}{a}$ and we have that the number of states equals the length in k space ($a/2\pi$) times the density of states ($2\pi/L$) for a given energy band. Therefore, there are a total of

$\frac{2\pi}{a} \cdot \frac{L}{2\pi} = \frac{L}{a} = M$ allowed Bloch wave functions or states in each band. The Pauli Exclusion

Principle says we can place 2 electrons in each Bloch state (2 because electrons have 2 possible spins). We therefore have $2M$ electron states in each band and each atom will provide 2 electron states to each band.

Discrete states separated by

$$\Delta k = \frac{2\pi}{Ma}, \quad Ma = L \quad (1.5)$$

As M is very large the band is quasi continuous and size of crystal affects the state density, Fig 1.3.

We can make the effective crystal very small and have only a few states per band. This is used for band gap engineering with thin layers or small structures.

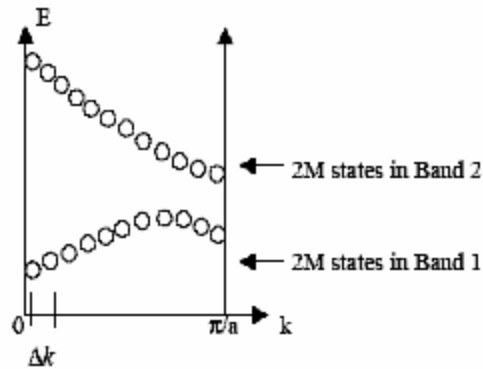


Figure 1.3: States in Bands

1-1.2.2 Insulators, metals and semiconductors

We have found that the electronic structure of a crystal consists of infinite number of energy bands, separated by band gaps in which there are no allowed energies. We also know: Each band can hold $2M$ electrons (M is the number of atoms in crystal)

- Bands fill up from bottom (lowest energy). As each atom has N electrons:
 - If N is even then, $N/2$ bands are completely full and the material is most likely an insulator.
 - If N is odd, the top band is half full and $(N-1)/2$ bands are completely full and the material is a conductor.

Consider carbon in the form of a diamond crystal (16 electrons), where the band structure depends on atomic spacing and therefore the crystal structure, Fig 1.4. Energy Band Structure: (we only show the top two bands). For carbon, the band gap between the first empty band and the last filled band is 5.4eV .

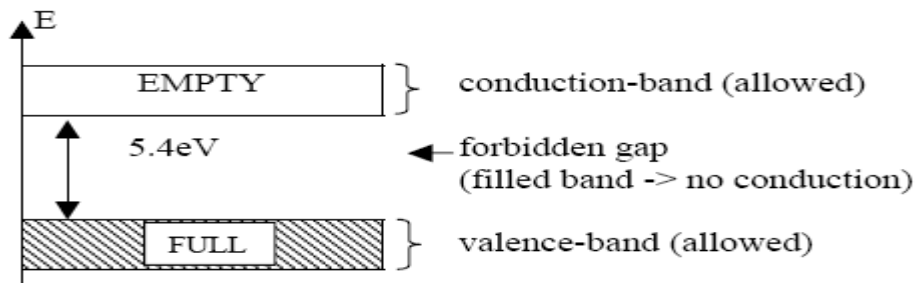


Figure 1.4: Energy band of carbon

Carbon has just enough electrons to fill the valence band (electrons tend to occupy lowest available levels) and seven lower bands, and the conduction band is empty. We have assumed that all electrons go to the lowest available E level. However, heat (temperature) is kinetic energy and some electrons will be raised to higher E states for $T > 0\text{K}$ (absolute 0). At room temperature, $kT \approx 0.0259\text{eV}$ and is the average kinetic energy of electrons (k is Boltzmann's constant $= 1.38 \times 10^{-23} \text{ J/K}$). Electron energies are described by an energy distribution function that is a function of T . The larger T the hotter (more energetic) the electrons are. As kT at room temperature $\ll E_g$ of carbon, we have that it is almost impossible for electron to acquire enough

thermal (kinetic) energy to be activated into the conduction band and carbon is an insulator.

Next consider silicon (14 electrons) Fig 1.5; same crystal structure as carbon, but smaller band gap.

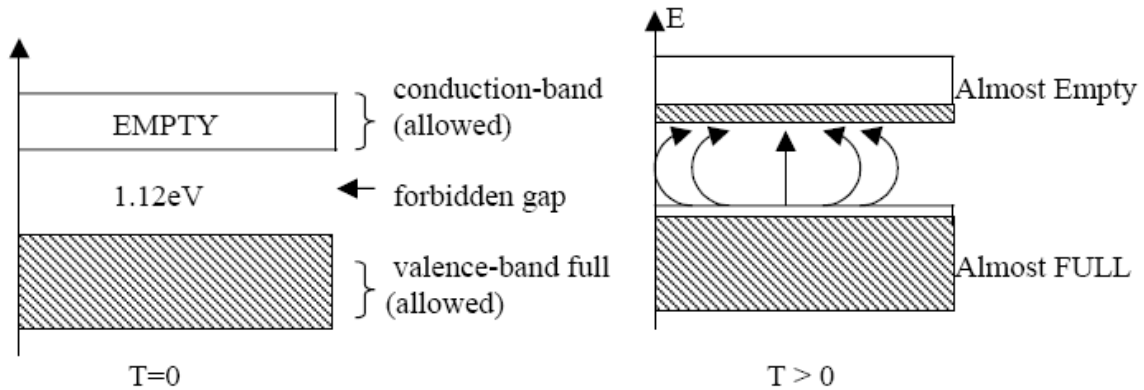


Figure 1.5: Energy band of silicon

Now there is a small (but finite) chance for a few electrons to be thermally excited from valence band to conduction band.

Finally consider an alkali metal Fig 1.6: Odd number of electrons: 1/2 filled band

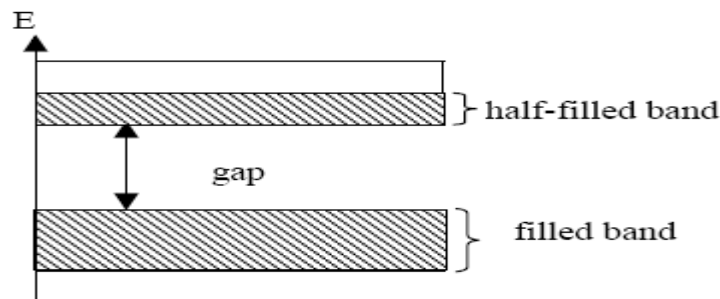


Figure 1.6: Energy band of an alkali metal

1-1.2.3 Conduction

We know that conduction is the movement, i.e, acceleration of an electron. To be accelerated an electron must gain energy and change its state/momentum. If a band is completely filled an electron cannot change state as all states are filled (grid lock!). Therefore, a band must be partially empty for conduction to occur. Electrons in the conduction band are free to move through crystal in response to electric field \vec{E} , as the conduction band is mostly empty. We need to find how nearly free electrons move in the conduction band.

Consider electron moving under influence of crystal potential ϕ , and externally applied electric field \vec{E} .

Example:

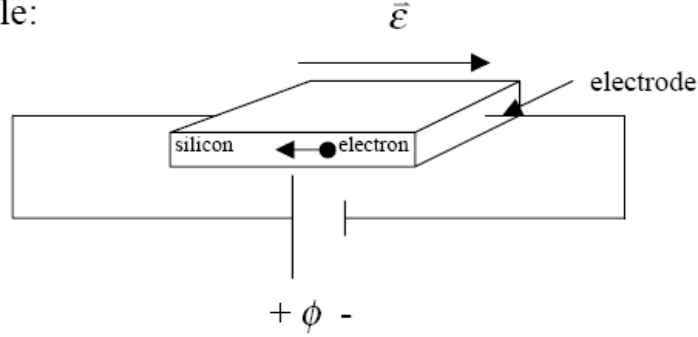


Figure 1.7: electrons moving under influence of crystal potential ϕ , and externally applied electric field $\vec{\varepsilon}$.

ϕ - electro-static potential and $\varepsilon = -\frac{d\phi}{dx}$

We model the electron as a wavepacket of Bloch waves (free electrons almost) centered round wave vector k . We expect an external field to accelerate the electron, increasing E and k and change the electron's state, Figure 1.8.

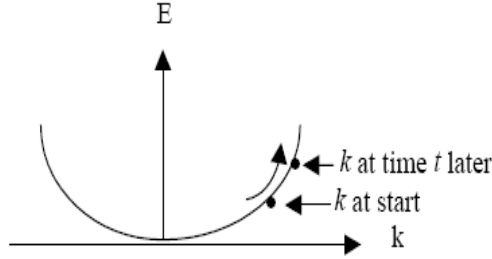


Figure 1.8: E-k curve

We can write:

$$E = -q\phi, \frac{dE}{dt} = -q \frac{d\phi}{dt}, \text{ and we know: } V_g = \frac{\Delta x}{\Delta t} = \frac{d\omega}{dk}$$

Rewriting we have:

$$\frac{dE}{dt} = -q \frac{d\phi}{dx} \cdot \frac{dx}{dt} = -q \frac{d\phi}{dx} \cdot V_g, V_g \text{ is the group velocity}$$

$$\therefore \frac{dE}{dk} \cdot \frac{dk}{dt} = -q \frac{d\phi}{dx} \cdot V_g, \text{ but } V_g = \frac{d\omega}{dk} = \frac{1}{\hbar} \frac{dE}{dk}$$

$$\therefore V_g \cdot \frac{d}{dt} \hbar k = -q \frac{d\phi}{dx} \cdot V_g, \text{ or } \frac{d}{dt} \hbar k = -q \frac{d\phi}{dx}$$

For a free electron, $\hbar k$ = momentum = p and $\frac{dp}{dt} = \frac{d}{dt}(\hbar k) = -q \frac{d\phi}{dx} = -q\varepsilon = F$, F is the electric

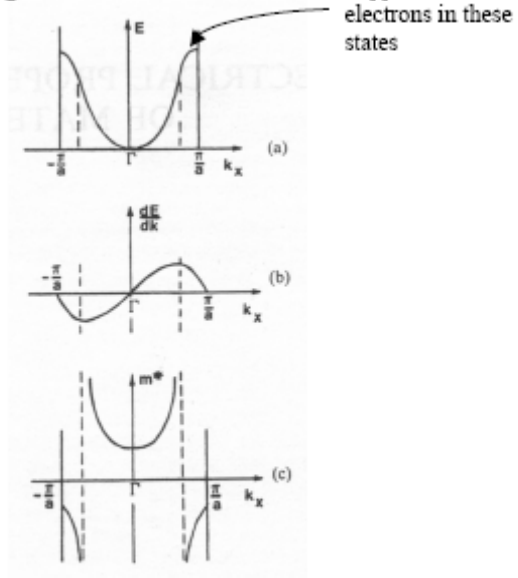
force. Therefore, for an electron in a periodic potential we identify $\hbar k$ as the “crystal momentum” which changes in response to external forces. We can also derive:

$$\frac{dV_g}{dt} = \frac{d}{dt} \left(\frac{1}{\hbar} \frac{dE}{dk} \right) = \frac{1}{\hbar} \frac{d}{dk} \left(\frac{dE}{dt} \right) = \frac{1}{\hbar} \frac{d}{dk} \left(\frac{dE}{dk} \right) \left(\frac{dk}{dt} \right) = \frac{1}{\hbar} \frac{d^2 E}{dk^2} \left(\frac{dk}{dt} \right)$$

$$\text{But } a = \frac{dV_g}{dt} = \left(\frac{1}{\hbar^2} \frac{d^2 E}{dk^2} \right) \frac{d}{dt} \hbar k = \left(\frac{1}{\hbar^2} \frac{d^2 E}{dk^2} \right) (-q\varepsilon)$$

This has the form $a = \frac{F}{m^*}$ where $m^* = \left(\frac{1}{\hbar^2} \frac{d^2 E}{dk^2} \right)^{-1}$ which relates the curvature of the band to mobility and to “effective mass”. We can show that a free electron “band” gives an effective mass equal to that of an electron. Therefore, electrons in a crystal are accelerated in response to an external force just as though they were free electrons with effective mass m^* .

Consider the following E-k curve:



$$\frac{d^2 E}{dk^2} < 0$$

Figure 1.9: E-k curves

These electrons have negative mass, move backwards in response to force. We interpret this as a +ve mass particle with a +ve charge and it's called a **hole**!

1-1.2.4 Electrons and holes

The energy band structure of real semiconductors can be very complex. However to understand many properties it is often enough to consider a band structure as shown in fig 1.10 below. We might regard this band representation as being typical for a semiconductor:

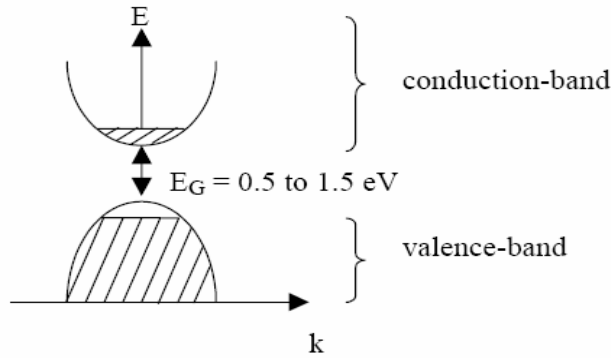


Figure 1.10: Band structure of a semiconductor

At 0K in a pure material, conduction band is completely empty, valence band completely filled, and no conduction (insulation). At 300K, a few electrons are promoted to the conduction band, leaving a few empty states near the top of the valence band.

At T'0K the charge carriers in a solid are in constant thermal motion. In the absence of external fields, this motion is quite random resulting from carriers scattered by collisions. As the motion is random there is no net displacement of charge in any direction and no net current flow ($I=0$). Clearly, with a field the electrons excited into the conduction band can contribute to current flow. How do the empty states in the valence band contribute to current flow?

As there are now empty states in the valence band, the electrons in this band can also respond to an applied field and contribute to the current flow. For the electrons occupying the vacant states,

$$\frac{1}{\hbar^2} \frac{d^2 E}{dk^2} < 0 \quad (1.6)$$

(Negative!) and the electrons will move in same direction as electric field (wrong way!)

Therefore, the electrons contribute to current as though they were positively charged, positive effective mass particles. It is therefore convenient to describe valence band as containing imaginary positive charge, positive mass particles called holes.

In a semi-conductor, we therefore have two charge carriers

- Electrons +ive mass, -ive charge (conduction band)
- Holes +ive mass, +ive charge (valence band)

1-1.2.5 Occupancy of available energy levels and filling the bands

We have determined $E(k)$ and a density of states. We also need a way of determining if a state is filled or not at a given temperature. The distribution of the energies of a large number of particles and its change with temperature can be calculated by means of statistical considerations. The probability of electrons occupying a level at energy E is given by:

$$f(E) = \frac{1}{1 + e^{(E-E_F)/k_s T}} \quad \text{Fermi-Dirac distribution function} \quad (1.7)$$

Where E_F is a parameter called the Fermi energy, which represents the highest energy that the electrons assume at $T = 0\text{ K}$. At any temperature above absolute zero ($T \neq 0$) the probability ($f(E)$) of occupation of the energy level at $E = E_F$ is $1/2$ (note there may not be an actual state at the Fermi energy!) The function is plotted in fig 1.11 below. The higher the T the more “smearing”.

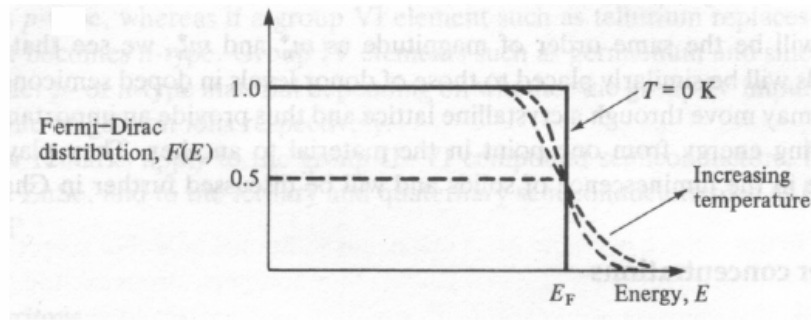


Figure 1.11: Probability ($f(E)$) of occupation of the energy level at $E = E_F$.

We can see from this figure that at $T = 0$ all levels that have an energy smaller than E_F are completely filled with electrons, whereas higher energy states are empty. Thus if an energy level is completely occupied by electrons, the Fermi distribution function $F(E)$ equals 1 (certainty); for an empty energy level one obtains $F(E) = 0$. The figure also shows that the probability of occupation of states above E_F is finite for $T > 0\text{ K}$ and that there is a corresponding probability that states below E_F are empty.

1-1.2.6 Calculating n

To calculate the carrier concentrations in energy bands we need to know the following parameters:

- (i) The *distribution of energy states* or levels as a function of energy within the energy band ($g(E)$).
- (ii) The *probability* of each of these states being occupied by an electron ($f(E)$).

So we can write:

$$n = \text{density of electrons} = \int_{E=0}^{\infty} f(E)g(E)dE \quad (1.8)$$

where: $f(E)$ - probability of occupancy at E and $g(E)$ - density of states in energy at E
The hole density is given by:

$$p = \text{density of electrons} = \int_{E=0}^{\infty} [1 - f(E)]g(E)dE \quad (1.9)$$

If the occupation probability of a level is given by $f(E)$, the probability that this level will not be occupied by an electron, i.e., the probability of occupation by a hole is:
 $[1 - f(E)]$ - probability of a missing electron at E (non occupancy)

1-1.2.7 Conduction band density of states

(Previously we calculated total number of states, not distribution.) Consider a semiconductor for which the bottom of the conduction band has spherical constant energy surfaces for free electrons with a mass m^* figure 1.12.

(Spherical is 3D version of parabolic (k) .)

$$E = \frac{\hbar^2 k^2}{2m^*}, \quad k = \frac{\sqrt{2m^* E}}{\hbar} \quad \text{or} \quad k^2 = k_x^2 = k_y^2 = k_z^2$$

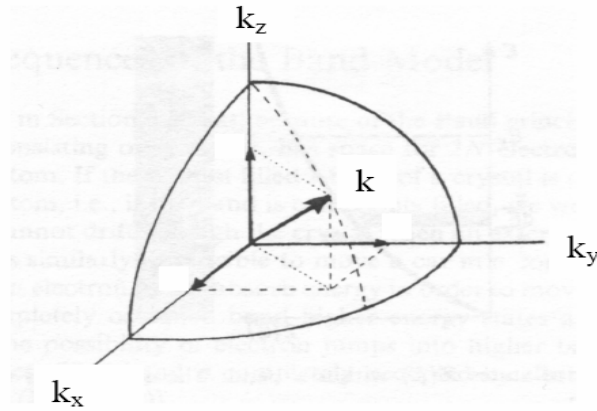


Figure 1.12: Semiconductor for which the bottom of the conduction band has spherical constant energy

A sphere of radius k in k -space contains $\left(\frac{4}{3}\pi k^3\right) \left(\frac{L^3}{(2\pi)^3}\right) \cdot 2$ allowed states (which are uniformly distributed in 3D k -space). Therefore the number of energy levels per unit volume (divide by L^3) is:

$$N(E) = 2 \cdot \frac{4\pi}{3} \cdot \frac{1}{(2\pi)^3} \cdot \frac{(2m^*)^{3/2}}{\hbar^3} E^{3/2} \quad (1.10)$$

The density of states in energy is therefore:

$$g(E) = \frac{d}{dE} N(E) = \frac{1}{2\pi^2} \cdot \frac{(2m^*)^{3/2}}{\hbar^3} E^{1/2} \quad (1.11)$$

Note that the density of states ($g(E)$) may be defined as the number of energy states per unit energy per unit volume. Returning to the calculation of carriers concentrations:

We have that the density of electrons in conduction band is:

$$\int_{E=0}^{\infty} f(E)g(E)dE$$

For $E \gg E_F$ we use the following approximation for $f(E)$ (Boltzmann distribution):

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

which is valid for the tail of the distribution when E is well away from E_F .
we get:

$$\therefore n = e^{-(E_C-E_F)/kT} \int_{E=0}^{\infty} g(E) e^{-(E-E_C)/kT} dE \quad (1.12)$$

Define:

$$N_C = \int_{E=0}^{\infty} g(E) e^{-(E-E_C)/kT} dE \quad (1.13)$$

The solution of this integral can be shown to be

$$N_C = \frac{1}{4} \left(\frac{2m^* kT}{\pi \hbar^2} \right)^{3/2} \quad (1.14)$$

with N_C - *effective density of states* in conduction band and thus giving:

$$n = N_C e^{-(E_C-E_F)/kT} \quad (1.15a)$$

$$\text{We can also get for } p: \quad p = N_V e^{-(E_F-E_V)/kT} \quad (1.15b)$$

With

$$N_C = \frac{1}{4} \left(\frac{2m^* kT}{\pi \hbar^2} \right)^{3/2} \quad (1.16)$$

SESSION 2-1: INTRINSIC SEMICONDUCTORS AND EXTRINSIC SEMICONDUCTORS

2-1.1 Intrinsic Semiconductors

A perfect semiconductor crystal containing no impurities or lattice defects is called an intrinsic semiconductor. As the carriers are generated in pairs, the concentration n of electrons in the conduction band equals the concentration p of holes in the valence band, $n = p = n_i$, (n_i - intrinsic carrier concentration) we can use $n = p$ to get:

$$\therefore N_C e^{-(E_C - E_F)/kT} = N_V e^{-(E_F - E_V)/kT} \quad (1.17)$$

and,

$$kT \ln \left(\frac{N_C}{N_V} \right) = E_C - E_F - E_F + E_V$$

giving,

$$E_F = \frac{E_C + E_V}{2} - \frac{kT}{2} \ln \left(\frac{N_C}{N_V} \right) \quad (1.18)$$

the second term on the right hand side is usually very small and $E_F \approx \frac{1}{2}E_g$.

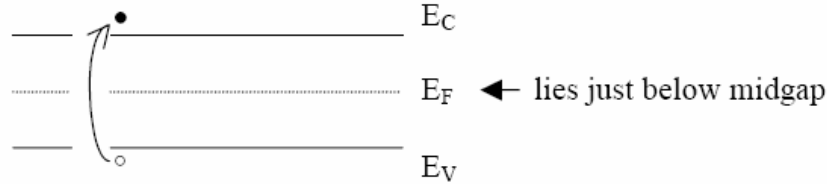


Figure 1.13: Energy band diagram

we also have:

$$np = n_i^2 \quad (1.19)$$

called the mass action law.

$$\text{And} \quad n_i = \sqrt{N_C N_V} e^{-(E_C - E_F)/kT} \quad (1.20a)$$

$$\text{Or} \quad n_i = \sqrt{N_C N_V} e^{-E_g/2kT} \quad (1.20b)$$

It is important in devices to control n and p concentrations and suppress the influence of intrinsic concentration. These equations are important in establishing upper limits in semiconductor operating temperature. We generally require $ni \ll$ (minimum doping density in device) and, practically, this means we need doping concentrations above 10^{14} cm^{-3}

2-1.2 Extrinsic Semiconductors

To engineer semiconductors we modify n and p by adding impurities. Suppose an Si atom is replaced with a phosphorus atom (an example of “doping”) We frequently use a diagram such as figure 1.14 to represent the atomic structure.

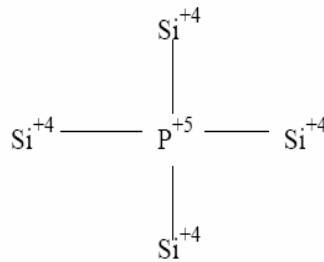


Figure 1.14: silicon crystal doped with phosphorus

We see that the phosphorus ion core has an extra +ve charge – this means that the periodic potential is disrupted and we get a localized energy level, E_D . The phosphorus also provides an extra electron to fill this level. We find that $E_C - E_D \approx 45$ meV that is very small, figure 1.15.

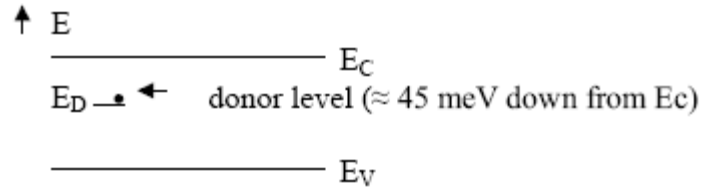


Figure 1.15: Energy levels of a doped silicon crystal.

This electron is easily thermally donated to the conduction band at room temperature as $E_C - E_D$ is small compared with kT . This means that at room temperature $n \approx N_D$ – we call this the complete ionization (only true if $n_i \ll N_D$). Therefore by doping Si crystal with phosphorus, we increase the free electron concentration – this is “n-type doping” E_F moves closer to E_C with increases in N_D , figure 1.16 (E_F concept for metals extended to other materials, probability of occupancy = 0.5)

We write

$$n = N_C e^{-(E_C - E_F)/kT}$$

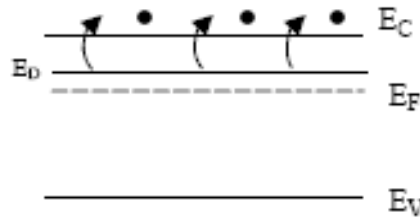


Figure 1.16: Energy band structure of a phosphorus doped silicon crystal.

Still have $np = n_i^2$ for the material without any applied bias. At constant temperature, p is reduced as n increases and $np = n_i^2 = \text{const}$. At low temperature, electrons return to donor levels; the dopant is frozen out. Next, suppose Si atom is replaced with Boron, again, we have a perturbed lattice and a local E-level created.

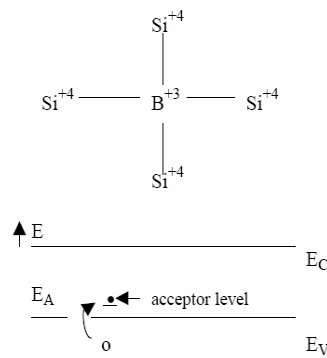


Figure 1.17: Silicon crystal doped with boron and Energy levels of a doped silicon crystal.

Boron is missing an electron and accepts an electron from valence band, creating a hole. Therefore doping with B increases hole concentration. We call this *p*-type doping, the electron concentration n is reduced fig 1.18. E_F moves closer to E_V .



Figure 1.18: Energy band structure of a boron doped silicon crystal.

$p \approx N_A$ if $p_i \ll N_A$ and at a reasonable temperature. At low temperatures, electrons return to valence band, holes are frozen out.

2-1.2.1 Temperature dependence

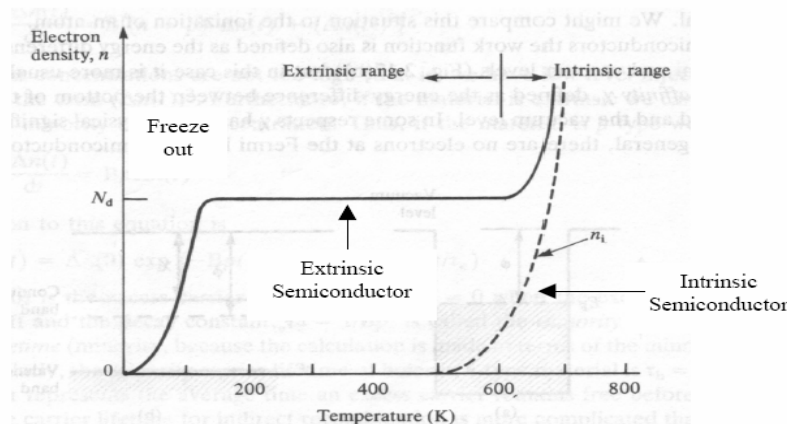


Figure 1.19: Plot of electron density against temperature.

- At low temperature electrons drop to lowest E state.
- At medium temperature, donors/acceptors activated.
- At high temperature large amount of excitation over the bandgap (n_i becomes large)

Calculation of temperature limits.

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

Look at two cases as limits:

- (1) $n = N_D/2$ – lower T limit (Impurities freezing out)
- (2) $n_i = N_D/2$ – upper T limit (Intrinsic concentration taking over)

Lower Limit

Define useful activation at 50% of donor levels occupied when $E_F = E_D$. We have typically $E_C - E_D = E_C - E_F = 0.05\text{eV}$

Therefore,
$$n = N_D / 2 = N_c \left(\frac{T}{300} \right)^{\frac{3}{2}} e^{-(E_c - E_F)/kT} \quad (1.21)$$

And

$$5 \times 10^{15} = 5.4 \times 10^{15} (T)^{\frac{3}{2}} e^{-580.4/T}$$

Giving, $T = 86 \text{ K}$

Upper Limit – when $n_i = N_D/2$

$$n_i = N_D / 2 = \sqrt{N_c N_v \left(\frac{T}{300} \right)^{\frac{3}{2}} e^{-E_g/2kT}} \quad (1.22)$$

$$5 \times 10^{15} = 3.33 \times 10^{15} (T)^{\frac{3}{2}} e^{-1.6 \times 10^{-19}/2kT}$$

$T = 675 \text{ K}$

Worked Examples

Useful equations:

$$n = N_c e^{-(E_c - E_F)/kT}, \quad p = N_v e^{-(E_F - E_v)/kT} \quad \text{or} \quad n = n_i e^{-(E_F - E_i)/kT}, \quad p = n_i e^{-(E_i - E_F)/kT} \quad \text{and} \quad np = n_i^2$$

Charge neutrality relationship: $p + N_A^+ = n + N_A^-$ or $p + N_D = n + N_A$ (total ionization)

An intrinsic Si sample (sample #1) at 300K, what n , p and E_F ?

Solution

$$\Rightarrow n = p = n_i = 1.5 \times 10^{10} \text{ cm}^{-3}, \quad E_F = E_i$$

If sample #1 is doped with $6 \times 10^{15} \text{ cm}^{-3}$ donors. What's n , p and E_F ?

Solution

$$\text{Since } N_D = 6 \times 10^{15} \text{ cm}^{-3}, \quad n_i = 1.5 \times 10^{10} \text{ cm}^{-3}, \quad N_D \gg n_i$$

$$n = N_D = 6 \times 10^{15} \text{ cm}^{-3}, \quad p = n_i^2 / n = (1.5 \times 10^{10})^2 / 6 \times 10^{15} = 3.75 \times 10^4 \text{ cm}^{-3}$$

$$E_F - E_i = kT \ln (n/n_i) = 0.0259 \ln \left(\frac{6 \times 10^{15}}{1.5 \times 10^{10}} \right) = 0.334 \text{ eV}$$

If sample #1 is also doped with $2 \times 10^{15} \text{ cm}^{-3}$ acceptors in addition to $6 \times 10^{15} \text{ cm}^{-3}$ donors, what is n , p , and E_F ?

Solution

$$N_D - N_A = 4 \times 10^{15} \text{ cm}^{-3} \text{ donors} \gg n_i = 1.5 \times 10^{10} \text{ cm}^{-3}$$

$$n = N_D - N_A = 4 \times 10^{15} \text{ cm}^{-3}$$

$$p = n_i^2 / n = 5.63 \times 10^4 \text{ cm}^{-3}$$

$$E_F - E_i = KT \ln (n/n_i) = 0.0259 \times \ln \left(\frac{4 \times 10^{15}}{1.5 \times 10^{10}} \right) = 0.324 \text{ eV}$$

If $N_D = N_A = 6 \times 10^{15} \text{ cm}^{-3}$, what is n , p , and E_F ?

Solution

$N_D - N_A = 0$, donors and acceptors negate each other

So $n = p = n_i$, $E_F = E_i$

The sample is compensated and looks like intrinsic

General solutions to problems above:

Assume a sample with intrinsic concentration n_i , doped with N_A & N_D . we want to calculate n , p and E_F .

Solution

$$p + N_D = n + N_A \quad (1)$$

$$np = n_i^2 \quad (2)$$

from (2), we have $p = n_i^2 / n$ (3) plug in (1)

$$\begin{aligned} \text{we have } n_i^2 / n + N_D &= n + N_A \\ \Rightarrow n^2 + (N_A - N_D) \cdot n - n_i^2 &= 0 \end{aligned}$$

$$\Rightarrow n = \frac{N_D - N_A}{2} + \sqrt{\left(\frac{N_D - N_A}{2} \right)^2 + n_i^2}$$

$$p = \frac{N_A - N_D}{2} + \sqrt{\left(\frac{N_D - N_A}{2} \right)^2 + n_i^2}$$

$$E_F - E_i = KT \ln (n/n_i)$$

Temperature Dependence

If for some reason n_i of sample #1 becomes 10^{17} cm^{-3} , and the sample is still doped with $6 \times 10^{15} \text{ cm}^{-3}$ donors. What does the sample look like? n-type? p-type? or intrinsic?

Solution

\Rightarrow Since $N_D \ll n_i$, $n = p = n_i$, $E_F = E_i$ so the sample is intrinsic.

CARRIER FLOW

SESSION 1-2: CONDUCTIVITY OF SEMICONDUCTOR AND DRIFT CURRENT

1-2.1 Conductivity Of Semiconductor

We know that Bloch waves (which describe electron flow) move freely through a perfect crystal. We also have the definition from classical physics:

$$\text{Acceleration} = \text{Force/mass ie} \quad a = \frac{F}{M} \quad (2.1)$$

If the force is a constant, velocity (dv/dt) goes to infinity and current will also.

A few free electrons or holes would give infinite conductivity! *What limits the velocity?* We do know that certain materials (metals, ceramics) under certain conditions (usually low temperature) show zero resistivity, however, most materials at reasonable temperatures show significant resistance to electron flow.

In practice, the crystal potential $V(r)$ is not perfectly periodic:

- lattice vibrates due to thermal energy (“phonons”).
- ion cores are displaced from their equilibrium positions.
- lattice contains impurities (doping) and imperfections.
- all of these effects “scatter” the electron waves and limit the packet velocity.

We find that a model for electron motion is a random walk figure 2.1 (not a straight trajectory)

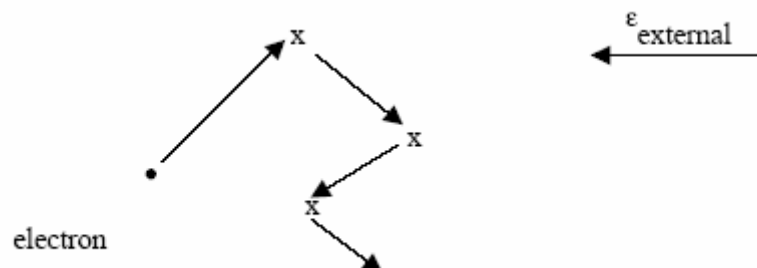


Figure 2.1: Model for electron motion

The electron is in random thermal motion (with a drift due to the ε' field). Mean time between collisions is τ ($\sim 10^{-13}$ s). There are in fact two mechanisms of current flow: (1) drift and (2) diffusion.

1-2.2 Drift Current

If an external field is present, it accelerates the electrons during time τ between scattering events figure 2.2.

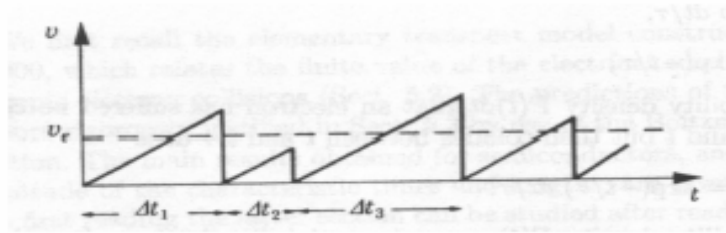


Figure 2.2: Movement of electrons between scattering events

The average velocity acquired from field is:

$$\bar{v} = \frac{qE}{m^*} \tau = \left(\frac{q\tau}{m^*} \right) E \quad (2.2)$$

Where $\left(\frac{q\tau}{m^*} \right)$ is defined as the electron mobility μ_n ($\text{cm}^2\text{V}^{-1}\text{s}^{-1}$) and relates field E to velocity as:

$$\bar{v} = \mu_{n,p} E \quad \text{and} \quad \langle x \rangle = \tau v = \frac{q}{m^*} \tau^2 E, \text{ the distance traveled at time } \tau. \quad (2.3)$$

The model breaks down at high field, when energy acquired by electrons between collisions becomes comparable to thermal energy. Then v is limited by v_{sat} the “scatter limited velocity” fig2.3.

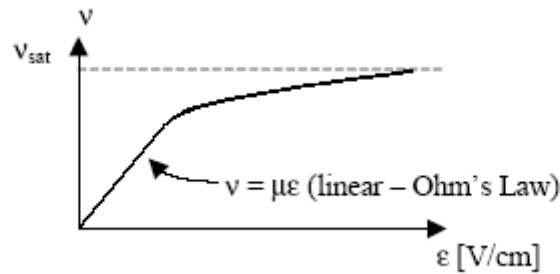


Figure 2.3: v is limited by v_{sat} the “scatter limited velocity”.

Many modern electronic devices work in the linear regime and it is important to have long τ as well as small m^* , for high μ . We must therefore have clean material and a good band structure (E-k relation).

To calculate the drift current carried by moving electrons; consider figure 2.4

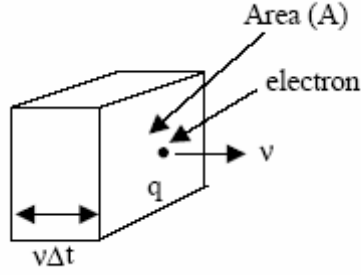


Figure 2.4: An electron with an area A

In time Δt , all electrons within value $v\Delta t$ cross area A where v is the electron's velocity. As n = density of electrons (cm^{-3}) we have $\Delta Q = n \cdot v\Delta t \cdot A \cdot q$ (q is the electron charge)

Defining J =current density $= \frac{\Delta Q}{\Delta t \cdot A} = qnv$ we get $\vec{J} = qn\vec{v}$ (2.4)

Therefore the Drift Currents in response to electric field is

$$\vec{J}_{n, drift} = q\mu_n n \vec{E} \quad \text{for electrons} \quad (2.5)$$

and

$$\vec{J}_{p, drift} = q\mu_p p \vec{E} \quad \text{for holes} \quad (2.6)$$

And the total drift current is

$$\vec{J}_{n,p, drift} = (q\mu_p p + q\mu_n n) \vec{E} \quad (2.7)$$

We can define a conductivity $\sigma [\Omega^{-1} \text{m}^{-1}]$ and $\vec{J} = \sigma \vec{E}$ (Ohms Law)

Thus for holes and electrons:

$$\sigma = q\mu_p p + q\mu_n n \quad (2.8)$$

Hence for n -type $\sigma = \frac{1}{\rho} = qn\mu_n = nq^2 \frac{\tau}{m^*}$ (p is small) and for p -type $\sigma = \frac{1}{\rho} = qp\mu_p = pq^2 \frac{\tau}{m^*}$,

(n is small)

Consider figure 2.5

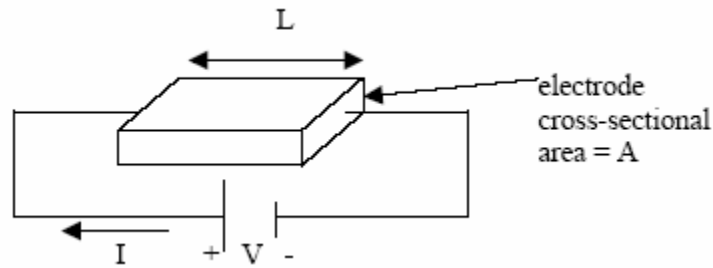


Figure 2.5: A piece of conductor

The total current is given by:

$$I = J \times A = \sigma \cdot A \cdot E, \quad \text{where } E = \left(\frac{V}{L} \right)$$

$$I = \sigma \cdot A \cdot \left(\frac{V}{L} \right) \quad \text{we can also define the resistivity } \rho = \frac{1}{\sigma} [\Omega \text{m}]$$

Resistivity

- Controlled by two factors:
 - Carrier density
 - Mobility: important also for frequency response f (frequency)
- Two types of carriers:

$$J_{total} = J_{electrons} + J_{holes}$$

$$J_T = q(n\mu_n + p\mu_p)E \quad \text{and} \quad \rho = \frac{1}{q(n\mu_n + p\mu_p)} \quad (2.9)$$

SESSION 2-2: DIFFUSION OF ELECTRONS AND HOLES AND GENERATION/RECOMBINATION

2-2.1 Diffusion Of Electrons And Holes

The random thermal motion of electrons causes a net flow of electrons from a region of high concentration to region of low concentration fig 2.6.

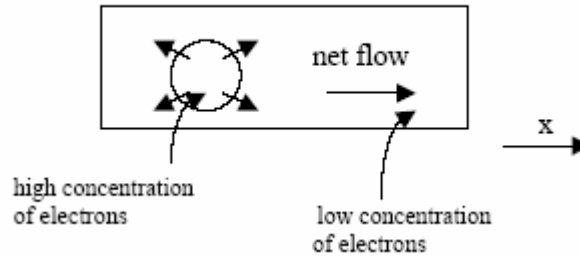


Figure 2.6: Random thermal motion of electrons

Mathematically we have

$$\bar{J}_n \text{diff} = qD_n \frac{dn}{dx} \quad (2.10a)$$

$$\bar{J}_p \text{diff} = qD_p \frac{dp}{dx} \quad (2.10b)$$

Where $D_{n,p}$ is the diffusion coefficient. Consider an example of a linear drop in concentration

$$\bar{J}_n \text{diff} = qD_n \frac{dn}{dx} = qD_n \frac{\Delta n}{L}$$

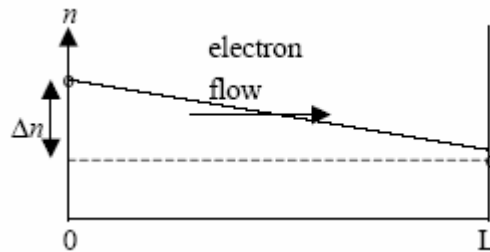


Figure 2.7: Linear drop in concentration

D and μ are related as $D = \frac{kT}{q} \mu$ which is the “Einstein relation” from kinetic theory both *diffusion* and *drift* depend on μ . The total current is sum of diffusion and drift in any situation ie.

$$\vec{J}_{total} = \vec{J}_{drift} + \vec{J}_{diffusion}$$

and
$$\vec{J}_{Total} = q(\mu_p p + \mu_n n) \vec{E} + q \left(D_p \frac{dp}{dx} + D_n \frac{dn}{dx} \right) \quad (2.11)$$

2-2.2 Generation/Recombination

2-2.2.1 Generation of electrons and holes

Generation: The movement of an electron from the valence band to the conduction band. This results in the creation of an electron-hole pair. We have talked about the thermal generation of holes and electrons but they can also be generated by light (Photogeneration).

Photogeneration

Remember that E - k diagrams of real crystals (whose structure is still periodic but more complex than a simple chain of atoms) are not simple parabolas.

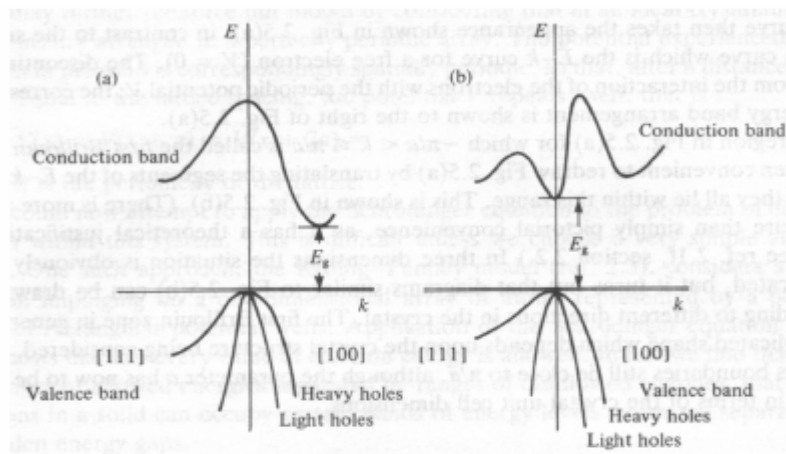


Figure 2.8: E-k relationship for real solids.

Figure 2.8 shows the E-k relationship for real solids (a) Silicon (which has an indirect bandgap) and (b) gallium arsenide (which has a direct bandgap).

One point that arises is that the maximum of the valence band does not always occur at the same k value as the minimum of the conduction band. We speak of a *direct* bandgap semiconductor when they do and an *indirect* bandgap semiconductor when they do not.

Analyzing the two cases:

- a) Direct gap material figure 2.9 (e.g. GaAs) –(min of conduction band at same k as max of valence band)

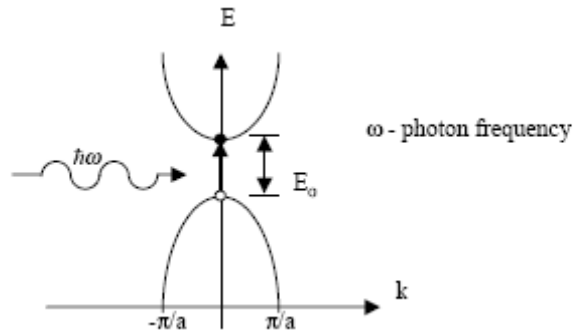


Figure 2.9: Direct gap material

($\Delta k=0$) Change of electron momentum is zero

A photon (a “particle of light”) with energy $E = \hbar\omega$ can be absorbed by promoting a valence band electron to the conduction band, creating an electron-hole pair. This is a two body collision (electron, photon). Direct band gap materials have strong light absorption and are modeled by an absorption parameter.

Model for light absorption fig 2.10: $I(x) = I_0 e^{-\alpha x}$

Where α - absorption coefficient [cm^{-1}], I_0 – incident intensity

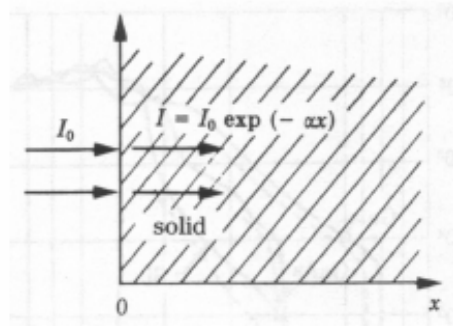


Figure 2.10: Model for light absorption

- b) Indirect gap material fig 2.11(e.g. Si) –(min of conduction band not! at same k as max of valence band)

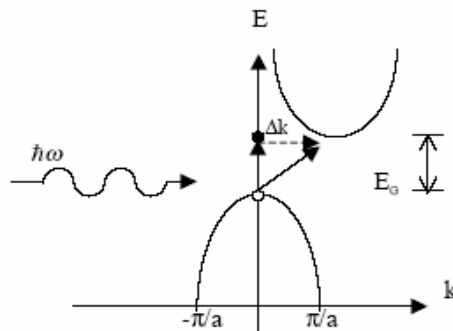


Figure 2.11: Indirect gap material

($\Delta k \neq 0$) Change of electron momentum is non zero.

A photon carries little momentum, for a transition we need Δk (a momentum change) and this can not be provided by the photon itself.

Promotion of valence band electron to conduction band requires momentum transfer from crystal lattice (interaction with a “phonon”); this is a relatively unlikely event as it is a three body collision (electron, photon, phonon). “phonon” = quantized lattice vibration (“particle” of heat)

\therefore light absorption is relatively weak.

Note: all semiconductors are effectively - transparent for light with $h\nu < E_G$ (no possible transition).

- absorb for $h\nu > E_G$ (transitions allowed)

2-2.2.2 Recombination of electrons and holes

Recombination: The movement of an electron from the conduction band to the valence band. This results in the destruction of an electron-hole pair.

Once created holes and electrons can recombine and be eliminated, in doing so, the energy released in the recombination is given up as light (radiative recombination) or heat (non radiative recombination).

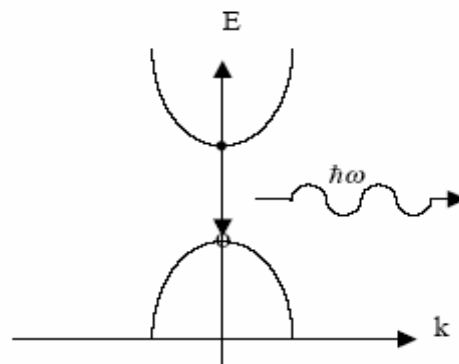


Figure 2.12: Direct radiative recombination

- a) Direct radiative recombination (e.g. GaAs)
The conduction band electron fills the valence band hole, releasing a photon with the energy $E_{\text{photon}} \sim E_G$. (figure 2.12)

This is the basis of semiconductor LEDs and lasers, and is a very unlikely process in indirect gap materials (need a phonon for momentum). Therefore can not make Si LEDs or lasers easily, without modifying material properties.

- b) Recombination through Midgap Energy levels (traps) (two stage processes).

Defect centers or traps are energy levels (ET) in the forbidden gap which are associated with defect states caused by the presence of impurities or lattice imperfections.

Consider figure 2.13

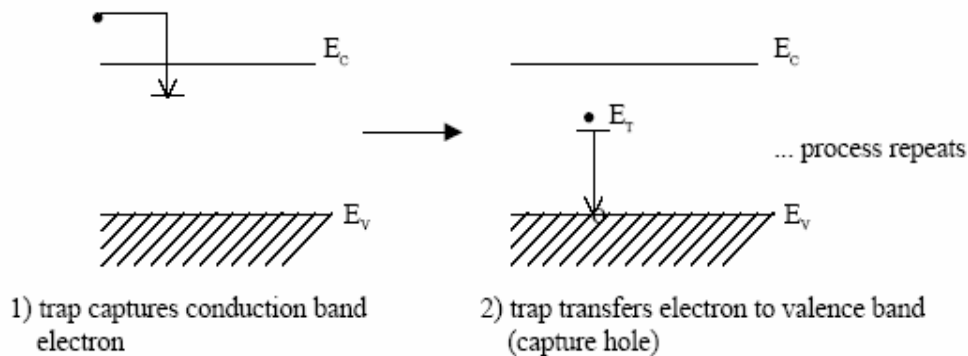


Figure 2.13: Recombination through Midgap Energy levels (traps)

- In silicon, midgap defect centers generally associated with transition metal contamination: Cu, Fe, etc like dopant.
- Traps are generally undesirable as we want long carrier lifetimes (elapsed time before recombination) in most devices.
- Traps shorten lifetime, produce multiple frequencies and energy pathways (contribute to non-radiative recombination).

Thermal Generation (trap aided)

Consider fig 2.14;

- heat (phonon) gives energy to electron.
- usually enabled by midgap energy levels (smaller energy).

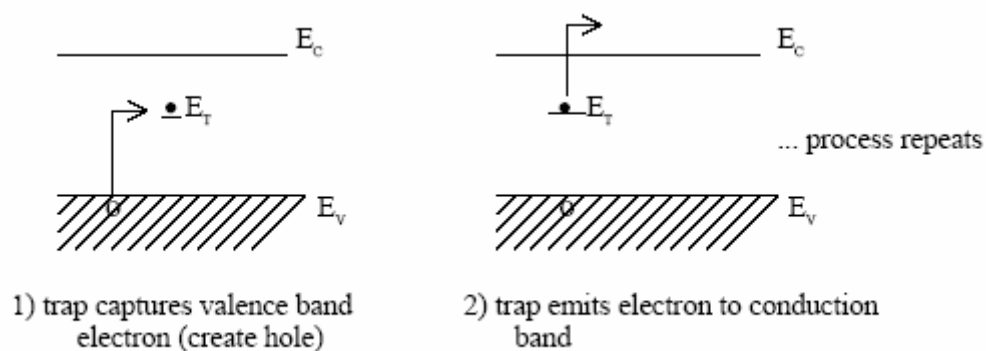


Figure 2.14: Thermal Generation (trap aided)

PN- DIODE DEVICES

SESSION 1-3: JUNCTION FORMATION, EQUILIBRIUM PROCESS AND BIASING

1-3.1 Junction Formation

Thought experiment - consider what happens if n and p type semiconductor samples are placed together, figure 3.1.

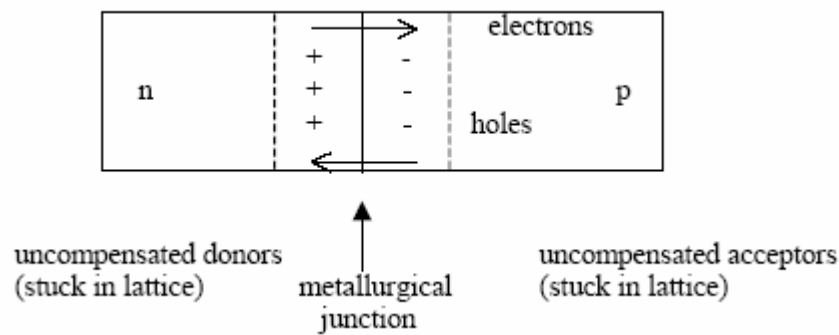


Figure 3.1: n and p type semiconductor samples placed together.

Due to diffusion we have two effects:

- electrons pour out of n -type material, leaving behind uncompensated donors (+ ions)
- holes pour out of p -type material, leaving behind uncompensated acceptors (- ions)

These electrons and holes then recombine, producing a region depleted of free carriers, figure 3.2; leaving only fixed charges (ionized donors and acceptors).

The diffusion process can not continue indefinitely as the space charge creates an electric field that opposes the diffusion of majority carriers (electrons in n -type and holes in p -type), though such diffusion is not prevented altogether.

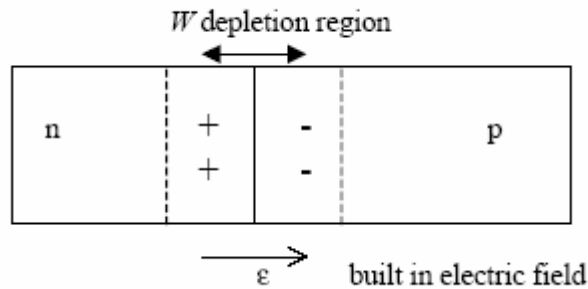


Figure 3.2: Electrons and holes recombine, producing a depletion region

The electric field will sweep minority carriers (holes in n type and electrons in p -type) across the junction so that there is a drift current of electrons from the p - to the n -type side and of holes from the n - to the p -type side which is in the opposite direction to the diffusion current. The junction field builds up until these two current flows are equal and equilibrium is achieved figure 3.3 (no *net* current flow).

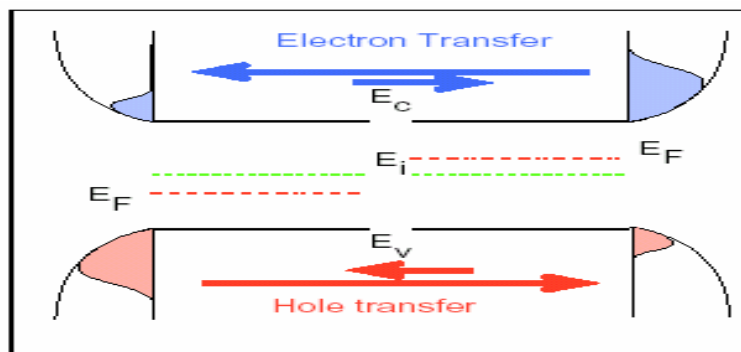


Figure 3.3: Equilibrium is achieved (no *net* current flow).

1-3.2 Equilibrium Process

It is a basic result of thermodynamics that in equilibrium the Fermi energy must be the same throughout the system. The induced electric field establishes a contact potential (ψ) between the two regions and the energy bands of the p -type side are displaced relative to those of the n -type side. This produces the Energy Band Diagram for a p - n Junction shown in figure 3.4.

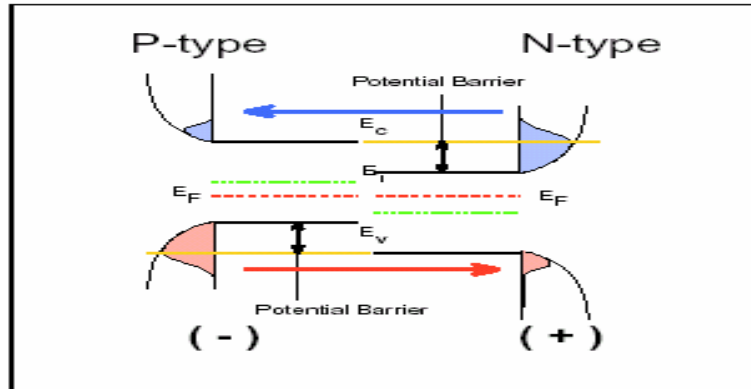


Figure 3.4: The same Fermi energy throughout the system at equilibrium

- Majority carriers transfer to the other side because of difference in Fermi level
- A voltage difference appears because of the charge transfer
- The voltage builds up until the currents are equilibrated
 - At equilibrium the currents of minority carriers must balance those of majority carriers.

1-3.2.1 After equilibrium

- Bands are continuous through the semiconductor
- Fermi level is constant
- The P side is charged negatively
- The N side is charged positively
- Equilibrium is based on balance of the majority and minority currents
- The device is inherently asymmetric

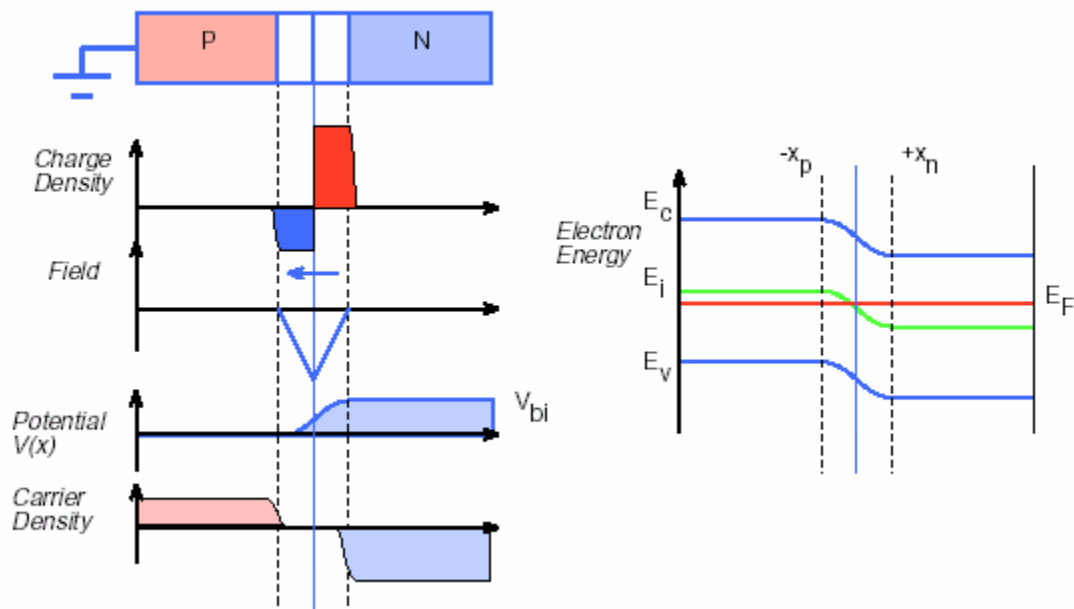


Figure 3.5: Charge Density, Field, Potential $V(x)$, carrier Density and Electron Energy Diagrams

1-3.2.2 Built-in-potential

Can be obtained directly from basic definitions without resorting to Poisson's equation and abrupt or non-abrupt case give same result. It is simply given as:

$$\text{Barrier} = E_i(P\text{-side}) - E_i(N\text{-side}) \quad (3.1)$$

Now consider semiconductor away from junction

$$\text{Left: } p \approx N_A; \quad E_i(x+) = E_F - KT \ln \frac{N_A}{n_i} \quad (3.2)$$

$$\text{Right: } n \approx N_D; \quad E_i(x-) = E_F + KT \ln \frac{N_D}{n_i} \quad (3.3)$$

- The built-in voltage is simply given by:

$$V_{bi} = \frac{1}{q} (E_i(x+) - E_i(x-)) = \frac{KT}{q} \ln \frac{N_A N_D}{n_i^2} \quad (3.4)$$

V_{bi} is called built-in potential

- OR V_{bi} can be obtained simply from using the bulk potentials we defined earlier:

$$qV_n = E_F - E_i = kT \ln \frac{N_D}{n_i} \quad (3.5)$$

$$qV_p = E_i - E_F = kT \ln \frac{n_i}{N_A} \quad (3.6)$$

Hence

$$qV_{bi} = qV_n - qV_p = kT \ln \frac{N_D N_A}{n_i^2} \quad (3.7)$$

And

$$V_{bi} = \frac{kT}{q} \ln \frac{N_D N_A}{n_i^2} \quad (3.8)$$

- Potential barrier acts like a dam
- Majority carriers confined to one side of junction
- Number of majority carriers able to overcome barrier equals number of minority carriers on other side
- We find that V_{bi} depends on doping and Temperature

Carriers and Potential Barrier

- A useful formula is derived from the usual carrier density relation:

$$n_p = n_i e^{[(E_i - E_F(p\text{-side})) / kT]} \quad \text{and} \quad n_n = n_i e^{[(E_F - E_i(n\text{-side})) / kT]} \quad (3.9)$$

1-3.2.3 Fields

- The depletion region gives rise to an electric field
- Fields goes from N to P region
- Force opposes motion of majority carriers

Poisson's Equation

We use the basic electrostatic equation (3.10) to determine the depletion width, built-in potential and the electric field.

$$\nabla \cdot E = \frac{\rho}{\epsilon_s \epsilon_o} \quad (3.10)$$

$$\rho(x) = q[N_A(x) - N_D(x) + p(x) - n(x)] \quad (3.11)$$

- In 1 dimension:
 - Constant charge density: $\rho(x) = \text{constant}$, $E(x) \propto x$
 - Zero charge density: zero field

Quantitative Relations

- Electric field (depletion extends from $-x_p$ to $+x_n$):

$$E_{left}(x) = -\frac{qN_A}{\epsilon_s \epsilon_o}(x_p + x) \quad (3.12)$$

$$E_{right}(x) = -\frac{qN_D}{\epsilon_s \epsilon_o}(x_n - x) \quad (3.13)$$

$$N_A x_p = N_D x_n \quad (3.14)$$

NB: Field is zero outside depletion region.

- Potential:

$$V_{left}(x) = -\frac{qN_A}{2\epsilon_s \epsilon_o}(x_p + x)^2 \quad (3.15)$$

$$V_{right}(x) = -\frac{qN_D}{2\epsilon_s \epsilon_o}(x_n - x)^2 \quad (3.16)$$

Bands Lineup

- Define Electrostatic Potential $V(x)$ as:

$$V(x) = \frac{1}{q}(E_F - E_i(x)) \quad (3.17)$$

- Define the field as the gradient of the potential:

$$E(x) = -\frac{dV_{bi}(x)}{dx} \quad (3.18)$$

- Define the charge density in regions of $E \neq 0$

Depletion Width

- Depletion area width can be computed to yield:

$$x_n = \sqrt{\frac{2\epsilon_s \epsilon_o N_A}{q N_D (N_A + N_D)}} V_{bi} \quad (3.19)$$

$$x_p = \sqrt{\frac{2\epsilon_s \epsilon_o N_D}{q N_A (N_A + N_D)}} V_{bi} \quad (3.20)$$

$$W = x_n + x_p = \sqrt{\frac{2\epsilon_s \epsilon_o N_A N_D}{q (N_A + N_D)}} V_{bi} \quad (3.21)$$

- Typical values: $\epsilon_s=11.8$, $\epsilon_o=8.85 \times 10^{-14}\text{F/cm}$
 $W \approx .2 - 20\mu\text{m}$

In practice, we often encounter one-sided (one side much more heavily doped than the other) abrupt junctions (abrupt transitions in doping from n to p) and the previous equation reduces to:

$$W = \sqrt{\frac{2\epsilon_s \epsilon_o}{q N_B}} V_{bi} \text{ , where } N_B=N_A \text{ if } N_D \gg N_A \text{ or } N_B=N_D \text{ if } N_A \gg N_D \quad (3.22)$$

Even for zero bias, there are electron and hole flows across the junction which exactly balance, and

$$J_n(\text{drift}) + J_n(\text{diffusion}) = 0 \text{ and } J_p(\text{drift}) + J_p(\text{diffusion}) = 0 \text{ for } V=0$$

Carriers and Potential Barrier

- A useful formula is derived from the usual carrier density relation:

$$\begin{aligned} n_p &= n_i \text{Exp}[(E_F - E_i(p\text{-side}))/kT] \\ n_n &= n_i \text{Exp}[(E_F - E_i(n\text{-side}))/kT] \end{aligned}$$

$$\frac{n_p(-x_p)}{n_n(+x_n)} = \exp[(E_i(n\text{-side}) - E_i(p\text{-side}))/KT] = e^{-qV_{bi}/KT} \quad (3.23)$$

And hence:

$$n_p(-x_p) = n_n(+x_n) e^{-qV_{bi}/KT} (= n_{p0}) \quad (3.24)$$

- This relation shows very directly it is the potential barrier that “holds back the majority carriers to completely spill out on the other side.

Carriers and Voltage

- Excess carriers are obtained from the barrier height:

$$V_{bi} \rightarrow V_{bi} - V$$

- Leading to ($-x_p$ is the edge of the depletion layer on the p-side, and x_n on the n-side):

$$n_p(-x_p) = n_{p0} e^{\frac{qV}{KT}} \quad (3.25)$$

$$n_p'(-x_p) = n_p(-x_p) - n_{p0} = n_{p0} \left(e^{\frac{qV}{KT}} - 1 \right) \quad (3.26)$$

$$p_n(x_n) = n_{n0} e^{\frac{qV}{KT}} \quad (3.27)$$

$$p_n'(x_n) = p_n(x_n) - p_{n0} = p_{n0} \left(e^{\frac{qV}{kT}} - 1 \right) \quad (3.28)$$

1-3.3 Biasing

Formulae remain valid. Simply:

$V_{bi} \rightarrow V_{bi} - V_{applied}$

- Depletion width:

$$W(V) = W(0) \sqrt{\frac{V_{bi} - V}{V_{bi}}} \approx 0.2 - 20 \mu\text{m} \quad (3.29)$$

- Electric field:

$$E_{\max} = E(0) = \left| \frac{qN_D}{\epsilon_s \epsilon_o} x_n \right| = \left| \frac{qN_A}{\epsilon_s \epsilon_o} x_p \right| \approx 10^3 - 10^5 \text{ V/cm} \quad (3.30)$$

- Definitions:

- Forward Bias: lower barrier, lower field, narrower depletion layer
- Reverse Bias: higher barrier, higher field, wider depletion layer

1-3.3.1 Reverse bias

If we apply a bias $V < 0$ fig 3.6, a very small current will flow.

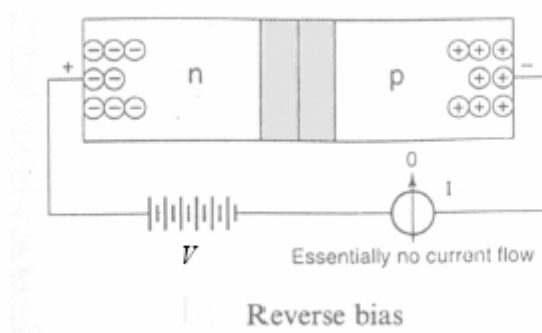


Figure 3.6: Reverse bias

Connecting the positive terminal of a d.c. source to the n side withdraws electrons and holes from the depletion area which becomes wider and the potential barrier grows higher figure 3.7.

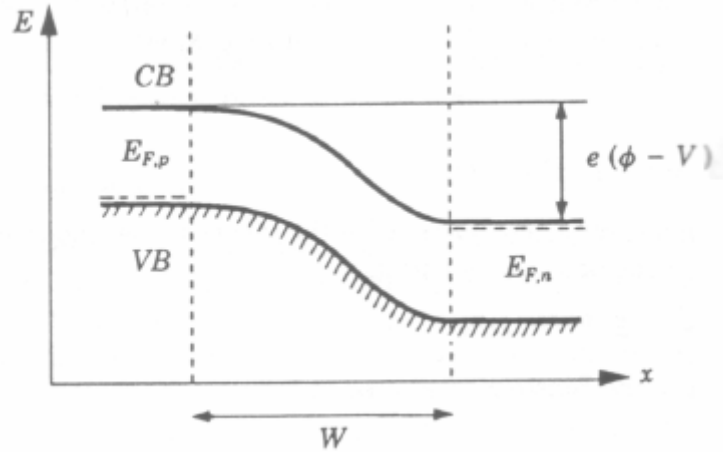


Figure 3.7: Depletion area becomes wider and the potential barrier grows higher.

We no longer have an equilibrium situation. The barrier is now so high that few electrons can cross from n -type to p type region reducing the diffusion current. However, electrons are still generated (thermally) in p -type region (not many it's p -type). And holes are still generated in n -type region (not many it's n -type).

There is also the generation of hole/electron pairs in depletion region and these cause small drift current. Therefore we have a small leakage current as these electrons and holes are swept across the depletion region (drift current).

$$J_o = q \left(\frac{D_p}{L_p} p_n + \frac{D_n}{L_n} n_p \right) \quad (3.31)$$

Where $D_{n,p}$ is the diffusion constant for electrons and holes. $L_{n,p}$ is the diffusion length, p_n is the hole concentration on the n -side and n_p is the electron concentration on the p -side. In order to keep this current small, the minority carriers (p_n and n_p) have to be kept at low levels. This can be accomplished by selecting semiconductors having a large energy gap and by high doping.

The drift current is relatively insensitive to the height of the potential barrier since all of the minority carriers generated may diffuse to the depletion region and be swept across it, whatever the size of the ϵ there.

1-3.3.2 Forward bias

Consider now the case when the n -side is connected to the negative terminal of the d.c. source figure 3.8, the depletion region narrows.

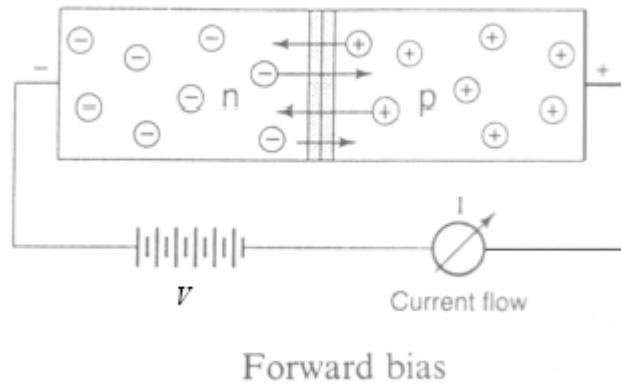


Figure 3.8: Forward bias

This has the effect of lowering the height of the potential barrier to $(\phi - V)$ figure 3.9. Consequently, majority carriers are able to surmount the potential barrier much more easily than in the equilibrium case so that the diffusion current becomes much larger than the drift current.

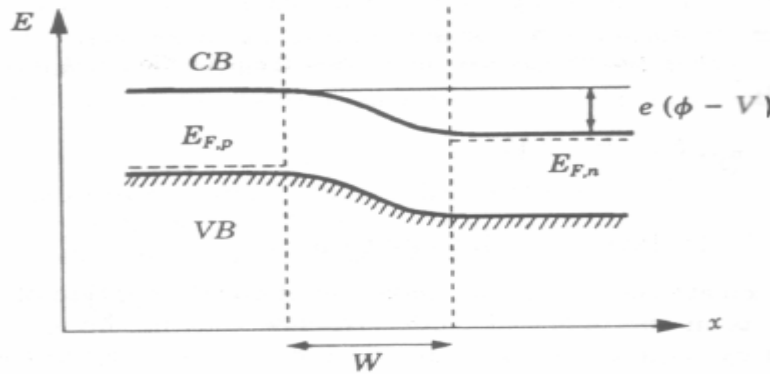


Figure 3.9: Forward bias lower the height of the potential barrier $(\phi - V)$

The diffusion current $J_{for} \propto e^{qV/kT}$ and the total current density is:

$$J = J_o (e^{qV/kT} - 1) \quad (3.32)$$

As the equation 3.32 indicates, this current is very strong temperature and bias dependent. We can also see that for forward bias (positive V) the net current increases exponentially with voltage. For reverse bias (negative V) the current is essentially constant and equal to $-J_o$.

The reverse bias generation current is small compared with the forward bias current. A typical I - V curve for a diode is shown in figure 3.10:

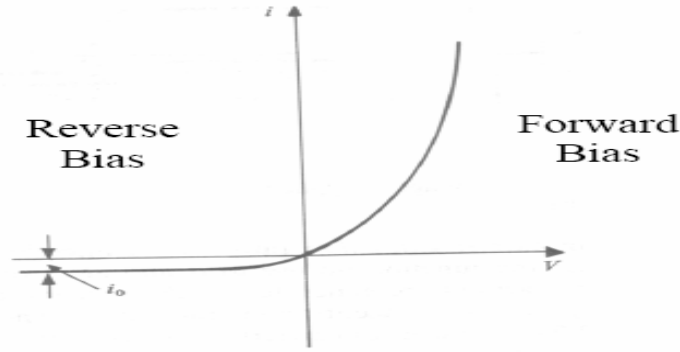


Figure 3.10: A typical I - V curve for a diode

Current-Voltage characteristic of a p - n junction diode. The reverse saturation current i_o is equal J_o multiplied by the junction cross-sectional area (A).

SESSION 2-3: DIODE CURRENT, DC MODEL, METAL-SEMICONDUCTOR JUNCTION AND "OHMIC" CONTACTS TO SEMICONDUCTORS

2-3.1 Diode Current

- Problem is easily solved by using the diffusion equation:

$$D_n \frac{\delta^2 n'_p}{\delta x^2} - \frac{\delta n'_p}{\delta \tau_n} = 0 \quad (3.33)$$

- The solution, for the current density, J_T :

$$J_T = J_n(-x_p) + J_p(+x_n) = \left[q \frac{D_n n_i^2}{L_n N_A} + q \frac{D_p n_i^2}{L_p N_D} \right] \left(e^{q \frac{V_A}{KT}} - 1 \right) \quad (3.34)$$

Where $L_n = \sqrt{D_n \tau_n}$ and $L_p = \sqrt{D_p \tau_p}$ are the diffusion lengths of n and p respectively

Or:

$$I_{sat} = \left[q \frac{D_n n_i^2}{L_n N_A} + q \frac{D_p n_i^2}{L_p N_D} \right] A \quad (3.35)$$

Where A is the cross sectional area

And
$$I_T = I_{sat} \left(e^{q \frac{V_A}{KT}} - 1 \right) \quad (3.36)$$

- Forward current increases sharply with bias voltage
- Reverse current remains constant until breakdown

Electrical Breakdown

- At large negative voltages the current increases sharply
- Cause is the generation process in the depletion region
- Minority carriers are accelerated to very large energies as they “slide down” the barrier
 - One electron (1e) loses its energy by creating an e-h pair (2e + 1h)
 - Each electron creates another e-h pair (4e+3h)
 - May become a runaway process: avalanching
- Current eventually limited by series resistor
- Breakdown may be a problem or a feature

Breakdown as a voltage regulator

Sharp current increase may be used to provide a voltage reference

Variations in load resistance or voltage supply end up in small output changes

Zener effect: special case for high-doping

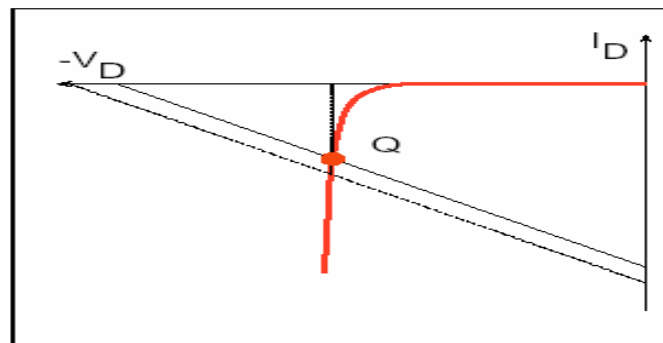


Figure 3.11: Diode in breakdown region

- **R_G currents** quite important in Silicon diodes
 - Carriers unable to overcome barrier may recombine in the depletion layers
 - Recombination process via traps

Forward current effect

- High-level injection:
- Series resistance of neutral regions and of contacts becomes relevant at high currents:

$$V_{Diode} = V_{bi} + R_s I_D \rightarrow V_{bi} = V_D - R_s I_D \quad (3.37)$$

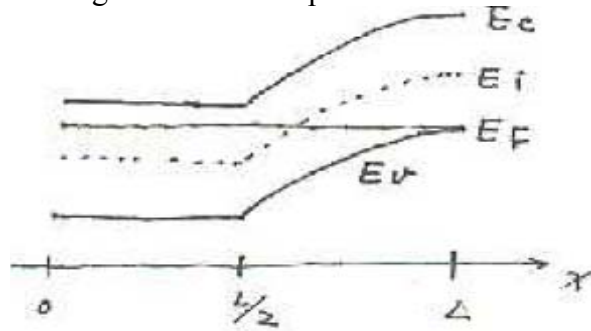
Conclusions

- Reviewed main aspects of diode response
- Forward case:
 - Series resistance
 - High-level injection
 - R-G currents

- Reverse case
 - Breakdown: avalanche, zener

Worked Examples

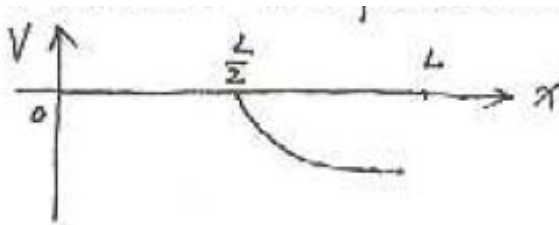
1. If the Energy Band Diagram of a Si sample is as follows



- a) Do equilibrium conditions prevail?
- b) Sketch $V(x)$
- c) Sketch $\varepsilon(x)$
- c) Sketch $n(x)$ and $p(x)$

Solution

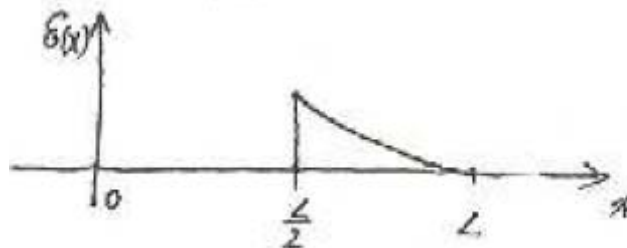
- a) Yes the E_F has the same energy value as a function of position
- b) If we assume the potential at $x = 0$ is $V(0)=0$, then



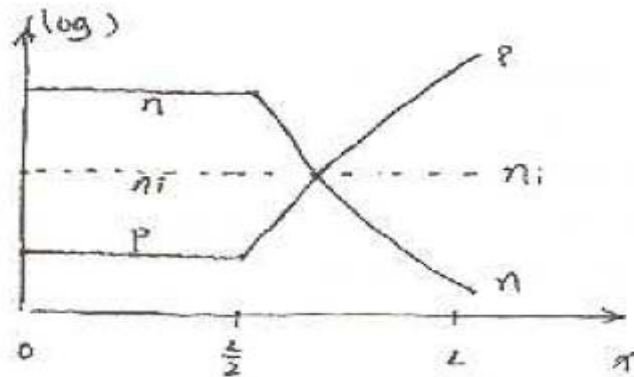
$V(x)$ has the same functional form as the “upside down” of $E_{c(x)}$

$$V(x) = -\frac{1}{8}(E_c(x) - E_{ref}), \text{ here we select } E_{ref} = E_c(0)$$

- c)
$$\varepsilon(x) = -\frac{dV(x)}{dx}$$



d) $n = n_i \exp[(E_F - E_i)/kT]$ $P = n_i \exp[(E_i - E_F)/kT]$



2. A Si step junction operated at 300K with $N_A = 10^{17} \text{ cm}^{-3}$ and $N_D = 10^{14} \text{ cm}^{-3}$:

Calculate the Fermi level positions in the p and n regions.

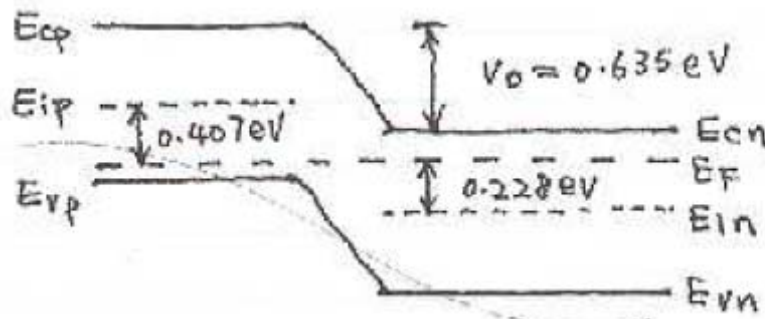
- Draw the band diagram for the junction and determine the contact potential V_o or V_{bi}
- Calculate V_o using equation. (5-8)
- Calculate x_{no}, X_{po} and ϵ_o
- Sketch $\epsilon(x)$ and charge density.

Solution

a) $E_{ip} - E_F = KT \cdot \ln \frac{p_p}{n_i} = 0.0259 \cdot \ln \frac{10^{17}}{1.5 \times 10^{10}} = 0.407 \text{ eV}$

$E_F - E_{in} = kT \cdot \ln \frac{n_n}{n_i} = 0.0259 \cdot \ln \frac{10^{14}}{1.5 \times 10^{10}} = 0.228 \text{ eV}$

b)



c) $V_{bi} = \frac{KT}{q} \ln \frac{N_A N_D}{n_i^2} = 0.0259 \ln \frac{10^{31}}{2.25 \times 10^{20}} = 0.635 \text{ V}$

d)



2-3.2 DC Model

- Applicable for very low frequencies where capacitance can be neglected
- Piecewise linear model, $V_T = 7V$ for Si, Slope determined by forward conductance
- Application: rectifier

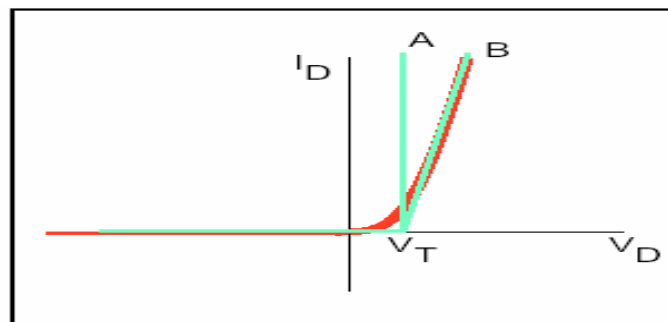


Figure 3.12: DC model

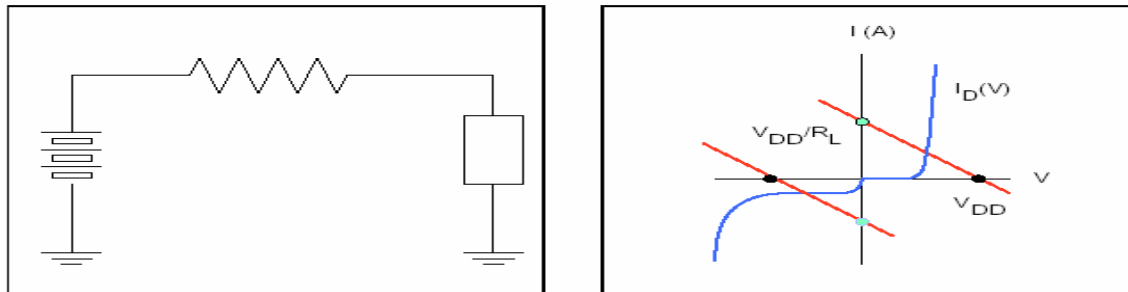


Figure 3.13: Circuit and Loadline

2-3.2.1 Small signal analysis

- The diode frequency response is controlled by G_0 - slope at Q-point
- For a simple circuit with a v_s AC signal over imposed on the operating supply:

$$v_D = V_D + \frac{Z_D(\omega)}{Z_D(\omega) + R_L} v_s \quad (3.38)$$

- Notice that the response may be highly asymmetric

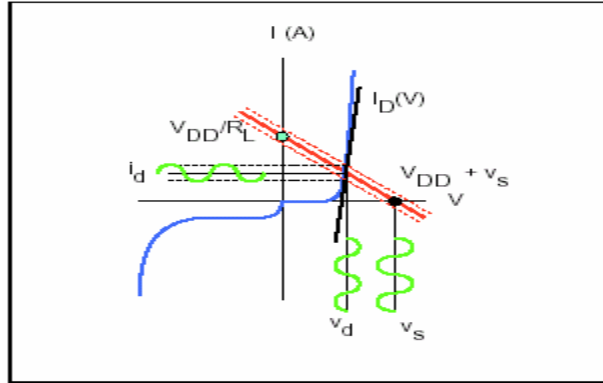


Figure 3.14: DC and AC loadlines

2-3.2.2 Small-signal model

- High-frequency model:
 - Capacitance due to the space charge C_D
 - Capacitance due to injected carriers: C_J
 - Conductance: G_D
- All are voltage dependent, i.e., depend on the Q-point

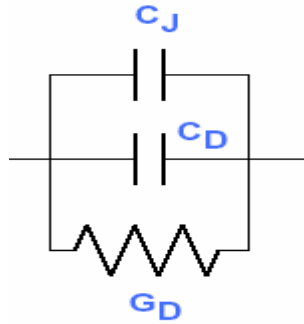


Figure 3.15: Small signal model

Small Signal Analysis

- Consider a diode biased at a Q-point (load-line analysis)
- A small signal is superimposed on the bias condition (Q-point):

$$v_D(t) = V_D(Q) + v_d(t) \quad (3.39)$$
- In general the system will exhibit a Q-point dependent admittance:

$$Y = G + j\omega C \quad (3.40)$$
- In reverse, $G \approx 0$ and only the reactive part is important

Reverse Admittance

- Small AC. signal $v_d(t)$ will modulate the width of the depletion layer, leading to a small-signal capacitance

$$C_J = \frac{\epsilon_s \epsilon_o}{W} A \quad (3.41)$$

- For an asymmetric junction $p^+ - n$ or $n^+ - p$:

$$W = \sqrt{\frac{2\epsilon_s \epsilon_o}{qN_B} (V_{bi} - V_A)} \quad (3.42)$$

- After some manipulation we have a simple expression for the capacitance of the junction:

$$C_J = \frac{C_{J0}}{\sqrt{1 - V_A/V_{bi}}} \quad (3.43)$$

- The capacitance is thus a direct function of applied voltage. This can be used in a circuit varactors
- The frequency dependence is also very important (microwave and fast logic)

2-3.3 Metal-Semiconductor Junction

2-3.3.1 Schottky barriers (metal-semiconductor junctions)

These are devices based on a metal-semiconductor junction. They are the oldest semiconductor devices used as rectifiers in “crystal” radios but were replaced by *pn*-junctions.

Band Structures

At surface of any solid there is an energy barrier holding in the electrons (work function).

2-3.3.1.1 metal band structure (half filled band).

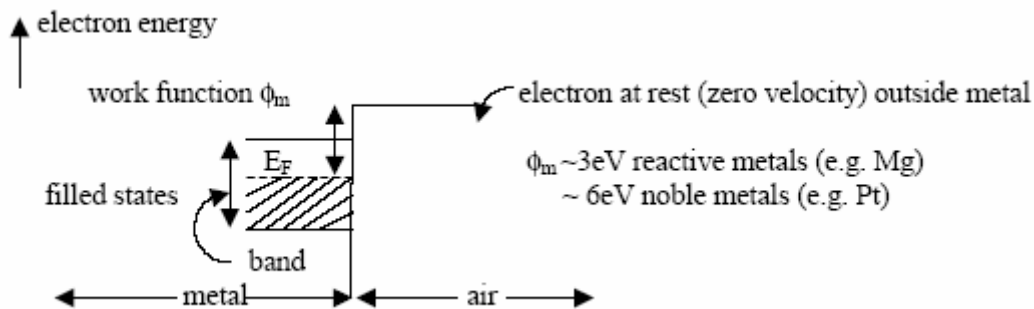


Figure 3.16: Metal band structure

ϕ_m is the energy to remove the electron from the Fermi level to vacuum level (ionization energy). The work function ϕ_m is strongly dependent on surface preparation and is not a precisely defined quantity (varies depending on surface).

Semiconductor Band Structure (conduction and valance band, *n*-type):

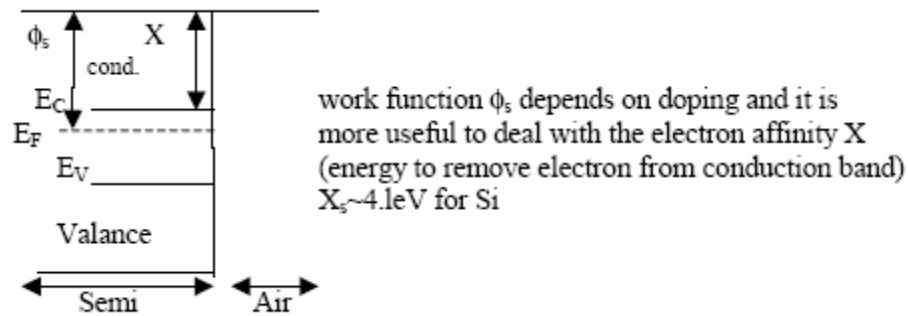


Figure 3.17: Semiconductor Band Structure

Both materials not touching ($\phi_m > \phi_s$):

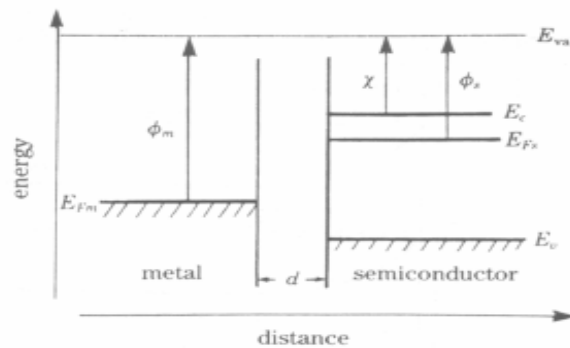


Figure 3.18: Metal and semiconductor in close proximity

After the metal and semiconductor have been brought into contact, electrons start to flow from the semiconductor “down” into the metal until the Fermi energies of both solids are equal, figure 3.19 (equilibrium condition).

- In semiconductor, a depletion region of width W is formed (uncompensated donors)
- In metal, electron current forms a negative surface charge layer
- An electric field is created
- An electric potential and thus a bending of the bands results

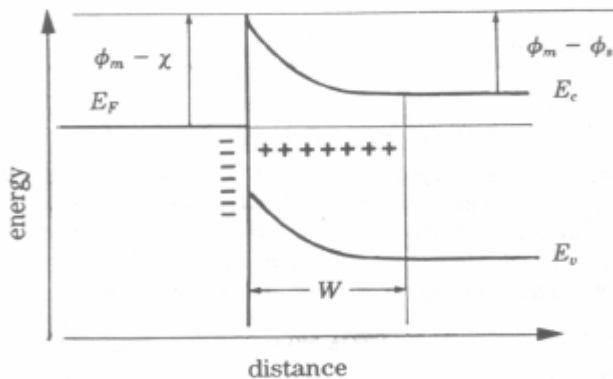


Figure 3.19: Metal and Semiconductor in contact (equilibrium condition).

We find that

$$V_0 = (\phi_m - \phi_s)/q = \frac{kT}{q} \ln\left(\frac{N_C}{N_D}\right), \text{ and } W = \sqrt{\frac{2\epsilon_{si}V_0}{qN_D}} \quad (3.44)$$

which mainly depend on doping concentration (N_D) and temperature. The equilibrium is reached when there are two equal and opposite electron flows over the barrier and $J=0$ at $V=0$.

2-3.3.1.2 reverse bias ($V<0$)

If the metal is connected to the negative terminal of a battery, the metal is charged even more negatively than without bias. Thus, the electrons in the semiconductor are repelled even more, figure 3.20.

- depletion region becomes wider
- potential barrier is increased.

The electron flows are unbalanced since the $J_{sm} \rightarrow 0$.

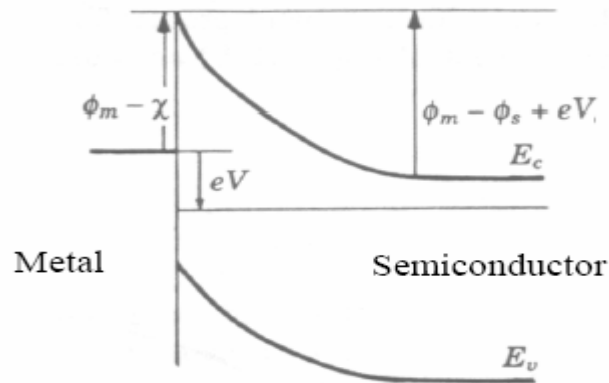


Figure 3.20: Reverse bias

A current flows as a result of a few electrons in the metal acquiring enough thermal energy to overcome barrier. This is a small current as barrier ($\phi_m - X$) not changed, and JM-S is as before. However, at high reverse bias, electrons can tunnel through barrier “field emission” and cause a charge current to flow.

2-3.3.1.3 forward bias

Now consider the case when the polarity of the battery is reversed (metal connected to the positive terminal of the battery). The depletion region is narrowed, figure 3.21. The potential barrier in the semiconductor is reduced and electrons “spill” from semiconductor into metal. We get a large current exponentially related to V flowing in the device. To compute the electron flow from semiconductor to metal we use thermionic emission theory.

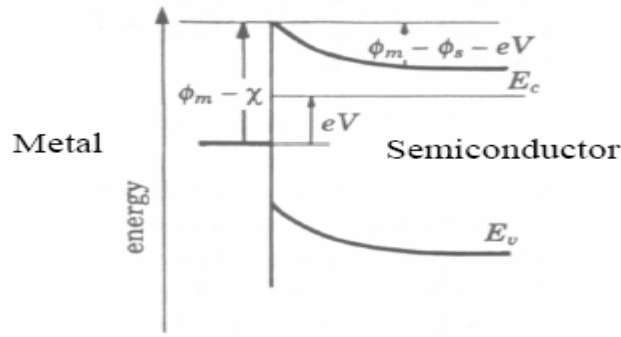


Figure 3.21: Forward bias

The current that flows from the metal into the semiconductor is

$$J_{M \rightarrow S} = \frac{q4\pi n^* k^2 T^2}{h^3} e^{-q(\phi_m - \chi)/kT} \quad (3.45)$$

Defining:

$C = \frac{q4\pi n^* k^2}{h^3}$ is Richardson's Constant, we have $J_{M \rightarrow S} = C e^{-q(\phi_m - \chi)/kT}$

The current flowing from the semiconductor into the metal is

$$J_{S \rightarrow M} = C e^{-q(\phi_m - \phi_s - V)/kT}$$

$$\text{The net current is given by: } J_{net} = J_s [e^{qV/kT} - 1] \quad (3.46)$$

$$\text{Just like diode, where } J_s = C T^2 e^{-q(\phi_m - \phi_s)/kT} \quad (3.47)$$

As typically there is very little minority carrier injection from semiconductor into metal; Schottkey diodes are said to be majority carrier devices.

2-3.4 “Ohmic” Contacts To Semiconductors

In Integrated Circuits the external connections are always metallic. A metal-semiconductor junction, which conducts current in only one direction, may be problematic. Need to make a low resistance contact with linear V-I characteristics.

e.g.

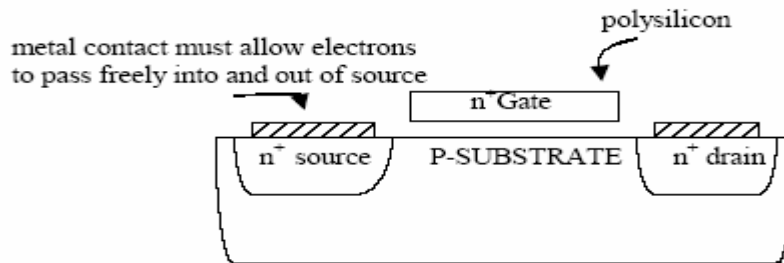


Figure 3.22: A metal-semiconductor junction

Consider now the case when $\phi_m < \phi_s$:

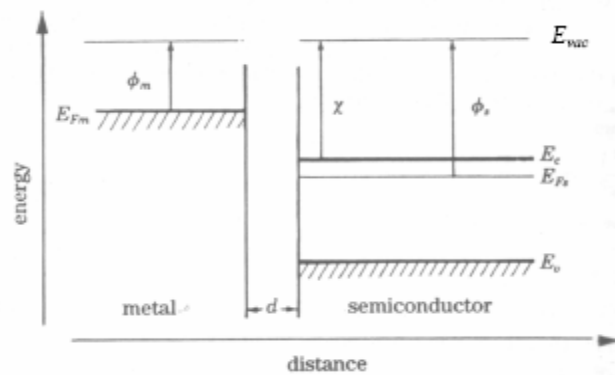


Figure 3.23: Bands of the metal and semiconductor placed together

Electrons flow from the metal into the semiconductor figure 3.23.

- In metal, electron current forms a positive surface charge layer
- An electric field is created (in the opposite direction as before)
- An electric potential and thus a bending of the bands results, figure 3.24

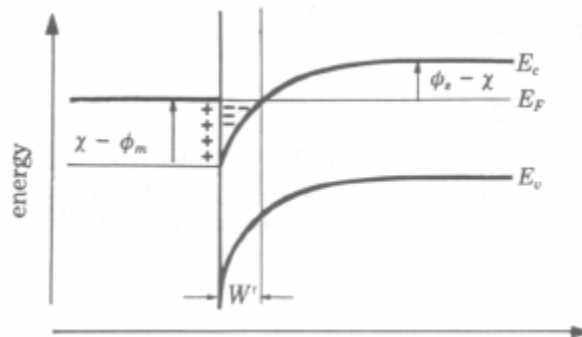


Figure 3.24: Bending of the bands

The bands of the semiconductor bend downward and no barrier exists for the flow of electrons in either direction. The current increases linearly with increasing voltage and is symmetric about the origin.

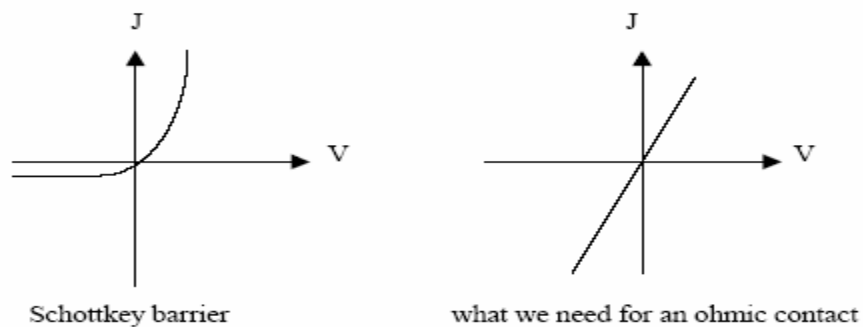


Figure 3.25: I-V characteristics of a Schottky barrier and an ohmic contact

Other applications:

Since only majority carriers are involved, no mutual annihilation (recombination) of electrons and holes can occur. This results in faster devices.

- The metal base provides better heat removal and is helpful in high power devices.
- Used for detection of long wavelength IR (low energy photons).
- This is useful in temperature sensors.

TRANSISTORS AND PHOTODETECTORS

SESSION 1-4: TRANSISTORS

Transistors are devices which control a large current with a smaller current or an applied voltage. They are typically used for amplifying a signal or for digital control (on/off switching). There are two basic types: bipolar structures which use a back to back diode structure and FET (field effect transistor) which is typically a surface controlled device such as JFETs, and MOSFETs.

1-4.1 The Junction FET

First FET device was the Junction FET (JFET) whose structure is shown in figure 5.1. Two junctions are formed on a Si substrate. Reverse bias causes the depletion areas to expand and reduces the size of conduction region.

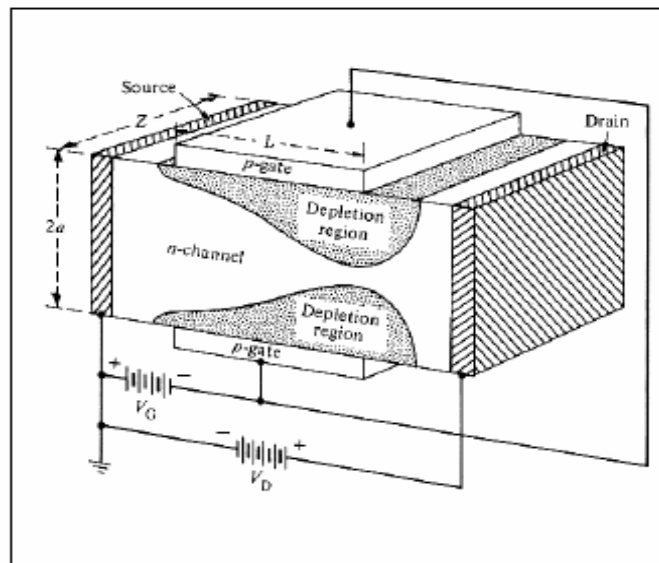


Figure 4.1: The Junction FET (JFET)

In the linear region there is a continuous path of charge (majority carriers) from drain to source: this is called the channel. As V_G is increased the width of the channel decreases because the depletion regions widen. We need to take into account the 2D character of the device. For $V_{DS} > 0$, the potential difference near the drain is given by $V_G - V_{DS}$. Increasing V_{DS} increases the

local reverse bias and thus narrows the channel further near the drain; source side is not affected. At some critical voltage V_P the two depletion layers will merge and the channel disappear; further increases in V_{DS} will not affect much the situation $V_{Dsat} = V_G - V_P$. V_P is the gate voltage for pinchoff when $V_{DS} = 0$.

1-4.1.1 Linear equations

Equations describing the $I-V$ characteristic are complicated. Compute the local cross-section and integrate:

$I_D(y) = q\mu_n N_D S(y)E = q\mu_n N_D Z a(1 - W(y)/a) \frac{dV}{dy}$ and the final expression looks like this:

$$I_D = I_D(V_{GS} - V_{DS})$$

$$= I_o \left\{ V_D - \frac{2}{3}(V_{bi} - V_P) \left[\left(\frac{V_D + V_{bi} - V_G}{V_{bi} - V_P} \right)^{3/2} - \left(\frac{V_{bi} - V_G}{V_{bi} - V_P} \right)^{3/2} \right] \right\} \quad (4.1)$$

This equation has the shape of an inverted parabola and reaches its maximum value at:

$$V_{Dsat} = V_G - V_P. \quad (4.2)$$

1-4.1.2 Saturation case

Saturation Voltage: $V_{Dsat} = V_G - V_P$.

Saturation Current: Simplified saturation current (Square law)

$$I_{Dsat} = I_o \left(1 - \frac{V_G}{V_P} \right)^2 \quad (4.3)$$

The JFET is a non-linear resistor where R changes with voltage:

- Two current sources in parallel
 - A conductance of value:

$$G = \frac{2qZ\mu_n N_d a}{L} \quad (4.4)$$

- A voltage-dependent source, with a relation:

$$I = -k(V_D - V_0)^{3/2} \quad (4.5)$$

The first corresponds to the resistance of the channel, the second stems from the areas that are lost to conduction because of depletion.

1-4.1.3 AC response

Small signal analysis:

$$I_D(V_{DS}, V_{DS}) = I_D(Q) + i_d(v_{gs}, v_{ds}) \quad (4.6)$$

$$I_D(V_{DS}, V_{DS}) = I_D(Q) + \frac{\partial I_D}{\partial V_{GS}} v_{gs} + \frac{\partial I_D}{\partial V_{DS}} v_{ds}$$

The small signal coefficients are thus:

$$g_d = \left(\frac{\partial I_D}{\partial V_{DS}} \right)_{V_{GS}} \quad \text{Drain conductance} \quad (4.7)$$

$$g_m = \left(\frac{\partial I_D}{\partial V_{GS}} \right)_{V_{DS}} \quad \text{- Transconductance} \quad (4.8)$$

Finally we get for the ac component: $i_D = g_m v_{gs} + g_d v_{ds}$ (figure 4.2) (4.9)

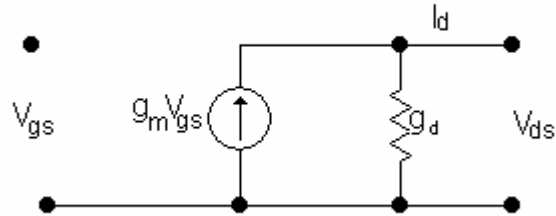


Figure 4.2: AC model

1-4.2 Bipolar Transistors

The structure is a back to back diode (ie npn or pnp). Two depletion regions are formed of exposed donors/acceptors with a large electric field in each causing band bending between the regions. There are three contacts an emitter, base, and collector figure 4.3. It is typically the base that is the controlling contact.

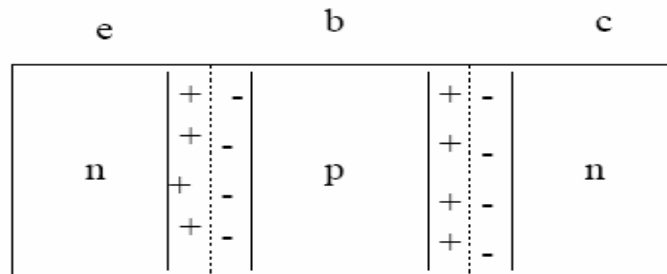


Figure 4.3: Basic npn structure

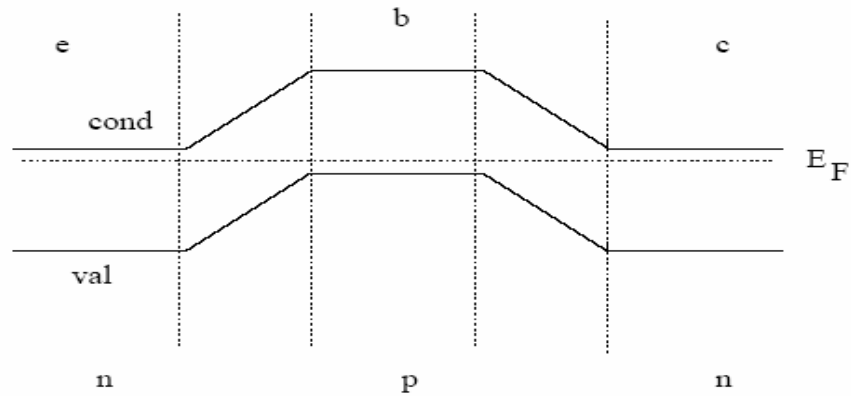


Figure 4.4: NPN band structure in equilibrium with no applied biases.

At equilibrium (no applied bias) the fermi levels of all three regions must line up figure 4.4. There is no **net** current through either junction. The junction widths and built in voltages are simply calculated using the diode expressions. Across each junction there are two exactly equal flows of both electrons and holes. As with the diode the built in voltages across each junction are precisely that needed to produce equal and opposite flows of both electrons and carriers and guarantee that no current is flowing.

There are many ways of biasing a BJT but the primary use is as an amplifier when biased in **forward active**. We will use this as an example of the current transport in a BJT. In forward active the Base-emitter junction is forward biased by V_{eb} and the base collector is reverse biased by V_{cb} figure 4.5. This results in the lowering of the barrier height between the emitter and the base and an increase in the barrier height between the base and collector.

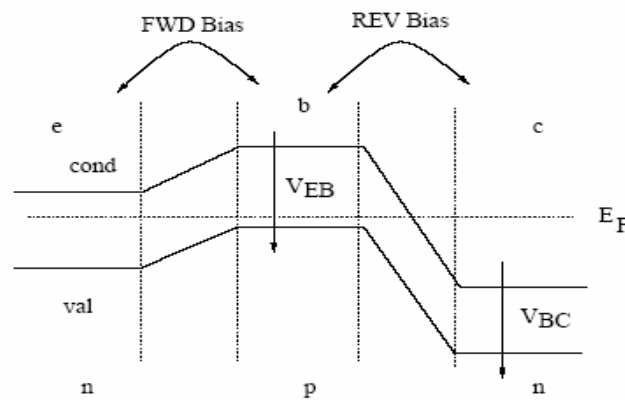


Figure 4.5: npn biased in forward active mode

The lowering of the base emitter junction barrier height will unbalance the electron flows across that junction, causing a significant **injection** of minority carriers (electrons) into the p type base. This injection of electrons is exponentially related to the applied base-emitter voltage.

At the collector junction, the reverse bias provides a means to **collect** electrons that are present in the base near the junction. Any electron that is in the base near the collector junction will **fall**

over the barrier and lower its energy. This results in very low electron (minority) concentrations in the base at the collector junction figure 4.6.

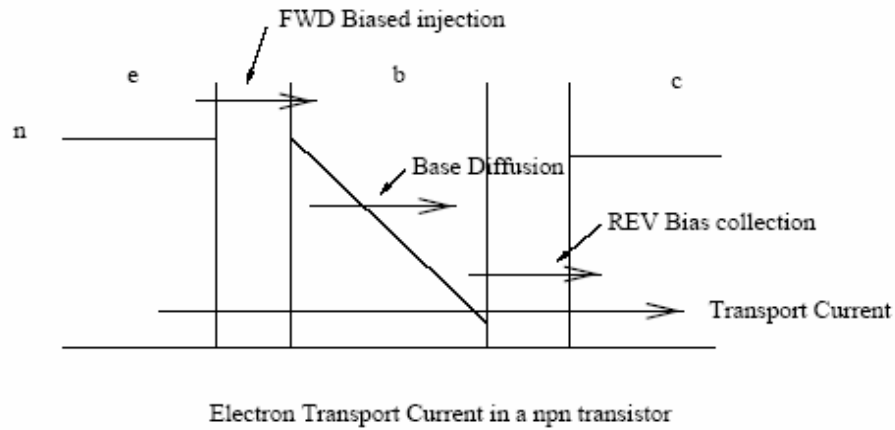


Figure 4.6: Electron transport current in a npn transistor

In forward active the basic mode of operation is therefore.

1. Electrons are injected into the base at the emitter junction. Causing a high concentration of electrons in the base at that junction, $n \approx e^{qV_{be}/kT}$

2. Electrons are collected at the collector junction as they fall over into the collector region. $n \approx 0$

3. A diffusional transport current is present from the emitter junction through the base to the collector current. If the base is thin then electron distribution is linear and $J \approx e^{qV_{be}/kT}$

The net result of this is a transport current that is controlled by the base emitter current. The equation of this transport current is: $I_C \approx I_E \approx I_0 e^{qV_{be}/kT}$

A closer analysis of the base emitter junction shows that there is a base current flowing that can be given by: $I_B \approx \frac{I_0}{\beta} e^{qV_{be}/kT}$ and we have: $I_C = \beta I_B$

A small base current can therefore control the collector current. β can be made large by choosing a good clean material and making the base width small (large diffusion current). This formula is only valid for forward active, but other modes of operation (cutoff, saturation, etc) can be understood in a similar manner.

1-4.3 MOSFET Transistors

The MOSFET (Metal Oxide Silicon Field Effect Transistor) is a device that controls a current between two contacts (Source and Drain) using a voltage contact (Gate). The device uses a surface effect to create an n-type region in a p-type substrate figure 4.7 (or the converse).

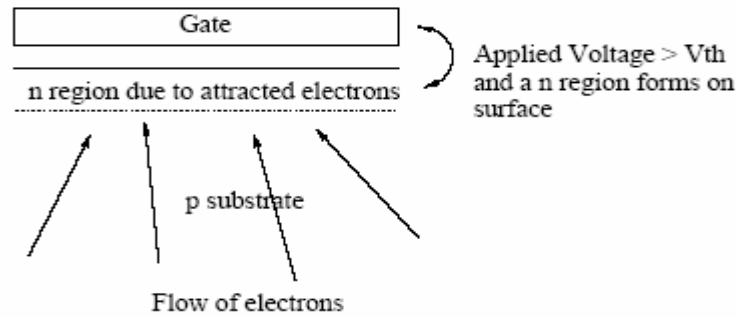


Figure 4.7: Flow of electrons

To understand this we take a simple capacitor structure using a p-type substrate an oxide layer and a metal gate. If we apply a positive potential to the gate (the substrate is grounded) electrons will be attracted to the gate and will pile up at the surface underneath the gate.

At some voltage V_{th} called the threshold voltage the region under the gate will have enough additional electrons that $n > p$ and the material will be n-type not p-type. The oxide is very important as it stops the current flow towards the gate and forces the electrons to “pile up” underneath the gate and turn the material n-type.

The basic MOSFET structure uses the capacitor structure figure 4.8 with n type regions placed at either edge, known as the source and drain.

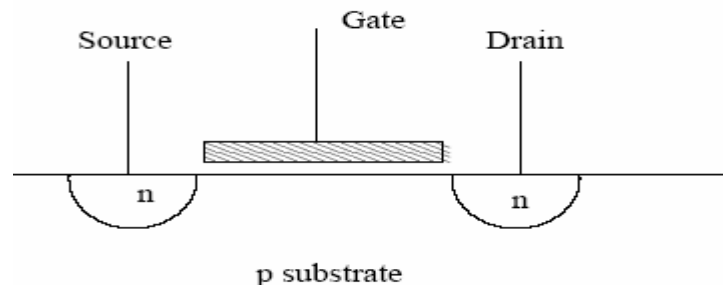


Figure 4.8: The NPN MOSFET

The basic operation of the device is to bias the gate with $V_G > V_{TH}$ and form an n-type region between the source and the drain. This provides a simple n-type path between the n-type source and drain regions for electrons to flow figure 4.9. This region is called a channel. Note that with out forming the channel there are two back to back diodes formed which will not allow appreciable current to flow.

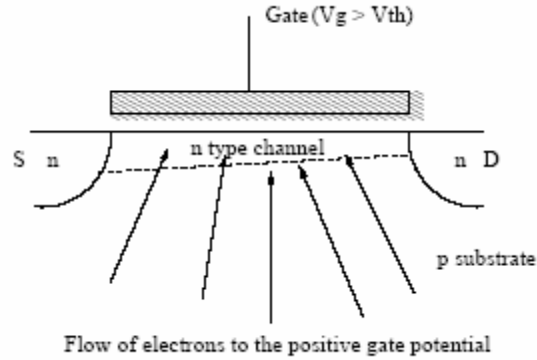


Figure 4.9: Flow of electrons to the positive gate potential

The formation of this channel provides a simple resistive path between the source and the drain. The thickness of the channel is function of the difference between the gate potential and the potential in the substrate near the surface. We can place a voltage between the source and the drain and cause a current to flow. Typically, we ground the source and bias the drain with V_{ds} . The ability to change the thickness of the channel using the gate potential provides a means of controlling the current from the source to the drain. We basically can form a voltage controlled resistor. One thing to note is that the application of a drain voltage raises the potential of the region of the substrate at the surface near the drain. This results in a thinner channel at that end (as the gate to substrate potential is reduced).

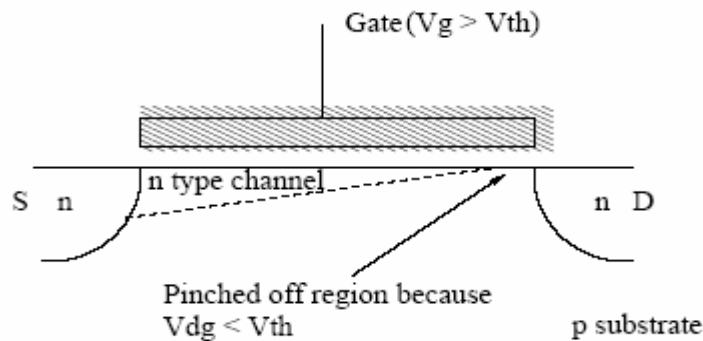


Figure 4.10: No inversion layer

A new condition arises if we increase the drain voltage substantially ($V_{DS} > V_{DS, sat}$, and $V_{DS, sat}$ is called the saturation voltage). The drain voltage becomes large enough that the gate to substrate potential at the drain is smaller than threshold. Therefore, the channel thickness at this end goes to zero figure 4.10. We call this pinch off. Electrically, the effect of pinch off is that the channel no longer acts like a simple resistor. The current becomes fixed (saturated) at the value just prior to pinch off.

1-4.3.1 The MOSFET current – voltage characteristics

The MOS is normally considered with source shorted to the bulk and grounded and drain at supply voltage. The behavior of the transistor is controlled by the voltages between gate and source (bulk) V_{GS} and between drain and source V_{DS} . The MOSFET is a symmetric device. There is no gate current I_G because of the oxide isolation. The drain current $I_{DS} = I_{DS}(V_{GS}, V_{DS})$ is a controlled voltage quantity. The threshold voltage V_{TH} is a critical parameter in the operation of the transistor. The MOSFET is therefore much more of a programmable resistor than the BJT.

The operation is very similar to JFET, three operational areas; namely Cutoff: $I_{DS} = 0$, Linear region: $I = I(V_{DS})$ and Saturation: where $I_{DS} = \text{constant}$.

At cutoff: The device is in accumulation and no conducting path exists. The current is the reverse current of the two junctions biased in reverse.

Linear region: In this region $V_{GS} > V_{TH}$ but the two dimensional nature of the device complicates the analysis. At any point under the gate, the surface potential is the sum of two contributions contribution from the gate voltage and that from ohmic drop along channel from current flowing. Thus surface potential varies from point to point under gate, the amount of charge and width is changing along the channel; and current is proportional to the amount of charge under the gate which is also proportional to $V_{GS} - V_{TH}$. That is, larger gate voltages produce more charge and hence more current but V_{DS} counteracts V_{TH} .

In Saturation: When the local voltage near the drain becomes less than V_{TH} the inversion disappears and the channel is pinched-off. The current continues to flow where the depletion region acts as a high resistance section of the channel. Further increases in V_{DS} simply increase the width of the Pinched-off area. If $L_{pinch\ off} \ll L$ the current saturates.

1-4.3.2 Computation of MOS currents

Total current (not density) is continuous, depends on how much charge is available to conduct. We need to decompose in x (across channel) and y (along channel). The x -field determines amount of charge and y -field determines current density.

The current density is given as in equation:

$$J_n \approx q\mu_n E = -q\mu_n n \frac{\delta\phi}{\delta y} \quad (4.10)$$

The total current is therefore given by equation where W is the width of the device

$$I_D = -W\mu_n Q_n \frac{\delta\phi}{\delta y} \quad (4.11)$$

But I_D is independent of y (charge conservation) and hence,

$$\int_0^L I_D dy = I_D L = -\mu_n W \int_0^{V_D} Q_n d\phi \quad (5.12)$$

And

$$I_D = -\frac{W\mu_n}{L} \int_0^{V_D} Q_n d\phi \quad (4.13)$$

But for simplicity we approximate the total charge as that under the gate which becomes:

$$Q_n(y) = -C_{ox}(V_{gs} - V_{th} - \Phi) \quad (4.14)$$

Where C_{ox} is the oxide capacitance per unit area, given by: $C_{ox} = \frac{\epsilon_{ox}\epsilon_o}{t_{ox}}$ and t_{ox} = oxide thickness.

By integration of I_D we obtain:

$$I_D = \frac{W\mu_n C_{ox}}{L} [(V_{GS} - V_{TH})V_{DS} - V_{DS}^2/2] \quad (4.15)$$

And the threshold voltage V_{TH} is given as:

$$V_{TH} = 2\Phi_B + \frac{\epsilon_s t_{ox}}{\epsilon_o} \sqrt{\frac{4qN_A}{\epsilon_s \epsilon_o}} \Phi_B \quad (4.16)$$

Example for $N_A=10^{15}$, $\Phi_B = 0.288$, $V_{th}=0.978V$

From equation (4.15), we note the following:

- Two terms in the current expression: one increases with V_{DS} , the other decrease.
- Maximum is reached at $V = V_{DS_{SAT}}$ and $V_{DS_{SAT}} = V_{GS} - V_{TH}$

$$\text{Then } I_{DS_{SAT}} = \frac{W\mu_n C_{OX}}{L} (V_{GS} - V_{TH})^2$$

- Example: $C_{OX} = 34.4 \times 10^{-9} F$, $V_{DS_{sat}} = 2.03V$, $I_{DS_{sat}} (V_{GS} = 3V) = 7.0 \times 10^{-5} A$

Therefore the equations for the MOSFET are:

1. The triode region— $V_{GS} > V_{th}$, $V_{DS} < V_{DS,sat}$

$$I_{DS} = I_0 [(V_{GS} - V_{th})V_{DS} - V_{DS}^2/2] \quad (4.17)$$

The $V_{DS}^2/2$ term takes the narrowing of the channel at the gate region into account as V_{DS} approaches $V_{DS,sat}$

2. The saturation region— $V_{GS} > V_{th}$, $V_{DS} > V_{DS,sat}$

$$\text{It can be shown that } V_{DS,sat} = V_{GS} - V_{th}, \text{ and } I_{DS,sat} = I_0 \frac{(V_{GS} - V_{th})^2}{2} \quad (4.18)$$

In saturation, the MOSFET acts like a nonlinear voltage controlled current.

1-4.3.3 Small signal analysis

Consider the device as a four terminal network, in the low frequency, input is a capacitor thus it is an open circuit at low frequency. The output is:

$$i_D(v_{DS}, v_{GS}) = I_D(V_{DS}, V_{GS}) + \left(\frac{\partial I_D}{\partial V_{DS}} \right)_{V_{GS}} v_{ds} + \left(\frac{\partial I_D}{\partial V_{GS}} \right)_{V_{DS}} v_{gs} \quad (4.19)$$

$$i_d(v_{DS}, v_{GS}) = g_d v_{ds} + g_m v_{gs} = \frac{1}{r_d} v_{ds} + g_m v_{gs} \quad (4.20)$$

Where g_d and g_m are the drain conductance and transconductance respectively; and r_d is the drain resistance shown in figure 4.11.

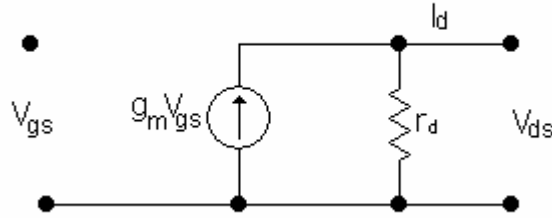


Figure 4.11: Small signal model of MOSFET

Below pinchoff:

$$g_d = \frac{W\mu_n C_{ox}}{L} (V_{gs} - V_{th} - V_{ds}) \quad (4.21)$$

$$g_m = \frac{W\mu_n C_{ox}}{L} V_{ds} \quad (4.22)$$

Above pinchoff:

$$g_d = 0 \quad (4.23)$$

$$g_m = \frac{W\mu_n C_{ox}}{L} (V_{gs} - V_{th}) = \frac{W\mu_n C_{ox}}{L} V_{ds(sat)} \quad (4.24)$$

Transit Time:

In digital logic the transit time may be more interesting than the cut off frequency:

$$t_t = \frac{L}{v_{drift}} = \frac{L}{\mu_n E} = \frac{L^2}{\mu_n V_{ds(sat)}} \quad (4.25)$$

Worked Examples

Example 1

Find the maximum width of the depletion region for an ideal MOS capacitor on p-type Si with $N_a = 10^{16} \text{ cm}^{-3}$.

Solution

The relative dielectric constant of Si is 11.8. We get Φ_F from

$$\Phi_F = \frac{kT}{q} \ln \frac{N_a}{n_i} = 0.0259 \ln \frac{10^{16}}{1.5 \times 10^{10}} = 0.347 \text{ V}$$

Thus

$$\begin{aligned} W_m &= 2 \sqrt{\frac{\epsilon_s \Phi_F}{q N_a}} = 2 \left[\frac{(11.8)(8.85 \times 10^{-14})(0.347)}{(1.6 \times 10^{-19})(10^{16})} \right]^{1/2} \\ &= 3.01 \times 10^{-5} \text{ cm} = 0.301 \mu\text{m} \end{aligned}$$

Example 2

Using the conditions of example 1 and a 100-Å thick SiO₂ layer, we can calculate major points on the C-V curve. The relative dielectric constant of SiO₂ is 3.9.

Solution

$$C_i = \frac{\epsilon_i}{d} = \frac{(3.9)(8.85 \times 10^{-14})}{10^{-6}} = 3.45 \times 10^{-7} \text{ F/cm}^2$$

$$Q_d = -q N_a W_m = -(1.6 \times 10^{-19})(10^{16})(0.301 \times 10^{-4})$$

$$= -4.82 \times 10^{-8} \text{ C/cm}^2$$

$$V_T = -\frac{Q_d}{C_i} + 2\phi_F = \frac{4.82 \times 10^{-8}}{3.45 \times 10^{-7}} + 2(0.347) = 0.834 \text{ V}$$

At maximum depletion

$$C_d = \frac{\epsilon_s}{W_m} = \frac{(11.8)(8.85 \times 10^{-14})}{0.301 \times 10^{-4}} = 3.47 \times 10^{-8} \text{ F/cm}^2$$

$$C_{\min} = \frac{C_i C_d}{C_i + C_d} = \frac{34.5 \times 34.7}{34.5 + 3.47} (10^{-8}) = 3.15 \times 10^{-8} \text{ F/cm}^2$$

Example 3

Calculate V_T for the MOS structure described in example 1 and 2, including the effects of flat band voltage. If n^+ polysilicon is used for the gate. $\Phi_{ms} = -0.95 \text{ V}$ for $N_a = 10^{16} \text{ cm}^{-3}$. Assuming an interface charge of $5 \times 10^{10} \text{ q(C/cm}^2\text{)}$.

Solution

$$\begin{aligned} V_T &= \Phi_{ms} + 2\phi_F - \frac{1}{C_i}(Q_i - Q_d) \\ &= -0.95 + 0.694 - \frac{(5 \times 10^{10} \times 1.6 \times 10^{-19}) - 4.82 \times 10^{-8}}{34.5 \times 10^{-8}} \\ &= -0.14 \text{ V} \end{aligned}$$

Example 4

Consider an n^+ polysilicon SiO_2 Si-p-channel device with $N_d = 10^{16} \text{ cm}^{-3}$ and $Q_i = 5 \times 10^{10} \text{ qC/cm}^2$. Calculate V_T for a gate oxide thickness of $0.01 \mu\text{m}$ and repeat for a field oxide thickness of $0.5 \mu\text{m}$.

Solution

Values of Φ_F , Q_i , and Q_d , can be obtained from examples 2 and 3 if we use appropriate signs. The value of C_i for the thin oxide case is same as in example 2. $\Phi_{ms} = -0.25 \text{ V}$.

$$V_T = -0.25 - 0.694 - \frac{8 \times 10^{-9} + 4.82 \times 10^{-8}}{34.5 \times 10^{-8}} = -1.1 \text{ V}$$

In the field region where $d = 0.5 \mu\text{m}$,

$$V_T = -0.944 - \frac{5.62 \times 10^{-8}}{6.9 \times 10^{-9}} = -9.1 \text{ V}$$

Example 5

For the p-channel transistor of example 4, calculate the boron ion dose F_B (B^+ ions/ cm^2) required to reduce V_T from -1.1 V to -0.5 V . Assume that the implanted acceptors form a sheet of negative charge just below the Si surface.

Solution

$$-0.5 = -1.1 + \frac{qF_B}{C_i}$$

$$F_B = \frac{3.45 \times 10^{-7}}{1.6 \times 10^{-19}} (0.6) = 1.3 \times 10^{12} \text{ cm}^{-2}$$

For a beam current of $10\mu\text{A}$ scanned over a 650 cm^2 target area,

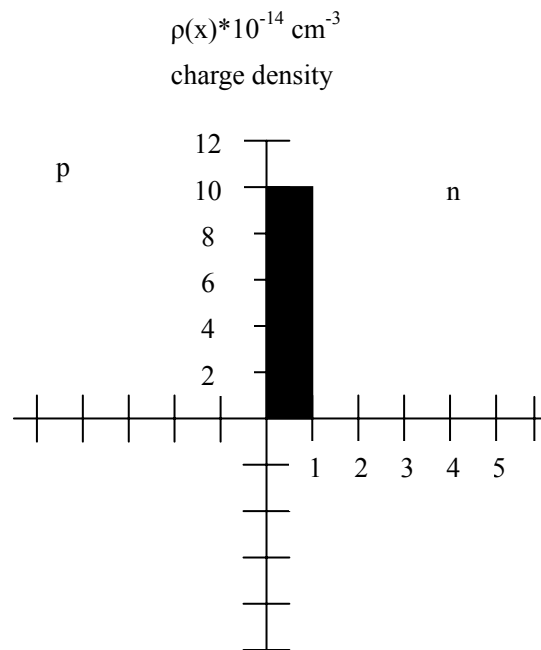
$$\frac{10^{-5} (\text{C} / \text{s})}{650 \text{ cm}^2} t(\text{s}) = 1.3 \times 10^{12} (\text{ions} / \text{cm}^2) \times 1.6 \times 10^{-19} (\text{C} / \text{ion})$$

The implanted time is $t = 13.5 \text{ s}$.

Practice Exercises

1.
 - a) Give two or three example of an n-type and a p-type dopant for Si
 n-type = _____ p-type = _____
 - b) Explain (briefly) why carrier mobility usually
 - i) decreases with increasing doping concentration
 - ii) decreases with increasing temperature
 - c) Briefly explain the difference between a shallow impurity and a deep level impurity in terms of how readily they ionize in Si
 - d) A bar of Si is doped with a donor concentration $N_D = 10^{16} \text{ cm}^{-3}$. Describe (and justify your answer) the hole and electron concentration at:
 - i) $T=0$
 - ii) $T=300^\circ\text{K}$
 - iii) $T=700^\circ\text{K}$

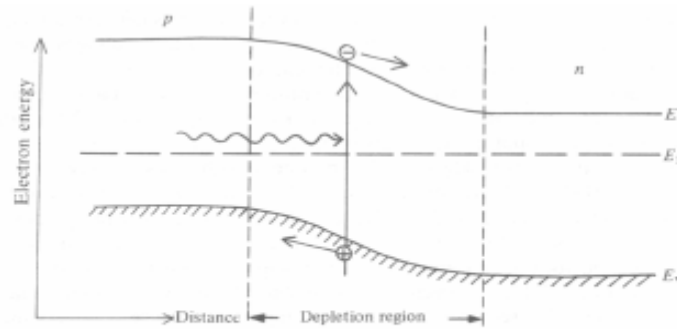
2. Consider a pn junction with $N_A = 2 \times 10^{14} \text{ cm}^{-3}$ and $N_D = 10^{15} \text{ cm}^{-3}$
 - a) Given the space charge distribution shown for the n-side, sketch the distribution for the p-side.
 - b) Sketch the electric field distribution, $E(x)$. Write the expression relating charge density, $\rho(x)$, to the electric field distribution at $x = 0$, $E(0)$. Determine the value of $E(0)$. Use identical graph as in (a)
 - c) Write the expression relating the built-in voltage, V_{bi} to the electric field distribution, $\mathcal{E}(x)$. Sketch the electrostatic potential, $V(x)$. Compute V_{bi} . Use identical graph as in (a)
 - d) Sketch the energy band diagram and label E_C , E_F , E_V , E_i .



3. Imagine that at time $t=0$ we sandwich a $2 \mu\text{m}$ thick undoped Si slab between an n-doped and a p-doped block of Si as shown. Assume complete ionization of the donor atoms, $\tau_p = \tau_n = 10^{-6} \text{ s}$, $D_n = 35 \text{ cm}^2/\text{s}$ and $D_p = 25 \text{ cm}^2/\text{s}$. What is the hole current density (magnitude and direction) across the undoped region at $t=0$? State any (reasonable) simplifying assumptions.

SESSION 2-4: PHOTODETECTORS

When a p - n junction is formed in a semiconductor material, a region depleted of mobile charge carriers is formed with a high internal electric field across it known as the depletion region. If an electron-hole pair is generated by photon absorption within this region, the internal field will cause the electron and hole to separate as shown in figure 5.1.



$$J_{ph} = -e\eta\phi$$

Figure 5.1: Separation of an electron and hole

Photons per second: ϕ

Quantum efficiency: $\eta = \frac{\text{photocarriers generated}}{\text{photons absorbed}}$

We may detect this charge separation in two ways. If the device is left on open circuit, an externally measurable potential will appear between the p and n regions. This is known as the *photovoltaic* mode of operation.

On the other hand, we may short circuit the device externally (usually operated under reverse bias) in which case an external current flows between the p and n regions figure 5.2. This is known as the *photoconductive* mode of operation.

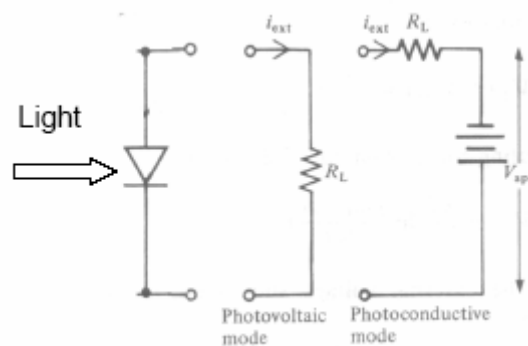


Figure 5.2: External current flows between the p and n regions

The current-voltage characteristics of a $p-n$ junction under various levels of illumination is shown in figure 5.3. The dark characteristic (absence of light) is that of an ordinary $p-n$ junction diode. Under increasing levels of illumination the curve is progressively shifted downwards.

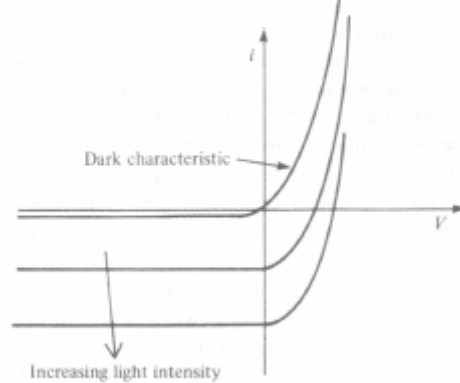


Figure 5.3: The current-voltage characteristics of a $p-n$ junction under various levels of illumination

The main drawback is the presence of a dark current which limits the ultimate sensitivity of the device.

2-4.1 Structure Of pn Photodiode

Simple pn photodiode

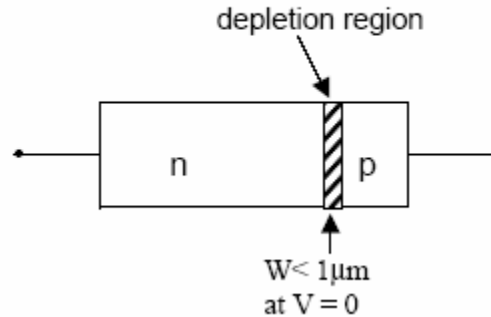


Figure 5.4: Simple pn photodiode structure

For high speed operation the depletion region must be kept thin to reduce the transit time of photocarriers. On the other hand, to increase the quantum efficiency, the depletion region must be sufficiently thick to allow a large fraction of the incident light to be absorbed. Thus there is a trade-off between the speed of response and quantum efficiency figure 5.4. The depletion width

$$W = \sqrt{\frac{2\epsilon_s(V_{bi} - V)}{qN_A}}$$

is therefore a strong determinant of efficiency and speed of the device.

By introducing an intrinsic layer of material in between the junction a $p-i-n$ diode is created figure 5.5. The depletion region thickness (the intrinsic layer) can be tailored to optimize the quantum efficiency and frequency response.

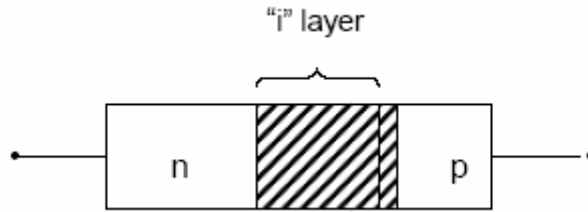


Figure 5.5: A p - i - n diode structure

We have a nominally intrinsic “i” layer, but in practice very lightly doped fully depleted region is formed. Photogenerated carriers are swept out of depletion region by high ϵ field and ideally, carriers move at $\sim V_{sat}$.

2-4.2 Avalanche Detectors

Useful internal amplification is achieved in the avalanche photodiode. In this device, a basic p - n structure is operated under very high (close to breakdown $\sim 100V$) reverse bias. In the very intense Electric field carriers can gain enough energy to enable other carriers to be excited across the energy gap by impact excitation. This results in an ‘avalanche’ of carrier multiplication. Thus, one photon produces many electrons/holes and lots of current (high sensitivity).

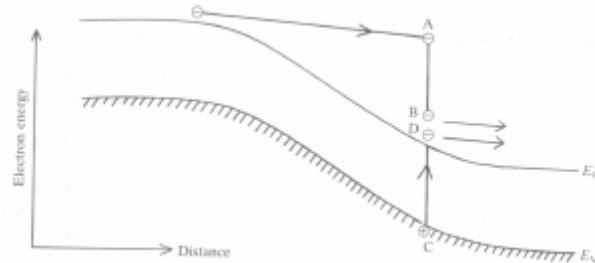


Figure 5.6: The principle of operation of an avalanche photodiode

Figure 5.6 illustrates the principle of operation of an avalanche photodiode. An electron having reached the point A has sufficient energy above the conduction band bottom to enable it to excite an electron from the valence band into the conduction band (C_D). In doing so, it falls from A to B.

2-4.3 Photodetector Materials

Silicon:

- band gap energy = 1.1 eV, sensitive for $\lambda < 1100$ nm
- indirect gap
- cheap, mature technology

Germanium:

- band gap energy = 0.7 eV, works in telecommunications range
- direct gap for $\lambda < 1550$
- difficult material

GaAs:

- band gap energy = 1.5 eV, sensitive for $\lambda < 860$ nm
- direct gap

InGaAsP:

- band gap energy = 0.7 eV, works in telecommunications range
- direct gap