ELECTRONIC DEVICES

About the Author

S. Salivahanan is the Principal of SSN College of Engineering, Chennai. He obtained his B.E. degree in Electronics and Communication Engineering from PSG College of Technology, Coimbatore, M.E. degree in Communication Systems from NIT, Trichy and Ph.D. in the area of Microwave Integrated Circuits from Madurai Kamaraj University. He has four decades of teaching, research, administration and industrial experience both in India and abroad. He has taught at NIT, Trichy, A.C. College of Engineering and Technology, Karaikudi, RV College of Engineering, Bangalore, and Mepco Schlenk Engineering College, Sivakasi. He has industrial experience as Scientist/Engineer at Space Applications Centre, ISRO, Ahmedabad, Telecommunication Engineer at State Organization of Electricity, Iraq and Electronics Engineer at Electric Dar Establishment, Kingdom of Saudi Arabia.

He is the author of 40 popular books which include all-time bestsellers such as *Basic Electrical and Electronics Engineering, Electronic Devices and Circuits, Linear Integrated Circuits*, and *Digital Signal Processing*, all published by McGraw Hill Education. He has also authored the books on *Digital Circuits and Design, Electromagnetic Field Theory, Circuit Theory*, *Network Analysis and Synthesis* and *Control Systems Engineering*. He has published several papers at national and international levels.

Professor Salivahanan is the recipient of Bharatiya Vidya Bhavan National Award for Best Engineering College Principal for 2011 from ISTE, and IEEE Outstanding Branch Counsellor and Advisor Award in the Asia-Pacific region for 1996-97. He was the Chairman of IEEE Madras Section for two years 2008 and 2009 and Syndicate Member of Anna University.

He is a Senior Member of IEEE, Fellow of IETE, Fellow of Institution of Engineers (India), Life Member of ISTE and Life Member of Society for EMC Engineers. He is also a member of IEEE societies in Microwave Theory and Techniques, Communications, Signal Processing, and Aerospace and Electronics.

ELECTRONIC DEVICES

S Salivahanan
Principal
SSN College of Engineering
Chennai, Tamil Nadu

McGraw Hill Education (India) Private Limited
CHENNAI

McGraw Hill Education Offices
Chennai New York St Louis San Francisco Auckland Bogotá Caracas
Kuala Lumpur Lisbon London Madrid Mexico City Milan Montreal
San Juan Santiago Singapore Sydney Tokyo Toronto

McGraw Hill Education (India) Private Limited

Published by McGraw Hill Education (India) Private Limited
444/1, Sri Ekambara Naicker Industrial Estate, Alapakkam, Porur, Chennai - 600 116

ISBN-13: 978-93-87432-37-6
ISBN-10: 93-87432-37-8

1 2 3 4 5 6 7 8 9 D101417 22 21 20 19 18

Printed and bound in India.

Managing Director: *Kaushik Bellani*

Director—Science & Engineering Portfolio: *Vibha Mahajan*
Senior Portfolio Manager—Science & Engineering: *Hemant K Jha*
Associate Portfolio Manager—Science & Engineering: *Vaishali Thapliyal*

Production Head: *Satinder S Baveja*
Assistant Manager—Production: *Jagriti Kundu*

General Manager—Production: *Rajender P Ghansela*
Manager—Production: *Reji Kumar*

Typeset at The Composers, 260, C.A. Apt., Paschim Vihar, New Delhi 110 063 and printed at

visit us at: www.mheducation.co.in

CONTENTS

PREFACE

Electronic Devices is designed specifically to cater to the needs of second semester ECE students. The book has a perfect blend of focused content and complete coverage. Solved university questions, which are tagged with specific topics, will be extremely helpful to students from the examination point of view. Simple, easy-to-understand and difficult-jargon-free text elucidates the fundamentals of electronics. Several solved examples, circuit diagrams and adequate questions further help students to understand and apply the concepts.

SALIENT FEATURES

- ✓ Comprehensive coverage of syllabus requirements
- ✓ Solutions of examination papers from 2010 to 2017 are present appropriately within the chapters:
  - ◆ Solved university questions as solved examples incorporated appropriately within each chapter
  - ◆ Theory questions are tagged within each chapter
- ✓ Rich exam-oriented pedagogy:
  - ◆ Solved Numerical Examples within chapters: 79
  - ◆ Theory questions tagged within chapters: 126
  - ◆ Two Mark Questions and Answers at the end of each chapter: 85
  - ◆ Unsolved Review Questions: 226

CHAPTER ORGANISATION

- ➤ **Chapter 1** covers PN junction diode, current equations, diffusion and drift current densities, forward and reverse bias characteristics, and switching characteristics.
- ➤ **Chapter 2** explains NPN-PNP junctions, early effect, current equations, input and output characteristics of CE, CB CC, Hybrid-p model, h-parameter model, Ebers Moll Model, Gummel Poon-model, Multi Emitter Transistor.
- ➤ **Chapter 3** deals with JFETs, drain and transfer characteristics, current equations, pinch off voltage and its significance, MOSFET characteristics, threshold voltage, channel length modulation, D-MOSFET, E-MOSFET, current equation, equivalent circuit model and its parameters.

- **Chapter 4** discusses metal-semiconductor junction, Gallium Arsenide devices, MESFET, Schottky barrier diode, Dual-Gate MOSFETs, FinFET, PIN-FET, CNTFET, Zener diode, Varactor diode, tunnel diode, Laser Diode, LDR.
- **Chapter 5** covers UJT, SCR, Diac, Triac, power BJT, power MOSFET, DMOS-VMOS, LED, LCD, phototransistor, optocoupler, solar cell, CCD.

ACKNOWLEDGEMENTS

I sincerely thank the management of SSN College of Engineering, Chennai for the constant encouragement, and for providing necessary facilities for completing this project.

I am highly appreciative of the editorial and production team of McGraw Hill Education (India) for their initiation and support to bring out this book in a short span of time.

I would like to take this opportunity to thank numerous reviewers of my earlier book on Electronic Devices. These include:

P L Ramesh	*Shanmuganathan Engineering College, Pudukkottai*
V Vanathe	*Dhirajlal Gandhi College of Technology, Salem, Pudukkottai*
D Arul Kumar	*Panimalar Institute of Technology, Chennai*
Balaji VR	*St. Joseph's Institute of Technology, Chennai*
N Nirmal Singh	*VV College of Engineering, Tirunelvelli*
K Rajesh	*Knowledge Institute of Technology, Salem*
S Karthie	*SSN College of Engineering, Chennai*

I would also like to thank **C Karthikeyini**, *KPR Institute of Engineering and Technology, Coimbatore* for her valuable feedback on the draft chapters of this book.

I am thankful to my wife Mrs. Kalavathy Salivahanan and sons Santhosh Kanna and Subadesh Kanna.

I welcome suggestions for the improvement of the book.

S. SALIVAHANAN

Publisher's Note

McGraw Hill Education (India) invites suggestions and comments from you, all of which can be sent to *info.india@mheducation.com* (kindly mention the title and author name in the subject line).

SEMICONDUCTOR DIODE

1.1 INTRODUCTION

Electronics has been defined as that branch of science and technology which relates to the conduction of electricity through vacuum by electrons alone or through gases by electrons and ions. Basically, it is a study of electron devices and their utilization. An electron device is that in which electrons flow through a vacuum or gas or semiconductor. In the beginning of 20th century, electronics began to take technological shape and it has enjoyed an explosive development in the last four decades.

Electronics has a wide range of applications, such as rectification, amplification, power generation, industrial control, photo-electricity, communications and so on. The electronic industry turns out a variety of items in the range of consumer electronics, control and industrial electronics, communication and broadcasting equipments, biomedical equipments, calculators, computers, microprocessors, aerospace and defence equipments and components.

Initially, the discovery of electricity motivated research that has led to the present day concept of an atom. Michael Faraday studied the passage of electricity through liquid solutions. The next step in the study of the electrical nature of matter was made by J.J. Thomson. He focused his attention on the passage of electricity through gases. The studies by Thomson of the electrical discharge through gases at low pressure established that an atom is composed of charged particles—electrons and protons. Historically, the study of electrical discharge through gases and the discovery of cathode rays marked the beginning of a new branch of Physics called Atomic Physics.

The *PN* junction diode is one of the semiconductor devices with two semiconductor materials in physical contact, one with excess of holes (*P*-type) and other with excess of electrons (*N*-type). A *PN* junction diode may be formed from a single-crystal intrinsic semiconductor by doping part of it with acceptor impurities and the remaining with donors. Such junctions can form the basis of very efficient rectifiers. The most important characteristic of a *PN* junction is its ability to allow the flow of current in only one direction. In the opposite direction, it offers very high resistance. The high-vacuum diode has largely been replaced by silicon and selenium rectifiers. Semiconductor diodes find wide applications in all phases of electronics viz., radio and TV, optoelectronics, power supplies, industrial electronics, instrumentation, computers, etc.

1.2 ATOMIC ENERGY LEVEL DIAGRAM

In the energy level diagram, the discrete energy states are represented by horizontal lines, and the height of the line represents the total energy E_n. Figure 1.1 shows the energy level diagram for hydrogen. The

number immediately to the right of a line gives the value of integer n, while the number to the left of each line gives the energy to this level in electron volts. The lowest energy level E_1 is called the normal or the ground state of the atom and the higher energy levels E_2, E_3, E_4, ... are called the excited states. As n increases, the energy levels crowd and tend to form a continuum.

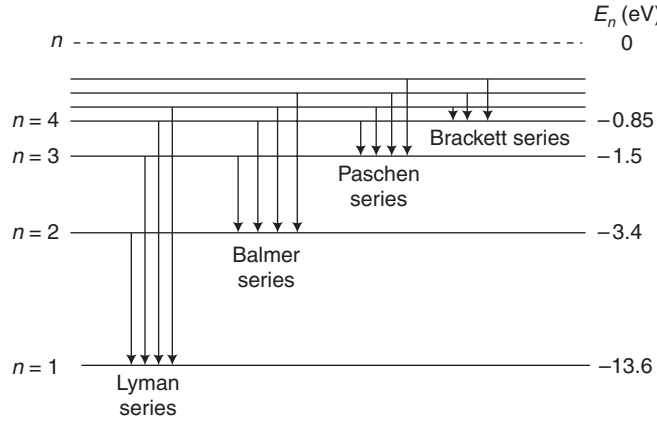


Fig. 1.1 *Energy level diagram of hydrogen*

Sometimes, it is more convenient to specify the emitted radiation by its wavelength, λ in Angstroms which can be written as,

$$\lambda = \frac{12,400}{E_2 - E_1} \tag{1.1}$$

where E_2 and E_1 are the energy levels in electron volts.

1.3 ENERGY-BAND STRUCTURES AND CONDUCTION IN INSULATORS, SEMICONDUCTORS AND METALS

A very poor conductor of electricity is called an insulator; an excellent conductor is a metal; and a material whose conductivity lies between these two extremes is a semiconductor. A material may be classified as one of these three depending upon its energy-band structure.

1.3.1 Insulator

An insulator is a material having extremely poor electrical conductivity. The energy-band structure at the normal lattice spacing is indicated schematically in Fig. 1.2(a). The forbidden energy gap is large; for diamond the gap energy is about 6 eV. If additional energy is given to an electron in the upper level of valence band, this electron attempts to cross the forbidden energy gap and enter the conduction band. However in an insulator, the additional energy which may ordinarily be given to an electron is, in general, much smaller than this high value of forbidden energy gap. Hence no electrical conduction is possible. The number of free electrons in an insulator is very small, roughly about 10^7 electrons/m^3.

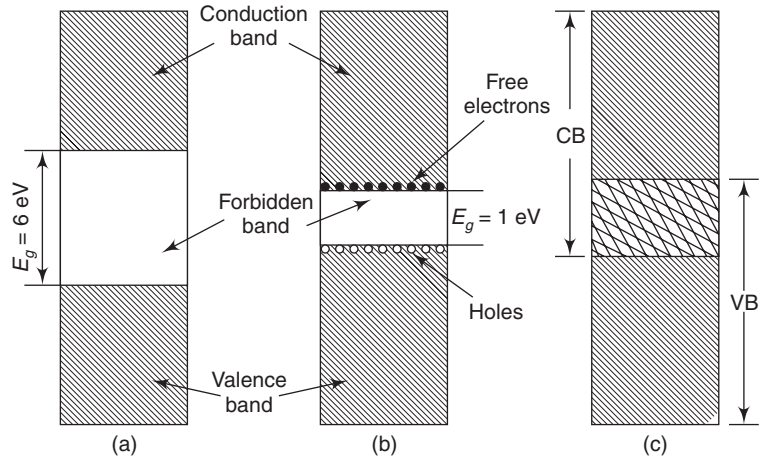


Fig. 1.2 *Energy-band gap in (a) Insulators, (b) Semiconductors, and (c) Metals*

1.3.2 Metal

The conduction in metals is only due to the electrons. A metal has overlapping valence and conduction bands. The valence band is only partially filled and the conduction band extends beyond the upper end of filled valence band. The outer electrons of an atom are as much associated with one ion as with another, so that the electron attachment to any individual atom is almost zero. The band occupied by the valence electrons may not be completely filled and that there are no forbidden levels at higher energies. Depending upon the metal, at least one, and sometimes two or three, electrons per atom are free to move throughout the interior of the metal under the action of applied fields. When an electric field is applied, few electrons may acquire enough additional energy and move to higher energy within the conduction band. Thus the electrons become mobile. Since the additional energy required for transfer of electrons from valence band to conduction band is extremely small, the conductivity of metal is excellent.

In *electron-gas theory* description of a metal, the metal is visualized as a region containing a periodic three-dimensional array of heavy, tightly bound ions permeated with a swarm of electrons that may move about quite freely. According to this theory, the electrons in a metal are continuously moving and the direction of flight changes whenever the electron collides with other electrons. The average distance travelled by an electron between successive collisions is called as *mean-free-path* of an electron. In the absence of any applied potential, the average current in a metal is zero because the number of electrons passing through unit area in any direction is almost same as the number of electrons passing through the same unit area in the opposite direction. This can be attributed to the random nature of motion of electrons.

When a constant electric field E (volts per metre) is applied to a metal, the electrons would be accelerated and the velocity would increase indefinitely with time. However, because of collision of electrons, electrons lose energy and a steady-state condition is reached where a finite value of drift velocity v_d is attained. The drift velocity, v_d is in the direction opposite to that of the electric field and its magnitude is proportional to E. Thus,

$$v_d = \mu E \tag{1.2}$$

where μ = mobility of the electron, m^2/volt-second. Due to the applied field, a steady-state drift velocity has been superimposed upon the random thermal motion of the electrons. Such a directed flow of electrons constitutes a current. If the concentration of free electrons is n (electrons per cubic meter), the current density J (amperes per square metre) is

$$J = nqv_d = nq\mu E = \sigma E \tag{1.3}$$

where

$$\sigma = nq\mu \tag{1.4}$$

σ is the *conductivity* of the metal in (ohm-metre)$^{-1}$. For a good conductor n is very large, approximately, 10^{28} electrons/m^3. Equation (1.3) can be recognized as Ohm's law which states that the conduction current density is proportional to the applied electric field. The energy acquired by the electrons from the applied field is given to the lattice ions as a result of collisions. Hence, power is dissipated within the metal by the electrons, and the power density (Joule heat) is given by

$$JE = \sigma E^2 \text{ watts/metre}^3 \tag{1.5}$$

1.3.3 Semiconductor

The conductivity of a material is proportional to the concentration of free electrons. The number of free electrons in a semiconductor lies between 10^7 and 10^{28} electrons/m^3. Thus, a semiconductor has conductivity much greater than that of an insulator but much smaller than that of a metal. Typically, semiconductor has forbidden energy gap of about 1 eV. The most important practical semiconductor materials are germanium and silicon, which have values of E_g of 0.785 and 1.21 eV, respectively, at 0 degree Kelvin. Energies of this magnitude normally cannot be acquired from an applied field. At low temperatures the valence band remains full, the conduction band empty, and these materials are insulators at low temperatures. The conductivity of these materials increases with temperature and hence these materials are called as intrinsic semiconductors. As the temperature is increased, some of the electrons in the valence band acquire thermal energy greater than the gap energy and move into the conduction band. These electrons are now free to move about under the influence of even a small applied field. These free electrons, also called as conduction electrons, constitute for conduction and the material becomes slightly conducting. The current density due to the motion of electrons is given by

$$J_n = n\mu_n qE = \sigma_n E \tag{1.6}$$

where, μ_n is the electron mobility, and the suffix 'n' represents that the respective terms are due to motion of electrons. The absence of an electron in the valence band is represented by a small circle and is called a hole. The hole may serve as a carrier of electricity whose effectiveness is comparable with the free electron. The hole conduction current density is given by

$$J_p = p\mu_p qE = \sigma_p E \tag{1.7}$$

where μ_p is the hole mobility and p is the hole concentration.

Hence, the total current density J in a semiconductor is given by

$$J = (n\mu_n + p\mu_p)qE = \sigma E \tag{1.8}$$

where $\sigma = (n\mu_n + p\mu_p)q$ is the total conductivity of a semiconductor. For a pure semiconductor (intrinsic semiconductor), the number of free electrons is exactly same as the number of holes. Thus, the total current density is

$$J = n_i(\mu_n + \mu_p)\, qE \tag{1.9}$$

where $n_i = n = p$ is the intrinsic concentration of a semiconductor.

The conductivity of an intrinsic semiconductor can be increased by introducing certain impurity atoms into the crystal. This results in allowable energy states which lie in the forbidden energy gap and these impurity levels also contribute to the conduction. Such a semiconductor material is called an extrinsic semiconductor.

Effective mass An electron travelling through a crystal under the influence of an externally applied field hardly notices the electrostatic field of the ions making up the lattice, i.e., it behaves as if the applied field were the only one present. This is the basis of the electron gas approximation and the assumption will now be looked at a little more closely. If the electrons were to experience only the applied field, E, then immediately after a collision it would accelerate in the direction of the field with an acceleration a, proportional to the applied force, i.e.,

$$qE = ma \tag{1.10}$$

where the constant of proportionality is the electron mass, m. When quantum theory is applied to the problem of an electron moving through a crystal lattice, it predicts that under the action of an applied field the electron acceleration will indeed be proportional to the field, but that the constant of proportionality will be different from the normal mass of an electron:

$$qE = m_n a \tag{1.11}$$

The field due to the lattice ions can, therefore, be ignored providing we treat the electrons inside the crystal as if they had a slightly different mass to the real electron mass or, to put it another way, the effect of the lattice ions on an electron is to make it behave as if it had a different mass. This new mass is called the effective mass of the electron, m_n, and the effective mass of a hole, m_p, can be similarly defined. In general, m_n and m_p are not the same. Usually, m_n is of the same order as m; in germanium, for instance, $m_n = 0.2$ m and in silicon, $m_n = 0.4$ m.

The value of effective mass is an important parameter for any semiconductor, especially from the device point of view. Certain semiconductors, for instance, have low electron effective mass and hence high electron mobility. This makes them very suitable for high-frequency devices. The III-V semiconductor GaAs comes into this category and some of the best microwave transistors are made from this material.

1.4 PRACTICAL SEMICONDUCTOR MATERIALS

All semiconductors have crystalline structure. The most commonly used semiconductor materials, germanium, silicon and gallium arsenide have practical applications in Electronics. The most frequently used semiconductors are germanium and silicon because the energy required to break their covalent bonds and release a free electron from their valence bands is lesser than that required for gallium arsenide. The energy required for releasing an electron from the valence band is 0.66 eV for germanium, 1.08 eV for silicon and 1.58 eV for gallium arsenide.

Germanium can be purified relatively well and crystallised easily. Germanium is an earth element and it is obtained from the ash of the certain coals or from the flue dust of the zinc smelters. The recovered germanium is in the form of germanium dioxide powder which is then reduced to pure germanium. Germanium diodes are used as infrared detectors in fibre-optic communication system because of narrower energy gap.

Silicon is an element found in most of the common rocks. Sand is silicon dioxide which is then reduced to 100% pure silicon. Silicon dioxide is a natural insulator which is useful in the fabrication of

semiconductor devices and integrated circuits. Silicon is largely preferred to germanium because of its large gap energy, which produces improved device properties at high temperatures. Silicon is a better thermal conductor and is required to remove unavoidable heat developed in the device.

Gallium arsenide has higher electron mobility, μ_n which leads to faster switching capabilities. It has high temperature operating capabilities because of its larger energy gap.

1.5 ELECTRON EMISSION FROM METALS

Electron emission is the process by which the free electrons escape from the surface of a substance. Metals are used for electron emission because a metal is made up of atoms bound in crystal lattices, of electrons bound to the atoms, and of many free electrons which are not bound to any particular location in the metal. The free electrons are always in motion and travel more or less freely throughout the body of the metal. However, these electrons are free only to the extent that they may transfer from one atom to another within the metal but they cannot leave the metal surface. If a certain amount of external energy is given to a free electron, its kinetic energy is increased and thus electron will cross over the surface barrier to escape from the metal. This amount of kinetic energy required at absolute zero temperature is known as *work function* of that metal. It is denoted by E_w, and expressed in electron-volt (eV). One eV, equal to 1.602×10^{-19} Joule, is the amount of energy acquired by an electron when it is accelerated through a potential difference of one volt (1 V).

The work function of pure metals varies approximately from 2 to 6 eV. The metal used for electron emission should have low work function so that a small amount of external energy is required to cause emission of electrons. Work function of a metal is principally determined by the spacing between its atoms. Wider spacing usually give lower values of work function. The work functions of some of the common metals are given in Table 1.1.

Table 1.1 *Work function of some of the common metals*

Metal	*Work function in eV*	*Metal*	*Work function in eV*
Ag	4.60	K	1.90
Al	3.00	Li	2.21
Ba	2.52	Na	2.00
Ca	3.00	Ni	5.00
Cs	1.67	Rb	1.82
Cu	4.30	Th	3.50
Fe	4.74	Zn	3.44

1.5.1 Types of Electron Emission

The four basic methods of obtaining electron emission from the surface of a metal are classified according to the type of additional energy (equal to the work function of the metal) supplied from the sources such as heat energy, energy stored in electric field, light energy and kinetic energy of the electric charges bombarding the metal surface, as discussed:

Thermionic emission When a metal is heated to sufficient temperature, enough thermal energy is imparted to electrons to enable them to escape from the metal surface. The higher the temperature, the greater is the emission of electrons. This method is known as *thermionic emission* and is employed in vacuum tubes.

Richardson, and later on Dushman, on the thermodynamic basis, derived the equation for thermionic emission at a certain temperature, T. This equation is given by

$$I_{th} = SA_oT^2\, e^{-E_w/kT} \tag{1.12}$$

where I_{th} = thermionic emission current in amperes
S = area of filament in m^2
A_o = constant
T = absolute temperature, K
k = Boltzmann constant, eV/K
E_w = work function, eV.

At high temperatures, the electrons are literally being 'boiled' from the metal surface. This is analogous to the boiling of water.

The commonly used thermionic emitters with low E_w, as given in Table 1.2, are tungsten in high voltage (kV) tubes, thoriated tungsten in high power (kW) tubes and oxide coated metals in low power electron tubes and Cathode Ray Tubes (CRT).

Table 1.2 *Properties of commonly used thermionic emitters*

Thermionic emitter	*Work function in eV*	*Range of operating temperature in K*	*Emission efficiency in mA/W*
Tungsten	4.52	2500–2600	2–10
Thoriated tungsten	2.63	1900–2000	50–100
Oxide-coated	1.10	900–1100	100–1000

Field emission When the potential difference between two electrodes is extremely high, electrons are emitted from the negative electrode, even at ordinary temperatures. If the electric field at the metallic surface of the negative electrode is of the order of 10^6 volts per metre, electrons are quite easily pulled out from the surface. This process of electron emission is known as *field emission*. Field emission does not depend on the temperature of the surface. This type of emission is employed in cold cathode devices and mercury and rectifier tube.

Photo-electric emission When the surface of certain metals are illuminated by a beam of light, electrons are ejected out of the metal by the light photons incident on the metal surface. Such an electron emission is known as *photo-electric emission* and is used in photo-electric cell. The greater the intensity of light beam falling on the metal surface, the greater is photo-electric emission.

Secondary emission When a beam of high velocity electrons strike a metal surface, the free electrons are ejected out of the metal. This process is known as *secondary emission*.

1.6 CLASSIFICATION OF SEMICONDUCTORS

Semiconductors are classified as (i) intrinsic (pure) and (ii) extrinsic (impure) types. The extrinsic semiconductors are of *N*-type and *P*-type.

Intrinsic semiconductor A pure semiconductor is called intrinsic semiconductor. Even at room temperature, some of the valence electrons may acquire sufficient energy to enter the conduction band to form free electrons. Under the influence of electric field, these electrons constitute electric current. A missing electron in the valence band leaves a vacant space there, which is known as a *hole*, as shown in Fig. 1.3. Holes also contribute to electric current.

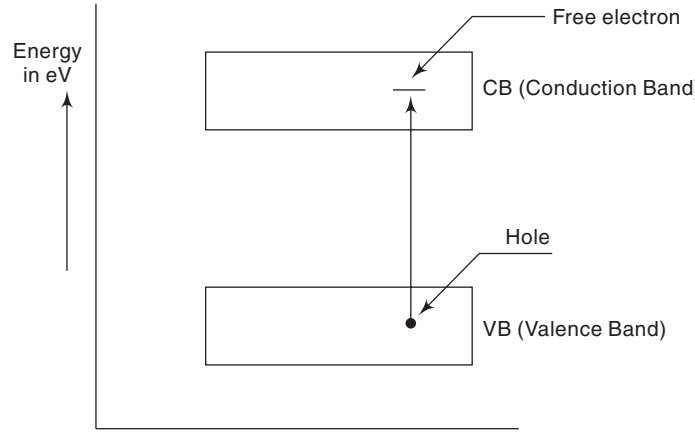


Fig. 1.3 *Creation of electron-hole pair in a semiconductor*

In an intrinsic semiconductor, even at room temperature, electron-hole pairs are created. When electric field is applied across an intrinsic semiconductor, the current conduction takes place due to free electrons and holes. Under the influence of electric field, total current through the semiconductor is the sum of currents due to free electrons and holes.

Though the total current inside the semiconductor is due to free electrons and holes, the current in the external wire is fully by electrons. In Fig. 1.4, holes being positively charged move towards the negative terminal of the battery. As the holes reach the negative terminal of the battery, electrons enter the semiconductor near the terminal (X) and combine with the holes. At the same time, the loosely held electrons near the positive terminal (Y) are attracted towards the positive terminal. This creates new holes near the positive terminal which again drift towards the negative terminal.

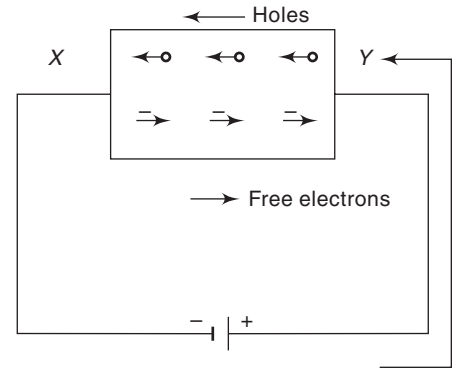


Fig. 1.4 *Current conduction in semiconductor*

Extrinsic semiconductor Due to the poor conduction at room temperature, the intrinsic semiconductor as such is not useful in the electronic devices. Hence, the current conduction capability of the intrinsic semiconductor should be increased. This can be achieved by adding a small amount of impurity to the intrinsic semiconductor, so that it becomes impure or extrinsic semiconductor. This process of adding impurity is known as *doping*.

The amount of impurity added is extremely small, say 1 to 2 atoms of impurity for 10^6 intrinsic atoms.

N-type semiconductor A small amount of pentavalent impurities such as arsenic, antimony or phosphorus is added to the pure semiconductor (germanium or silicon crystal) to get *N*-type semiconductor.

Germanium atom has four valence electrons and antimony has five valence electrons. As shown in Fig. 1.5, each antimony atom forms a covalent bond with surrounding four germanium atoms. Thus, four valence electrons of antimony atom form covalent bond with four valence electrons of individual germanium atom and fifth valence electron is left free which is loosely bound to the antimony atom.

This loosely bound electron can be easily excited from the valence band to the conduction band by the application of electric field or increasing the thermal energy. Thus, every antimony atom contributes one conduction electron without creating a hole. Such pentavalent impurities are called donor impurities because it donates one electron for conduction. On giving an electron for conduction, the donor atom becomes positively charged ion because it loses one electron. But it cannot take part in conduction because it is firmly fixed in the crystal lattice.

Thus, the addition of pentavalent impurity (antimony) increases the number of electrons in the conduction band, thereby increasing the conductivity of *N*-type semiconductor. As a result of doping, the number of free electrons far exceeds the number of holes in an *N*-type semiconductor. Hence electrons are called majority carriers and holes are called minority carriers.

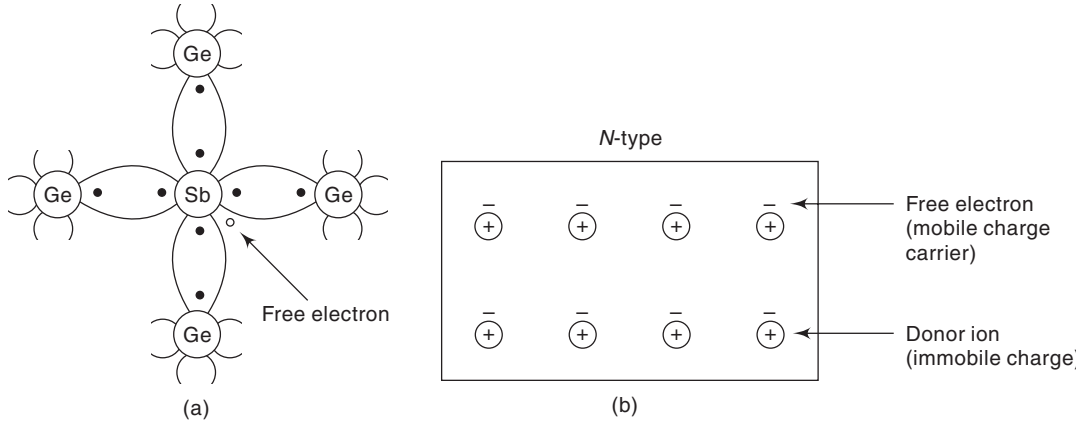


Fig. 1.5 *N-type semiconductor: (a) Formation of covalent bonds, (b) Charged carriers*

P-type semiconductor A small amount of trivalent impurity such as aluminium or boron is added to the pure semiconductor to get the *P*-type semiconductor. Germanium (Ge) atom has four valence electrons and boron has three valence electrons as shown in Fig. 1.6. Three valence electrons in boron form covalent bond with four surrounding atoms of Ge leaving one bond incomplete which gives rise to a hole. Thus trivalent impurity (boron) when added to the intrinsic semiconductor (germanium) introduces a large number of holes in the valence band. These positively charged holes increase the conductivity of *P*-type semiconductor. Trivalent impurities such as boron is called acceptor impurity because it accepts free electrons in the place of holes. As each boron atom donates a hole for conduction, it becomes a negatively charged ion. As the number of holes is very much greater than the number of free electrons in a *P*-type material, holes are termed as majority carriers and electrons as minority carriers.

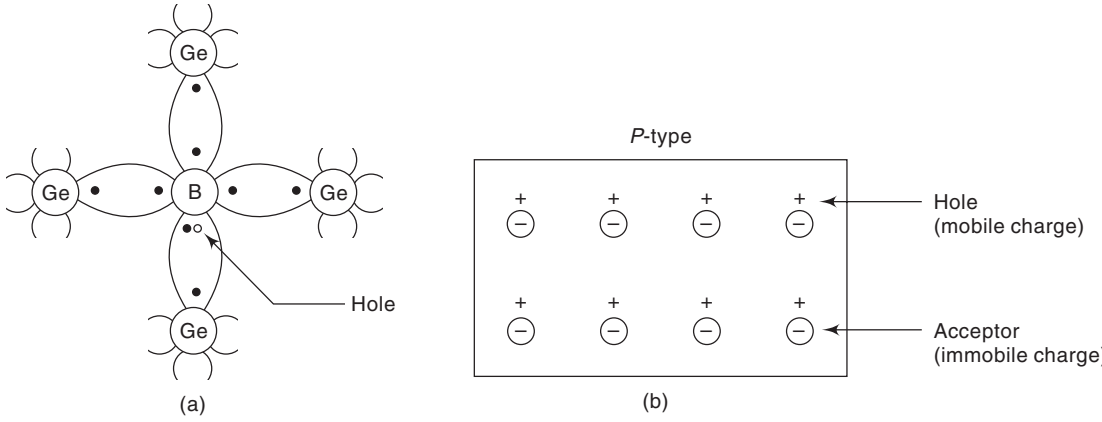


Fig. 1.6 *P-type semiconductor: (a) Formation of covalent bonds, (b) Charged carriers*

1.7 CONDUCTIVITY OF SEMICONDUCTOR

In a pure semiconductor, the number of holes is equal to the number of electrons. Thermal agitation continues to produce new electron-hole pairs and the electron-hole pair disappears because of recombination. With each electron-hole pair created, two charge-carrying particles are formed. One is negative which is the free electron with mobility μ_n. The other is positive, i.e., the hole with mobility μ_p. The electrons and holes move in opposite directions in an electric field E, but since they are of opposite sign, the current due to each is in the same direction. Hence the total current density J within the intrinsic semiconductor is given by

$$\begin{aligned} J &= J_n + J_p \\ &= q \cdot n \cdot \mu_n E + q \cdot p \cdot \mu_p \cdot E \\ &= (n\mu_n + p\mu_p) qE \\ &= \sigma E \end{aligned} \tag{1.13}$$

where J_n = electron drift current density
J_p = hole drift current density
n = number of electrons per unit volume, i.e., magnitude of free-electron (negative) concentration
p = number of holes per unit volume, i.e., magnitude of hole (positive) concentration
E = applied electric field strength, V/m
q = charge of electron or hole, Coulomb

Hence, σ is the conductivity of a semiconductor which is equal to $(n\,\mu_n + p\,\mu_p)q$. The resistivity (ρ) of a semiconductor is the reciprocal of conductivity, i.e., $\rho = \dfrac{1}{\sigma}$.

It is evident from the above equation that current density within a semiconductor is directly proportional to the applied electric field.

For pure (intrinsic) semiconductor, $n = p = n_i$, where n_i is the intrinsic carrier concentration.

Therefore, $$J = n_i (\mu_n + \mu_p)\, qE$$

and conductivity of an intrinsic semiconductor is $\sigma_i = q \cdot n_i(\mu_n + \mu_p)$. Hence, it is clear that conductivity of an intrinsic semiconductor depends upon its intrinsic concentration (n_i) and the mobility of electrons (μ_n) and holes (μ_p). The intrinsic conductivity of germanium and silicon increase by approximately 5 per cent per °C and 7 per cent per °C rise in temperature respectively due to the influence of n_i.

Conductivity of N- and P-type semiconductors The conductivity of an intrinsic semiconductor,

$$\sigma_i = q \cdot n_i(\mu_n + \mu_p) = q \cdot (n\,\mu_n + p\mu_p) \tag{1.14}$$

For *N*-type semiconductor, as $n>>p$, then the conductivity, $\sigma = q \cdot n \cdot \mu_n$.

For *P*-type semiconductor, as $p>>n$, the conductivity, $\sigma = q \cdot p \cdot \mu_p$.

EXAMPLE 1.1

The mobility of free electrons and holes in pure germanium are 3800 and 1800 cm^2 /V-s respectively. The corresponding values for pure silicon are 1300 and 500 cm^2/V-s, respectively. Determine the values of intrinsic conductivity for both germanium and silicon. Assume $n_i = 2.5 \times 10^{13}$ cm^{-3} for germanium and $n_i = 1.5 \times 10^{10}$ cm^{-3} for silicon at room temperature.

Solution (i) The intrinsic conductivity for germanium,

$$\begin{aligned}\sigma_i &= qn_i(\mu_n + \mu_p)\\ &= (1.602 \times 10^{-19})\,(2.5 \times 10^{13})\,(3800 + 1800)\\ &= 0.0224 \text{ S/cm}\end{aligned}$$

(ii) The intrinsic conductivity for silicon,

$$\begin{aligned}\sigma_i &= qn_i\,(\mu_n + \mu_p)\\ &= (1.602 \times 10^{-19})\,(1.5 \times 10^{10})\,(1300 + 500)\\ &= 4.32 \times 10^{-6} \text{ S/cm}\end{aligned}$$

1.8 ENERGY DISTRIBUTION OF ELECTRONS

An electron inside the metal must possess an energy level that is at least greater than the surface barrier energy E_B, so as to escape to a higher level. It is therefore important to know about the energies possessed by the electrons in a metal. This is given by the energy distribution function.

Energy density The distribution in energy of the free electrons in a metal is given by

$$dn = \rho dE$$

where dn is the number of electrons per cubic metre whose energies lie in the energy interval dE and ρ is the density of electrons in a given energy interval.

It is assumed that there are no potential variations within the metal, since only free electrons are considered. Hence there must be the same number of electrons in each cubic metre of the metal, i.e.,

the density in space (electrons per cubic metre) is a constant. However, there will be electrons having all possible energies within each unit volume of the metal. This distribution in energy is expressed by

$$\rho = f(E)\, N(E)$$

where $N(E)$ is the density of states in the conductance band, and $f(E)$ is the probability that a quantum state with energy E is occupied by an electron. Therefore,

$$N(E) = \gamma E^{\frac{1}{2}}$$

Here, γ is a constant defined by

$$\gamma = \frac{4\pi}{h^3}(2m_n)^{\frac{3}{2}}(1.602 \times 10^{-19})^{\frac{3}{2}} = 6.82 \times 10^{27}$$

where the dimensions of γ are $(m^{-3})\,(eV)^{-\frac{3}{2}}$, m_n is the mass of the electron in kg and h is the Planck's constant in Joule-second.

Fermi–Dirac function The Fermi-Dirac probability function $f(E)$ specifies the fraction of all states at energy E (in eV) occupied under conditions of thermal equilibrium. From quantum physics,

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

where k is the Boltzman constant in eV/K, T is the temperature in K and E_F is the Fermi level or characteristic energy for the crystal in eV.

The Fermi level represents the energy state with 50% probability of being filled if no forbidden band exists. That is, if $E = E_F$, then $f(E) = \frac{1}{2}$ for any value of temperature.

The plots of $f(E)$ vs $(E - E_F)$ and $(E - E_F)$ vs $f(E)$ are shown in Fig. 1.7(a) and Fig. 1.7(b) respectively for $T = 0$°K and larger values of temperature.

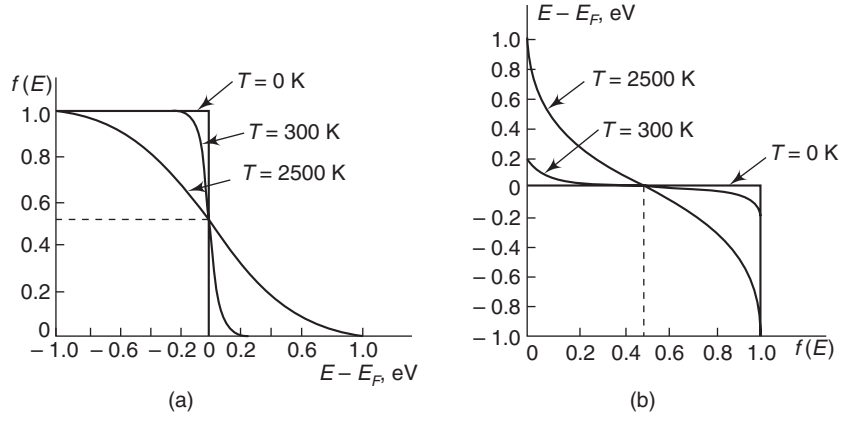


Fig. 1.7 *Fermi-Dirac distribution function f(E) gives the probability that a state of energy E is occupied*

At $T = 0$ K, the following conditions exist:

(i) If $E > E_F$, the exponential term becomes infinite and $f(E) = 0$. Consequently, there is no probability of finding an occupied quantum state of energy greater than E_F at absolute zero temperature.

(ii) If $E < E_F$, the exponential becomes zero and $f(E) = 1$. All quantum levels with energies less than E_F will be occupied at $T = 0$ K.

From the above equations, we get at absolute zero temperature,

$$\rho = \begin{cases} \gamma E^{\frac{1}{2}}; & \text{for } E < E_F \\ 0 \quad ; & \text{for } E > E_F \end{cases}$$

It implies that there are no electrons at 0 K which have energies in excess of E_F. Therefore, the Fermi energy is the maximum energy that any electron may possess at absolute zero temperature.

The relationship given by the above equation is called the *completely degenerate energy distribution function*. In fact, all particles should have zero energy at 0 K. Based on Pauli's exclusion principle, it is also to be mentioned that, since no two electrons has the same set of quantum numbers, not all the electrons can have the same energy even at 0 K.

Fermi level An expression for E_F may be obtained on the basis of the completely degenerate function. The total number of free electrons is given by

$$n = \int_0^{E_F} \gamma E^{\frac{1}{2}}\, dE = \frac{2}{3}\gamma E_F^{\frac{3}{2}}$$

i.e. $$E_F = \left(\frac{3n}{2\gamma}\right)^{\frac{2}{3}}, \quad \text{where } \gamma = 6.82 \times 10^{27}$$

Therefore, $$E_F = 3.64 \times 10^{-19}\, n^{\frac{2}{3}}$$

Since the density of free electrons, n, varies from metal to metal, E_F will also vary among metals. Generally, the numerical value of E_F is less than 10 eV.

1.9 CARRIER CONCENTRATION IN INTRINSIC SEMICONDUCTOR

To calculate the conductivity of a semiconductor, the concentration of free electrons n and the concentration of free holes p must be known.

$$dn = N(E)\, f(E)\, dE$$

where dn represents the number of conduction electrons per cubic metre whose energies lie between E and $E + dE$ and $N(E)$ is the density of states. In a semiconductor, the lowest energy in the conduction band is E_C and hence,

$$N(E) = \gamma (E - E_C)^{\frac{1}{2}}$$

The Fermi Dirac probability function $f(E)$ is given by

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

where E_F is the Fermi level or characteristic energy for the crystal in eV.

The concentration of electrons in the conduction band is,

$$n = \int_{E_C}^{\infty} N(E)\, f(E)\, \mathrm{d}E$$

For $E \geq E_C$, $E - E_F >> kT$,

$$f(E) = e^{-(E-E_F)/kT}$$

and

$$n = \int_{E_C}^{\infty} \gamma (E - E_C)^{\frac{1}{2}} e^{-(E-E_F)/kT}\, dE$$

Substitute $(E - E_C) = x^2$, i.e., $E = x^2 + E_C$ and $dE = 2\, xdx$

At $x = 0$, $E = E_C$

At $x = \infty$, $E = \infty$

Therefore,

$$n = \int_0^{\infty} \gamma x e^{\left(-\frac{x^2 + E_C - E_F}{kT}\right)} [2xdx]$$

$$= 2\gamma e^{-\left(\frac{E_C - E_F}{kT}\right)} \int_0^{\infty} x^2 e^{-\frac{x^2}{kT}} dx$$

We know that, $\int_0^{\infty} x^{2n} e^{-x2/a^2}\, dx = \sqrt{\pi}\, \frac{2n!}{n!} \left(\frac{a}{2}\right)^{2n+1}$

Here, $n = 1$ and $a = \sqrt{kT}$

Therefore,

$$n = 2\gamma e^{-\left(\frac{E_C - E_F}{kT}\right)} \times 2\sqrt{\pi} \left(\frac{\sqrt{kT}}{2}\right)^3$$

Substituting $\gamma = \frac{4\pi}{h^3} (2m_n)^{\frac{3}{2}} (1.602 \times 10^{-19})^{\frac{3}{2}}$, we have

$$n = 2 \times \frac{4\pi}{h^3} (2m_n)^{\frac{3}{2}} (1.602 \times 10^{-19})^{\frac{3}{2}} \times \frac{\sqrt{\pi}\, (kT)^{\frac{3}{2}}}{4} \times e^{-\left(\frac{E_C - E_F}{kT}\right)}$$

$$= 2 \left(\frac{2m_n \pi kT}{h^2}\right)^{\frac{3}{2}} \times (1.602 \times 10^{-19})^{\frac{3}{2}} e^{-\left(\frac{E_C - E_F}{kT}\right)} = N_C\, e^{-\left(\frac{E_C - E_F}{kT}\right)}$$

where $N_C = 2 \left(\frac{2\pi m_n kT}{h^2}\right)^{\frac{3}{2}} (1.602 \times 10^{-19})^{\frac{3}{2}}$, where m_n is the effective mass of an electron.

When the maximum energy in the valence band is E_V, the density of states is given by

$$N(E) = \gamma (E_V - E)^{\frac{1}{2}}$$

The Fermi function of a hole is $[1 - f(E)]$ and is given by

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

where $E_F - E >> kT$ for $E \leq E_V$.

The concentration of holes in the valence band is,

$$p = \int_{-\infty}^{E_V} \gamma (E_V - E)^{1/2}\, e^{-(E_F - E)/kT}\, dE$$

This integral evaluates to

$$p = N_V e^{-(E_F - E_V)/kT}$$

where

$$N_V = 2\left(\frac{2\pi m_p kT}{h^2}\right)^{3/2} (1.602 \times 10^{-19})^{\frac{3}{2}}, \text{ where } m_p \text{ is the effective mass of a hole.}$$

Fermi level in an intrinsic semiconductor In the case of intrinsic material, the crystal must be electrically neutral.

$$n_i = p_i$$

Therefore, $N_C e^{-(E_C - E_F)/kT} = N_V e^{-(E_F - E_V)/kT}$

Taking the logarithm on both sides,

$$\ln \frac{N_C}{N_V} = \frac{E_C + E_V - 2E_F}{kT}$$

$$E_F = \frac{E_C + E_V}{2} - \frac{kT}{2} \ln \frac{N_C}{N_V}$$

If the effective masses of a free electron and hole are the same,

$$N_C = N_V$$

Then,

$$E_F = \frac{E_C + E_V}{2}$$

From the above equation, at the centre of the forbidden energy band, Fermi level is present.

Donor and acceptor impurities If a pentavalent substance (antimony, phosphorous or arsenic) is added as an impurity to pure germanium, four of the five valence electrons of the impurity atoms will occupy covalent bonds and the fifth electron will be available as a carrier of current. These impurities donate excess electron carriers and are hence called *donor* or *N*-type impurities.

If a trivalent impurity (boron, gallium or indium) is added to an intrinsic semiconductor, only three covalent bonds are filled, and the vacancy in the fourth bond constitutes a hole. These impurities are known as *acceptor* or *P*-type impurities.

Fermi level in a semiconductor having impurities The Fermi level in an *N*-type material is given by

$$E_F = E_C - kT \ln \frac{N_C}{N_D}$$

where $N_D = N_C e^{-(E_C - E_F)/kT}$, the concentration of donor atoms.

The Fermi level in a *P*-type material is given by

$$E_F = E_V + kT \ln \frac{N_V}{N_A}$$

where $N_A = N_V e^{-(E_F - E_V)/kT}$, the concentration of acceptor atoms. The change in the position of Fermi level in *N*- and *P*-type semiconductors is shown in Fig. 1.8.

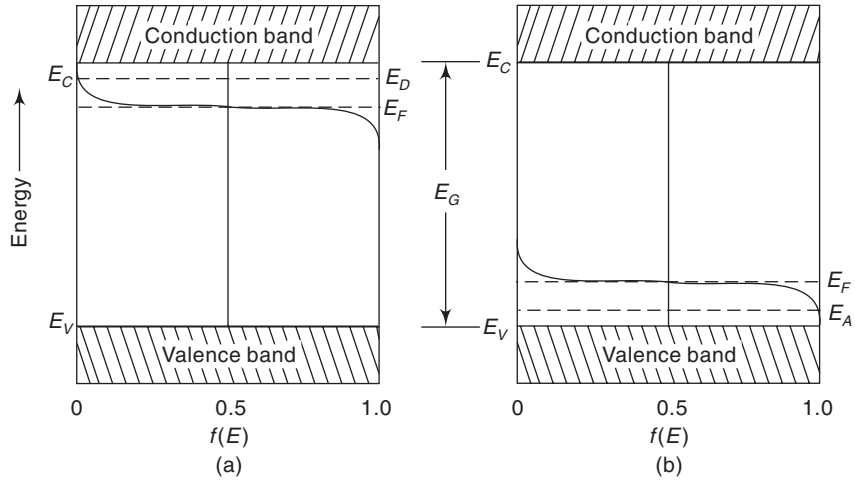


Fig. 1.8 *Positions of Fermi level in (a) N-type, and (b) P-type semiconductors*

Movement of E_F with temperature In an *N*-type semiconductor, as temperature T increases, more number of electron-hole pairs are formed. At very high temperature T, the concentration of thermally generated electrons in the conduction band will be far greater than the concentration of donor electrons. In such a case, as concentration of electrons and holes become equal, the semiconductor becomes essentially intrinsic and E_F returns to the middle of the forbidden energy gap. Hence, it is concluded that as the temperature of the *P*-type and *N*-type semiconductor increases, E_F progressively moves towards the middle of the forbidden energy gap.

EXAMPLE 1.2

In an *N*-type semiconductor, the Fermi level is 0.3 eV below the conduction level at a room temperature of 300 K. If the temperature is increased to 360 K, determine the new position of the Fermi level.

Solution The Fermi level in an *N*-type material is given by

$$E_F = E_C - kT \ln \frac{N_C}{N_D}$$

Therefore,
$$(E_C - E_F) = kT \ln \frac{N_C}{N_D}$$

At $T = 300$ K,
$$0.3 = 300\, k \ln \frac{N_C}{N_D} \quad (1)$$

Similarly,
$$E_C - E_{F1} = 360\, k \ln \frac{N_C}{N_D} \quad (2)$$

Equation (2) divided by Eq. (1) gives

$$\frac{E_C - E_{F1}}{0.3} = \frac{360}{300}$$

Therefore, $$E_C - E_{F1} = \frac{360}{300} \times 0.3 = 0.36 \text{ eV}$$

Hence, the new position of the Fermi level lies 0.36 eV below the conduction level.

EXAMPLE 1.3

In a *P*-type semiconductor, the Fermi level is 0.3 eV above the valance band at a room temperature of 300 K. Determine the new position of the Fermi level for temperatures of (a) 350 K, and (b) 400 K.

Solution The Fermi level in a *P*-type material is given by

$$E_F = E_V + kT \ln \frac{N_V}{N_A}$$

Therefore, $$(E_F - E_V) = kT \ln \frac{N_V}{N_A}$$

At $T = 300$ K, $0.3 = 300\, k \ln \frac{N_V}{N_A}$

(a) At $T = 350$ K, $(E_{F1} - E_V) = 350k \ln \frac{N_V}{N_A}$

Hence, from the above equation

$$\frac{E_{F1} - E_V}{0.3} = \frac{350}{300}$$

Therefore, $$E_{F1} - E_V = \frac{350}{300} \times 0.3 = 0.35 \text{ eV}$$

(b) At $T = 400$ K, $(E_{F2} - E_V) = 400\, k \ln \frac{N_V}{N_A}$

Hence, from the above equation,

$$\frac{E_{F2} - E_V}{0.3} = \frac{400}{300}$$

Therefore, $$E_{F2} - E_V = \frac{400}{300} \times 0.3 = 0.4 \text{ eV}$$

EXAMPLE 1.4

In an *N*-type semiconductor, the Fermi level lies 0.2 eV below the conduction band. Find the new position of Fermi level if the concentration of donor atoms is increased by a factor to (a) 4 and (b) 8. Assume $kT = 0.025$ eV.

Solution In an *N*-type material, the concentration of donor atoms is given by

$$N_D = N_C\, e^{-(E_C - E_F)/kT}$$

Let initially $N_D = N_{DO}$, $E_F = E_{FO}$ and $E_C - E_{FO} = 0.2$ eV

Therefore, $$N_{DO} = N_C\, e^{-0.2/0.025} = N_C\, e^{-8}$$

(a) When $N_D = 4N_{DO}$ and $E_F = E_{F1}$, then

$$4N_{DO} = N_C\, e^{-(E_C - E_{F1})/0.025} = N_C\, e^{-40\,(E_C - E_{F1})}$$

Therefore, $4 \times N_C e^{-8} = N_C e^{-40(E_C - E_{F1})}$

Therefore, $4 = e^{-40(E_C - E_{F1}) + 8}$

Taking natural logarithm on both sides, we get

$$\ln 4 = -40(E_C - E_{F1}) + 8$$

$$1.386 = -40(E_C - E_{F1}) + 8$$

Therefore, $E_C - E_{F1} = 0.165$ eV

(b) When $N_D = 8N_{DO}$ and $E_F = E_{F2}$, then

$$\ln 8 = -40(E_C - E_{F2}) + 8$$

$$2.08 = -40(E_C - E_{F2}) + 8$$

Therefore, $E_C - E_{F2} = 0.148$ eV

EXAMPLE 1.5

In a *P*-type semiconductor, the Fermi level lies 0.4 eV above the valence band. Determine the new position of Fermi level if the concentration of acceptor atoms is multiplied by a factor of (a) 0.5 and (b) 4. Assume $kT = 0.025$ eV.

Solution In a *P*-type material, the concentration of acceptor atoms is given by

$$N_A = N_V e^{-(E_F - E_V)/kT}$$

Let initially $N_A = N_{AO}$, $E_F = E_{FO}$ and $E_{FO} - E_V = 0.4$ eV

Therefore, $N_{AO} = N_V e^{-0.4/0.025} = N_V e^{-16}$

(a) When $N_A = 0.5$, N_{AO} and $E_F = E_{F1}$, then

$$0.5N_{AO} = N_V e^{-(E_{F1} - E_V)/0.025} = N_V e^{-40(E_{F1} - E_V)}$$

Therefore, $0.5 \times N_V e^{-16} = N_V e^{-40(E_{F1} - E_V)}$

Therefore, $0.5 = e^{-40(E_{F1} + E_V) + 16}$

Taking natural logarithm on both sides, we get

$$\ln (0.5) = -40(E_{F1} - E_V) + 16$$

Therefore, $E_{F1} - E_V = 0.417$ eV

(b) When $N_A = 4N_{AO}$ and $E_F = E_{F2}$, then

$$\ln 4 = -40(E_{F2} - E_V) + 16$$

Therefore, $E_{F2} - E_V = 0.365$ eV

1.10 MASS-ACTION LAW

If a pure semiconductor is doped with *N*-type impurities, the number of electrons in the conduction band increases above a level and the number of holes in the valence band decreases below a level, which would be available in the intrinsic (pure) semiconductor. Similarly, the addition of *P*-type impurities to a pure semiconductor increases the number of holes in the valence band above a level and decreases the number of electrons in the conduction band below a level, which would have been

available in the intrinsic semiconductor. This is because the rate of recombination increases due to the presence of a large number of free electrons (or holes).

Further, the experimental results state that under thermal equilibrium for any semiconductor, the product of the number of holes and the number of electrons is constant and is independent of the amount of donor and acceptor impurity doping. This relation is known as *mass-action law* and is given by

$$n \,.\, p = n_i^2 \tag{1.15}$$

where n is the number of free electrons per unit volume, p the number of holes per unit volume and n_i the intrinsic concentration.

While considering the conductivity of the doped semiconductors, only the dominant majority charge carriers have to be considered.

Charge densities in N-type and P-type semiconductors The law of mass-action has given the relationship between free electron concentration and hole concentration. These concentrations are further related by the law of Electrical Neutrality as explained below.

Let N_D be the concentration of donor atoms in an N-type semiconductor. In order to maintain the electric neutrality of the crystal, we have

$$\begin{aligned} n_N &= N_D + p_N \\ &\approx N_D \end{aligned}$$

where n_N and p_N are the electron and hole concentration in the N-type semiconductor. The value of p_N is obtained from the relations of mass-action law as

$$\begin{aligned} p_N &= \frac{n_i^2}{n_N} \\ &\approx \frac{n_i^2}{N_D}, \text{ which is } << n_N \quad \text{or} \quad N_D. \end{aligned}$$

Similarly, in a P-type semiconductor we have

$$\begin{aligned} p_P &= N_A + n_P \\ &\approx N_A \end{aligned}$$

where N_A, p_P and n_P are the concentrations of acceptor impurities, holes and electrons respectively in a P-type semiconductor.

From mass-action law, $\quad n_P = \dfrac{n_i^2}{p_p}$

Therefore, $\quad n_P = \dfrac{n_i^2}{N_A}$, which is $<< p_P$ or N_A

Extrinsic conductivity The conductivity of an N-type semiconductor is given by

$$\sigma_N = q n_N \mu_n \approx q N_D \mu_n, \text{ since } n_N \approx N_D.$$

The conductivity of a P-type semiconductor is given by

$$\sigma_P = q p_P \mu_p \approx q N_A \mu_p, \text{ since } p_P \approx N_A.$$

The doping of intrinsic semiconductor considerably increases its conductivity.

If the concentration of donor atoms added to a *P*-type semiconductor exceeds the concentration of acceptor atoms, i.e., $N_D >> N_A$, then the semiconductor is converted from a *P*-type to *N*-type. Similarly, a large number of acceptor atoms added to an *N*-type semiconductor can convert it to a *P*-type semiconductor if $N_A >> N_D$. This concept is precisely used in the fabrication of *PN* junction, which is an essential part of semiconductor devices and integrated circuits.

EXAMPLE 1.6

Consider a silicon *PN* junction at $T = 300$ K so that $n_i = 1.5 \times 10^{10}$ cm^{-3}. The *N* type doping is 1×10^{10} cm^{-3} and a forward bias of 0.6 V is applied to the *PN* junction. Calculate the minority hole concentration at the edge of the space charge region. (AU April/May 2015)

Solution Given $T = 300$K, $n_i = 1.5 \times 10^{10}$ cm^{-3}, $n \times 1 \times 10^{10}$ cm^{-3} and $V_F = 0.6$ V.

Mass Action law given by Eq. (1.15) is

$$n \cdot p = n_i^2$$

Therefore, the concentration of holes, $p = \dfrac{n_i^2}{n}$

$$= \frac{(1.5 \times 10^{10})^2}{1 \times 10^{10}} = 2.25 \times 10^{10} \text{ cm}^{-3}$$

EXAMPLE 1.7

Find the conductivity of silicon (a) in intrinsic condition at a room temperature of 300 K, (b) with donor impurity of 1 in 10^8, (c) with acceptor impurity of 1 in 5×10^7 and (d) with both the above impurities present simultaneously. Given that n_i for silicon at 300 K is 1.5×10^{10} cm^{-3}, $\mu_n = 1300$ cm^2/V-s, $\mu_p = 500$ cm^2/V-s, number of Si atoms per cm^3 = 5×10^{22}.

Solution (a) In intrinsic condition, $n = p = n_i$

Hence, $\sigma_i = qn_i(\mu_n + \mu_p)$

$$= (1.602 \times 10^{-19})(1.5 \times 10^{10})(1300 + 500)$$

$$= 4.32 \times 10^{-6} \text{ S/cm}$$

(b) Number of silicon atoms/cm^3 $= 5 \times 10^{22}$

Hence, $N_D = \dfrac{5 \times 10^{22}}{10^8} = 5 \times 10^{14}$ cm^{-3}

Further, $n \approx N_D$

Therefore, $p = \dfrac{n_i^2}{n} \approx \dfrac{n_i^2}{N_D}$

$$= \frac{(1.5 \times 10^{10})^2}{5 \times 10^{14}} = 0.46 \times 10^6 \text{ cm}^{-3}$$

Thus, $p << n$. Hence, p may be neglected while calculating the conductivity.

Hence, $\sigma = nq\mu_n = N_D q\mu_n$

$= (5 \times 10^{14})\ (1.602 \times 10^{-19})\ (1300)$

$= 0.104$ S/cm

(c) $N_A = \dfrac{5 \times 10^{22}}{5 \times 10^{7}} = 10^{15}\ \text{cm}^{-3}$

Further, $p \approx N_A$

Hence, $n = \dfrac{n_i^2}{p} \approx \dfrac{n_i^2}{N_A}$

$= \dfrac{(1.5 \times 10^{10})^2}{10^{15}} = 2.25 \times 10^{5}\ \text{cm}^{-3}$

Thus, $p >> n$. Hence, n may be neglected while calculating the conductivity.

Hence, $\sigma = pq\mu_p = N_A q\mu_p$

$= (10^{15} \times 1.602 \times 10^{-19} \times 500)$

$= 0.08$ S/cm

(d) With both types of impurities present simultaneously, the net acceptor impurity density is,

$N'_A = N_A - N_D = 10^{15} - 5 \times 10^{14} = 5 \times 10^{14}\ \text{cm}^{-3}$

Hence, $\sigma = N'_A\, q\mu_p$

$= (5 \times 10^{14})\ (1.602 \times 10^{-19})\ (500)$

$= 0.04$ S/cm

EXAMPLE 1.8

Determine the resistivity of germanium (a) in instrinsic condition at 300 K, (b) with donor impurity of 1 in 10^7, (c) with acceptor impurity of 1 in 10^8, (d) with both the above impurities simultaneously. Given that for germanium at room temperature $n_i = 2.5 \times 10^{13}/\text{cm}^3$, $\mu_n = 3800\ \text{cm}^2/\text{V-Vs}$, $\mu_p = 1800\ \text{cm}^2/\text{V-Vs}$ and a number of Germanium atoms/$\text{cm}^3 = 4.4 \times 10^{22}$.

Solution (a) $n = p = n_i = 2.5 \times 10^{13}\ \text{cm}^{-3}$

Therefore, conductivity, $\sigma = qn_i(\mu_n + \mu_p)$

$= (1.602 \times 10^{-19})(2.5 \times 10^{13})(3800 + 1800)$

$= 0.0224$ S/cm

Hence, resistivity, $\rho = \dfrac{1}{\sigma} = \dfrac{1}{0.0224} = 44.64\ \Omega\text{-cm}$

(b) $N_D = \dfrac{4.4 \times 10^{22}}{10^{7}} = 4.4 \times 10^{15}\ \text{cm}^{-3}$

Also, $n = N_D$

Therefore, $p = \dfrac{n_i^2}{n} = \dfrac{n_i^2}{N_D}$

$$= \frac{(2.5 \times 10^{13})^2}{4.4 \times 10^{15}} = 1.42 \times 10^{11} \text{ holes/cm}^3$$

Here, as $n >> p$, p can be neglected.

Therefore, conductivity, $\sigma = nq\mu_n = N_D q\mu_n$

$$= (4.4 \times 10^{15})\,(1.602 \times 10^{-19})\,(3800) = 2.675 \text{ S/cm}$$

Hence, resistivity, $\rho = \frac{1}{\sigma} = \frac{1}{2.675} = 0.374\ \Omega\text{-cm}$

(c) $$N_A = \frac{4.4 \times 10^{22}}{10^8} = 4.4 \times 10^{14} \text{ cm}^{-3}$$

Also, $p = N_A$

Therefore, $$n = \frac{n_i^2}{p} = \frac{n_i^2}{N_A} = \frac{(2.5 \times 10^{13})^2}{4.4 \times 10^{14}} = 1.42 \times 10^{12} \text{ electrons/cm}^3$$

Here, as $p >> n$, n may be neglected. Then

Conductivity, $\sigma = pq\mu_p = N_A q\mu_p$

$$= (4.4 \times 10^{14})\,(1.602 \times 10^{-19})\,(1800) = 0.1267 \text{ S/cm}$$

Hence, resistivity, $\rho = \frac{1}{\sigma} = \frac{1}{0.1267} = 7.89\ \Omega\text{-cm}$

(d) With both p and n type impurities present,

$$N_D = 4.4 \times 10^{15} \text{ cm}^{-3} \text{ and } N_A = 4.4 \times 10^{14} \text{ cm}^{-3}$$

Therefore, the net donor density N_D' is

$$N_D' = (N_D - N_A) = (4.4 \times 10^{15} - 4.4 \times 10^{14}) = 3.96 \times 10^{15} \text{ cm}^{-3}$$

Therefore, effective $n = N'_D = 3.96 \times 10^{15} \text{ cm}^{-3}$

$$p = \frac{n_i^2}{N_D'} = \frac{(2.5 \times 10^{13})^2}{3.96 \times 10^{15}} = 1.578 \times 10^{11} \text{ cm}^{-3}$$

Here, again $p\left(= \frac{n_i^2}{N_D'}\right)$ is very small compared with N_D' and may be neglected in calculating the effective conductivity.

Therefore, $\sigma = N_D' q\mu_n$

$$= (3.96 \times 10^{15})\,(1.6 \times 10^{-19})\,(3800) = 2.408 \text{ S/cm}$$

Hence, resistivity $\rho = \frac{1}{\sigma} = \frac{1}{2.408} = 0.415\ \Omega\text{-cm}$

EXAMPLE 1.9

A sample of silicon at a given temperature T in intrinsic condition has a resistivity of 25×10^4 Ω-cm. The sample is now doped to the extent of 4×10^{10} donor atoms/cm^3 and 10^{10} acceptor atoms/cm^3. Find the total conduction current density if an electric field of 4 V/cm is applied across the sample. Given that $\mu_n = 1250$ cm^2/V-s, $\mu_p = 475$ cm^2/V-s at the given temperature.

Solution $\sigma_i = qn_i(\mu_n + \mu_p) = \dfrac{1}{25 \times 10^4}$

Therefore, $n_i = \dfrac{\sigma_i}{q(\mu_n + \mu_p)} = \dfrac{1}{(25 \times 10^4)\,(1.602 \times 10^{-19})\,(1250 + 475)}$

$= 1.45 \times 10^{10}\ \text{cm}^{-3}$

Net donor density $N_D\ (= n) = (4 \times 10^{10} - 10^{10}) = 3 \times 10^{10}\ \text{cm}^{-3}$

Hence, $p = \dfrac{n_i^2}{N_D} = \dfrac{(1.45 \times 10^{10})^2}{3 \times 10^{10}} = 0.7 \times 10^{10}\ \text{cm}^{-3}$

Hence, $\sigma = q(n\mu_n + p\mu_p)$

$= (1.602 \times 10^{-19})\,(3 \times 10^{10} \times 1250 + 0.7 \times 10^{10} \times 475)$

$= 6.532 \times 10^{-6}\ \text{S/cm}$

Therefore, total conduction current density,

$J = \sigma E = 6.532 \times 10^{-6} \times 4 = 26.128 \times 10^{-6}\ \text{A/cm}^2$

EXAMPLE 1.10

Find the concentration (densities) of holes and electrons in *N*-type silicon at 300 K, if the conductivity is 300 S/cm. Also find these values for *P*-type silicon. Given that for Silicon at 300 K, $n_i = 1.5 \times 10^{10}/\text{cm}^3$, $\mu_n = 1300\ \text{cm}^2/\text{V-s}$ and $\mu_p = 500\ \text{cm}^2/\text{V-s}$.

Solution (a) Concentration in *N*-type silicon

The conductivity of an *N*-type silicon is $\sigma = qn\mu_n$

Concentration of electrons, $n = \dfrac{\sigma}{q\mu_n}$

$= \dfrac{300}{(1.602 \times 10^{-19})\,(1300)} = 1.442 \times 10^{18}\ \text{cm}^{-3}$

Hence concentration of holes, $p = \dfrac{n_i^2}{n} = \dfrac{(1.5 \times 10^{10})^2}{1.442 \times 10^{18}} = 1.56 \times 10^2\ \text{cm}^{-3}$

(b) Concentration in *P*-type silicon

The conductivity of a *P*-type silicon is $\sigma = qp\mu_p$

Hence, concentration of holes $p = \dfrac{\sigma}{q\mu_p}$

$= \dfrac{300}{(1.602 \times 10^{-19})\,(500)} = 3.75 \times 10^{18}\ \text{cm}^{-3}$

and concentration of electrons, $n = \dfrac{n_i^2}{p} = \dfrac{(1.5 \times 10^{10})^2}{3.75 \times 10^{18}} = 0.6 \times 10^2\ \text{cm}^{-3}$

EXAMPLE 1.11

A specimen of pure germanium at 300 K has a density of charge carriers $2.5 \times 10^{19}/m^3$. It is doped with donor impurity atoms at the rate of one impurity atom every 106 atoms of germanium. All impurity atoms are supposed to be ionised. The density of germanium atom is 4.2×10^{28} atoms/m^3. Calculate the resistivity of the doped germanium if electron mobility is 0.38 m^2/V-s.

If the Germanium bar is 5×10^{-3} m long and has a cross sectional area of $(5 \times 10^{-6})^2$ m^2, determine its resistance and the voltage drop across the semiconductor bar for a current of 1 μA flowing through it.

Solution Density of added impurity atoms is

$$N_D = \frac{4.2 \times 10^{28}}{10^6} = 4.2 \times 10^{22} \text{ atoms/m}^3$$

Also, $n \approx N_D$

Therefore, $$p = \frac{n_i^2}{n} = \frac{n_i^2}{N_D} = \frac{(2.5 \times 10^{19})^2}{4.2 \times 10^{22}} = 1.488 \times 10^6 \text{ m}^{-3}$$

Here, as $p << n$, p may be neglected.

Therefore, $$\sigma = qN_D \mu_n$$
$$= (1.602 \times 10^{-19})\ (4.2 \times 10^{22})\ (0.38) = 2.554 \times 10^3 \text{ S/m}$$

Therefore, resistivity, $$\rho = \frac{1}{\sigma} = \frac{1}{2.554 \times 10^3} = 0.392 \times 10^{-3} \ \Omega\text{-m}$$

Resistance, $$R = \frac{\rho L}{A} = \frac{0.392 \times 10^{-3} \times 5 \times 10^{-3}}{(5 \times 10^{-6})^2} = 78.4 \text{ k}\Omega$$

Voltage drop, $$V = RI = 78.4 \times 10^3 \times 10^{-6} = 78.4 \text{ mV}$$

EXAMPLE 1.12

The resistivities of the two sides of an abrupt germanium diode are 2 Ω cm (*P* side) and 1 Ωcm (*N* side) at 300 K. The mobility of electrons and holes in germanium are μ_n = 3800 cm^2/V sec and μ_p = 1800 cm^2/V sec respectively. Calculate the height E_0 of the potential-energy barrier. (AU May/June 2012)

Solution Since the conductivities of *P* and *N* are given by $\sigma_{p0} = q\mu_p p_{p0}$ and $\sigma_{n0} = q\mu_n n_{n0}$, the concentration of holes and electrons in the *P*- and *N*-type germanium junction is given by

$$p_{p0} = \frac{\sigma_{p0}}{q\mu_p} = \frac{\frac{1}{2}\Omega \text{ cm}}{(1.602 \times 10^{-19})(1800 \text{ cm}^2/\text{V sec})}$$
$$= 1.74 \times 10^{15} \text{ cm}^{-3}$$

$$n_{n0} = \frac{\sigma_{n0}}{q\mu_n} = \frac{\frac{1}{2}\Omega \text{ cm}}{(1.602 \times 10^{-19})(3800 \text{ cm}^2/\text{V sec})}$$
$$= 1.64 \times 10^{15} \text{ cm}^{-3}$$

The concentration of hole in *N*-side of the function is given by

$$p_{n0} = \frac{n_i^2}{n_{n0}} = \frac{6.25 \times 10^{26}}{1.64 \times 10^{15}} = 3.8 \times 10^{11} \text{ cm}^{-3}$$

We know that, the contact potential is

$$V_0 = \frac{E_0}{q} = \frac{kT}{q} \ln \frac{p_{p0}}{p_{n0}}$$

Therefore,

$$E_0 = 0.0259 \times \ln\left(\frac{1.74 \times 10^{15}}{3.8 \times 10^{11}}\right) = 0.218 \text{ eV}$$

$$V_0 = \frac{E_0}{q} = \frac{0.218 \text{ eV}}{q} = 0.218 \text{ V}$$

1.11 PROPERTIES OF INTRINSIC SEMICONDUCTORS

Table 1.3 gives the important properties of intrinsic silicon and germanium at room temperature (300 K).

Table 1.3 *Properties of some common semiconductors at room temperature*

	Si	*Ge*	*GaAs*
Atomic number	14	32	–
Atomic weight	28.09	72.59	–
Atomic density, (m^{-3})	5.02×10^{28}	4.42×10^{28}	–
Lattice constant, a (nm)	0.543	0.565	0.563
Relative permittivity, ε_r	11.8	16.0	13.5
Density, g/cm^3	2.33	5.32	–
Energy gap, E_G (eV)	1.08	0.66	1.58
Electron mobility, μ_n (m^2/V–s)	0.13	0.38	0.85
Hole mobility, μ_p (m^2/V–s)	0.05	0.18	0.04
Intrinsic concentration, n_i (m^{-3})	1.38×10^{16}	2.5×10^{19}	9×10^{12}
Electron diffusion constant, D_n (m^2/S) = $\mu_n V_T$	0.0034	0.0099	0.020
Hole diffusion constant, D_P (m^2/S) = $\mu_p V_T$	0.0013	0.0047	–
Density of states at conduction band edge, N_C (m^{-3})	2.8×10^{25}	1.0×10^{25}	4.7×10^{23}
Density of states at valence band edge, N_V (m^{-3})	1.0×10^{25}	6.0×10^{24}	7.0×10^{24}
Intrinsic resistivity, Ω-cm	23×10^4	45	–
Melting point	1420	936	1250

1.12 VARIATION IN SEMICONDUCTOR PARAMETERS WITH TEMPERATURE

As the semiconductors operate on a wide range of temperature, the variation of semiconductor parameters such as intrinsic concentration (n_i), mobility (μ), conductivity (σ) and energy gap (E_G) with temperature are important and are discussed below.

Intrinsic concentration (n_i) We know that the electron concentration,

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

and hole concentration,

$$p = N_V\, e^{-(E_F - E_V)/kT}$$

Therefore, the product of electron-hole concentrations,

$$np = N_C\, N_V\, e^{-(E_C - E_V)/kT}$$
$$= N_C\, N_V\, e^{-(E_G)/kT}$$

According to mass-action law, $np = n_i^2$. Hence, it is clear that the intrinsic concentration, n_i, is independent of the Fermi level but depends on temperature and energy gap, $E_G = E_C - E_V$. The above equation is valid for either an extrinsic or intrinsic material.

Also, in an intrinsic semiconductor, as temperature rises the intrinsic concentration (n_i) increases and the conductivity (σ) also increases. The intrinsic concentration n_i is very sensitive to temperature and is given by

$$n_i^2 = A_o\, T^3\, e^{-E_{GO}/kT}$$

where A_o = a constant, independent of temperature

T = temperature (in K)

E_{GO} = forbidden energy gap at K (in eV) and

k = Boltzmann constant (in eV/K).

The increase in n_i^2 with temperature has an effect on the charge densities in extrinsic semiconductor. In an *N*-type semiconductor, the number of free electrons (n) does not change appreciably with the increase in temperature, while the number of holes increases. In a *P*-type semiconductor, the number of free electrons (n) increases with the increase in temperature, while the number of holes (p) remains constant.

Mobility (μ) The mobility (μ) of an intrinsic semiconductor varies as T^{-m} over a temperature (T) range of 100 to 400 K. For silicon, $m = 2.5$ for electrons and 2.7 for holes. Similarly, for germanium $m = 1.66$ for electrons and 2.33 for holes. The mobility of an intrinsic semiconductor decreases with the rise in temperature. Also, the mobility is a function of electric field intensity (EV) and it remains constant only if $E < 10^3$ m in *N*-type silicon. For $10^3 < E < 10^4$ V/m, μ_n varies approximately as $E^{-1/2}$. For higher fields, $\mu_n = \dfrac{V_d}{E}$ and hence, μ_n varies inversely as E and the carrier speed approaches a constant value of about 10^5 m/s.

Conductivity (σ) The conductivity (σ) of an intrinsic semiconductor depends upon the number of hole-electron pairs and mobility. The number of hole-electron pairs increases with the rise in

temperature, while its mobility decreases. However, the conductivity of an intrinsic semiconductor increases with the increase in temperature. The conductivity at any temperature (T K) is given by

$$\sigma = \sigma_o [1 + \alpha (T - T_o)]$$

where α is the temperature coefficient.

The conductivity of extrinsic semiconductor decreases with the rise in temperature, as the number of majority carriers is almost constant and mobility decreases.

Energy gap The energy gap decreases with the increase in temperature and is given by

$$E_G(T) = E_{Go} - \beta T$$

where β = a constant, whose value depends upon the nature of the material. Its value for silicon is 3.6×10^{-4} and for germanium 2.23×10^{-4}, and E_{Go} = energy gap at 0 K. Its value for silicon is 1.21 eV and for germanium, 0.785 eV.

1.13 DRIFT CURRENT AND DRIFT CURRENT DENSITY

[AU April 2015, 12 marks]

The flow of charge, i.e., current, through a semiconductor material are of two types, namely drift and diffusion. The net current that flows through a *PN* junction diode also has two components, viz. (i) drift current and (ii) diffusion current.

Drift current When an electric field is applied across the semiconductor material, the charge carriers attain a certain drift velocity v_d, which is equal to the product of the mobility of the charge carriers and the applied electric field intensity, E. The holes move towards the negative terminal of the battery and electrons move towards the positive terminal. This combined effect of movement of the charge carriers constitutes a current known as the *drift current*. Thus the drift current is defined as the flow of electric current due to the motion of the charge carriers under the influence of an external electric field.

Drift current density Figure 1.9 shows that a conducting wire of length l cm contains N electrons. If an electron travels a distance of l cm in the conductor in time T sec, the total number of electrons passing through any cross section of wire per second is $\frac{N}{T}$.

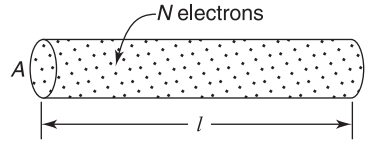


Fig. 1.9 *Conducting wire to determine drift current density*

The total current flowing through the wire with area of cross-section A is given by

$$I = \frac{Nq}{T} = \frac{Nqv_d}{l}$$

where v_d is the drift velocity and time $T = \frac{l}{v_d}$. The drift current density J is defined as the current per unit area of the conducting medium.

i.e., $$J = \frac{I}{A} = \frac{Nqv_d}{lA} \text{ A/cm}^2$$

Here, lA is the volume containing N electrons and the electron concentration, $n = \frac{N}{lA}$.

Therefore, $$J = nqv_d = \rho v_d$$

where $\rho = nq$ is the charge density, in coulombs per cubic centimeter. Generally, the above current density can also be written as

$$J = nqv_d = nq\mu E = \sigma E$$

where E is the applied electric field intensity in V/cm, μ is the mobility of electrons in cm^2/V-s, the drift velocity, $v_d = \mu E$ in cm/s, and the conductivity, $\sigma = nq\mu$ in $\text{S} \cdot \text{m}^{-1}$. The equation $J = \sigma E$ is also called Ohm's law.

The drift current density due to the charge carriers such as free electrons and holes are the current passing through a square centimeter perpendicular to the direction of flow. The equation for the drift current density, J_n, due to free electrons is given by

$$J_n = qn\mu_n E \text{ A/cm}^2$$

and the drift current density, J_p, due to holes is given by

$$J_p = qp\mu_p E \text{ A/cm}^2$$

where n = number of free electrons per cubic centimetre

p = number of holes per cubic centimetre

μ_n = mobility of electrons in cm^2/V-s

μ_p = mobility of holes in cm^2/V-s

E = applied electric field intensity in V/cm

q = charge of an electron = 1.602×10^{-19} coloumb.

1.14 DIFFUSION CURRENT AND DIFFUSION CURRENT DENSITY

It is possible for an electric current to flow in a semiconductor even in the absence of the applied voltage provided a concentration gradient exists in the material. A concentration gradient exists if the number of either electrons or holes is greater in one region of a semiconductor as compared to the rest of the region. In a semiconductor material, the charge carriers have the tendency to move from the region of higher concentration to that of lower concentration of the same type of charge carriers. Thus, the movement of charge carriers takes place resulting in a current called *diffusion current*. The diffusion current depends on the material of the semiconductor, type of charge carriers and the concentration gradient.

As indicated in Fig. 1.10(a), the hole concentration $p(x)$ in a semiconductor bar varies from a high value to a low value along the x-axis and is constant in the y- and z-directions.

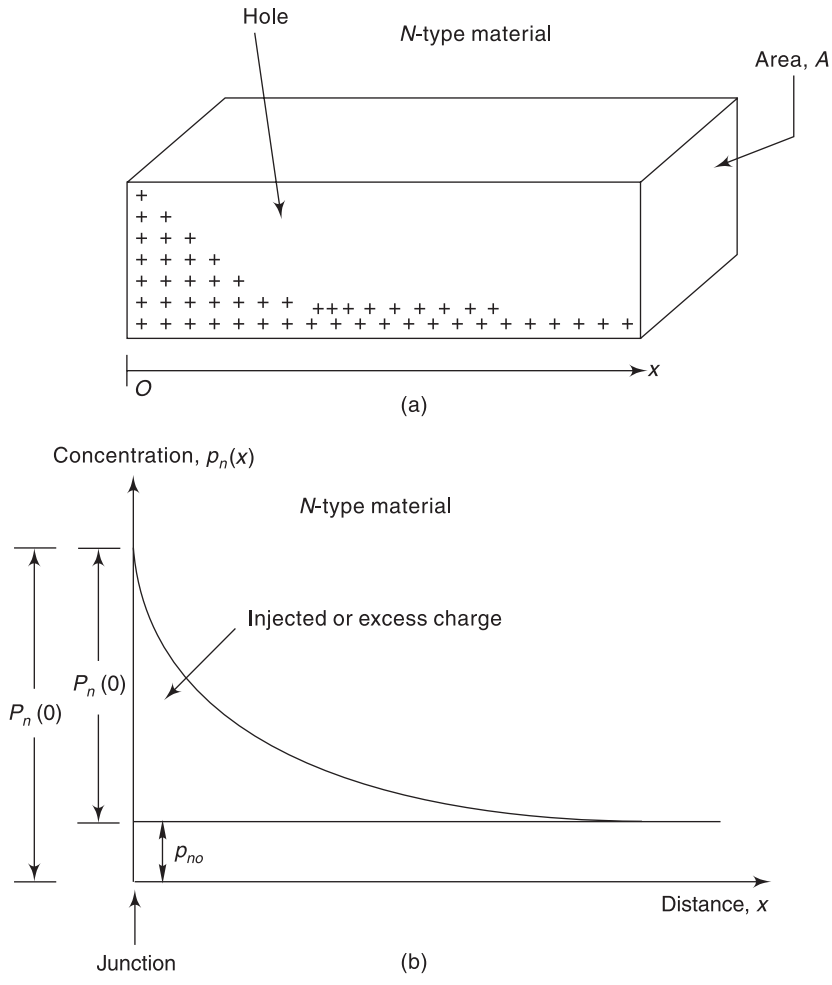


Fig. 1.10 *(a) Excess hole concentration varying along the axis in an N-type semiconductor bar, (b) The resulting diffusion current*

Diffusion current density The diffusion current density J_p due to holes is given by

$$J_p = -qD_p \frac{dp}{dx} \text{ A/cm}^2 \tag{1.16a}$$

Since the hole density $p(x)$ decreases with increasing x as shown in Fig. 1.10(b), dp/dx is negative and the minus sign in the above equation is needed in order that J_p has a positive sign in the positive x-direction.

The diffusion current density due to the free electrons, J_n, is given by

$$J_n = qD_n \frac{dn}{dx} \text{ A/cm}^2 \tag{1.16b}$$

where dn/dx and dp/dx are the concentration gradients for electrons and holes respectively, in the x-direction and D_n and D_p are the diffusion coefficients expressed in cm^2/s for electrons and holes, respectively.

Total current The total current in a semiconductor is the sum of drift current and diffusion current. Therefore, for a *P*-type semiconductor, the total current per unit area, i.e., the total current density is given by

$$J_p = qp\mu_p E - qD_p \frac{dp}{dx}$$

Similarly, the total current density for an *N*-type semiconductor is given by

$$J_n = qn\,\mu_n\,E + qD_n \frac{dn}{dx}$$

1.14.1 Einstein Relationship for Semiconductor

There exists a definite relationship between the mobility and diffusion coefficient of a particular type of charge carrier in the same semiconductor. The higher the value of mobility of a charge carrier, the greater will be its tendency to diffuse. The equation which relates the mobility μ and the diffusion coefficient D is known as the *Einstein Relationship*. The Einstein relationship is expressed as

$$\frac{D_p}{\mu_p} = \frac{D_n}{\mu_n} = \frac{kT}{q} = V_T$$

The importance of Einstein relationship is that it can be used to determine D_p (or D_n), if the mobility of holes (or electrons) is measured experimentally. For an intrinsic silicon, D_p = 13 cm^2/s and D_n = 34 cm^2/s. For an intrinsic germanium, D_p = 47 cm^2/s and D_n = 99 cm^2/s.

Diffusion length (L) As shown in Fig. 1.10, the excess hole or electron densities fall off exponentially with distance as a result of the recombination of these excess minority carriers with the majority carriers of the semiconductor. Here, the excess charge carriers have a finite life time, τ, before they are totally destroyed by recombination. The average distance that on excess charge carrier can diffuse during its life time is called the diffusion length L, which is given by

$$L = \sqrt{D\tau}$$

where D is the diffusion coefficient that may be related to the drift mobility, μ, through the Einstein relation as

$$D = \mu \frac{kT}{q}$$

If the transverse length of the semiconductor is greater than the diffusion length L, then the terminal currents are the recombination currents arising out of the recombination, as every electron lost by recombination is supplanted by the terminal electrode to maintain the charge neutrality.

1.15 CARRIER LIFE TIME

We know that in an intrinsic semiconductor the number of holes is equal to the number of free electrons. However, thermal agitation generates new electron-hole pairs. The electrons have a limited life time (τ_n) in the conduction band and periodically fall back to the valance band, so that a mobile electron-hole pair disappears in the recombination process with the energy of excitation appearing as heat energy. On an average, an electron will exist for τ_n seconds and a hole for τ_p seconds before

recombination. The *mean life times* τ_n and τ_p of electrons and holes are very important parameters as they indicate the time required for the excessive electron and hole concentrations to return to their equilibrium values.

Thus, the carrier life time is defined as the time for which, on an average, a charge carrier will exist before recombination with a carrier of opposite charge. Its value varies from nanoseconds to hundreds of micro-seconds and depends on temperature and impurity concentration in the semiconductor material. Gold is extensively used as recombination agent by the manufacturers of semiconductor devices.

Consider an *N*-type semiconductor having thermal equilibrium concentration p_o and n_o of holes and electrons, respectively. When the specimen is illuminated by light or injection of carriers, additional electron-hole pairs are generated uniformly throughout the medium. This causes the concentration of holes and electrons to increase from p_o and n_o to new values. The rate of change of hole density is equal to their rate of generation minus the rate at which they recombine with electrons. Thus,

$$\frac{dp}{dt} = G - R$$

where p is the minority hole concentration at any time t, and G and R are the generation and recombination rates for the minority carriers.

The generation rate, G, is a function of temperature only since charge carriers are produced only by thermal excitation in the absence of current flow. Hence, at a constant temperature, $G(t)$ is constant.

Assuming that the mean life time of hole τ_p is independent of the magnitude of the hole concentration, we get

$$R = \frac{p}{\tau_p} = \text{decrease in hole concentration per second due to recombination.}$$

Also, from the definition of the generation rate, we get

G = increase in hole concentration per second due to thermal generation.

Since charge can neither be created nor destroyed, the rate dp/dt, at every instant of time, equals the algebraic sum of the rates given in the above equations of R and G. Thus,

$$\frac{dp}{dt} = G - \frac{p}{\tau_p}$$

under steady state conditions, $\frac{dp}{dt} = 0$ and with no radiation falling on the semiconductor, the hole concentration p reaches its equilibrium value p_o. Therefore, $G = \frac{p_o}{\tau_p}$

Hence,
$$\frac{dp}{dt} = \frac{p_o - p}{\tau_p}$$

A similar equation can be derived in an exactly similar manner for the change in excess electron density with time in the *P*-type semiconductor and is given by

$$\frac{dn}{dt} = \frac{n_o - n}{\tau_n}$$

1.16 CONTINUITY EQUATION

The fundamental law governing the flow of charge is called the *continuity equation*. The continuity equation as applied to semiconductors describes how the carrier concentration in a given elemental volume of the crystal varies with time and distance. The variation in density is attributable to two basic causes, viz. (i) the rate of generation and loss by recombination of carriers within the element, and (ii) drift of carriers into or out of the element. The continuity equations enable us to calculate the excess density of electrons or holes in time and space.

As shown Fig. 1.11, consider an infinitesimal *N*-type semiconductor bar of volume of area A and length dx and the average minority carrier (hole) concentration P, which is very small compared with the density of majority carriers. At time t, if minority carriers (holes) are injected, the minority current entering the volume at x is I_p and leaving at $x + dx$ is $I_p + dI_p$ which is predominantly due to diffusion. The minority carrier concentration injected into one end of the semiconductor bar decreases exponentially, with distance into the specimen, as a result of diffusion and recombination. Here, dI_p is the decrease in number of coulombs per second within the volume.

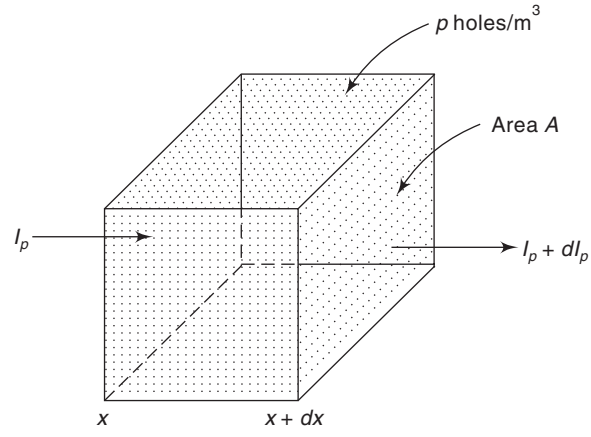


Fig. 1.11 *Relating to continuity equation*

Since the magnitude of the carrier charge is q, then $\frac{dI_p}{p}$ equals the decrease in the number of holes per second within the elemental volume $A \propto x$. As the current density $J_p = \frac{I_p}{A}$, we have

$$\frac{1}{q_A} \cdot \frac{dI_p}{dx} = \frac{1}{q} \cdot \frac{dJ_p}{dx} = \text{decrease in hole concentration per second, due to current } I_p.$$

We know that there is an increase of holes per unit volume per second given by $G = p_o/\tau_p$ due to thermal generation. Further, there is a decrease of holes per unit volume per second given by $R = p/\tau_p$ due to recombination but charge can neither be created nor destroyed. Hence, increase in holes per unit volume per second, dp/dt, must equal the algebraic sum of all the increases in hole concentration. Thus,

$$\frac{\partial p}{\partial t} = -\frac{p - p_o}{\tau p} - \frac{1}{q}\frac{\partial J_p}{\partial x}$$

where

$$J_p = -qD_p\frac{dp}{dx} + qp\mu_p E$$

Therefore,

$$\frac{\partial p}{\partial t} = -\frac{p - p_o}{\tau_p} + D_p\frac{d^2p}{dx^2} - \mu_p\frac{d(pE)}{dx}$$

This is the *continuity equation* or *equation of conservation of charge* for holes stating the condition of dynamic equilibrium for the density of mobile carrier holes. Here, partial derivatives have been used since both p and J_p are functions of both t and x.

Similarly, the continuity equation for electrons states the condition of dynamic equilibrium for the density of mobile carrier electrons and is given by

$$\frac{\partial n}{\partial t} = \frac{n_o - n}{\tau n} - \frac{1}{q}\frac{\partial J_n}{\partial x}$$

where $$J_n = -qD_n\frac{dn}{dx} + qn\,\mu_n E$$

Therefore, $$\frac{\partial n}{\partial t} = -\frac{n - n_o}{\tau_n} + D_n\frac{d^2n}{dx^2} - \mu_n\frac{d(nE)}{dx}$$

We now consider three special cases of continuity equation.

Concentration independent of distance with zero electric field For this special case, the continuity equation can be changed into

$$\frac{\partial p}{\partial t} = -\frac{p - p_o}{\tau_p}$$

Solving the above equation, we get

$$p - p_o = A_1\, e^{-t/\tau_p} \text{ where } A \text{ is a constant}$$

Concentration independent of time with zero electric field For this special case, the continuity equation can be changed into

$$0 = -\frac{p - p_o}{\tau_p} + D_p\frac{d^2p}{dx^2}$$

$$\frac{d^2p}{dx^2} = \frac{p - p_o}{\tau_p D_p}$$

Solving the above equation, we get

$$p - p_o = A_1\, e^{-x/L_p} + A_2\, e^{x/L_p}$$

where A_1 and A_2 are constants.

$L_p = \sqrt{D_p\,\tau_p}$ = diffusion length for holes

Concentration varies sinusoidally with time and with zero electric field Let $P(x, t) = P(x)\, e^{j\omega t}$

For this special case the continuity equation can be changed into

$$j\omega P(x) = -\frac{P(x)}{\tau p} + D_p\frac{d^2\,P(x)}{dx^2}$$

$$\frac{d^2P}{dx^2} = \frac{(1 + j\omega\tau p)}{L_p^2}P$$

At $\omega = 0$,

$$\frac{d^2\,P}{dx^2} = \frac{P}{L_p^2}$$

The above equation is the same as that of the second special case.

Minority carrier injection in homogeneous semiconductors Consider a semiconductor bar shown in Fig. 1.10(a). This bar is uniformly doped with donar atoms so that the charge concentration $n = N_D$ is uniform throughout the bar on which radiation falls on one end of the bar at $x = 0$. Near the illuminated surface, the bound electrons in the covalent bonds capture some of the photons. This energy transfer results in breaking of covalent bonds and generation of hole-electron pairs.

The minority carrier (hole) concentration P is very small compared with the doping level, i.e., $P << n$. The condition $p = P + p_0 << n$ which states that the minority concentration is much smaller than the majority concentration is called the low-level injection. The controlling differential equation for p is

$$\frac{d^2p}{dx^2} = \frac{p - p_0}{\tau_p D_p}$$

The diffusion length for holes L_p is given by

$$L_p = \sqrt{D_p \tau_p}$$

The differential equation for the injected concentration $P = p - p_0$ becomes

$$\frac{d^2P}{dx^2} = \frac{P}{L_p^2}$$

The solution of the equation is

$$P(x) = A_1 e^{-x/AL_p} + A_2 e^{x/L_p}$$

when $x \to \infty$, $A_2 = 0$. At $x = 0$, the injected concentration $P(0)$ to satisfy this boundary condition, $A_1 = P(0)$. Therefore,

$$P(x) = P(0) e^{-x/L_p} = p(x) - p_0$$

Here, the hole concentration decreases exponentially with distance as shown in Fig. 1.10(b).

Diffusion current The minority (hole) diffusion current is $I_p = AJ_p$, where A is the area of cross-section of the bar. Therefore,

$$I_p(x) = \frac{Aq D_p P(0)}{L_p} e^{-x/L_p}$$
$$= \frac{Aq D_p}{L_p} (P(0) - p(0)) e^{-x/L_p}$$

This current decreases exponentially with distance x as that of minority carrier concentration.

The majority (electron) diffusion current is

$$Aq D_n \frac{dn}{dx} = Aq D_n \frac{dp}{dx} = -\frac{D_n}{D_p} I_p$$

where $I_p = -Aq D_n \frac{dp}{dx}$. The magnitude of ratio of majority to minority diffusion current is $D_n/D_p \approx 3$ for Si and 2 for Ge.

Drift current For an open circuit semiconductor bar, the sum of the hole and electron currents should be zero everywhere. Therefore, a majority (electron) drift current I_{nd} exists such that

$$I_p + \left(I_{nd} - \frac{D_n I_p}{D_p} \right) = 0$$

Therefore,

$$I_{nd} = \left(\frac{D_n}{D_p} - 1 \right) I_p$$

The hole drift current I_{pd} is given by

$$I_{pd} = Aqp\mu_p E = \frac{p}{n} \frac{\mu_p}{\mu_n} \left(\frac{D_n}{D_p} - 1 \right) I_p$$

where the electric field, $$E = \frac{1}{Aqn\,\mu_n} \left(\frac{D_n}{D_p} - 1 \right) I_p$$

Here as $p << n$, then $I_{pd} << I_p$, i.e., the hole drift current is negligible compared to the hole diffusion current.

1.17 FORWARD AND REVERSE BIAS CHARACTERISTICS OF *PN* JUNCTION DIODE

[AU Dec 2015, Nov 2014, Nov 2010 and June 2010, 12 marks]

1.17.1 *PN* Junction Diode in Equilibrium with no Applied Voltage

In a piece of semiconductor material, if one half is doped by *P*-type impurity and the other half is doped by *N*-type impurity, a *PN* junction is formed. The plane dividing the two halves or zones is called *PN* junction. As shown in Fig. 1.11, the *N*-type material has high concentration of free electrons, while *P*-type material has high concentration of holes. Therefore, at the junction there is a tendency for the free electrons to diffuse over to the *P*-side and holes to the *N*-side. This process is called *diffusion.* As the free electrons move across the junction from *N*-type to *P*-type, the donor ions become positively charged. Hence a positive charge is built on the *N*-side of the junction. The free electrons that cross the junction uncover the negative acceptor ions by filling in the holes. Therefore, a net negative charge is established on the *P*-side of the junction. This net negative charge on the *P*-side prevents further diffusion of electrons into the *P*-side. Similarly, the net positive charge on the *N*-side repels the hole crossing from *P*-side to *N*-side. Thus a barrier is set-up near the junction which prevents further movement of charge carriers, i.e., electrons and holes. As a consequence of the induced electric field across the depletion layer, an electrostatic potential difference is established between *P*- and *N*-regions, which is called the potential barrier, junction barrier, diffusion potential, or contact potential, V_o. The magnitude of the contact potential V_o varies with doping levels and temperature. V_o is 0.3 V for germanium and 0.72 V for silicon.

The electrostatic field across the junction caused by the positively charged *N*-type region tends to drive the holes away from the junction and negatively charged *P*-type region tends to drive the electrons away from the junction. The majority holes diffusing out of the *P*-region leave behind negatively charged acceptor atoms bound to the lattice, thus exposing negative space charge in a previously neutral region. Similarly, electrons diffusing from the *N*-region expose positively ionised donor atoms, and a double space charge layer builds up at the junction as shown in Figs. 1.12(a) and (c).

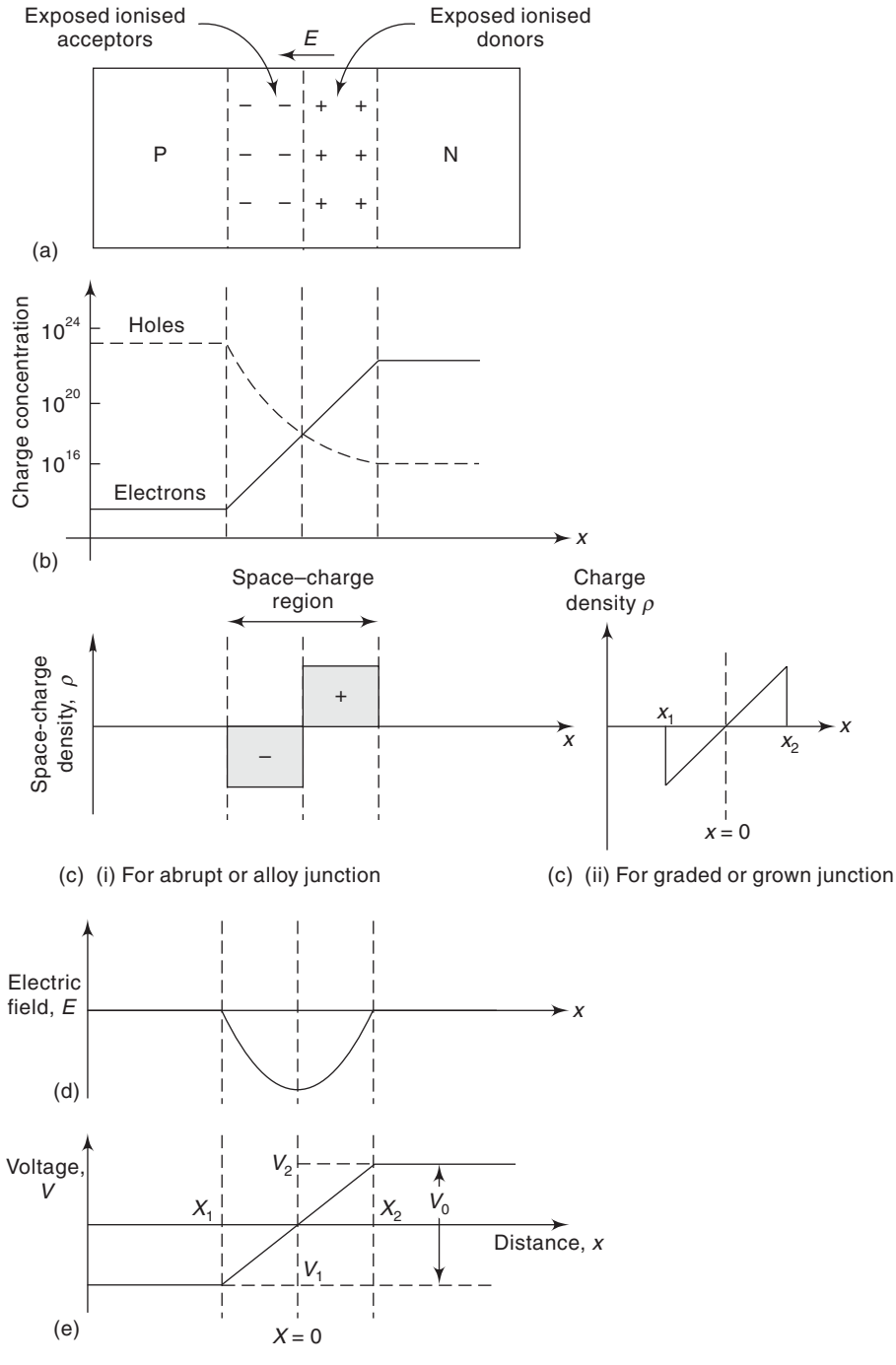


Fig. 1.12 *Formation of PN junction*

It is noticed that the space-charge layers are of opposite sign to the majority carriers diffusing into them, which tends to reduce the diffusion rate. Thus, the double space of the layer causes an electric field to be set up across the junction directed from *N*- to *P*-regions, which is in such a direction to inhibit diffusion of majority electrons and holes, as illustrated in Figs. 1.12(a) and (d). The shape of the charge density, ρ, depends upon how the diode is doped, Thus, the junction region is depleted of mobile charge carriers. Hence, it is called the depletion region (layer), the space charge region, or the transition region. The depletion region is of order 0.5 μm thick. There are no mobile carriers in this very narrow depletion layer. Hence no current flows across the junction and the system is in equilibrium. To the left of this depletion layer, the carrier concentration is $p \approx N_A$, and to its right it is $n \approx N_D$.

Calculation of depletion width Let us now consider the width of the depletion region in the junction of Fig. 1.12. The region contains space charge due to the fact that, donors on the *N*-side and acceptors on the *P*-side have lost their accompanying electrons and holes. Hence, an electric field is established which, in turn, causes a difference in potential energy, qV_o, between the two parts of the specimen. Thus, a potential is built up across the junction and Fig. 1.12(e) represents the variation in potential. Here, *P*-side of the junction is at a lower potential than the *N*-side which means that the electrons on the *P*-side have a great potential energy.

In this analysis, let us consider an *alloy junction* in which there is an abrupt change from acceptor ions on *P*-side to donor ions on *N*-side. Assume that the concentration of electrons and holes in the depletion region is negligible and that all of the donors and acceptors are ionised. Hence, the regions of space charge may be described as

$$\rho = \begin{cases} -qN_A, & 0 > x > X_1 \\ -qN_D, & X_2 > x > 0 \\ 0, & \text{elsewhere} \end{cases}$$

where ρ is the space charge density, as indicated in Fig. 1.12(c)(i). The axes have been chosen in Fig. 1.12(e) in such a way that V_1 and X_1 have negative values. The potential variation in the space charge region can be calculated by using Poisson's equation, which is given by

$$\nabla^2 V = -\frac{\rho(x, y, z)}{\varepsilon_0 \varepsilon_r}$$

where ε_r is the relative permittivity. The relevant equation for the required one-dimensional problem is

$$\frac{d^2 V}{dx^2} = -\frac{\rho}{\varepsilon_0 \varepsilon_r}$$

Applying the above equation to the *P*-side of the junction, we get

$$\frac{d^2 V}{dx^2} = \frac{qN_A}{\varepsilon_0 \varepsilon_r}$$

Integrating twice, we get

$$V = \frac{qN_A x^2}{2\varepsilon_0 \varepsilon_r} + Cx + D$$

where C and D are the constants of integration.

From the Fig. 1.12(e), we have $V = 0$ at $x = 0$, and hence $D = 0$. When $x < X_1$ on the *P*-side, the potential is constant, so that $\frac{dV}{dx} = 0$ at $x = X_1$. Hence,

$$C = -\frac{qN_A}{\varepsilon_o \varepsilon_r} \cdot X_1$$

Therefore,
$$V = \frac{qN_A x^2}{2\varepsilon_o \varepsilon_r} - \frac{qN_A}{\varepsilon_o \varepsilon_r} \cdot X_1 \cdot x$$

i.e.
$$V = \frac{qN_A}{\varepsilon_o \varepsilon_r}\left(\frac{x^2}{2} - X_1 \cdot x\right)$$

As $V = V_1$ at $x = X_1$, we have

$$V_1 = -\frac{qN_A}{2\varepsilon_o \varepsilon_r} \cdot X_1^2$$

If we apply the same procedure to the *N*-side, we get

$$V_2 = \frac{qN_D}{2\varepsilon_o \varepsilon_r} \cdot X_2^2$$

Therefore, the total built-in potential or the contact potential is V_o, where

$$V_o = V_2 - V_1 = \frac{q}{2\varepsilon_o \varepsilon_r}\left(N_A X_1^2 + N_D X_2^2\right)$$

We know the fact that the positive charge on the *N*-side must be equal in magnitude to the negative charge on the *P*-side for the neutral specimen. Hence,

$$N_A X_1 = -N_D X_2$$

and substituting this relationship in the above equation and using the fact that X_1 is a negative quantity, we get

$$X_1 = -\left[\frac{2\varepsilon_o \varepsilon_r V_o}{qN_A\left(1 + \frac{N_A}{N_D}\right)}\right]^{1/2}$$

Similarly,

$$X_2 = \left[\frac{2\varepsilon_o \varepsilon_r V_o}{qN_D\left(1 + \frac{N_D}{N_A}\right)}\right]^{1/2}$$

The total depletion width, $W = X_2 - X_1$ and hence, $W^2 = X_1^2 + X_2^2 - 2X_1 X_2$, and then substituting for X_1 and X_2 from the above equations, we find

$$W = \left[\frac{2\varepsilon_o \varepsilon_r V_o}{q}\left(\frac{N_A + N_D}{N_A N_D}\right)\right]^{1/2}$$

Here, in an *alloy junction*, the depletion width W is proportional to $(V_o)^{1/2}$.

In a *Grown Junction,* the charge density (ρ) varies linearly with distance (x) as shown in Fig. 1.12(c)(ii). If a similar analysis is carried for this junction, it is found that W varies as $(V_o)^{1/3}$ instead of $(V_o)^{1/2}$.

1.17.2 Under Forward Bias Condition

[AU May 2017, Dec 2016, June 2016, Dec 2015, Nov 2014 and May 2013, 8 marks]

When positive terminal of the battery is connected to the *P*-type and negative terminal to the *N*-type of the *PN* junction diode, the bias applied is known as forward bias.

Operation As shown in Fig. 1.13, the applied potential with external battery acts in opposition to the internal potential barrier and disturbs the equilibrium. As soon as equilibrium is disturbed by the application of an external voltage, the Fermi level is no longer continuous across the junction. Under the forward bias condition, the applied positive potential repels the holes in *P*-type region so that the holes move towards the junction and the applied negative potential repels the electrons in the *N*-type region and the electrons move towards the junction. Eventually, when the applied potential is more than the internal barrier potential, the depletion region and internal potential barrier disappear.

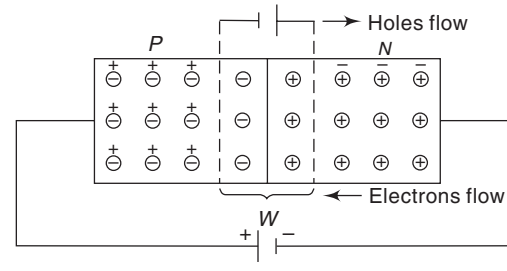


Fig. 1.13 *PN junction under forward bias*

V–I Characteristics of a diode under forward bias Under forward bias condition, the *V–I* characteristics of a *PN* unction diode are shown in Fig. 1.14. As the forward voltage (V_F) is increased, for $V_F < V_O$, the forward current I_F is almost zero (region OA) because the potential barrier prevents the holes from *P*-region and electrons from *N*-region to flow across the depletion region in the opposite direction.

For $V_F > V_O$, the potential barrier at the junction completely disappears and hence, the holes cross the junction from *P*-type to *N*-type and the electrons cross the junction in the opposite direction, resulting in relatively large current flow in the external circuit.

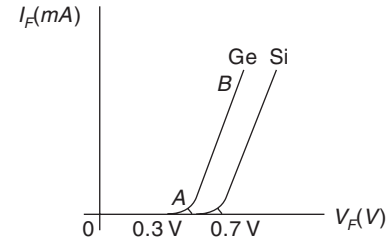


Fig. 1.14 *V-I characteristics of a diode under forward bias condition*

A feature worth to be noted in the forward characteristics shown in Fig. 1.14 is the cut in or threshold voltage (V_r) below which the current is very small. It is 0.3 V and 0.7 V for germanium and silicon, respectively. At the cut in voltage, the potential barrier is overcome and the current through the junction starts to increase rapidly.

1.17.3 Under Reverse Bias Condition

When the negative terminal of the battery is connected to the *P*-type and positive terminal of the battery is connected to the *N*-type of the *PN* junction, the bias applied is known as reverse bias.

Operation Under applied reverse bias as shown in Fig. 1.15, holes which form the majority carriers of the *P*-side move towards the negative terminal of the battery and electrons which form the majority

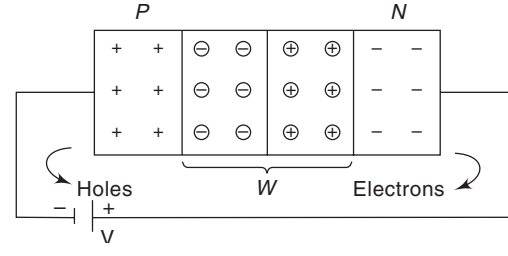


Fig. 1.15 *PN junction under reverse bias*

carrier of the *N*-side are attracted towards the positive terminal of the battery. Hence, the width of the depletion region which is depleted of mobile charge carriers increases. Thus, the electric field produced by applied reverse bias, is in the same direction as the electric field of the potential barrier. Hence, the resultant potential barrier is increased which prevents the flow of majority carriers in both directions; the depletion width, *W*, is proportional to $\sqrt{V_o}$ under reverse bias. Therefore, theoretically no current should flow in the external circuit. But in practice, a very small current of the order of a few microamperes flows under reverse bias as shown in Fig. 1.16. Electrons forming covalent bonds of the semiconductor atoms in the *P*- and *N*-type regions may absorb sufficient energy from heat and light to cause breaking of some covalent bonds. Hence electron-hole pairs are continually produced in both the regions. Under the reverse bias condition, the thermally generated holes in the *P*-region are attracted towards the negative terminal of the battery and the electrons in the *N*-region are attracted towards the positive terminal of the battery. Consequently, the minority carriers, electrons in the *P*-region and holes in the *N*-region, wander over to the junction and flow towards their majority carrier side giving rise to a small reverse current. This current is known as *reverse saturation current, I_o*. The magnitude of reverse saturation current mainly depends upon junction temperature because the major source of minority carriers is thermally broken covalent bonds.

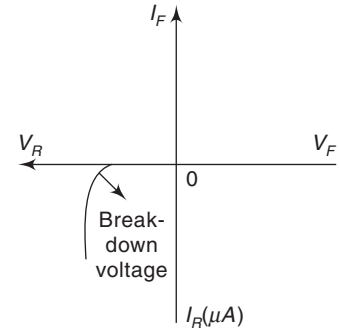


Fig. 1.16 *V-I characteristics under reverse bias*

For large applied reverse bias, the free electrons from the *N*-type moving towards the positive terminal of the battery acquire sufficient energy to move with high velocity to dislodge valence electrons from semiconductor atoms in the crystal. These newly liberated electrons, in turn, acquire sufficient energy to dislodge other parent electrons. Thus, a large number of free electrons are formed which is commonly called as an avalanche of free electrons. This leads to the breakdown of the junction leading to very large reverse current. The reverse voltage at which the junction breakdown occurs is known as *Breakdown Voltage*, V_{BD}.

1.17.4 *PN* Junction as a Diode

[AU April 2015 and June 2011, 8 marks]

Figure 1.17 shows the current-voltage characteristics of *PN* junction. The characteristics of the *PN* junction vary enormously depending upon the polarity of the applied voltage. For a forward-bias voltage, the current increases exponentially with the increase of voltage. A small change in the forward-bias voltage increases the corresponding forward-bias current by orders of magnitude and hence the forward-bias *PN* junction will have a very small resistance. The level of current flowing across a forward-biased *PN* junction largely depends upon the junction area. In the reverse-bias direction, the current remains small, i.e., almost zero, irrespective of the magnitude of the applied voltage and hence the reverse-bias *PN* junction will have a high resistance. The reverse bias current depends on the area, temperature and type of semiconductor material.

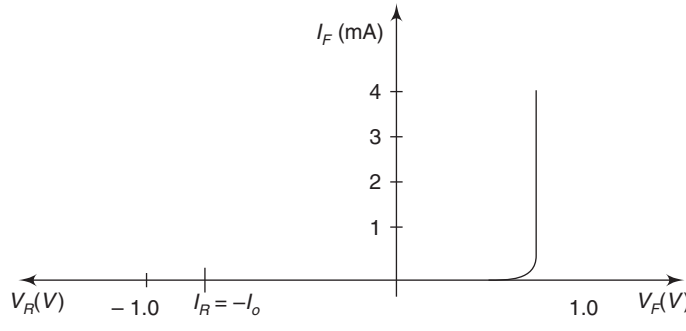


Fig. 1.17 *Ideal I-V characteristics of a PN junction diode*

The semiconductor device that displays these *I-V* characteristics is called a *PN* junction diode. Figure 1.18 shows the *PN* junction diode with forward-bias and reverse-bias and their circuit symbols. The metal contacts are indicated with which the homogeneous *P*-type and *N*-type materials are provided. Thus two metal-semiconductor junctions, one at each end of the diode, are introduced. The contact potential across these junctions is approximately independent of the direction and magnitude of the current. A contact of this type is called an *ohmic contact*, which has low resistance. In the forward-bias, a relatively large current is produced by a fairly small applied voltage. In the reverse-bias, only a very small current, ranging from nanoamps to microamps is produced. The diode can be used as a voltage controlled switch, i.e., OFF for a reverse-bias voltage and ON for a forward-bias voltage.

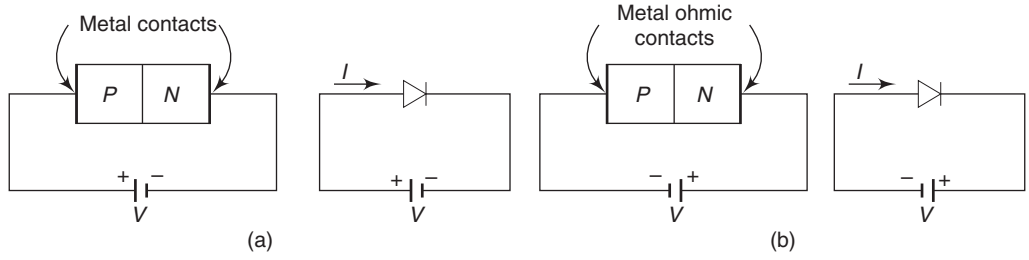


Fig. 1.18 *(a) Forward-biased PN junction diode and its circuit symbol, (b) Reverse-biased PN junction diode and its circuit symbol*

When a diode is reverse-biased by atleast 0.1V, the diode current is $I_R = -I_o$. As the current is in the reverse direction and is a constant, it is called the diode *reverse saturation current.* Real diodes exhibit reverse-bias current that are considerably larger than I_o. This additional current is called a *generation current* which is due to electrons and holes being generated within the space-charge region. A typical value of I_o may be 10^{-14} A and a typical value of reverse-bias current may be 10^{-9} A.

1.17.5 *PN* junction Diode as Rectifier

A *PN* junction diode is a two terminal device that is polarity sensitive. When the diode is forward biased, the diode conducts and allows current to flow through it without any resistance, i.e., the diode is ON. When the diode is reverse biased, the diode does not conduct and no current flows through it, i.e., the diode is OFF, or providing a blocking function. Thus an ideal diode acts as a switch, either open or closed, depending upon the polarity of the voltage placed across it. The ideal diode has zero resistance under forward bias and infinite resistance under reverse bias.

1.17.6 Diode Ratings or Limiting Values of *PN* Junction Diode

The *PN* junction diode will perform satisfactorily only if it is operated within certain limiting values. They are the following:

(a) Maximum Forward Current It is the highest instantaneous current under forward bias condition that can flow through the junction.

(b) Peak Inverse Voltage (PIV) It is the maximum reverse voltage that can be applied to the *PN* junction. If the voltage across the junction exceeds PIV, under reverse bias condition, the junction gets damaged.

(c) Maximum Power Rating It is the maximum power that can be dissipated at the junction without damaging the junction. Power dissipation is the product of voltage across the junction and current through the junction.

(d) Maximum Average Forward Current It is usually given at a special temperature, usually 25°C, (77°F) and refers to the maximum amount of average current that can be permitted to flow in the forward direction. If this rating exceeds its limit, then the structure breakdown can occur.

(e) Repetitive Peak Forward Current It is the maximum peak current that can be permitted to flow in the forward direction in the form of recurring pulses. The limiting value of this current is 450 mA.

(f) Maximum Surge Current It is the maximum current permitted to flow in the forward direction in the form of nonrecurring pulses. The current should not equal this value for more than a few milliseconds.

The above diode rating are subject to change with temperature variations. If the operating temperature is more than that stated for the rating, then the ratings must be decreased.

1.18 ENERGY-BAND STRUCTURE OF OPEN-CIRCUITED *PN* JUNCTION

[AU May 2012, 6 marks]

Consider that a *PN* junction has *P*-type and *N*-type materials in close physical contact at the junction on an atomic scale. Hence, the energy band diagrams of these two regions undergo relative shift to equalise the Fermi level. The Fermi level E_F should be constant throughout the specimen at equilibrium. The distribution of electrons or holes in allowed energy states is dependent on the position of the Fermi level. If this is not so, electrons on one side of the junction would have an average energy higher than those on the other side, and this causes transfer of electrons and energy until the Fermi levels on the two sides get equalised. However, such a shift does not disturb the relative position of the conduction band, valence band and Fermi level in any region. Equalisation of Fermi levels in the *P* and *N* materials of a *PN* junction is similar to equalisation of levels of water in two containers on being joined together.

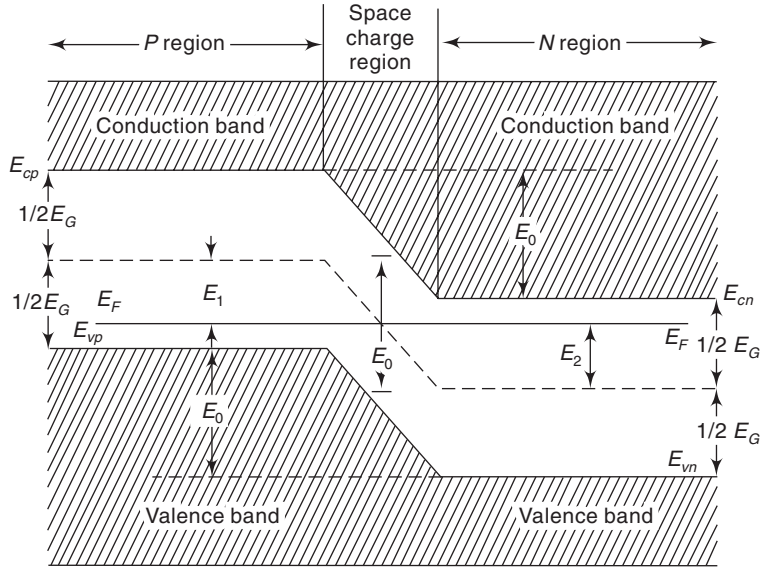


Fig. 1.19 *Energy-band structure*

The energy band diagram for a *PN* junction is shown in Fig. 1.19, where the Fermi level E_F is closer to the conduction band edge E_{cn} in the *N*-type material while it is closer to the valence band edge E_{vp} in the *P*-type material. It is clear that the conduction band edge E_{cp} in the *P*-type material is higher than the conduction band edge E_{cn} in the *N*-type material. Similarly, the valence band edge E_{vp} in the *P*-type material is higher than the valence band edge E_{vn} in the *N*-type material. As illustrated in Fig. 1.19, E_1 and E_2 indicate the shifts in the Fermi level from the intrinsic conditions in the *P* and *N* materials respectively. Then the total shift in the energy level E_0 is given by

$$E_0 = E_1 + E_2 = E_{cp} - E_{cn} = E_{vp} - E_{vn} \tag{1.17}$$

This energy E_0 (in eV) is the potential energy of the electrons at the *PN* junction, and is equal to qV_0, where V_0 is the contact potential (in volts) or contact difference of potential or the barrier potential.

Contact difference of potential A contact difference of potential exists across an open circuited *PN* junction. We now proceed to obtain an expression for E_0. From Fig. 1.19, we find that

$$E_F - E_{vp} = \frac{1}{2} E_G - E_1 \tag{1.18}$$

$$E_{cn} - E_F = \frac{1}{2} E_G - E_2 \tag{1.19}$$

Combining Eqs (1.18) and (1.19), we get

$$E_0 = E_1 + E_2 = E_G - (E_{cn} - E_F) - (E_F - E_{vp}) \tag{1.20}$$

We know that

$$np = N_C N_V e^{-E_G/kT}$$

and

$$np = n_i^2 \text{ (Mass-action law)}$$

From the above equations, we get

$$E_G = kT \ln \frac{N_C N_V}{n_i^2} \tag{1.21}$$

We know that for *N*-type material $E_F = E_C - kT \ln \frac{N_C}{N_D}$. Therefore, from this equation, we get

$$E_{cn} - E_F = kT \ln \frac{N_C}{n_n} = kT \ln \frac{N_C}{N_D} \tag{1.22}$$

Similarly for *P*-type material $E_F = E_V + kT \ln \frac{N_V}{N_A}$. Therefore, from this equation, we get

$$E_F - E_{vp} = kT \ln \frac{N_V}{p_p} = kT \ln \frac{N_V}{N_A} \tag{1.23}$$

Substituting from Eqs (1.21), (1.22) and (1.23) into Eq. (1.20), we get

$$E_0 = kT \left[\ln \frac{N_C N_V}{n_i^2} - \ln \frac{N_C}{N_D} - \ln \frac{N_V}{N_A} \right]$$

$$= kT \ln \left[\frac{N_C N_V}{n_i^2} \times \frac{N_D}{N_C} \times \frac{N_A}{N_V} \right]$$

$$= kT \ln \frac{N_D N_A}{n_i^2} \tag{1.24}$$

As $E_0 = qV_o$, then the contact difference of potential or barrier voltage is given by

$$V_o = \frac{kT}{q} \ln \frac{N_D N_A}{n_i^2} \tag{1.25}$$

In the above equations, E_0 in electron volts and k is in electron volt per degree Kelvin. The contact difference of potential V_o is expressed in volt and is numerically equal to E_0. From Eq. (1.24), we note that E_o (hence V_o) depends upon the equilibrium concentrations and not on the charge density in the transition region.

An alternative expression for E_0 may be obtained by substituting the equations of $n_n \approx N_D$, $p_n = \frac{n_i^2}{N_D}$, $n_n p_p = n_i^2$, $p_p \approx N_A$ and $n_p = \frac{n_i^2}{N_A}$ into Eq. (1.24). Then we get

$$E_0 = kT \ln \frac{p_{po}}{p_{no}} = kT \ln \frac{n_{no}}{n_{po}} \tag{1.26}$$

where subscript 0 represents the thermal equilibrium condition.

EXAMPLE 1.13

(a) The resistivities of the *P*-region and *N*-region of a germanium diode are 6 Ω-cm and 4 Ω-cm, respectively. Calculate the contact potential V_o and potential energy barrier E_o. (b) If the doping densities of both *P* and *N*-regions are doubled, determine V_o and E_o. Given that $q = 1.602 \times 10^{-19}$ C, $n_i = 2.5 \times 10^{13}/\text{cm}^3$, $\mu_p = 1800$ cm^2/V-s, $\mu_n = 3800$ cm^2/V-s and $V_T = 0.026$ V at 300 K.

Solution (a) Resistivity, $\rho = \frac{1}{\sigma} = \frac{1}{N_A q \mu_p} = 6\ \Omega\text{-cm}$

Therefore, $N_A = \frac{1}{6q\mu_p} = \frac{1}{6 \times 1.602 \times 10^{-19} \times 1800} = 0.579 \times 10^{15}/\text{cm}^3$

Similarly, $N_D = \frac{1}{4q\mu_n} = \frac{1}{4 \times 1.602 \times 10^{-19} \times 3800} = 0.411 \times 10^{15}/\text{cm}^3$

Therefore, $V_0 = V_T \ln \frac{N_D N_A}{n_i^2} = 0.026 \ln \frac{0.579 \times 0.411 \times 10^{30}}{(2.5 \times 10^{13})^2} = 0.1545$ V

Hence $E_0 = 0.1545$ eV

(b) $V_0 = 0.026 \ln \frac{2 \times 0.579 \times 10^{15} \times 2 \times 0.411 \times 10^{15}}{(2.5 \times 10^{13})^2} = 0.1906$ V

Therefore, $E_0 = 0.1906$ eV

EXAMPLE 1.14

Consider a germanium *PN* junction at 300 K with doping concentration $N_A = 1.5 \times 10^{15}$ cm^{-3} in the *P* and *N* sides of the junction respectively. Assuming the intrinsic carrier concentration of germanium $n_i = 2.5 \times 10^{13}$ cm^{-3} at 300 K, determine the contact potential across the junction. (AU May/June 2012)

Solution The potential energy E_0 is given by

$$E_0 = kT \ln\left[\frac{N_A N_D}{n_i^2}\right]$$

$$= (0.0259\,\text{eV}) \ln\left[\frac{1.5\times10^{18}\times2\times10^{15}}{6.25\times10^{26}}\right]$$

$$= 0.398 \text{ eV}$$

Since $E_0 = q \cdot V_0$, the contact potential across the *PN* junction is

$$V_0 = \frac{E_0}{q} = \frac{0.398\times1.602\times10^{-19}}{1.602\times10^{-19}} = 0.398 \text{ V}$$

EXAMPLE 1.15

Calculate the built in potential barrier in a *PN* junction. Consider a silicon *PN* junction at 300 K with doping densities $N_A = 1 \times 10^{18}$ cm^{-3} and $N_D = 1 \times 10^{15}$ cm^{-3}. Assume $n_i = 1.5 \times 10^{10}$ cm^{-3}. (AU Nov/Dec 2014)

Solution Given $T = 300$K, $N_A = 1 \times 10^{18}$ cm^{-3}, $N_D = 1 \times 10^{15}$ cm^{-3} and $n_i = 1.5 \times 10^{10}$ cm^{-3}.

The built in potential in a *PN* junction is given by

$$V_0 = \frac{kT}{q} \ln\left(\frac{N_A N_D}{n_i^2}\right), \text{ where } \frac{kT}{q} = 0.026 \text{ V}$$

Substituting the given values in the above equation, we get

$$V_0 = 0.026 \ln\left[\frac{1\times10^{18}\times1\times10^{15}}{(1.5\times10^{10})^2}\right] = 0.3505 \text{ V}$$

EXAMPLE 1.16

Consider a Si *PN* junction at $T = 300$K with doping concentrations of $N_A = 10^{16}$ cm^{-3} and $N_D = 10^{15}$ cm^{-3}. Assume that $m = 1.5 \times 10^{10}$ cm^{-3}. Calculate width of the space charge region in a *PN* junction, when a reverse bias voltage $V_R = 5$ V is applied. (AU Nov/Dec 2014)

Solution Given $N_A = 10^{16}$ cm^{-3}, $N_D = 10^{15}$ cm^{-3}, $m = 1.5 \times 10^{10}$ cm^{-3} and $V_R = 5$ V.

The width of space charge region for *PN* junction is given by

$$W = \left[\frac{2\varepsilon_o\varepsilon_r(V_o - V_R)}{q}\left(\frac{N_A + N_D}{N_A N_D}\right)\right]^{\frac{1}{2}}$$

Substituting the given values, we get

$$W = \left[\frac{2\times8.854\times10^{-12}\times12\times(0.7-(-5))}{1.602\times10^{-19}}\left(\frac{10^{16}+10^{15}}{10^{16}\times10^{15}}\right)\right]^{\frac{1}{2}} = 0.8\ \mu\text{m}$$

1.19 QUANTITATIVE THEORY OF *PN* DIODE CURRENTS

[AU Dec 2016, Dec 2015, May 2014, Nov 2012, Nov 2011, June 2011 and June 2010, 16 marks]

Let us now derive the expression for the total current as a function of applied voltage assuming that the width of the depletion region is zero. When a forward bias is applied to a diode, holes are injected from the *P*-side into the *N*-side. Due to this, the concentration of holes in the *N*-side (p_n) is increased from its thermal equilibrium value (p_{no}) and injected hole concentration [$P_n(x)$] decreases exponentially with respect to distance (x).

$$P_n(x) = p_n - p_{no} = P_n(0)\, e^{-x/L_p}$$

where L_p is the diffusion length for holes in the *N*-material.

$$p_n(x) = p_{no} + P_n(0)\, e^{-x/L_p} \tag{1.27}$$

Injected hole concentration at $x = 0$ is

$$P_n(0) = p_n(0) - p_{no} \tag{1.28}$$

These several components of hole concentration in the *N*-side of a forward biased diode are shown in Fig. 1.10, in which the density $p_n(x)$ decreases exponentially with distance (x).

Let p_p and p_n be the hole concentration at the edges of the space charge in the *P*- and *N*-sides, respectively. Let V_B (= $V_o - V$) be the effective barrier potential across the depletion layer. Then

$$p_p = P_n\, e^{V_B/V_T} \tag{1.29}$$

where V_T is the volt-equivalent of temperature.

This is the Boltzmann's relation of kinetic gas theory. This equation is valid as long as the hole current is small compared with diffusion or drift current. This condition is called low level injection.

Under open circuit condition (i.e., $V = 0$), $p_p = p_{po}$, $p_n = p_{no}$ and $V_B = V_o$. Equation (1.23) can be changed into

$$p_{po} = p_{no}\, e^{V_o/V_T} \tag{1.30}$$

Under forward bias condition let V be the applied voltage, then the effective barrier voltage

$$V_B = V_o - V$$

The hole concentration throughout the *P*-side is constant and equal to the thermal equilibrium value ($p_p = p_{po}$). The hole concentration varies exponentially with distance into the *N*-side.

At $x = 0$, $\quad p_n = p_n(0)$

Equation (1.29) can be changed into

$$p_{po} = p_n(0)\, e^{(V_o - V)/V_T} \tag{1.31}$$

Comparing Eqs (1.30) and (1.31),

$$p_n(0) = p_{no}\, e^{V/V_T}$$

This boundary condition is called the *law of the junction*. Substituting this into Eq. (1.28), we get

$$P_n(0) = p_{no}\left(e^{V/V_T} - 1\right) \tag{1.32}$$

The diffusion hole current in the *N*-side is

$$I_{pn}(x) = -Aq\,D_p\frac{dp_n(x)}{dx}$$

$$= -Aq\,D_p\frac{d}{dx}\left[p_{no} + P_n(0)\,e^{-x/L_p}\right]$$

$$= \frac{AqD_p\,P_n(0)}{L_p}\,e^{-x/L_p}$$

From this equation, it is evident that the injected hole current decreases exponentially with distance.

Forward currents The hole current crossing the junction into the *N*-side with $x = 0$ is

$$I_{pn}(0) = \frac{AqD_p\,P_n(0)}{L_p} = \frac{AqD_p\,p_{no}}{L_p}\left(e^{V/V_T} - 1\right)$$

The electron current crossing the junction into the *P*-side with $x = 0$ is

$$I_{np}(0) = \frac{AqD_n\,N_p(0)}{L_n} = \frac{AqD_n\,n_{po}}{L_n}\left(e^{V/V_T} - 1\right)$$

The total diode current,

$$I = I_{pn}(0) + I_{np}(0) = I_o\left(e^{V/V_T} - 1\right)$$

where $I_o = \dfrac{Aq\,D_p\,p_{no}}{L_p} + \dfrac{AqD_n\,n_{po}}{L_n}$ = reverse saturation current.

If we consider carrier generation and recombination in the space-charge region, the general equation of the diode current is approximately given by

$$I = I_o\left[e^{(V/\eta\,V_T)} - 1\right]$$

where V = external voltage applied to the diode, and η = a constant, 1 for germanium and 2 for silicon.

Reverse saturation currents We know that $p_n = \dfrac{n_i^2}{N_D}$ and $n_p = \dfrac{n_i^2}{N_A}$. Applying these relationships in the above equation of reverse saturation current, I_o, we get

$$I_o = Aq\left[\frac{D_p}{L_p\,N_D} + \frac{D_n}{L_n\,N_A}\right]n_i^2$$

where $n_i^2 = A_o\,T^3\,e^{-E_{G0}/kT} = A_o\,T^3\,e^{-V_{G0}/V_T}$, where V_{Go} is a voltage which is numerically equal to the forbidden gap energy E_{Go} in electron volts.

For a germanium diode, the diffusion constants D_p and D_n vary approximately inversely proportional to T. Hence, the temperature dependence of I_o is

$$I_o = K_1\,T^2\,e^{\frac{-V_{Go}}{V_T}}$$

where K_1 is a constant independent of temperature.

For a silicon diode, I_o is proportional to n_i instead of n_i^2. Hence,

$$I_o = K_2 T^{\frac{3}{2}} e^{\frac{-V_{Go}}{2V_T}}$$

where K_2 is a constant independent of temperature.

1.20 DIODE CURRENT EQUATION

The diode current equation relating the voltage V and current I is given by

$$I = I_o \left[e^{(V/\eta V_T)} - 1\right]$$

where I = diode current
I_o = diode reverse saturation current at room temperature
V = external voltage applied to the diode
η = a constant, 1 for germanium and 2 for silicon
V_T = $kT/q = T/11600$, volt-equivalent of temperature, i.e., thermal voltage

where k = Boltzmann's constant (1.38×10^{-3} J/K)
q = charge of the electron (1.602×10^{-19} C)
T = temperature of the diode junction (K) = (°C + 273°)

At room temperature, (T = 300 K), V_T = 26 mV. Substituting this value in the current equation, we get

$$I = I_o[e^{(40\ V/\eta)} - 1]$$

Therefore, for germanium diode, $I = I_o\ [e^{40V} - 1]$, since $\eta = 1$ for germanium. For silicon diode, $I = I_o[e^{20V} - 1]$, since $\eta = 2$ for silicon.

If the value of applied voltage is greater than unity, then the equation of diode current for germanium,

$$I = I_o\ (e^{40V})$$

and for silicon,

$$I = I_o\ (e^{20V})$$

When the diode is reverse biased, its current equation may be obtained by changing the sign of the applied voltage V. Thus, the diode current with reverse bias is

$$I = I_o \left[e^{(-V/\eta\ V_T)} - 1\right]$$

If $V >> V_T$, then the term $e^{(-V/\eta\ V_T)} << 1$, therefore $I \approx -I_o$, termed as reverse saturation current, which is valid as long as the external voltage is below the breakdown value.

EXAMPLE 1.17

When a reverse bias is applied to a germanium *PN* junction diode, the reverse saturation current at room temperature is 0.3 μA. Determine the current flowing in the diode when 0.15 V forward bias is applied at room temperature.

Solution Given $I_o = 0.3 \times 10^{-6}$ A and $V_F = 0.15$ V.

The current flowing through the *PN* diode under forward bias is

$$I = I_o\left(e^{40\ V_F} - 1\right)$$

$$= 0.3 \times 10^{-6} (e^{40 \times 0.15} - 1)$$

$$= 120.73\ \mu A$$

EXAMPLE 1.18

The reverse saturation current of a silicon *PN* junction diode is 10 μA. Calculate the diode current for the forward-bias voltage of 0.6 V at 25 °C. (AU May/June 2016)

Solution Given $V_F = 0.6$ V, $T = 273 + 25 = 298$ K

$I_o = 10\ \mu A = 1 \times 10^{-5}$ A and $\eta = 2$ for silicon.

The volt-equivalent of the temperature (T) is

$$V_T = \frac{T}{11{,}600} = \frac{298}{11{,}600} = 25.7 \times 10^{-3}\ \text{V}$$

Therefore, the diode current, $I = I_o \left(e^{\frac{V_F}{\eta V_T} - 1} \right)$

$$I = 10^{-5} \left(e^{\frac{0.6}{2 \times 25.7 \times 10^{-3}}} - 1 \right) = 1.174\ \text{A}$$

EXAMPLE 1.19

The diode current is 0.6 mA when the applied voltage is 400 mV, and 20 mA when the applied voltage is 500 mV. Determine η. Assume $\frac{kT}{q} = 25$ mV. (AU May/June 2016)

Solution

The diode current, $I = I_o \left(e^{\frac{qV}{\eta kT}} - 1 \right)$

Therefore, $0.6 \times 10^{-3} = I_o \left(e^{\frac{qV}{\eta kT}} - 1 \right) = I_o\, e^{\frac{qV}{\eta kT}}$

$$= I_o \cdot e^{\frac{400}{25\eta}} = I_o \cdot e^{\frac{16}{\eta}} \quad (1)$$

Also, $20 \times 10^{-3} = I_o \cdot e^{\frac{500}{25\eta}} = I_o \cdot e^{\frac{20}{\eta}}$ (2)

Dividing Eq. (2) by Eq. (1), we get

$$\frac{20 \times 10^{-3}}{0.6 \times 10^{-3}} = \frac{I_o \cdot e^{\frac{20}{\eta}}}{I_o \cdot e^{\frac{16}{\eta}}}$$

Therefore, $\frac{100}{3} = e^{\frac{4}{\eta}}$

Taking natural logarithms on both sides, we get

$$\log_e \frac{100}{3} = \frac{4}{\eta}$$

$$3.507 = \frac{4}{\eta}$$

Therefore, $$\eta = \frac{4}{3.507} = 1.14$$

EXAMPLE 1.20

Find the voltage at which the reverse current in a germanium *PN* junction diode attains a value of 90% of its saturation value at room temperature.

Solution We know that the current of a *PN* junction diode is

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

Therefore, $$-0.90\, I_o = I_o\left(e^{\frac{V}{V_T}} - 1\right)$$

where $$V_T = \frac{T}{11{,}600} = 26 \text{ mV}$$

$$-0.9 = \left(e^{\frac{V}{0.026}} - 1\right)$$

$$0.1 = e^{\frac{V}{0.026}}$$

Therefore, $$V = -0.06 \text{ V}$$

EXAMPLE 1.21

Determine the ideal reverse saturation current density in a silicon *PN* junction at $T = 300$ K. Consider the following parameters in the silicon *PN* junction:
$N_A = N_D = 10^{16}$ cm^{-3}, $n_i = 1.5 \times 10^{10}$ cm^{-3}, $D_n = 25$ cm^2/s, $T_{po} = T_{no} = 5 \times 10^{-7}$ s, $D_p = 10$ cm^2/s, $\varepsilon_r = 11.7$. Comment on the result. **(AU April/May 2015)**

Solution Given, $T = 300$ K, $N_A = N_D = 10^{16}$ cm^{-3}, $n_i = 1.5 \times 10^{10}$ cm^{-3}, $D_n = 25$ cm^2/s, $T_{po} = T_{no} = 5 \times 10^{-7}$ s, $D_p = 10$ cm^2/s

The reverse saturation current is given by

$$I_o = Aq\left[\frac{D_p}{L_p \cdot N_D} + \frac{D_n}{L_n \cdot N_A}\right] n_i^2$$

We know that, $J_o = \frac{I_o}{A}$

Therefore, reverse saturation current density is,

$$J_o = q\left[\frac{D_p}{L_p \cdot N_D} + \frac{D_n}{L_n \cdot N_A}\right] n_i^2$$

where $$L_p = \sqrt{D_p \cdot T_{po}} = \sqrt{10 \times 10^{-4} \times 5 \times 10^{-7}} = 2.236 \times 10^{-5}$$

and $$L_n = \sqrt{D_n \cdot T_{no}} = \sqrt{25 \times 10^{-4} \times 5 \times 10^{-7}} = 3.535 \times 10^{-5}$$

Hence, $$J_o = 1.602 \times 10^{-19}\left[\frac{10 \times 10^{-4}}{2.236 \times 10^{-5} \times 10^{22}} + \frac{25 \times 10^{-4}}{3.535 \times 10^{-5} \times 10^{22}}\right](1.5 \times 10^{-16})^2$$

$$= 0.416\ \mu A/s$$

Ideally, the reverse saturation current in the *PN* junction should be zero. But practically, due to the flow of minority charge carriers, there will be a leakage current in the order of *nA*.

1.21 IDEAL VERSUS PRACTICAL–RESISTANCE LEVELS (STATIC AND DYNAMIC)

An ideal diode should offer zero resistance in forward bias and infinite resistance in the reverse bias. But in practice no diode can act as an ideal diode, i.e., an actual diode does not behave as a perfect conductor when forward biased and as a perfect insulator when reverse biased. Let us consider four resistances of the diode (a) DC or static resistance, (b) AC or dynamic resistance, (c) average AC resistance and (d) reverse resistance.

DC or static resistance (R_F) It is defined as the ratio of the voltage to the current, V/I, in the forward bias characteristics of the *PN* junction diode. In the forward bias characteristics of the diode as shown in Fig. 1.20, the DC or static resistance (R_F) at the operating point can be determined by using the corresponding levels of voltage V and current I, i.e., $R_F = \frac{V}{I}$. Here, the DC resistance is independent of the shape of the characteristics in the region surrounding the point of interest. The DC resistance levels at the knee and below will be greater than the resistance levels obtained for the characteristics above the knee. Hence, the DC resistance will be low when the diode current is high. As the static resistance varies widely with V and I, it is not a useful parameter.

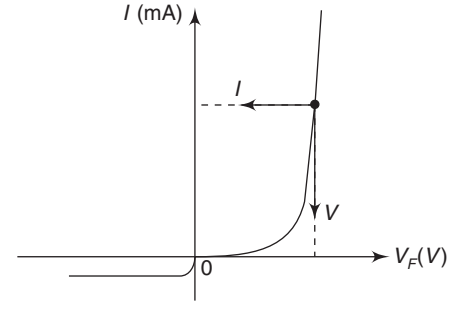


Fig. 1.20 *Forward biasing of a diode*

AC or dynamic resistance (r_f) It is defined as the reciprocal of the slope of the volt-ampere characteristics.

$$r_f = \frac{\text{change in voltage}}{\text{resulting change in current}} = \frac{\Delta V}{\Delta I}$$

A straight line drawn tangent to the curve through the Quiescent Point (*Q*-point) as shown in Fig. 1.21(a) will define a specific change in voltage and current which may be used to determine the AC or dynamic resistance for this region of the diode characteristics. As shown in Fig. 1.21(b), for a small change in voltage, there will be a corresponding change

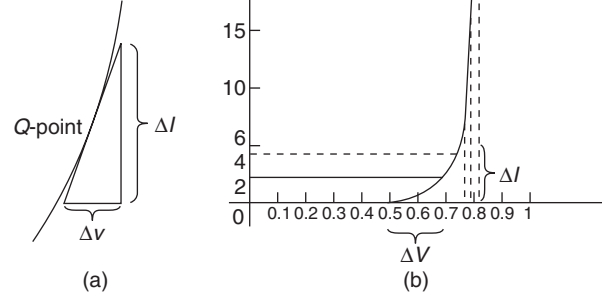


Fig. 1.21 *Dynamic resistance*

in current, which is equidistant to either side of the *Q*-point. Hence, the AC or dynamic resistance is determined as $r_f = \frac{\Delta V}{\Delta I}$.

The derivative of a function at a point is equal to the slope of the tangent line drawn at the point. The Schockley's equation for the forward and reverse bias regions is defined by

$$I = I_o (e^{V/\eta V_T} - 1)$$

Taking the derivative of the above equation w.r.t. the applied voltage, V, we get

$$\frac{dI}{dV} = \frac{d}{dV}[I_o (e^{V/\eta V_T} - 1)]$$

$$= I_o \left[\frac{1}{\eta V_T} \cdot e^{V/\eta V_T}\right]$$

$$= \frac{I_o e^{V/\eta V_T}}{\eta V_T}$$

$$= \frac{I + I_o}{\eta V_T}$$

Generally $I >> I_o$ in the vertical-slope section of the characteristics. Therefore,

$$\frac{dI}{dV} \cong \frac{I}{\eta V_T}$$

Therefore,

$$\frac{dV}{dI} = r_f = \frac{\eta V_T}{I}$$

The dynamic resistance varies inversely with current, i.e., $r_f = \frac{\eta V_T}{I}$, where $V_T = T/1600$, the volt equivalent of temperature (T) of the diode junction (K) and η is a constant whose value is equal to 1 for Germanium and 2 for Silicon diodes. At room temperature $V_T = 26$ mV.

AC resistance of a diode is the sum of bulk resistance r_b and junction resistance r_j. Bulk resistance (r_b) is the sum of ohmic resistance of the *P*- and *N*-type semiconductors.

Average AC resistance It is the resistance associated with the device for the region if the input signal is sufficiently large to produce a wide range of the characteristics as shown in Fig. 1.22. Therefore,

$$r_{\text{av}} = \left.\frac{\Delta V}{\Delta I}\right|_{\text{point to point}}$$

As with the DC and AC resistance levels, the lower the level of currents used to determine the average, the higher is the resistance level.

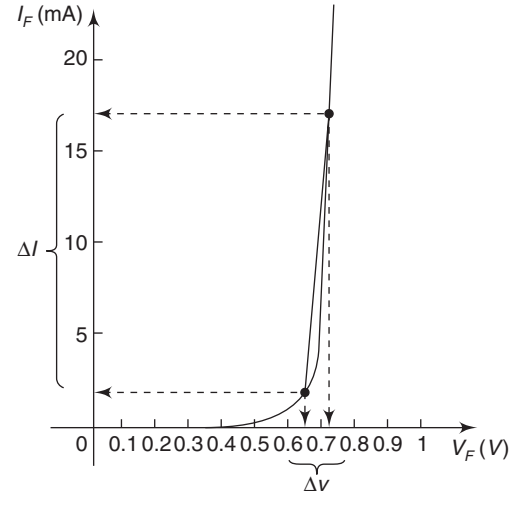


Fig. 1.22 *Average AC resistance*

Reverse resistance It is the resistance offered by the *PN* junction diode under reverse bias condition. It is very large compared to the forward resistance, which is in the range of several MΩ.

EXAMPLE 1.22

Determine the forward resistance of a *PN* junction diode when the forward current is 5 mA at $T = 300$ K. Assume silicon diode.

Solution Given, for a silicon diode, the forward current, $I = 5$ mA, $T = 300$ K.

Forward resistance of a *PN* junction diode, $r_f = \dfrac{\eta V_T}{I}$, where $V_T = \dfrac{T}{11{,}600}$ and $\eta = 2$ for silicon

Therefore, $$r_f = \frac{2 \times \dfrac{T}{11{,}600}}{5 \times 10^{-3}} = \frac{2 \times 300}{11{,}600 \times 5 \times 10^{-3}} = 10.34\ \Omega$$

EXAMPLE 1.23

Find the value of DC resistance and AC resistance of a germanium junction diode at 25°C with $I_o = 25$ μA and at an applied voltage of 0.2 V across the diode.

Solution Given $I_o = 25\ \mu A$, $T = 25°C = 298$ K and $V = 0.2$ volt.

$$I = I_o\left(\frac{V}{e^{\eta V_T} - 1}\right) = 25 \times 10^{-6}\left(\frac{0.2}{e^{26 \times 10^{-3}} - 1}\right) = 54.79\ \text{mA}$$

DC resistance, $$R_F = \frac{V}{I} = \frac{0.2}{54.79 \times 10^{-3}} = 3.65\ \Omega$$

For germanium, $$\eta = 1,\ V_T = \frac{KT}{q} = 25.71\ \text{mV}$$

AC resistance, $$r_f = \frac{\eta V_T}{I} = \frac{25.71 \times 10^{-3}}{54.79 \times 10^{-3}} = 0.47\ \Omega$$

EXAMPLE 1.24

Calculate the dynamic forward and reverse resistance of a *PN* junction diode when the applied voltage is 0.25 V at $T = 300$ K given $I_o = 2$ μA.

Solution Given $V = 0.25$ V, $T = 300$ K, $I_o = 2\ \mu A$.

At $T = 300$ K, $V_T = 26$ mV.

Assuming it to be silicon diode, $\eta = 2$

Therefore, $$I = I_o\left(e^{\frac{V}{\eta V_T}} - 1\right) = 2 \times 10^{-6}\left(\frac{0.25}{e^{2 \times 26 \times 10^{-3}} - 1}\right) = 0.24\ \text{mA}$$

$$r_f = \frac{\eta V_T}{I} = \frac{2 \times 26 \times 10^{-3}}{0.24 \times 10^{-3}} = 216.67\ \Omega$$

For germanium diode, $\eta = 1$.

$$I = I_o\left(\frac{V}{e^{\eta V_T}-1}\right) = 2 \times 10^{-6}\left(\frac{0.25}{e^{26\times10^{-3}}-1}\right) = 0.03 \text{ A}$$

$$r_f = \frac{\eta V_T}{I} = \frac{26 \times 10^{-3}}{0.03} = 0.867\ \Omega$$

Reverse resistance $\frac{V}{I_o} = \frac{0.25}{2 \times 10^{-6}} = 125\ \text{k}\Omega$

EXAMPLE 1.25

A *PN*-junction diode has a reverse saturation current of 30 μA at a temperature, of 125°C. At the same temperature, find the dynamic resistance for 0.2 V bias in forward and reverse directions.

Solution Given, the reverse saturation current, $I_o = 30 \times 10^{-6}$ A and $V = 0.2$ V.

We know that the dynamic resistance $= \frac{\eta V_T}{I_o e^{V/\eta V_T}}$

Here, $\eta = 1$ for germanium and $V_T = \frac{T}{11{,}600} = \frac{125 + 273}{11{,}600} = 34.3$ mV

Therefore, forward dynamic resistance $r_f = \frac{34.3 \times 10^{-3}}{30 \times 10^{-6}\,(e^{0.2/34.3\times10^{-3}})} = 3.356\ \Omega$

Reverse dynamic resistance, $r_r = \frac{\eta V_T}{I_o e^{-V/\eta V_T}} = \frac{34.3 \times 10^{-3}}{30 \times 10^{-6}\,(e^{-0.2/34.3 \times 10^{-3}})} = 389.5\ \text{k}\Omega$

EXAMPLE 1.26

If two similar Germanium diodes are connected back to back and the voltage V is impressed upon, calculate the voltage across each diode and current through each diode. Assume similar value of I_o = 1 μA for both the diodes and η = 1.

Solution The arrangement is shown in the Fig. 1.23.

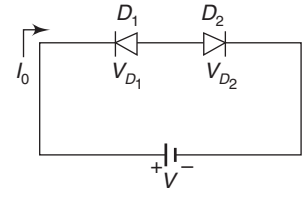


Fig. 1.23

As D_1 is reverse biased, the total current flowing in the circuit is $I_0 = 1\mu$A. The diode D_2 is forward biased and its forward current is equal to the reverse current $I_0 = 1$ μA, which can flow as D_1 is reverse biased.

For diode D_2, $I = I_0 = 1$ μA and voltage across D_2 is V_{D2}.

Therefore, $I = I_0\,[e^{V/\eta V_T} - 1]$

or $I_0 = I_0\,[e^{V_{D_2}/\eta V_T} - 1]$

Hence, $e^{V_{D_2}/\eta V_T} = 1 + 1 = 2$

$$\frac{V_{D2}}{\eta V_T} = \ln 2$$

$$V_{D_2} = \eta V_T \times \ln 2 = 1 \times 26 \times 10^{-3} \times 0.6931 = 0.01802 \text{ V}$$

Therefore, $V_{D_1} = V - V_{D_2} = V - 0.01802$ V

The current through the diodes D_1 and D_2 is $I = I_0 = 1$ μA

EXAMPLE 1.27

Determine the forward resistance of a *PN* junction diode, when the forward current is 5 mA at $T = 300$ K. Assume silicon diode.

Solution Given, for a silicon diode, the forward current, $I = 5$ mA. $T = 300$ K.

Forward resistance of a *PN* junction diode, $r_f = \dfrac{\eta V_T}{I}$ where $V_T = \dfrac{T}{11{,}600}$ and $\eta = 2$ for silicon

Therefore, $$r_f = \frac{2 \times \dfrac{T}{11{,}600}}{5 \times 10^{-3}} = \frac{2 \times 300}{11{,}600 \times 5 \times 10^{-3}} = 10.34\ \Omega$$

1.22 TRANSITION OR SPACE CHARGE (OR DEPLETION REGION) CAPACITANCE (C_T)

[AU Nov 2013 and Nov 2010, 16 marks]

Under reverse bias condition, the majority carriers move away from the junction, thereby uncovering more immobile charges. Hence the width of the space–charge layer at the junction increases with reverse voltage. This increase in uncovered charge with applied voltage may be considered a capacitive effect. The parallel layers of oppositely charged immobile ions on the two sides of the junction form the capacitance, C_T, which is expressed as

$$C_T = \left|\frac{dQ}{dV}\right|$$

where dQ is the increase in charge caused by a change in voltage dV. A change in voltage dV in a time dt will result in a current $I = dQ/dt$ given by

$$I = C_T \frac{dV}{dt}$$

Therefore C_T is important while considering a diode or a transistor as a circuit element. The quantity C_T is called the transition, space–charge, barrier or depletion region capacitance.

1.22.1 Step-graded Junction

A *PN* junction is formed from a single-crystal intrinsic semiconductor by doping part of it with acceptor impurities and the remaining with donors. A junction between *P*-type and *N*-type materials may be fabricated in a variety of ways.

The change in impurity concentration from *P*- to *N*-type semiconductor occurs in a very short length, typically much less than 1 μm. In an abrupt *PN* junction, there is a sudden step change from acceptor ions on one side to donor ions on the other side. Such a junction is fabricated by placing trivalent indium against *N*-type germanium and heating the combination to a high temperature for a short time. Since some of the indium dissolves into the germanium, the *N*-type germanium is changed into *P*-type at the junction. Such a step-graded junction is called an *alloy*, or *fussion junction*. A step-graded junction is also formed between emitter and base of an integrated transistor.

A diffused junction is graded in which case the donor and acceptor concentrations are functions of distance across the junction. Then the acceptor density, N_A, gradually decreases and the donor density, N_D, gradually increases till $N_A = N_D$ is reached. Therefore, N_D increases and N_A decreases to zero.

It is not necessary for the abrupt junction to be symmetrical, that is, the doping concentrations at either side of the junction are dissimilar.

As shown in Fig. 1.24, consider a *PN* diode which is asymmetrically doped at the junction. Since the net charge is zero, then $qN_AW_p = qN_DW_n$.

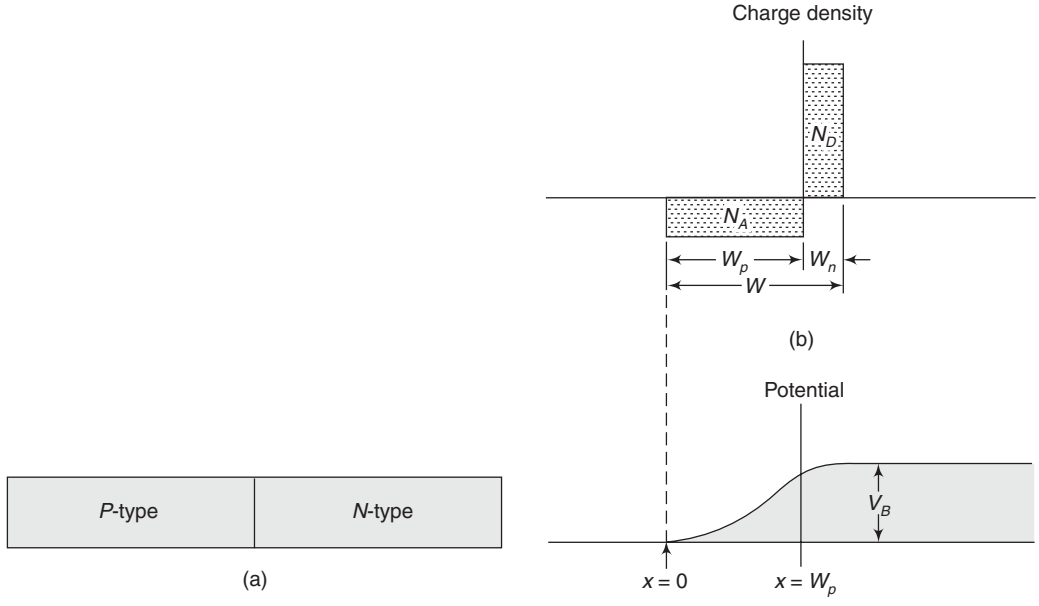


Fig. 1.24 *Charge density and potential variation at an alloy PN junction*

If $N_A >> N_D$, then $W_p << W_n \approx W$. The relationship between potential and charge density is given by the Poisson's equation, $\frac{d^2V}{dx^2} = \frac{qN_A}{\varepsilon}$.

Integrating the above equation twice, we get

$$\iint d^2V = \iint \frac{qN_A}{\varepsilon} dx^2$$

Therefore,
$$V = \frac{qN_Ax^2}{2\varepsilon}$$

At $x = W_p \approx W$, $V = V_B$, the barrier potential that appears across the uncovered acceptor ions. Thus

$$V_B = \frac{qN_A W^2}{2\varepsilon} \tag{1.33}$$

Here $V_B = V_0 - V$, where V is a negative number for an applied reverse bias and V_0 is the contact potential. Hence, the width of the depletion layer increases with applied reverse voltage, i.e., $V_B \propto W^2$. Therefore, $W \propto \sqrt{V_B}$.

The total charge density of a *P*-type material with area of the junction *A* is given by

$$Q = qN_A WA$$

Differentiating the above equation w.r.t. *V*, we get

$$C_T = \left|\frac{dQ}{dV}\right| = AqN_A \left|\frac{dW}{dV}\right| \tag{1.34}$$

Differentiating Eq. (1.33) w.r.t. *V*, we get

$$1 = \frac{qN_A 2W}{2\varepsilon}\left|\frac{dW}{dV}\right|$$

Therefore,
$$\left|\frac{dW}{dV}\right| = \frac{\varepsilon}{qN_A W} \tag{1.35}$$

Substituting Eq. (1.35) in Eq. (1.34), we get

$$C_T = \left|\frac{dQ}{dV}\right| = AqN_A \frac{\varepsilon}{qN_A W}$$

Therefore,
$$C_T = \frac{\varepsilon A}{W} \tag{1.36}$$

Here ε is the permittivity of the material, *A* the cross-sectional area of the junction and *W* is the width of the depletion layer over which the ions are uncovered. The depletion width, *W*, is given by

$$W = \left[\frac{2\varepsilon_o \varepsilon_r (V_o - V)}{q}\left(\frac{N_A + N_D}{N_A N_D}\right)\right]^{1/2} \tag{1.37}$$

where *V* is the applied voltage and V_o is the barrier potential, or the contact potential.

When no external voltage is applied, i.e., $V = 0$, the width of the depletion region of a *PN* junction diode is of the order of 0.5 microns. The movement of majority carriers across the junction causes opposite charges to be stored at this distance *W* apart. This depletion region acts as a dielectric between the two conducting *P*- and *N*-regions. Therefore, these regions act as a parallel plate capacitor whose transition capacitance C_T is approximately 20 pF with no external bias.

When forward bias $+V$ is applied, the effective barrier potential, $V_B = [V_o - (+V)]$, is lowered and hence the width of the depletion region *W* decreases and C_T increases. Under reverse bias condition, the majority carriers move away from the junction, thereby uncovering more immobile charges. Now the effective barrier potential, $V_B = [V_o - (-V)]$, is increased and hence, *W* increases with reverse voltage and C_T decreases correspondingly. The values of C_T range from 5 to 200 pF, the larger values being for the high power diodes. This property of voltage variable capacitance with the reverse bias appears in varactors, vari-caps or volta-caps.

1.23 DIFFUSION (OR STORAGE) CAPACITANCE (C_D)

[AU May 2012 and Nov 2010, 16 marks]

The capacitance that exists in a forward biased junction is called a diffusion or storage capacitance (C_D), whose value is usually much larger than C_T, which exists in a reverse-biased junction. This is

also defined as the rate of change of injected charge with applied voltage, i.e., $C_D = \frac{dQ}{dV}$, where dQ represents the change in the number of minority carriers stored outside the depletion region when a change in voltage across the diode, dV, is applied.

Calculation of C_D Let us assume that the P material in one side of the diode is heavily doped in comparison with the N side. Since the holes move from the P to the N side, the hole current $I \approx I_{pn}(0)$.

The excess minority charge Q existing on the N side is given by

$$Q = \int_0^\infty AqP_n(0)\, e^{-x/L_p}\, dx = \left[\frac{AqP_n(0)e^{-x/L_p}}{-1/L_p}\right]_0^\infty = L_p\, AqP_n(0)$$

Differentiating the above equation, we get

$$C_D = \frac{dQ}{dV} = AqL_p \frac{d[P_n(0)]}{dV} \tag{1.38}$$

We know that the diffusion hole current in the N-side is $I_{pn}(x) = AqD_pP_n(0)/L_p\, e^{-x/L_p}$. The hole current crossing the junction into the N-side with $x = 0$ is $I_{pn}(0) = \frac{AqD_p\, P_n(0)}{L_p}$.

Therefore,
$$I = \frac{AqD_p\, P_n(0)}{L_p}$$
$$P_n(0) = \frac{IL_p}{AqD_p}$$

Differentiating the above equation w.r.t. "V", we get

$$\frac{d[P_n(0)]}{dV} = \frac{dI}{dV}\frac{L_p}{AqD_p}$$

Upon substituting in Eq. (1.38), we have

$$C_D = \frac{dQ}{dV} = \frac{dI}{dV}\frac{L_p^2}{D_p}$$

Therefore, $C_D = g\tau$, where $g = \frac{dI}{dV}$ is the diode conductance and $\tau = \frac{L_p^2}{D_p}$ is the mean life time of holes in the N-region.

From diode current equation, $g = \frac{I}{\eta V_T}$.

Therefore,
$$C_D = \frac{\tau I}{\eta V_T}.$$

where τ is the mean life time for holes and electrons.

The diffusion capacitance C_D increases exponentially with forward bias or, alternatively, that it is proportional to diode forward current, I. The values of C_D range from 10 to 1000 pF, the larger values being associated with the diode carrying a larger anode current, I.

The effect of C_D is negligible for a reverse-biased PN junction. As the value of C_D is inversely proportional to frequency, it is high at low frequencies and it decreases with the increase in frequency.

1.24 TEMPERATURE DEPENDENCE OF *V-I* CHARACTERISTICS OF DIODES

[AU June 2013 and June 2011, 8 marks]

The reverse saturation current I_o is temperature dependent while voltage equivalent of temperature V_T is also temperature dependent. Hence, the diode current involving I_o and V_T is temperature dependent. The overall diode characteristics depends on the temperature.

The dependence of I_o on temperature T is given by

$$I_o = KT^m e^{-V_{Go}/\eta V_T} \tag{1.39}$$

where K = constant independent of temperature (not the Boltzmann's constant)
m = 2 for Ge and 1.5 for Si

and V_{Go} = forbidden energy gap = 0.785 V for Ge and 1.21 V for Si.

As temperature increases, the value of I_o increases and hence the diode current increases. To keep diode current constant, it is necessary to reduce the applied voltage V of the diode.

Let us calculate, the rate of change of the applied voltage to keep the diode current constant. For a constant diode current, $\frac{dI}{dT} = 0$. Hence, a change in voltage has to be calculated.

A diode current equation is given by

$$I = I_o (e^{V/\eta V_T} - 1)$$

Since $I >> I_o$ for a forward characteristics, we have

$$I = I_o e^{V/\eta V_T} \tag{1.40}$$

Substituting Eq. (1.39) into Eq. (1.40), we get

$$I = KT^m e^{-V_{GO}/\eta V_T} \cdot e^{V/\eta V_T}$$
$$= KT^m e^{(V-V_{GO})/\eta V_T} \tag{1.41}$$

Since $V_T = kT$, where k is Boltzmann's constant,

$$I = KT^m e^{(V-V_{GO})/\eta k_T}$$

For a constant diode current, $dI/dT = 0$. Hence, differentiating the above equation with respect to T, we get

$$\frac{dI}{dT} = K\left[mT^{m-1} e^{(V-V_{Go})/\eta kT} + T^m e^{(V-VGo)/\eta kT} \cdot \frac{d}{dT}\left(\frac{V-V_{Go}}{\eta kT}\right)\right]$$

$$= K e^{(V-V_{Go})/\eta kT}\left[mT^{m-1} + \frac{T^m}{\eta k}\left(\frac{T\frac{dV}{dT} - (V-V_{Go}) \times 1}{T^2}\right)\right]$$

$$= K e^{(V-V_{Go})/\eta kT}\left[\frac{mT^m}{T} + \frac{T^m}{\eta kT^2}\left(T\frac{dV}{dT} - (V-V_{Go})\right)\right]$$

Note that V_{Go} is forbidden energy gap at 0K and hence it is a constant from differentiation point of view.

Taking T^m outside and rearranging the above equation, we get

$$\frac{dI}{dT} = K e^{(V-V_{Go})/\eta kT} \times T^m \left[\frac{m\eta kT + \left(T\frac{dV}{dT} - (V - V_{Go})\right)}{\eta kT^2}\right]$$

$$= K e^{(V-V_{Go})/\eta kT} \times \frac{T^m}{\eta kT^2}\left[m\eta kT + \left(T\frac{dV}{dT} - (V - V_{Go})\right)\right]$$

Replacing kT with V_T, we get

$$\frac{dI}{dT} = K e^{(V-V_{Go})/\eta kT} \times \frac{T^{m-1}}{\eta V_T}\left[m\eta V_T + \left(T\frac{dV}{dT} - (V - V_{Go})\right)\right]$$

Now $\frac{dI}{dT} = 0$ for constant diode current. Hence, equating the above equation to zero, we get

$$m\eta V_T + T\frac{dV}{dT} - (V - V_{Go}) = 0$$

$$T\frac{dV}{dT} = V - V_{Go} - m\eta V_T$$

$$\frac{dV}{dT} = \frac{V - (V_{Go} + m\eta V_T)}{T}$$

This is the required change in voltage necessary to keep diode current constant.

Hence for germanium, at cut-in voltage $V = V_\gamma = 0.2$ V and with $m = 2$, $\eta = 1$, $T = 30$ K and $V_{Go} = 0.785$ V in the above equation, we get

$$\frac{dV}{dT} = \frac{0.2 - (0.785 + 2 \times 1 \times 26 \times 10^{-3})}{300} = -2.12 \text{ mV/°C for Ge}$$

The negative sign indicates that the voltage must be reduced at a rate of 2.12 mV per degree change in temperature to keep diode current constant.

Similarly, $\frac{dV}{dT} = -2.3$ mV/°C, for Si

Practically, the value of $\frac{dV}{dT}$ is assumed to be –2.5 mV/°C for either Ge or Si at room temperature.

Thus,
$$\frac{dV}{dT} = -2.5 \text{ mV/°C} \tag{1.42}$$

The negative sign indicates that dV/dT decreases with increase in temperature.

1.24.1 Effect of Temperature on Reverse Saturation Current

To study by what rate I_o changes with respect to temperature, consider Eq. (1.39) again. That is,

$$I_o = KT^m e^{-V_{Go}/\eta V_T}$$

Taking logarithm on both sides, we get

$$\ln (I_o) = \ln (KT^m e^{-V_{Go}/\eta V_T)}$$

$$= \ln K + \ln T^m - \frac{V_{Go}}{\eta V_T}$$

$$= \ln K + m \ln T - \frac{V_{Go}}{\eta V_T}$$

Substituting $V_T = kT$, we get

$$\ln (I_o) = \ln K + m \ln T - \frac{V_{Go}}{\eta kT}$$

Differentiating this equation with respect to T, we get

$$\frac{d[\ln I_o]}{dT} = 0 + \frac{m}{T} - \frac{V_{Go}}{\eta k} \cdot \left(\frac{-1}{T^2}\right) = \frac{m}{T} + \frac{V_{Go}}{\eta k T^2}$$

Replacing kT with V_T we have

$$\frac{d[\ln I_o]}{dT} = \frac{m}{T} + \frac{V_{Go}}{\eta T V_T}$$

For germanium, substituting the values of various terms terms at room temperature, we get

$$\frac{d[\ln I_o]}{dT} = \frac{2}{300} + \frac{0.785}{1 \times 300 \times 26 \times 10^{-3}} = 0.11 \text{ per °C}$$

This indicates that I_o increases by 11 % per degree rise in temperature. For silicon, we get

$$\frac{d[\ln I_o]}{dT} = 0.08 \text{ per °C}$$

This indicates that I_o increases by 8 % per degree rise in temperature.

Practically it is found that the reverse saturation current I_o increases by 7 % per °C change in temperature for both silicon and germanium diodes. If at T°C is 1 μA, then at $(T + 1)$ °C, it becomes 1.07 μA and so on. From this, it can be concluded that reverse saturation current approximately doubles, i.e., 1.07^{10} for every 10°C rise in temperature.

The above result can be mathematically represented as,

$$I_{o2} = \left(2^{\frac{T_2 - T_1}{10}}\right) I_{o1} = \left(2^{\frac{\Delta T}{10}}\right) I_{01}$$

where I_{o2} is the reverse saturation current at T_2 and I_{o1} is the reverse saturation current at T_1.

1.24.2 Temperature Dependence of *V–I* Characteristics

The rise in temperature increases the generation of electron–hole pairs in semiconductors and increases their conductivity. As a result, the current through the *PN* junction diode increases with temperature as given by the diode current equation,

$$I = I_o [e^{(V/\eta VT)} - 1]$$

The reverse saturation current I_o of diode increases approximately 7 percent/°C for both germanium and silicon. Since $(1.07)^{10} \approx 2$, reverse saturation current approximately doubles for every 10°C rise in temperature. Hence, if the temperature is incrcased at fixed voltage, the current I increases. To bring the current I to its original value, the voltage V has to be reduced. It is found that at room temperature for either germanium or silicon, $\frac{dV}{dT} \approx -2.5$ mV/°C in order to maintain the current I to a constant value.

At room temperature, i.e., at 300 K, the value of barrier voltage or cut-in voltage is about 0.3 V for germanium and 0.7 V for silicon. The barrier voltage is temperature dependent and it decreases by 2 mV/°C for both germanium and silicon. This fact may be expressed in mathematical form, which is given by

$$I_{o2} = I_{o1} \times 2^{(T_2 - T_1)/10}$$

where I_{o1} = saturation current of the diode at temperature (T_1), and I_{o2} = saturation current of the diode at temperature (T_2).

Figure 1.25 shows the effect of increased temperature on the characteristic curve of a *PN* junction diode. A germanium diode can be used up to a maximum of 75°C and a silicon diode to a maximum of 175°C.

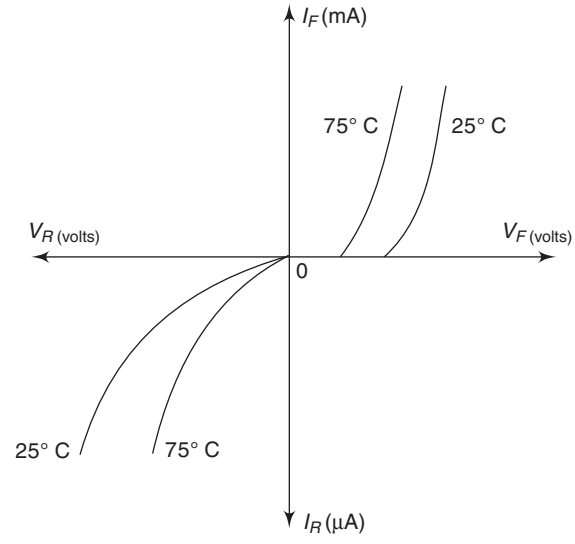


Fig. 1.25 *Effect of temperature on the diode characteristics*

EXAMPLE 1.28

The voltage across a silicon diode at room temperature (300 K) is 0.7 volts when 2 mA current flows through it. If the voltage increases to 0.75 V, calculate the diode current (assume V_T = 26 mV).

Solution Given data

Room temperature = 300 K

Voltage across a silicon diode, V_{D1} = 0.7 V

Current through the diode I_{D1} = 2 mA

When the voltage increases to V_{D2} = 0.75 V, then

$$\frac{I_{D2}}{I_{D1}} = \frac{I_o(e^{V_{D2}/V_T\eta} - 1)}{I_o(e^{V_{D1}/V_T\eta} - 1)} = \frac{e^{0.75/26\times10^{-3}} \times 2 - 1}{e^{0.7/26\times10^{-3}} \times 2 - 1} = 2.615$$

Therefore,

$$I_{D2} = 2.615 \times I_{D1}$$
$$= 2.615 \times 2 \times 10^{-3} = 5.23 \text{ mA}$$

EXAMPLE 1.29

A silicon diode has a saturation current of 7.5 μA at room temperature 300 K. Calculate the saturation current at 400 K.

Solution Given $I_{o1} = 7.5 \times 10^{-6}$ A at $T_1 = 300$ K = 27°C and $T_2 = 400$ K = 127°C

Therefore, the saturation current at 400 K is

$$\begin{aligned} I_{o2} &= I_{o1} \times 2^{(T_2 - T_1)/10} \\ &= 7.5 \times 10^{-6} \times 2^{(127-27)/10} \\ &= 7.5 \times 10^{-6} \times 2^{10} = 7.68 \text{ mA} \end{aligned}$$

EXAMPLE 1.30

The reverse saturation current of the Ge transistor is 2 μA at room temperature of 25°C and increases by a factor of 2 for each temperature increase of 10°C. Find the reverse saturation current of the transistor at a temperature of 75°C.

Solution Given $I_{o1} = 2$ μA at $T_1 = 25$°C, $T_2 = 75$°C

Therefore, the reverse saturation current of the transistor at $T_2 = 75$°C is

$$\begin{aligned} I_{o2} &= I_{o1} \times 2^{(T_2 - T_1)/10} = 2 \times 10^{-6} \times 2^{\left(\frac{75-25}{10}\right)} \\ &= 2 \times 10^{-6} \times 2^5 = 64 \text{ μA} \end{aligned}$$

EXAMPLE 1.31

A diode with 700 mW power dissipation at 25°C had a 5 mW/°C derating factor. If the forward voltage drop remains constant at 0.7 V, calculate the maximum forward current at 25°C and at 65°C temperatures. (AU May/June 2011)

Solution At 25°C, the maximum forward current is

$$I_F = \frac{P}{V_F} = \frac{700 \text{ mW}}{0.7 \text{ V}} = 1 \text{ A}$$

At 65°C,

$$\begin{aligned} P &= (P \text{ at } T_1) - [\Delta T \times \text{derating factor}] \\ &= 700 \text{ mW} - 40 \times 5 \text{ mW} = 500 \text{ mW} \end{aligned}$$

Therefore, $$I_F = \frac{P}{V_F} = \frac{500 \text{ mW}}{0.7 \text{ V}} = 714 \text{ mA}$$

1.25 JUNCTION DIODE SWITCHING CHARACTERISTICS

[AU May 2017, Dec 2015, Nov and May 2014, 8 marks]

Diodes are often used in switching mode as shown Fig. 1.26. When a forward voltage is applied to the diode, the diode is in ON-state. When the applied bias voltage to the *PN* diode is suddenly reversed

in the opposite direction, the minority charge carriers in the *P*- and *N*-sides of the diode are not instantaneously removed and hence the diode is not switched to OFF-state immediately. The diode response reaches a steady state only after an interval of time, called the *recovery time*. The forward recovery time, t_{fr}, is defined as the time required for forward voltage or current to reach a specified value (time interval between the instant of 10% diode voltage to the instant this voltage reaches within 10% of its final value) after switching diode from its reverse-to forward-biased state. Fortunately, the forward recovery time possesses no serious problem. Therefore, only the reverse recovery time, t_{rr}, has to be considered in practical applications.

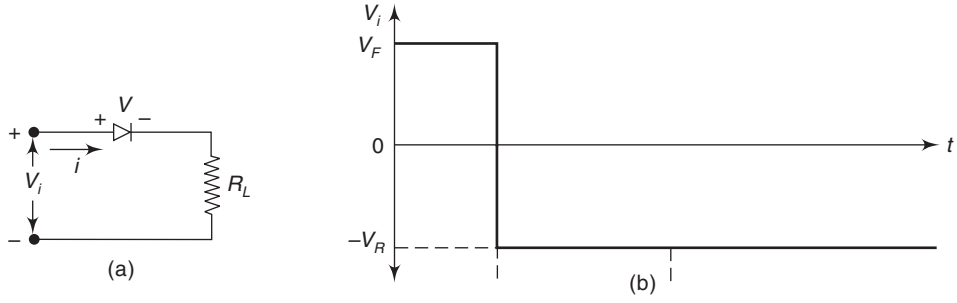


Fig. 1.26 *Forward biased diode (a) circuit and (b) input voltage characteristics*

It is assumed that the *N*-side of the diode is heavily doped than the *P*-side. The distribution of carrier concentration over a distance x from the junction by neglecting the space charge region is drawn as shown in Fig. 1.27(a). Here, n_{po}, p_{no} are minority charge carriers and n_{no}, p_{po} are majority charge carriers.

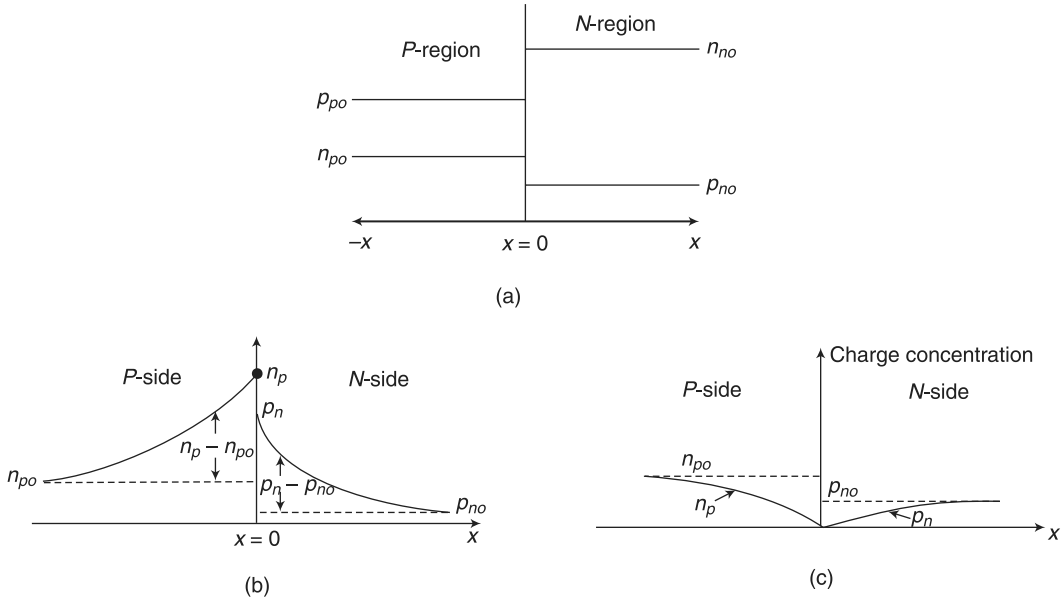


Fig. 1.27 *Charge carrier concentration with respect to distance x*

When the diode is forward biased, the positive terminal is connected to the *P*-side and the negative terminal is connected to the *N*-side. The holes in the *P*-side are repelled and they recombine with electrons in the N-side and vice versa for the electrons in the *N*-side. Therefore, at the junction $x = 0$ the concentration of specific charge carriers will be maximum but decreases exponentially on entering the oppositely charged region until it reaches the concentration of the minority charge carriers of the same type in that region. It is represented in Fig. 1.27(b).

The number of excess electrons in the *P*-side at a distance x is $n_p - n_{po}$ and the number of excess holes in the *N*-side at a distance x is $p_n - p_{no}$. Both n_p and p_n decrease exponentially with increase in distance from the junction. On application of reverse bias, the voltage source is connected in the opposite direction to that of the forward bias. Majority charge carriers in the diode will be attracted by its opposite terminals and will not cross the junction, and hence the width of the barrier will increase. However, the minority carriers on each side will cross the junction and hence current flows. But this current will be very less in magnitude, and will be in the range of nA for Si, and μA for Ge. This is known as reverse saturation current. Ideally, this must be zero, but due to manufacturing defects, minority charge carriers will always be present, resulting in larger than zero reverse saturation current. At the junction $x = 0$, the number of electrons in the *P*-side becomes zero and the same for the holes in the *N*-side. But at greater distances, it will reach the saturation values of the minority charge carriers of that side. It is represented in Fig. 1.27(c).

When the *PN* junction diode is forward biased, the minority electron concentration in the *P*-region is approximately linear. If the junction is suddenly reverse biased, at t_1, then because of this stored electronic charge, the reverse current (I_R) is initially of the same magnitude as the forward current (I_F). The diode will continue to conduct until the injected or excess minority carrier density ($p - p_o$) or ($n - n_o$) has dropped to zero. However, as the stored electrons are removed into the *N*-region and the contact, the available charge quickly drops to an equilibrium level and a steady current eventually flows corresponding to the reverse bias voltage as shown in Fig. 1.28 (b).

As shown in Fig. 1.28(a), the applied voltage $V_i = V_F$ for the time up to t_1 is in the direction to forward-bias the diode. The resistance R_L is large so that the drop across R_L is large when compared to the drop across the diode. Then the current is $I \approx \dfrac{V_F}{R_L} = I_F$. Then, at time $t = t_1$, the input voltage is suddenly reversed to the value of $-V_R$. Due to the reasons explained above, the current does not become zero and has the value $I = \dfrac{V_R}{R_L} = -I_R$ until the time $t = t_2$. At $t = t_2$, when the excess minority carriers have reached the equilibrium state, the magnitude of the diode current starts to decrease, as shown in Fig. 1.23(c).

During the time interval from t_1 to t_2, the injected minority carriers have remained stored and hence this time interval is called the *storage time* (t_s).

After the instant $t = t_2$, the diode gradually recovers and ultimately reaches the steady-state. The time interval between t_2 and the instant t_3 when the diode has recovered nominally, is called the *transition time*, t_t. The recovery is said to have completed (i) when even the minority carriers remote from the junction have diffused to the junction and crossed it, and (ii) when the junction transition capacitance, C_T, across the reverse-biased junction has got charged through the external resistor R_L to the voltage $-V_R$.

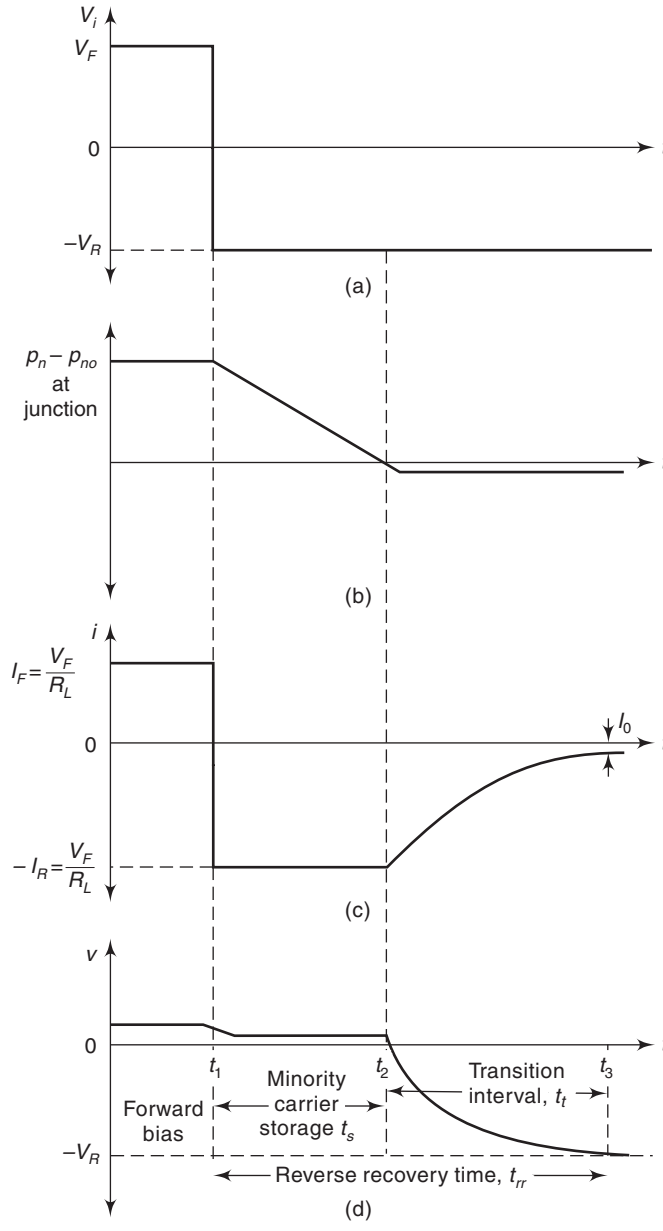


Fig. 1.28 *Switching characteristics of PN junction diode*

As shown in Fig. 1.28 (d), the reverse recovery time (or turn-off time) of a diode, t_{rr}, is the interval from the current reversal at $t = t_1$ until the diode has recovered to a specified extent in terms either of the diode current or of the diode resistance, i.e., $t_{rr} = t_s + t_t$.

For commercial switching type diodes, the reverse recovery time, t_{rr}, ranges from less than 1 ns up to as high as 1 μs. This switching time obviously limits the maximum operating frequency of the device. If the

time period of the input signal is such that $T = 2 \cdot t_{rr}$, then the diode conducts as much in reverse as in the forward direction. Hence it does not behave as a one way device. In order to minimise the effect of the reverse current, the time period of the operating frequency should be a minimum of approximately 10 times t_{rr}. For example, if a diode has t_{rr} of 2 ns, its maximum operating frequency is

$$f_{\max} = \frac{1}{T} = \frac{1}{10 \times t_{rr}} = \frac{1}{10 \times 2 \times 10^{-9}} = 50 \text{ MHz}$$

The t_{rr} can be reduced by shortening the length of the *P*-region in a *PN* junction diode. The stored charge and, consequently, the switching time can also be reduced by introduction of gold impurities into the junction diode by diffusion. The gold dopant, some times called a *life time killer*, increases the recombination rate and removes the stored minority carriers. This technique is used to produce diodes and other active devices for high speed applications.

1.26 BREAKDOWN IN *PN* JUNCTION DIODES

The diode equation predicts that, under reverse bias conditions, a small constant current, the saturation current, I_o, flows due to minority carriers, which is independent of the magnitude of the bias voltage. But this prediction is not entirely true in practical diodes. There is a gradual increase of reverse current with increasing bias due to the ohmic leakage currents around the surface of the junction. Also, there is a sudden increase in reverse current due to some sort of *breakdown*, when the reverse bias voltage approaches a particular value called breakdown voltage, V_{BD}, as shown in Fig. 1.29. Once breakdown occurs, the diode is no longer blocking current and the diode current can be controlled only by the resistance of the external circuit.

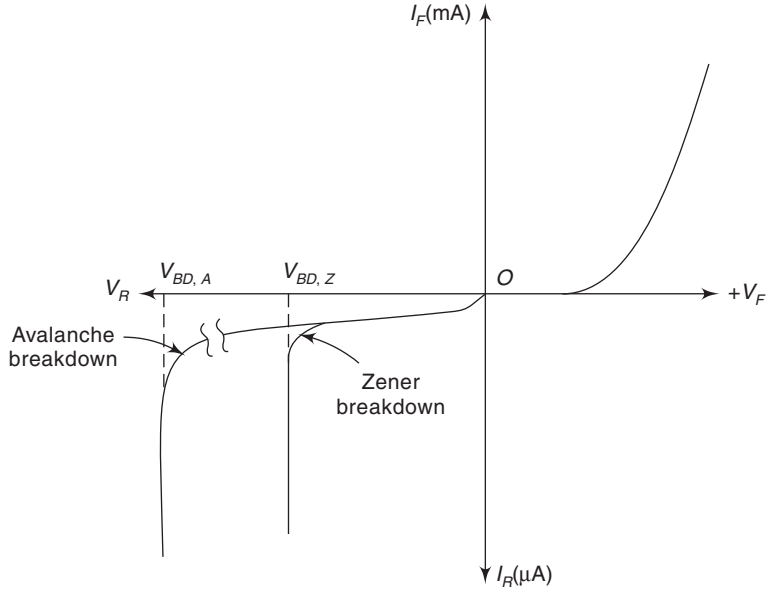


Fig. 1.29 *Breakdown in PN junction diodes*

The breakdown occurs due to avalanche effect in which thermally generated minority carriers cross the depletion region and acquire sufficient kinetic energy from the applied potential to produce

new carriers by removing valence electrons from their bonds. These new carriers will in turn collide with other atoms and will increase the number of electrons and holes available for conduction. This multiplication effect of free carriers may be represented by the following equation:

$$M = \frac{1}{1 - \left(\frac{V}{V_{BD}}\right)^n}$$

where M = carrier multiplication factor, which is the ratio of the total number of electrons leaving the depletion region to the number entering the region
V = applied reverse voltage
V_{BD} = reverse breakdown voltage
n = empirical constant, which depends on the lattice material and the carrier type, for N-type silicon, $n \approx 4$ and for P-type, $n \approx 2$

It is evident from the above equation that M is being very small for $V = 0.9\ V_{BD}$. But when $V > 0.9\ V_{BD}$, M is very large and the resulting reverse current is given by $I_R = MI_o$, where I_o is the reverse saturation current before breakdown.

As V approaches the breakdown voltage V_{BD}, the value of M will become infinite and there is a rapid increase in carrier density and a corresponding increase in current. Because of the cumulative increase in carrier density after each collision, the process is known as *avalanche breakdown*.

Even if the initially available carriers do not gain enough energy to disrupt bonds, it is possible to initiate breakdown through a direct rupture of the bonds because of the existence of strong electric field. Under these circumstances, the breakdown is referred to as *Zener breakdown*.

Voltage reference diodes that utilise the almost constant voltage characteristics in the breakdown region are also called *avalanche diodes* or sometimes *Zener diodes.* The Zener effect is in diodes with breakdown voltage below 6 V. The operating voltage in avalanche breakdown are from several volts to several hundred volts with power ratings up to 50 W.

True Zener diode action displays a negative temperature coefficient, i.e., breakdown voltage decreases with increasing temperature. True avalanche diode action exhibits a positive temperature coefficient, i.e., breakdown voltage increases with increasing temperature.

It is clear that the breakdown voltage for a particular diode can be controlled during manufacture by altering the doping levels in the junction. The breakdown voltage for silicon diodes can be made to occur at a voltage as low as 5 V with 10^{17} impurity atoms per cubic cm or as high as 1000 V when doped to a level of only 10^{14} impurity atoms per cubic cm.

1.27 DIODE AS A CIRCUIT ELEMENT

The *PN* junction diode is considered as a circuit element. The basic diode circuit shown in Fig. 1.30 consists of a DC voltage V_S which is supplied across a resistor and a diode. In order to find the instantaneous diode voltage V and current I, the circuit can be analysed when the instantaneous source voltage is V_S. From Kirchhoff's voltage law (KVL), the instantaneous diode voltage is

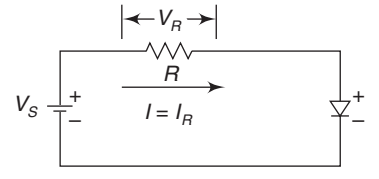


Fig. 1.30 *Simple diode circuit*

$$V_S = IR + V \tag{1.43}$$

which can be expressed as

$$I = \frac{V_S}{R} - \frac{V}{R} \tag{1.44}$$

The ideal diode current equation relating the diode voltage V_D and the current I_D is

$$I = I_0 \left[e^{(V/\eta V_T)} - 1 \right] \tag{1.45}$$

where I_o is the diode reverse saturation current at room temperature, η is a constant (1 for Ge and 2 for Si) and V_T is the thermal voltage, i.e., $V_T = \frac{T}{11{,}600}$ in which T is the temperature of the diode junction (300 K) or $V_T = 26$ mV at room temperature.

Substituting Eq. (1.45) into Eq. (1.43), we get

$$V_S = I_0 R\left[e^{(V/\eta V_T)} - 1 \right] + V \tag{1.46}$$

which has only one unknown variable V. Since Eq. (1.46) is a transcendental equation, this equation cannot be solved directly.

As these equations contain both linear and exponential terms, it is difficult to solve by hand. The iteration (trial and error) technique can be used to find a solution to this equation. The graphical analysis technique involves plotting two simultaneous equations and locating their point of intersection, which is called the quiescent point, or the Q-point.

1.28 *PN* DIODE APPLICATIONS

An ideal *PN* junction diode is a two terminal polarity sensitive device that has zero resistance (diode conducts) when it is forward biased and infinite resistance (diode does not conduct) when reverse biased. Due to this characteristic, the diode finds number of applications as given below.

(i) rectifiers in DC power supplies
(ii) switch in digital logic circuits used in computers
(iii) clamping network used as DC restorer in TV receivers and voltage multipliers
(iv) clipping circuits used as wave shaping circuits used in computers, radars, radio and TV receivers
(v) demodulation (detector) circuits.

The same *PN* junction with different doping concentration finds special applications as follows:

(i) detectors (APD, PIN photo diode) in optical communication circuits
(ii) Zener diodes in voltage regulators
(iii) varactor diodes in tuning sections of radio and TV receivers
(iv) light emitting diodes in digital displays
(v) LASER diodes in optical communications
(vi) Tunnel diodes as a relaxation oscillator at microwave frequencies.

Two Mark Questions and Answers

1. Find the voltage at which the reverse current in a germanium PN junction diode attains a value of 90% of its saturation value at room temperature. (April/May 2017)

Refer to Example 1.20.

2. What is meant by doping in a semiconductor? (Dec 2015/Jan 2016)

The process of adding impurity to intrinsic semiconductor is known as *doping*. As a result of doping, an extrinsic semiconductor is formed.

3. Define storage time. (May/June 2016)

When the *PN* junction diode switches from ON state to OFF state, the injected minority carriers remain stored in an equilibrium state across the junction during a time interval which is called *storage time*.

4. What is barrier potential? (Nov/Dec 2016)

With no applied voltage across the *PN* junction diode, a barrier is set-up across the junction which prevents movement of charge carriers i.e., electrons and holes. This leads to induced electric field across the depletion layer, and an electrostatic potential difference is established between the P- and N-regions. Such potential is called the barrier potential, junction barrier, diffusion potential, or contact potential, V_o.

5. What is recovery time? Give its types. (Nov/Dec 2016)

When the applied bias voltage to the *PN* diode is suddenly reversed in the opposite direction, the minority charge carriers in the *P*- and *N*- sides of the diode are not instantaneously removed and hence, the diode is not switched to OFF-state immediately. The diode response reaches a steady state only after an interval of time, called the *recovery time*. There are two types of recovery time, namely (i) forward recovery time and (ii) reverse recovery time.

6. Define electron volt. (May/June 2014)

One eV is the amount of energy acquired by an electron when it is accelerated through a potential difference of one volt (1 V). The value of 1 eV is equal to 1.602×10^{-19} joule.

7. Sketch the forward bias characteristics of the *PN* junction diode. (April/May 2015)

Refer to Fig. 1.15.

8. Define diffusion current and drift current. (Nov/Dec 2014, April/May2015)

(a) Drift current: The drift current is defined as the flow of electric current due to the motion of the charge carriers under the influence of an external electric field. The drift current is given by

$$I_n = Aqn\mu_0 E \qquad \text{(due to free electrons)}$$

$$I_p = Aqp\mu_0 E \qquad \text{(due to holes)}$$

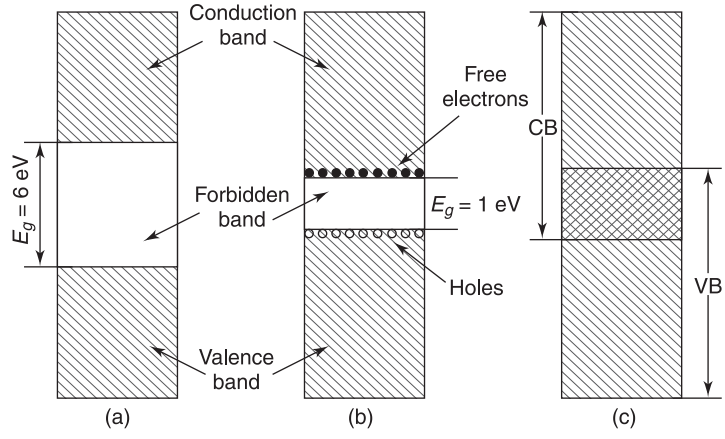
(b) Diffusion current: A concentration gradient exists if the number of either electrons or holes is greater in one region of a semiconductor, compared to the rest of the region. The charge carriers tend to move from the region of higher concentration to that of lower counteraction

of the same type of charge carriers. There, the movement of charge carriers takes place resulting in a current called diffusion current. The diffusion current is given by

$$I_n = -A \cdot qD_p \frac{dp}{dx} \qquad \text{(due to holes)}$$

$$I_n = A \cdot qD_n \frac{dn}{dx} \qquad \text{(due to electrons)}$$

9. Draw energy band diagram of semiconductor. **(Nov/Dec 2013)**



Energy band diagram of (a) Insulator, (b) Semiconductor and (c) Metal

10. Define mass action law. **(Nov/Dec 2010, May/June 2014)**

Under thermal equilibrium for any semiconductor, the product of the number of holes and the number of electrons is constant and is independent of the amount of donor and acceptor impurity doping. This relation is known as mass-action law and is given by

$$n \cdot p = n_i^2$$

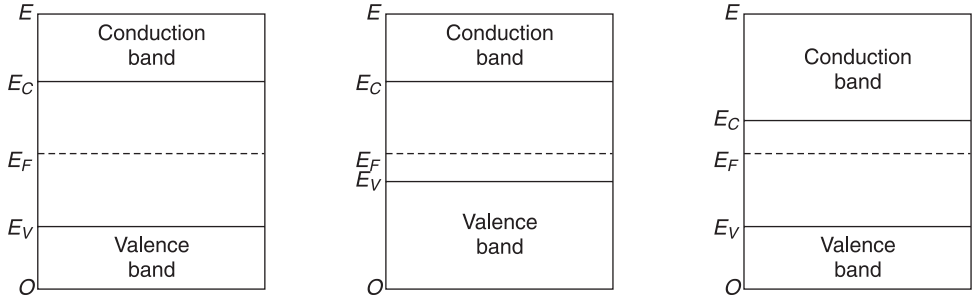
11. Differentiate between intrinsic and extrinsic semiconductors. **(May/June 2011, May/June 2013)**

	Intrinsic Semiconductor	Extrinsic Semiconductor
1	It is pure semi-conducting material and no impurity atoms are added to it.	It is prepared by doping a small quantity of impurity atoms of the pure semi-conducting materials.
2	The number of free electrons in the conduction band and the number of holes in the valence band are exactly equal and very small indeed.	The number of free electrons and holes are never equal. There are excess of electrons in *N*-type semi-conductors and excess of holes in *P*-type semi-conductors.
3	Its electrical conductivity is a function of temperature alone.	Its electrical conductivity depends upon the temperature as well as on the quantity of impurity atoms doped in the structure.

(*Contd.*)

4	Its electrical conductivity is low.	Its electrical conductivity is high
5	*Examples*: Crystalline forms of pure silicon and germanium.	*Examples*: Silicon "Si" and germanium "Ge" crystals with impurity atoms of As, Sb, P etc. or In B, Al, etc.

12. Draw the energy band diagrams for intrinsic and extrinsic semiconductors.



Energy band diagrams for (a) intrinsic semiconductor, (b) P-type and (c) N-type extrinsic semiconductors

13. What is the principle of operation of a *PN* junction diode in reverse bias condition?
(May/June 2014)

As shown in Fig. 1.15, when reverse bias is applied to the *PN* junction, the holes in the *P*-side more towards the negative terminal of the battery and the electrons are attracted towards the positive terminal of the battery. This increases the depletion width and hence, a very small current of the order of few microamperes flows across the junction.

14. What is meant by PIV (Peak Inverse Voltage) of a *PN* junction diode? (May/June 2012)

The maximum reverse voltage by which a *PN* junction diode can withstand without breakdown of the junction is called peak inverse voltage.

15. State the effect of temperature of *PN* junction diode. (Nov/Dec 2012)

The current through the *PN* junction diode increases with temperature as given by the diode current equation,

$$I = I_0 \left[e^{\frac{V}{\eta V_T}} - 1 \right]$$

The reverse saturation current I_0 of diode increases approximately 7 percent /°C, for both germanium and silicon.

Refer to Fig. 1.25 which shows the effect of temperature on the diode characteristics.

16. What is diffusion capacitance? (Nov/Dec 2010, 2012, 2013)

The capacitance that exists in a forward bias junction is called a diffusion or storage capacitance. It is defined as the rate of change of injected charge with applied voltage.

Referring to Section 1.23, the diffusion capacitance, $C_D = \frac{dQ}{dV} = \frac{\tau I}{nV_T}$

17. Give the expression for transition capacitance and diffusion capacitance of a *PN* diode. **(May/June 2010, 2013)**

Refer Section 1.22, the transition capacitance, $C_T = \frac{\varepsilon A}{W}$

Refer Section 1.23, the diffusion capacitance, $C_D = \frac{\tau I}{n_{VT}}$

18. Define avalanche breakdown. **(May/June 2010)**

Under reverse bias condition, when the reverse voltage approaches a particular value called breakdown voltage, the reverse current suddenly increases because the velocity of thermally generated minority carriers increases and they cross the depletion region and acquire sufficient kinetic energy from the applied potential to produce new charge carriers by removing valence electrons from their bonds. These new carriers will in turn collide with other atoms and will increases the number of electron-holes available for conduction. Because of the cumulative increase in carrier density after each collision, this process is called *avalanche breakdown*.

19. Write any two applications of Zener diode. **(May/June 2014)**

Zener diode is used as

(a) Voltage regulator,
(b) Voltage clipper circuits and
(c) For controlling the output amplitude.

20. What is Zener breakdown? **(Nov/Dec 2011)**

When the *P* and *N* regions are heavily doped, direct rupture of covalent bonds takes place because of the strong electric field at the junction of *PN* diode. The new electron-hole pair, so created, increases the reverse current in a reverse biased *PN* diode. As a result of heavy doping of *P* and *N* regions, the depletion width becomes very small and the field across the depletion region becomes very high and it is due to ruptures of the covalent bond. This breakdown is known as *Zener breakdown*.

Review Questions

1. Define the term 'Electronics.'
2. List the major areas of applications of electronics.
3. Describe the energy band structures of an insulator, a metal and a semiconductor.
4. Define mean-free-path of an electron.
5. Define drift velocity of an electron.
6. What is meant by effective mass of electron? Explain.

7. What are the three commonly used semiconductors?
8. What are the band gap energies of germanium, silicon and gallium arsenide?
9. Explain how electrons are emitted from metals.
10. Explain the term work function.
11. Discuss briefly the different types of electron emission.
12. Give the Richardson–Dushman equation and explain the effect of various factors in it.
13. What is meant by intrinsic semiconductor?
14. Explain the differences between intrinsic and extrinsic semiconductors.
15. Explain what a hole is. How do they move in intrinsic semiconductor?
16. What is meant by doping in a semiconductor?
17. Discuss the following with respect to semiconductor: (i) doping (ii) dopant (iii) donor and (iv) acceptor.
18. Explain "majority and minority carriers" in a semiconductor.
19. What is meant by *N*-type impurity in a semiconductor?
20. What is meant by *P*-type impurity in a semiconductor?
21. Define the terms conductivity and mobility in a semiconductor.
22. Derive the conductivity equation for an *N*-type and *P*-type semiconductor.
23. Prove that the conductivity of a semiconductor is given by, $\sigma = q(P\,\mu_p + n\,\mu_n)$.
24. Describe the phenomenon of diffusion of charge carriers in semiconductors.
25. In an *N*-type semiconductor, the Fermi level lies 0.4 eV below the conduction band at 300 K. Determine the new position of the Fermi level if (a) the temperature is increased to 400 K and (b) the concentration of donor atoms is increased by a factor of 6. Assume $kT = 0.03$ eV.

    (*Ans*. (a) 0.533 eV and (b) 0.3463 eV below the conduction band)
26. In a *P*-type semiconductor, the Fermi level lies 0.4 eV above the valence band at 300 K. Determine the new position of the Fermi level (a) at 450 K and (b) if the concentration of acceptor atoms is multiplied by a factor of 2. Assume $kT = 0.03$ eV. (*Ans*. (a) 0.6 eV and (b) 0.38 eV above the valence band)
27. The mobility of electrons and holes in a sample of intrinsic germanium at room temperature are 0.36 m^2/V-s and 0.17 m^2/V-s, respectively. If the electron and hole densities are each equal to $2.5 \times 10^{19}/m^3$, calculate the conductivity. (*Ans*. $\sigma_i = 2.12$ S/m)
28. Compute the conductivity of a silicon semiconductor which is doped with acceptor impurity to a density of 10^{22} atoms/m^3. Given that $n_i = 1.4 \times 10^{16}/m^3$, $\mu_n = 0.145\ m^2$/V-s and $\mu_p = 0.05\ m^2$/V-s. (*Ans.* 80 S/m)
29. The conductivity of a pure silicon at room temperature is 5×10^{-4} S/m. How many aluminium atoms per m^3 are required so that a saturation conductivity of 200 S/m could be achieved in silicon using aluminium as an impurity? Given that the mobility of holes in Silicon is 0.05 m^2/V-s and the mobility of electrons is 0.13 m^2/V-s. (*Ans*. $2.5 \times 10^{22}/m^3$)
30. Calculate the conductivity of a pure silicon at room temperature of 300 K. Given that $n_i = 1.5 \times 10^{16}/m^3$, $\mu_n = 0.13\ m^2$/V-s, $\mu_p = 0.05\ m^2$/V-s and $q = 1.602 \times 10^{-19}$ C. Now the silicon is doped 2 in 10^8 of a donor impurity. Calculate its conductivity if there are 5×10^{28} silicon atoms/m^3. By what factor has the conductivity increase? (*Ans.* 4.32×10^{-4} S/m; 20.8 S/m; ≈ 48,000)
31. The mobility of free electrons and holes in pure silicon are 0.13 and 0.05 m^2/V-s and the corresponding values for pure germanium are 0.38 and 0.18 m^2/V-s respectively. Determine the values of intrinsic conductivity for both silicon and germanium. Given that $n_i = 2.5 \times 10^{19}/m^3$ for germanium and $n_i = 1.5 \times 10^{16}/m^3$ for silicon at room temperature. (*Ans.* 0.43 S/m; 2.24 S/m)
32. (a) A crystal of pure germanium has sufficient antimony (*N*-type or donor impurities) added to produce 1.5×10^{22} antimony atoms/m^3. The electron and hole mobility are 0.38 m^2/V-s and 0.18 m^2/V-s respectively, and the

intrinsic charge carrier density is $2.5 \times 10^{19}/m^3$. Calculate (i) the density of electrons and holes in the crystal, and (ii) the conductivity.

(b) A second Germanium crystal is produced which is doped with 2.5×10^{22} indium (*P*-types or acceptor impurities) atoms/m^3. Repeat the calculations listed in part (a).

(c) A *PN* junction is made by joining the two crystals described above. Calculate its barrier voltage at 300 K.

(*Ans.* (a) $n = 1.5 \times 10^{22}/m^3$, $p = 4.167 \times 10^{16}/m^3$, $\sigma = 912/\Omega$-m
(b) $n = 2.5 \times 10^{16}/m^3$, $p = 2.5 \times 10^{22}/m^3$, $\sigma = 720/\Omega$-m (c) 0.335 V)

33. Explain the drift and diffusion currents for a semiconductor.
34. State and explain Mass-action law.
35. What is Einstein relationship in a *PN* junction?
36. Derive the continuity equation from the first principle.
37. With reference to charge carriers in a semiconductor, define and explain the terms (i) mobility and (ii) life time.
38. What is a *PN* junction? How is it formed?
39. Explain the formation of depletion region in a *PN* junction.
40. Draw the energy band diagram of a *PN* junction and explain the working of a diode.
41. (a) The resistivities of the *P*-region and *N*-region of a silicon diode are 6 Ω-cm and 4 Ω-cm respectively. Calculate the contact potential V_o and potential energy barrier E_o. (b) if the doping densities of both *P*- and *N*-regions are tripled, determine V_o and E_o. Given that $q = 1.602 \times 10^{-19}$ C, $n_i = 1.5 \times 10^{10}/cm^3$, $\mu_p = 500\ cm^2/V\text{-}s$, $\mu_n = 1300\ cm^2/V\text{-}s$ and $V_T = 0.026$ V at 300 K.

(*Ans.* (a) $V_o = 0.6$ *V*, $E_o = 0.6$ eV (b) $V_o = 0.66$V, $E_o = 0.66$ eV)

42. Show that in an intrinsic semiconductor, the Fermi level is located at the middle of the unallowable energy gap.
43. Sketch the conduction and valence bands before and after diffusion of carriers in a *PN* junction.
44. Explain how a barrier pointial is developed at the *PN* junction.
45. Describe the action of *PN* junction diode under forward bias and reverse bias.
46. Show that the *PN* diode works as a rectifier.
47. Explain how unidirectional current flow is possible through a *PN* junction diode.
48. Explain *V–I* characteristics of a *PN* junction diode.
49. Indicate the differences between the characteristics of silicon and germanium diodes and state approximately their cut-in voltages.
50. Explain the following terms in a *PN* junction diode:
    (a) Maximum forward current
    (b) Peak inverse voltage, and
    (c) Maximum power rating
51. Explain the terms: (i) static resistance, (ii) dynamic resistance, (iii) junction resistance, and (iv) reverse resistance of a diode.
52. Write the volt-ampere equation for a *PN* diode. Give the meaning of each symbol.
53. What are the factors governing the reverse saturation current in a *PN* junction diode?
54. Determine the forward bias voltage applied to a silicon diode to cause a forward current of 10 mA and reverse saturation current, $I_o = 25 \times 10^{-7}$ A at room temperature. (*Ans*. 0.4 V)
55. The reverse saturation current I_o in a germanium diode is 6 μA. Calculate the current flowing through the diode when the applied forward bias voltages are 0.2, 0.3 and 0.4 V at room temperature.

(*Ans.* 13.15 mA, 21.5 mA, 28.8 mA)

56. Define the term transition capacitance C_T of a *PN* diode.
57. Explain the term diffusion capacitance C_D of a forward biased diode.
58. Explain the effect of temperature of a diode.

59. Distinguish between avalanche and Zener mechanisms.
60. Can an ordinary rectifier diode be used as a Zener diode? Explain.
61. Mention the applications of *PN* junction diode.
62. Give the ideal diode current-voltage relationship. Describe the meaning of I_o and $V_{T.}$

BIPOLAR JUNCTION TRANSISTOR

2

2.1 INTRODUCTION

A Bipolar Junction Transistor (BJT) is a three-terminal semiconductor device in which the operation depends on the interaction of both majority and minority carriers and hence the name *bipolar*. The BJT is analogous to a vacuum triode and is comparatively smaller in size. It is used in amplifier and oscillator circuits, and as a switch in digital circuits. It has wide applications in computers, satellites and other modern communication systems.

This chapter deals with construction and operation of *NPN* and *PNP* transistors, types of configuration, breakdown in transistors, Ebers–Moll model, *h*-parameters model, hybrid-π model, Gummel–Poon model and the multi-emitter transistor.

2.2 CONSTRUCTION

[AU May 2012, 6 marks]

The BJT consists of a silicon (or germanium) crystal in which a thin layer of *N*-type silicon is sandwiched between two layers of *P*-type silicon. This transistor is referred to as *PNP*. Alternatively, in a *NPN* transistor, a layer of *P*-type material is sandwiched between two layers of *N*-type material. The two types of the BJT are represented in Fig. 2.1.

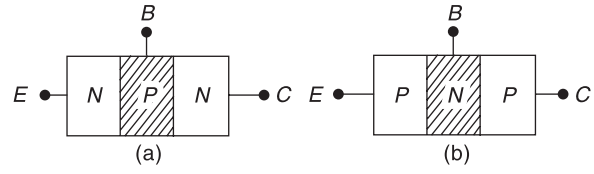


Fig. 2.1 *Transistor: (a) NPN (b) PNP*

The symbolic representation of the two types of the BJT is shown in Fig. 2.2. The three portions of the transistor are Emitter, Base and Collector, shown as *E*, *B* and *C*, respectively. The arrow on the emitter specifies the direction of current flow when the *EB* junction is forward biased.

The emitter is heavily doped so that it can inject a large number of charge carriers into the base. Base is lightly doped and very thin. It passes most of the injected charge carriers from the emitter into the collector. Collector is moderately doped.

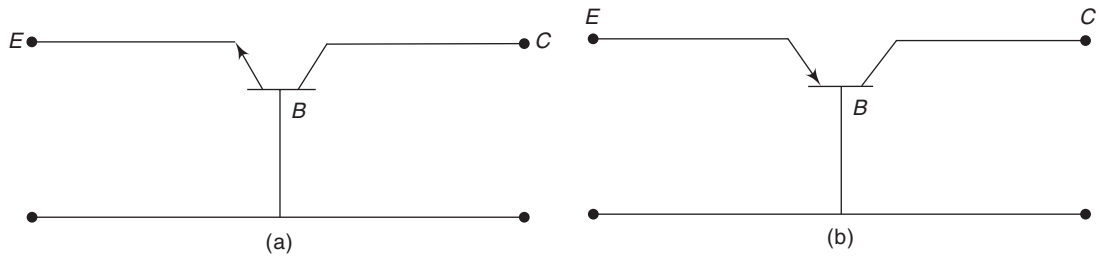


Fig. 2.2 *Circuit symbol: (a) NPN transistor (b) PNP transistor*

2.3 TRANSISTOR BIASING

As shown in Fig. 2.3, usually the emitter-base junction is forward biased and collector-base junction is reverse biased. Due to the forward bias on the emitter-base junction an emitter current flows through the base into the collector. Though, the collector-base junction is reverse biased, almost the entire emitter current flows through the collector circuit.

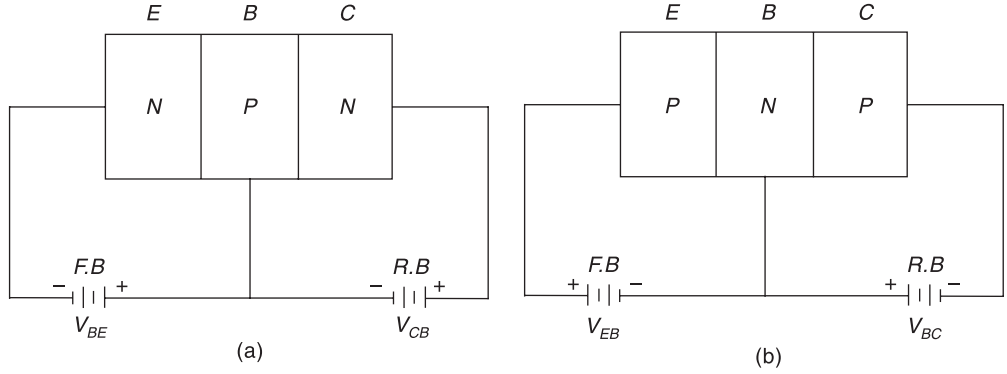


Fig. 2.3 *Transistor biasing: (a) NPN transistor (b) PNP transistor*

2.4 OPERATION OF AN *NPN* TRANSISTOR

[AU Nov 2013 and May 2012, 16 marks]

As shown in Fig. 2.4, the forward bias applied to the emitter base junction of an *NPN* transistor causes a lot of electrons from the emitter region to crossover to the base region. As the base is lightly doped with *P*-type impurity, the number of holes in the base region is very small and hence the number of electrons that combine with holes in the *P*-type base region is also very small. Hence a few electrons combine with holes to constitute a base current I_B. The remaining electrons (more than 95%) crossover into the collector region to constitute a collector current I_C. Thus the base and collector current summed up gives the emitter current, i.e., $I_E = -(I_C + I_B)$.

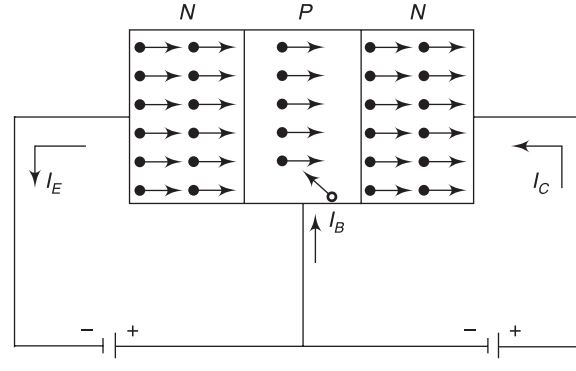


Fig. 2.4 *Current in an NPN transistor*

In the external circuit of the *NPN* bipolar junction transistor, the magnitudes of the emitter current I_E, the base current I_B and the collector current I_C are related by $I_E = I_C + I_B$.

2.5 OPERATION OF A *PNP* TRANSISTOR

[AU April 2015 and May 2012, 10 marks]

As shown in Fig. 2.5, the forward bias applied to the emitter-base junction of a *PNP* transistor causes a lot of holes from the emitter region to crossover to the base region as the base is lightly doped with *N*-types impurity. The number of electrons in the base region is very small and hence the number of holes combined with electrons in the *N*-type base region is also very small. Hence, a few holes combined with electrons to constitute a base current I_B. The remaining holes (more than 95%) crossover into the collector region to constitute a collector current I_C. Thus the collec- tor and base current when summed up gives the emitter current, i.e., $I_E = -(I_C + I_B)$.

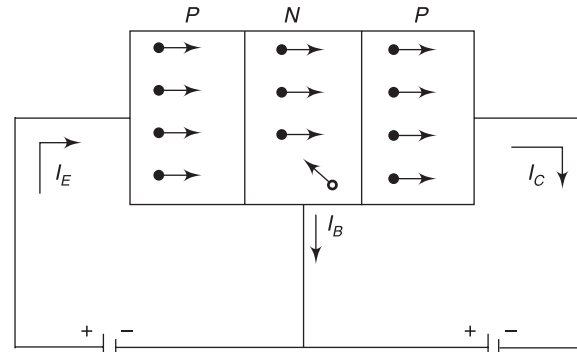


Fig. 2.5 *Current in a PNP transistor*

In the external circuit of the PNP bipolar junction transistor, the magnitudes of the emitter current I_E, the base current I_B and the collector current I_C are related by

$$I_E = I_C + I_B \tag{2.1}$$

This equation gives the fundamental relationship between the currents in a bipolar transistor circuit. Also, this fundamental equation shows that there are current amplification factors α and β in common base transistor configuration and common emitter transistor configuration respectively for the static (DC) currents, and for small changes in the currents.

Large-signal current gain (α) The large signal current gain of a common base transistor is defined as the ratio of the negative of the collector-current increment to the emitter-current change from cut-off ($I_E = 0$) to I_E, i.e.,

$$\alpha = -\frac{(I_C - I_{CBO})}{I_E - 0} \tag{2.2}$$

where I_{CBO} (or I_{CO}) is the reverse saturation current flowing through the reverse biased collector-base junction, i.e., the collector to base leakage current with emitter open. As the magnitude of I_{CBO} is negligible when compared to I_E, the above expression can be written as

$$\alpha = -\frac{I_C}{I_E} \tag{2.3}$$

Since I_C and I_E are flowing in opposite directions, α is always positive. Typical value of α ranges from 0.90 to 0.995. Also, α is not a constant but varies with emitter current I_E, collector voltage V_{CB} and temperature.

General transistor equation In the active region of the transistor, the emitter is forward biased and the collector is reverse biased. The generalised expression for collector current I_C for collector junction voltage V_C and emitter current I_E is given by

$$I_C = -\alpha I_E + I_{CBO}(1 - e^{V_C/V_T}) \tag{2.4}$$

If V_C is negative and $|V_c|$ is very large compared with V_T, then the above equation reduces to

$$I_C = -\alpha I_E + I_{CBO} \tag{2.5}$$

If V_C, i.e., V_{CB}, is few volts, I_C is independent of V_C. Hence the collector current I_C is determined only by the fraction α of the current I_E flowing in the emitter.

Relation among I_C, I_B and I_{CBO} From Eq. (2.5), we have

$$I_C = -\alpha I_E + I_{CBO}$$

Since I_C and I_E are flowing in opposite directions,

$$I_E = -(I_C + I_B)$$

Therefore,

$$I_C = -\alpha[-(I_C + I_B)] + I_{CBO}$$

$$I_C - \alpha I_C = \alpha I_B + I_{CBO}$$

$$I_C(1-\alpha) = \alpha I_B + I_{CBO}$$

$$I_C = \frac{\alpha}{1-\alpha} I_B + \frac{I_{CBO}}{1-\alpha}$$

Since

$$\beta = \frac{\alpha}{1-\alpha}, \tag{2.6}$$

the above expression becomes

$$I_C = (1 + \beta) I_{CBO} + \beta I_B \tag{2.7}$$

Relation among I_C, I_B and I_{CEO} In the common-emitter (CE) transistor circuit, I_B is the input current and I_C is the output current. If the base circuit is open, i.e., $I_B = 0$, then a small collector current flows from the collector to emitter. This is denoted as I_{CEO}, the collector-emitter current with base open. This current I_{CEO} is also called the collector to emitter leakage current.

In this CE configuration of the transistor, the emitter-base junction is forward-biased and collector-base junction is reverse-biased and hence the collector current I_C is the sum of the part of the emitter current I_E that reaches the collector, and the collector-emitter leakage current I_{CEO}. Therefore, the part of I_E, which reaches collector is equal to $(I_C - I_{CEO})$.

Hence, the *large-signal current gain* (β) is defined as,

$$\beta = \frac{(I_C - I_{CEO})}{I_B} \tag{2.8}$$

From the equation, we have

$$I_C = \beta I_B + I_{CEO} \tag{2.9}$$

Relation between I_{CBO} and I_{CEO}

Comparing Eqs (2.7) and (2.9), we get the relationship between the leakage currents of transistor common-base (CB) and common-emitter (CE) configurations as

$$I_{CEO} = (1 + \beta) I_{CBO} \tag{2.10}$$

From this equation, it is evident that the collector-emitter leakage current (I_{CEO}) in CE configuration is $(1 + \beta)$ times larger than that in CB configuration. As I_{CBO} is temperature-dependent, I_{CEO} varies by large amount when temperature of the junction changes.

Expression for emitter current The magnitude of emitter-current is

$$I_E = I_C + I_B$$

Substituting Eq. (2.7) in the above equation, we get

$$I_E = (1 + \beta)\, I_{CBO} + (1 + \beta)\, I_B \tag{2.11}$$

Substituting Eq. (2.6) into Eq. (2.11), we have

$$I_E = \frac{1}{1 - \alpha} I_{CBO} + \frac{1}{1 - \alpha} I_B \tag{2.12}$$

DC current gain (β_{DC} or h_{FE}) The DC current gain is defined as the ratio of the collector current I_C to the base current I_B. That is,

$$\beta_{\text{DC}} = h_{FE} = \frac{I_C}{I_B} \tag{2.13}$$

As I_C is large compared with I_{CEO}, the large signal current gain (β) and the DC current gain (h_{FE}) are approximately equal.

2.6 TYPES OF CONFIGURATION [AU May 2014, 16 marks]

When a transistor is to be connected in a circuit, one terminal is used as an input terminal, the other terminal is used as an output terminal and the third terminal is common to the input and output. Depending upon the input, output and common terminal, a transistor can be connected in three configurations. They are: (i) Common base (CB) configuration, (ii) Common emitter (CE) configuration, and (iii) Common collector (CC) configuration.

(i) CB configuration This is also called grounded base configuration. In this configuration, emitter is the input terminal, collector is the output terminal and base is the common terminal.

(ii) CE configuration This is also called grounded emitter configuration. In this configuration, base is the input terminal, collector is the output terminal and emitter is the common terminal.

(iii) CC configuration This is also called grounded collector configuration. In this configuration, base is the input terminal, emitter is the output terminal and collector is the common terminal.

The supply voltage connections for normal operation of an *NPN* transistor in the three configurations are shown in Fig. 2.6.

2.6.1 CB Configuration [AU Nov 2013, May 2013 and June 2011, 8 marks]

The circuit diagram for determining the static characteristics curves of an *NPN* transistor in the common base configuration is shown in Fig. 2.7.

Input characteristics To determine the input characteristics, the collector-base voltage V_{CB} is kept constant at zero volt and the emitter current I_E is increased from zero in suitable equal steps by increasing V_{EB}. This is repeated for higher fixed values of V_{CB}. A curve is drawn between emitter current I_E

and emitter-base voltage V_{EB} at constant collector-base voltage V_{CB}. The input characteristics thus obtained are shown in Fig. 2.8.

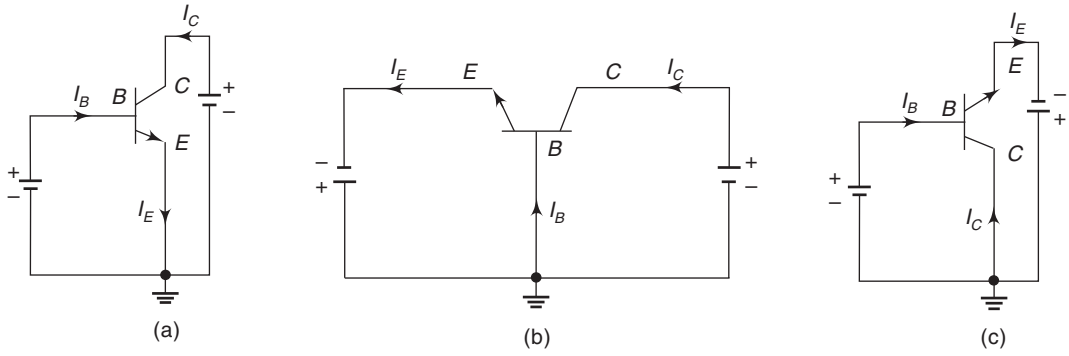


Fig. 2.6 *Transistor configuration: (a) Common base, (b) Common emitter, (c) Common collector*

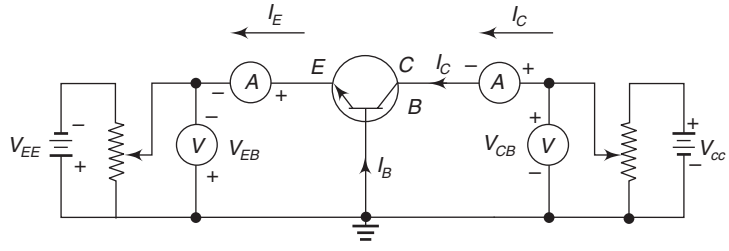


Fig. 2.7 *Circuit to determine CB static characteristics*

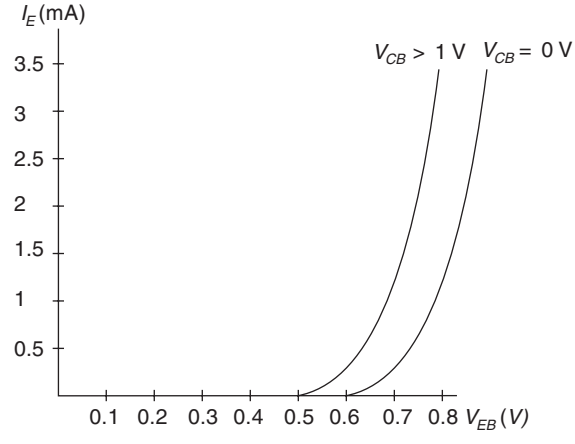


Fig. 2.8 *CB input characteristics*

When V_{CB} is equal to zero and the emitter-base junction is forward biased as shown in the characteristics, the junction behaves as a forward biased diode so that emitter current I_E increases rapidly with small increase in emitter-base voltage V_{EB}. When V_{CB} is increased keeping V_{EB} constant, the width of the base region will decrease. This effect results in an increase of I_E. Therefore, the curves shift towards the left as V_{CB} is increased.

Output characteristics To determine the output characteristics, the emitter current I_E is kept constant at a suitable value by adjusting the emitter-base voltage V_{EB}. Then V_{CB} is increased in suitable equal steps and the collector current I_C is noted for each value of I_E. This is repeated for different fixed values of I_E. Now the curves of I_C versus V_{CB} are plotted for constant values of I_E and the output characteristics thus obtained is shown in Fig. 2.9.

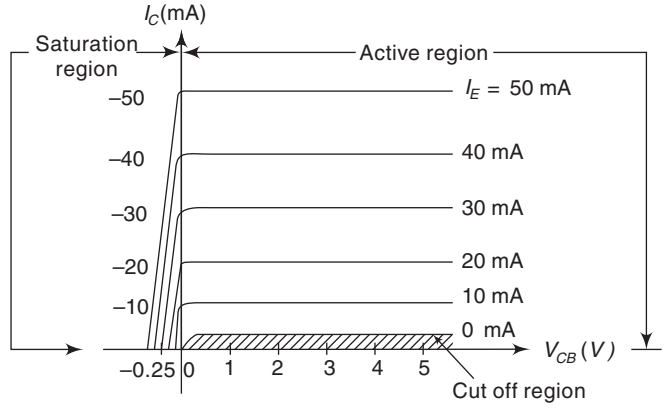


Fig. 2.9 *CB output characteristics*

From the characteristics, it is seen that for a constant value of I_E, I_C is independent of V_{CB} and the curves are parallel to the axis of V_{CB}. Further, I_C flows even when V_{CB} is equal to zero. As the emitter-base junction is forward biased, the majority carriers, i.e., electrons, from the emitter are injected into the base region. Due to the action of the internal potential barrier at the reverse biased collector-base junction, they flow to the collector region and give rise to I_C even when V_{CB} is equal to zero.

Early effect or base-width modulation **[AU Nov 2016, April 2015 and Nov 2014, 8 marks]**

As the collector voltage V_{CC} is made to increase the reverse bias, the space charge width between collector and base tends to increase, with the result that the effective width of the base decreases. This dependency of base-width on collector-to-emitter voltage is known as the *Early effect*. This decrease in effective base-width has three consequences:

(i) There is less chance for recombination within the base region. Hence, α increases with increasing $|V_{CB}|$.

(ii) The charge gradient is increased within the base, and consequently, the current of minority carriers injected across the emitter junction increases.

(iii) For extremely large voltages, the effective base-width may be reduced to zero, causing voltage breakdown in the transistor. This phenomenon is called the *punch through*.

For higher values of V_{CB}, due to Early effect, the value of α increases. For example, α changes, say from 0.98 to 0.985. Hence, there is a very small positive slope in the CB output characteristics and hence the output resistance is not zero.

Transistor parameters

The slope of the CB characteristics will give the following four transistor parameters. Since these parameters have different dimensions, they are commonly known as common base *hybrid parameters* or *h-parameters*.

(a) Input impedance (h_{ib}) It is defined as the ratio of the change in (input) emitter voltage to the change in (input) emitter current with the (output) collector voltage V_{CB} kept constant. Therefore,

$$h_{ib} = \frac{\Delta V_{EB}}{\Delta I_E},\ V_{CB} \text{ constant} \tag{2.14}$$

It is the slope of CB input characteristics I_E versus V_{EB} as shown in Fig. 2.8. The typical value of h_{ib} ranges from 20 Ω to 50 Ω.

(b) Output admittance (h_{ob}) It is defined as the ratio of change in the (output) collector current to the corresponding change in the (output) collector voltage with the (input) emitter current I_E kept constant. Therefore,

$$h_{ob} = \frac{\Delta I_C}{\Delta V_{CB}},\ I_E \text{ constant} \tag{2.15}$$

It is the slope of CB output characteristics I_C versus V_{CB} as shown in Fig. 2.9. The typical value of this parameter is of the order of 0.1 to 10 μ mhos.

(c) Forward current gain (h_{fb}) It is defined as a ratio of the change in the (output) collector current to the corresponding change in the (input) emitter current keeping the (output) collector voltage V_{CB} constant. Hence,

$$h_{fb} = \frac{\Delta I_C}{\Delta I_E},\ V_{CB} \text{ constant.} \tag{2.16}$$

It is the slope of I_C versus I_E curve. Its typical value varies from 0.9 to 1.0.

(d) Reverse voltage gain (h_{rb}) It is defined as the ratio of the change in the (input) emitter voltage and the corresponding change in (output) collector voltage with constant (input) emitter current, I_E. Hence,

$$h_{rb} = \frac{\Delta V_{EB}}{\Delta V_{CB}},\ I_E \text{ constant} \tag{2.17}$$

It is the slope of V_{EB} versus V_{CB} curve. Its typical value is of the order of 10^{-5} to 10^{-4}.

2.6.2 CE Configuration

[AU Nov 2016, May 2016, Dec 2015, May 2015, May 2014, Nov 2012 and June 2010, 10 marks]

Input characteristics To determine the input characteristics, the collector to emitter voltage is kept constant at zero volt and base current is increased from zero in equal steps by increasing V_{BE} in the circuit shown in Fig. 2.10.

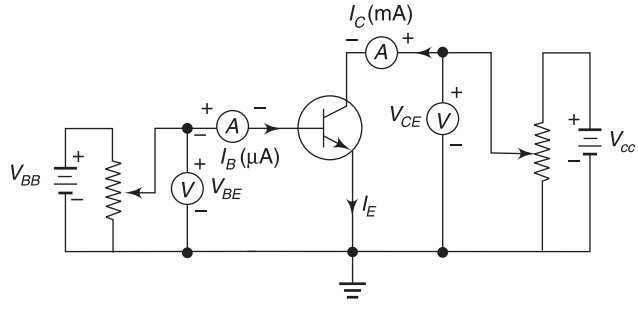


Fig. 2.10 *Circuit to determine CE static characteristics*

The value of V_{BE} is noted for each setting of I_B. This procedure is repeated for higher fixed values of V_{CE}, and the curves of I_B V_s. V_{BE} are drawn. The input characteristics thus obtained are shown in Fig. 2.11.

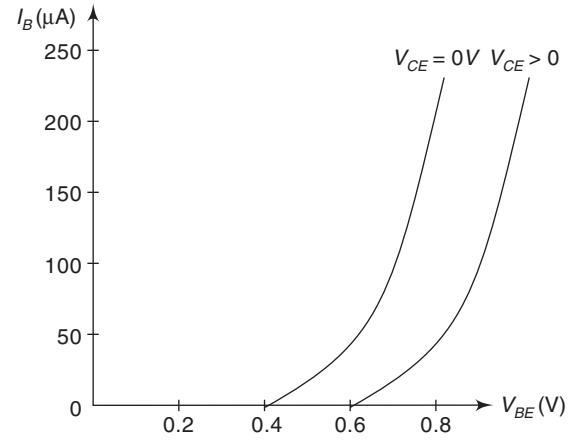


Fig. 2.11 *CE input characteristics*

When $V_{CE} = 0$, the emitter-base junction is forward biased and the junction behaves as a forward biased diode. Hence the input characteristic for $V_{CE} = 0$ is similar to that of a forward-biased diode. When V_{CE} is increased, the width of the depletion region at the reverse biased collector-base junction will increase. Hence the effective width of the base will decrease. This effect causes a decrease in the base current I_B. Hence, to get the same value of I_b as that for $V_{CE} = 0$, V_{BE} should be increased. Therefore, the curve shifts to the right as V_{CE} increases.

Output characteristics To determine the output characteristics, the base current I_B is kept constant at a suitable value by adjusting base-emitter voltage, V_{BE}. The magnitude of collector-emitter voltage V_{CE} is increased in suitable equal steps from zero and the collector current I_C is noted for each setting V_{CE}. Now the curves of I_C versus V_{CE} are plotted for different constant values of I_B. The output characteristics thus obtained are shown in Fig. 2.12.

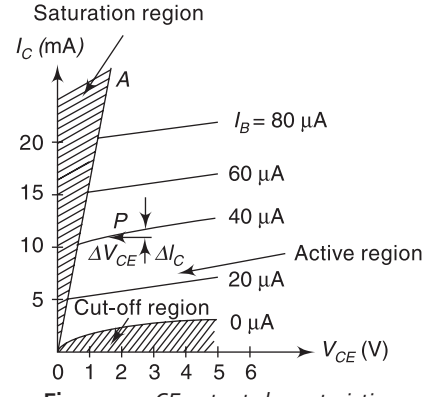


Fig. 2.12 *CE output characteristics*

From Eqs. (2.6) and (2.7), we have

$$\beta = \frac{\alpha}{1-\alpha} \quad \text{and} \quad I_C = (1+\beta)\, I_{CBO} + \beta I_B$$

For larger values of V_{CE}, due to Early effect, a very small change in a is reflected in a very large change in β. For example, when $\alpha = 0.98$, $\beta = \dfrac{0.98}{1-0.98} = 49$. If α increases to 0.985, then $\beta = \dfrac{0.985}{1-0.985} = 66$. Here, a slight increase in α by about 0.5% results in an increase in β by about 34%. Hence, the output characteristics of CE configuration show a larger slope when compared with CB configuration.

The output characteristics have three regions, namely, saturation region, cut-off region and active region. The region of curves to the left of the line *OA* is called the *saturation region* (hatched), and the line *OA* is called the saturation line. In this region, both junctions are forward biased and an increase in the base current does not cause a corresponding large change in I_C. The ratio of $V_{CE(\text{sat})}$ to I_C in this region is called saturation resistance.

The region below the curve for $I_B = 0$ is called the *cut-off region* (hatched). In this region, both junctions are reverse biased. When the operating point for the transistor enters the cut-off region, the transistor is OFF. Hence, the collector current becomes almost zero and the collector voltage almost equals V_{CC}, the collector supply voltage. The transistor is virtually an open circuit between collector and emitter.

The central region where the curves are uniform in spacing and slope is called the *active region* (unhatched). In this region, emitter-base junction is forward biased and the collector-base junction is reverse biased. If the transistor is to be used as a linear amplifier, it should be operated in the active region.

If the base current is subsequently driven large and positive, the transistor switches into the saturation region via the active region, which is traversed at a rate that is dependent on factors such as gain and frequency response. In this ON condition, large collector current flows and collector voltage falls to a very low value, called $V_{CE\text{sat}}$, typically around 0.2 V for a silicon transistor. The transistor is virtually a short circuit in this state.

High speed switching circuits are designed in such a way that transistors are not allowed to saturate, thus reducing switching times between ON and OFF times.

Transistor parameters **[AU Nov 2016, 8 marks]**

The slope of the CE characteristics will give the following four transistor parameters. Since these parameters have different dimensions, they are commonly known as common emitter *hybrid parameters* or *h-parameters.*

(a) Input impedance (h_{ie}) It is defined as the ratio of the change in (input) base voltage to the change in (input) base current with the (output) collector voltage V_{CE} kept constant. Therefore,

$$h_{ie} = \frac{\Delta V_{BE}}{\Delta I_B},\ V_{CE}\ \text{constant} \tag{2.18}$$

It is the slope of CE input characteristics I_B versus V_{BE} as shown in Fig. 2.11. The typical value of h_{ie} ranges from 500 to 2000 Ω.

(b) Output admittance (h_{oe}) It is defined as the ratio of change in the (output) collector current to the corresponding change in the (output) collector voltage with the (input) base current I_B kept constant. Therefore,

$$h_{oe} = \frac{\Delta I_C}{\Delta V_{CE}},\ I_B\ \text{constant} \tag{2.19}$$

It is the slope of CE output characteristic I_C versus V_{CE} as shown in Fig. 2.12. The typical value of this parameter is of the order of 0.1 to 10 μ mhos.

(c) Forward current gain (h_{fe}) It is defined as a ratio of the change in the (output) collector current to the corresponding change in the (input) base current keeping the (output) collector voltage V_{CE} constant. Hence,

$$h_{fe} = \frac{\Delta I_C}{\Delta I_B},\ V_{CE}\ \text{constant} \tag{2.20}$$

It is the slope of I_C versus I_B curve. Its typical value varies from 20 to 200.

(d) Reverse voltage gain (h_{re}) It is defined as the ratio of the change in the (input) base voltage and the corresponding change in (output) collector voltage with constant (input) base current, I_B. Hence,

$$h_{re} = \frac{\Delta V_{BE}}{\Delta V_{CE}},\ I_B\ \text{constant} \tag{2.21}$$

It is the slope of V_{BE} versus V_{CE} curve. Its typical value is of the order of 10^{-5} to 10^{-4}.

Derivation of analytical expression for CE output characteristics of BJT

The output characteristics of BJT in CE configuration are shown in Fig. 2.12. We know that, for a *PNP* transistor, the collector-emitter voltage V_{CE} is given by

$$V_{CE} = V_C - V_E$$

where $V_C = V_T \ln\left(1 - \frac{\alpha_N I_E + I_C}{I_{CO}}\right)$ and $V_E = V_T \ln\left(1 - \frac{\alpha_1 I_C + I_E}{I_{EO}}\right)$.

Therefore,

$$\begin{aligned} V_{CE} &= V_T \ln\left(1 - \frac{\alpha_N I_E + I_C}{I_{CO}}\right) - V_T \ln\left(1 - \frac{\alpha_1 I_C + I_E}{I_{EO}}\right) \\ &= V_T \ln\left(\frac{I_{CO} - \alpha_N I_E - I_C}{I_{CO}}\right) - V_T \ln\left(\frac{I_{EO} - \alpha_1 I_C - I_E}{I_{EO}}\right) \\ &= V_T \ln\left(\frac{I_{CO} - \alpha_N I_E - I_C}{I_{EO} - \alpha_1 I_C - I_E}\right)\left(\frac{I_{EO}}{I_{CO}}\right) \end{aligned}$$

We know that, $I_E = -(I_B + I_C)$.

Hence,

$$\begin{aligned} V_{CE} &= V_T \ln\left(\frac{\mathrm{I_{CO}} + \alpha_N I_B + \alpha_N I_C - I_C}{I_{EO} - \alpha_1 I_C + I_B + I_C}\right)\left(\frac{\alpha_1}{\alpha_N}\right) \qquad \left(\text{since } \frac{I_{EO}}{I_{CO}} = \frac{\alpha_1}{\alpha_N}\right) \\ &= V_T \ln\left(\frac{I_{CO} + \alpha_N I_B + (\alpha_N - 1) I_C}{I_{EO} + I_B + (1 - \alpha_1) I_C}\right)\left(\frac{\alpha_1}{\alpha_N}\right) \\ &= V_T \ln\frac{\alpha_1}{\alpha_N} + V_T \ln\left(\frac{I_{CO} + \alpha_N I_B - I_C(1 - \alpha_N)}{I_{EO} + I_B + I_C(1 - \alpha_1)}\right) \end{aligned}$$

If $I_B >> I_{EO}$ and $I_B >> \frac{I_{CO}}{\alpha_N}$, we have

$$\begin{aligned} V_{CE} &= V_T \ln\left[\left(\frac{\alpha_1}{\alpha_N}\right)\left(\frac{\alpha_N I_B - I_C(1 - \alpha_N)}{I_B + I_C(1 - \alpha_1)}\right)\right] \\ &= V_T \ln\left[\frac{I_B - \left(\frac{1 - \alpha_N}{\alpha_N}\right) I_C}{\frac{I_B}{\alpha_1} + \left(\frac{1 - \alpha_1}{\alpha_1}\right) I_C}\right] \end{aligned}$$

Now dividing both numerator and denominator by I_B,we get

$$V_{CE} = V_T \ln \left[\frac{1 - \left(\frac{1-\alpha_N}{\alpha_N} \right) \frac{I_C}{I_B}}{\frac{1}{\alpha_1} + \left(\frac{1-\alpha_1}{\alpha_1} \right) \frac{I_C}{I_B}} \right]$$

$$= V_T \ln \left[\frac{1 - \frac{1}{\beta_N} \frac{I_C}{I_B}}{\frac{1}{\alpha_1} + \frac{1}{\beta_1} \frac{I_C}{I_B}} \right] \qquad \left(\text{since} \quad \beta = \frac{\alpha}{1-\alpha} \right)$$

or

$$V_{CE} = -V_T \ln \left[\frac{\frac{1}{\alpha_1} + \frac{1}{\beta_1} \frac{I_C}{I_B}}{1 - \frac{1}{\beta_N} \frac{I_C}{I_B}} \right]$$

Here, if $I_C = 0$, $V_{CE} = -V_T \ln \left(\frac{1}{\alpha_1} \right)$. This shows that V_{CE} is not equal to zero even if I_C becomes zero and hence, the common emitter characteristic curves do not pass through the origin.

2.6.3 CC Configuration

The circuit diagram for determining the static characteristics of an *NPN* transistor in the common collector configuration is shown in Fig. 2.13.

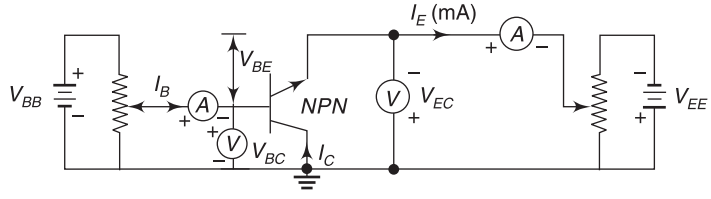


Fig. 2.13 *Circuit to determine CC static characteristics*

Input characteristics To determine the input characteristics, V_{EC} is kept at a suitable fixed value. The base-collector voltage V_{BC} is increased in equal steps and the corresponding increase in I_B is noted. This is repeated for different fixed values of V_{EC}. Plots of V_{BC} versus I_B for different values of V_{EC} shown in Fig. 2.14 are the input characteristics.

Output characteristics The output characteristics shown in Fig. 2.15 are the same as those of the common emitter configuration.

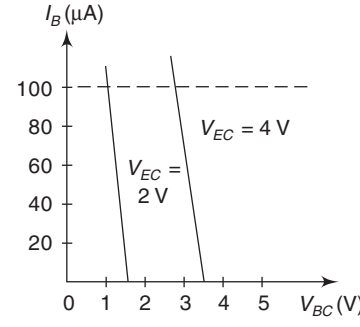


Fig. 2.14 *CC input characteristics*

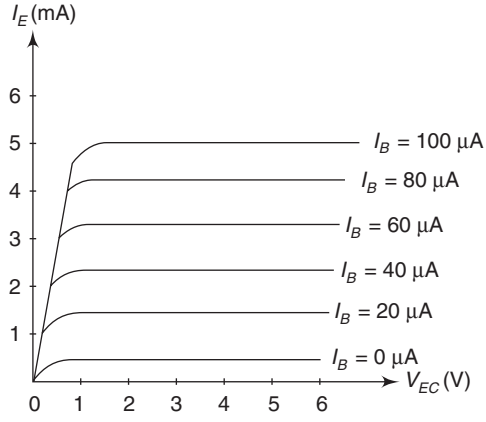


Fig. 2.15 *CC output characteristics*

2.6.4 Comparison [AU May 2014, May 2013, Nov 2012 and Nov 2011 , 6 marks]

Table 2.1 *A comparison of CB, CE and CC configurations*

Property	*CB*	*CE*	*CC*
Input resistance	Low (about 100 Ω)	Moderate (about 750 Ω)	High (about 750 kΩ)
Output resistance	High (about 450 kΩ)	Moderate (about 45 kΩ)	Low (about 25 Ω)
Current gain	1	High	High
Voltage gain	About 150	About 500	Less than 1
Phase shift between input and output voltages	0 or 360°	180°	0 or 360°
Applications	For high frequency circuits	For audio frequency circuits	For impedance matching

2.6.5 Current Amplification Factor [AU May 2017, Dec 2015, 8 marks]

In a transistor amplifier with AC input signal, the ratio of change in output current to the change in input current is known as the current amplification factor.

In the CB configuration the current amplification factor, $\alpha = \dfrac{\Delta I_C}{\Delta I_E}$ (2.22)

In the CE configuration the current amplification factor, $\beta = \dfrac{\Delta I_C}{\Delta I_B}$ (2.23)

In the CC configuration the current amplification factor, $\gamma = \dfrac{\Delta I_E}{\Delta I_B}$ (2.24)

Relationship between α and β We know that $\Delta I_E = \Delta I_C + \Delta I_B$

By definition, $\Delta I_C = \alpha \Delta I_E$

Therefore, $\Delta I_E = \alpha \Delta I_E + \Delta I_B$

i.e., $\Delta I_B = \Delta I_E (1 - \alpha)$

Dividing both sides by ΔI_C, we get

$$\frac{\Delta I_B}{\Delta I_C} = \frac{\Delta I_E}{\Delta I_C}(1 - \alpha)$$

Therefore,
$$\frac{1}{\beta} = \frac{1}{\alpha}(1 - \alpha)$$
$$\beta = \frac{\alpha}{(1 - \alpha)}$$

Rearranging, we also get $\alpha = \frac{\beta}{(1 + \beta)}$, or $\frac{1}{\alpha} - \frac{1}{\beta} = 1$ (2.25)

From this relationship, it is clear that as α approaches unity, β approaches infinity. The CE configuration is used for almost all transistor applications because of its high current gain, β.

Relation among α, β and γ In the CC transistor amplifier circuit, I_B is the input current and I_E is the output current.

From Eq. (2.22), $\gamma = \frac{\Delta I_E}{\Delta I_B}$

Substituting $\Delta I_B = \Delta I_E - \Delta I_C$, we get

$$\gamma = \frac{\Delta I_E}{\Delta I_E - \Delta I_C}$$

Dividing the numerator and denominator on RHS by ΔI_E, we get

$$\gamma = \frac{\frac{\Delta I_E}{\Delta I_E}}{\frac{\Delta I_E}{\Delta I_E} - \frac{\Delta I_C}{\Delta I_E}} = \frac{1}{1 - \alpha}$$

Therefore, $\gamma = \frac{1}{1 - \alpha} = (\beta + 1)$ (2.26)

2.7 TRANSISTOR AS AN AMPLIFIER

(a) CE transistor as an amplifier **[AU May 2017, 6 marks, Nov 2011, 12 marks]**

Figure 2.16(a) shows an amplifier circuit using CE transistor configuration. In this circuit, an NPN transistor is used in CE configuration. Here, V_{BB} supply will forward bias the emitter-base junction and V_{CC} supply will reverse bias the collector-base junction. This biasing arrangement makes the transistor to operate in the active region. The magnitude of the input ac signal v_i always forward bias the emitter-base junction regardless of the polarity of the signal.

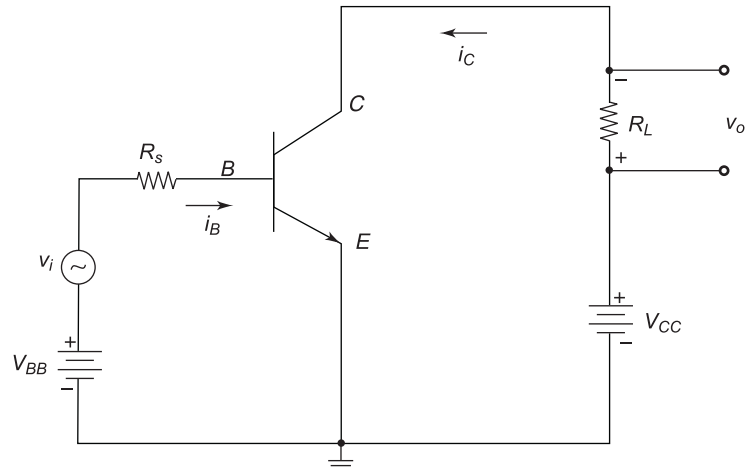


Fig. 2.16(a) *CE transistor as an amplifier*

During the positive half cycle of the input signal v_i, the forward bias across the emitter-base junction is increased. As a result, more electrons are injected into the base and reaches the collector, resulting in an increase in collector current i_c. This increase in collector current produces a greater voltage drop across the load resistance R_L.

However, during the negative half cycle of the input signal v_i, the forward bias across the emitter-base junction is decreased, resulting in a decrease in collector current i_c. This decrease in collector current produces a smaller voltage drop across the load resistance R_L. Hence, it is clear that a small change in the input ac signal in CE transistor amplifier produces a large change at the output with a voltage gain of around 500 and a phase shift of 180°. Here, the voltage gain is the ratio of output voltage to input voltage. Comparing to CB and CC transistor configurations, this CE transistor configuration is widely used in amplifier circuits due to its high voltage gain.

(b) CB transistor as an amplifier

A load resistor R_L is connected in series with the collector supply voltage V_{CC} of CB transistor configuration as shown in Fig. 2.16(b).

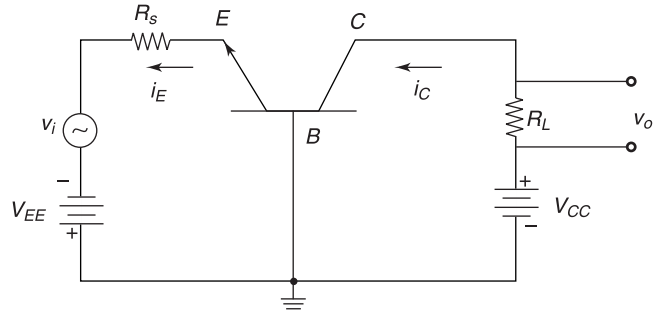


Fig. 2.16(b) *CB transistor as an amplifier*

A small change in the input voltage between emitter and base, say ΔV_i, causes a relatively larger change in emitter current, say ΔI_E. A fraction of this change in current is collected and passed through R_L and

is denoted by symbol α'. Therefore, the corresponding change in voltage across the load resistor R_L due to this current is $\Delta V_0 = \alpha' R_L \Delta I_E$.

In CB configuration, the voltage amplification $A_v = \dfrac{\Delta V_0}{\Delta V_i}$ is around 150 with zero phase shift and thus the transistor acts as an amplifier.

2.8 LARGE SIGNAL, DC AND SMALL SIGNAL CE VALUES OF CURRENT GAIN

We know from the characteristics of CE configuration, the current amplification factor is $\beta \equiv \dfrac{\alpha}{1-\alpha}$,

and $$I_C = (1 + \beta) I_{CBO} + \beta I_B$$

The above equation can be expressed as

$$I_C - I_{CBO} = \beta I_{CBO} + \beta I_B$$

Therefore, $$\beta = \frac{I_C - I_{CBO}}{I_B - (-I_{CBO})}$$

The output characteristics of CE configuration show that in the *cut-off* region, the values $I_E = 0$, $I_C = I_{CBO}$ and $I_B = -I_{CBO}$. Therefore, the above equation gives the ratio of the collector-current increment to the base-current change from cut-off to I_B, and hence β *is called the large-signal current gain of common-emitter transistor*.

The DC current gain of the transistor is given by

$$\beta_{\text{DC}} \equiv h_{FE} \equiv \frac{I_C}{I_B}$$

Based on this h_{FE} value, we can determine whether the transistor is in saturation or not. For any transistor, in general, I_B is large compared to I_{CBO}. Under this condition, the value of $h_{FE} \approx \beta$.

The small-signal CE forward short-circuit gain β' is defined as the ratio of a collector-current increment ΔI_C for a small base-current change ΔI_B, at a fixed collector-to-emitter voltage V_{CE}.

i.e., $$\beta' \equiv \left.\frac{\partial I_C}{\partial I_B}\right|_{V_{CE}}$$

If β is independent of currents, then $\beta' = \beta \approx h_{FE}$. However, β is a function of current, then $\beta' = \beta + (I_{CBO} + I_B)\dfrac{\partial \beta}{\partial I_B}$. By using $\beta' = h_{fe}$ and $\beta \approx h_{FE}$. Therefore, the above equation becomes

$$h_{fe} = \frac{h_{FE}}{1 - (I_{CBO} + I_B)\dfrac{\partial h_{FE}}{\partial I_C}}$$

In Fig. 2.17, the h_{FE} versus I_C shows a maximum and hence $h_{fe} > h_{FE}$ for smaller currents, and $h_{fe} < h_{FE}$ for larger currents. Therefore, the above equation is valid only for the active region.

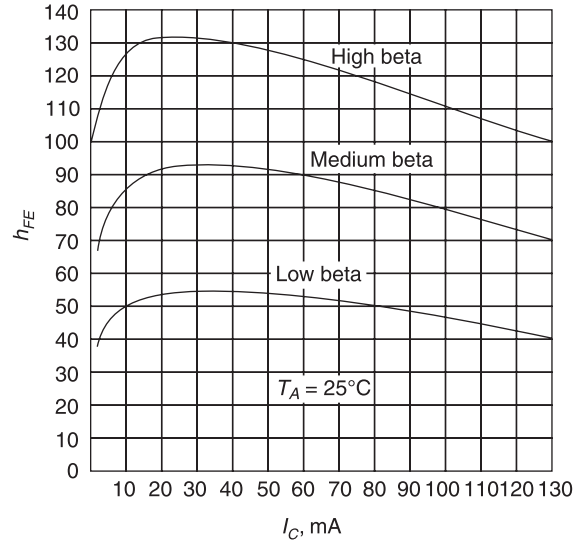


Fig. 2.17 *Characteristic curves of DC current gain h_{FE} (at $V_{CE} = -0.25$ V) versus collector current for low, medium and high beta values*

EXAMPLE 2.1

In a common-base transistor circuit, the emitter current I_E is 10 mA and the collector current I_C is 9.8 mA. Find the value of the base current I_B.

Solution Given

$$I_E = 10 \text{ mA and } I_C = 9.8 \text{ mA}$$

We know that emitter current is

$$I_E = I_B + I_C$$

i.e., $$10 \times 10^{-3} = I_B + 9.8 \times 10^{-3}$$

Therefore, $$I_B = 0.2 \text{ mA}$$

EXAMPLE 2.2

In a common-base connection, the emitter current I_E is 6.28 mA and the collector current I_C is 6.20 mA. Determine the common-base DC current gain.

Solution Given $$I_E = 6.28 \text{ mA and } I_C = 6.20 \text{ mA}$$

We know that common-base DC current gain,

$$\alpha = \frac{I_C}{I_E} = \frac{6.20 \times 10^{-3}}{6.28 \times 10^{-3}} = 0.987$$

EXAMPLE 2.3

The common-base DC current gain of a transistor is 0.967. If the emitter current is 10 mA, what is the value of base current?

Solution Given $\alpha = 0.967$ and $I_E = 10$ mA

The common-base DC current gain (α) is

$$\alpha = 0.967 = \frac{I_C}{I_E} = \frac{I_C}{10 \times 10^{-3}}$$

Therefore, $I_C = 0.967 \times 10 \times 10^{-3} = 9.67$ mA

The emitter current $I_E = I_B + I_C$

i.e., $10 \times 10^{-3} = I_B + 9.67 \times 10^{-3}$

Therefore, $I_B = 0.33$ mA

EXAMPLE 2.4

The transistor has $I_E = 10$ mA and $\alpha = 0.98$. Determine the values of I_C and I_B.

Solution Given $I_E = 10$ mA and $\alpha = 0.98$

The common-base DC current gain, $\alpha = \frac{I_C}{I_E}$

i.e., $0.98 = \frac{I_C}{10 \times 10^{-3}}$

Therefore, $I_C = 0.98 \times 10 \times 10^{-3} = 9.8$ mA

The emitter current $I_E = I_B + I_C$

i.e., $10 \times 10^{-3} = I_B + 9.8 \times 10^{-3}$

Therefore, $I_B = 0.2$ mA

EXAMPLE 2.5

If a transistor has a α of 0.97, find the value of β. If $\beta = 200$, find the value of α.

Solution If $\alpha = 0.97$, $\beta = \frac{\alpha}{1-\alpha} = \frac{0.97}{1-0.97} = 32.33$

If $\beta = 200$, $\alpha = \frac{\beta}{\beta+1} = \frac{200}{200+1} = 0.995$

EXAMPLE 2.6

A transistor has $\beta = 100$. If the collector current is 40 mA, find the value of emitter current.

Solution Given $\beta = 100$ and $I_C = 40$ mA

$$\beta = 100 = \frac{I_C}{I_B} = \frac{40 \times 10^{-3}}{I_B}$$

Therefore, $I_B = 40 \times 10^{-3}/100 = 0.4$ mA and

$$I_E = I_B + I_C = (0.4 + 40) \times 10^{-3} = 40.4 \text{ mA}$$

EXAMPLE 2.7

A transistor has $\beta = 150$. Find the collector and base currents, if $I_E = 10$ mA.

Solution Given $\beta = 150$ and $I_E = 10$ mA

The common-base current gain, $\alpha = \dfrac{\beta}{\beta + 1} = \dfrac{150}{150 + 1} = 0.993$

Also, $\alpha = \dfrac{I_C}{I_E}$

i.e., $0.993 = \dfrac{I_C}{10}$

Therefore, $I_C = 0.993 \times 10 \times 10^{-3} = 9.93$ mA

The emitter current $I_E = I_B + I_C$

i.e., $10 \times 10^{-3} = I_B + 9.93 \times 10^{-3}$

Therefore, $I_B = (10 - 9.93) \times 10^{-3} = 0.07$ mA

EXAMPLE 2.8

Determine the values of I_B and I_E for the transistor circuit if $I_C = 80$ mA and $\beta = 170$.

Solution Given $\beta = 170$ and $I_C = 80$ mA

We know that (β), $\beta = 170 = \dfrac{I_C}{I_B} = \dfrac{80 \times 10^{-3}}{I_B}$

Therefore, $I_B = \dfrac{80 \times 10^{-3}}{170} = 0.47$ mA

and $I_E = I_B + I_C = (0.47 + 80)$ mA $= 80.47$ mA

EXAMPLE 2.9

Determine the values of I_C and I_E for the transistor circuit of $\beta = 200$ and $I_B = 0.125$ mA.

Solution Given $I_B = 0.125$ mA and $\beta = 200$

Therefore, $\beta = 200 = \dfrac{I_C}{I_B} = \dfrac{I_C}{0.125 \times 10^{-3}}$

Therefore, $I_C = 200 \times 0.125 \times 10^{-3} = 25$ mA

and $I_E = I_B + I_C = (0.125 + 25) \times 10^{-3} = 25.125$ mA

EXAMPLE 2.10

Determine the values of I_C and I_B for the transistor circuit of $I_E = 12$ mA and $\beta = 100$.

Solution Given $I_E = 12$ mA and $\beta = 100$

We know that base current, $I_B = \dfrac{I_E}{1+\beta} = \dfrac{12 \times 10^{-3}}{1 + 100} = 0.1188$ mA

and collector current, $I_C = I_E - I_B = (12 - 0.1188) \times 10^{-3} = 11.8812$ mA

EXAMPLE 2.11

A transistor has $I_B = 100$ μA and $I_C = 2$ mA. Find (a) β of the transistor, (b) α of the transistor, (c) emitter current I_E, (d) if I_B changes by + 25 μA and I_C changes by + 0.6 mA, find the new value of β.
(AU April/May 2017)

Solution Given $I_B = 100\ \mu\text{A} = 100 \times 10^{-6}$ A and $I_C = 2$ mA $= 2 \times 10^{-3}$ A.

(a) To find β of the transistor

$$\beta = \frac{I_C}{I_B} = \frac{2 \times 10^{-3}}{100 \times 10^{-6}} = 20$$

(b) To find α of the transistor

$$\alpha = \frac{\beta}{\beta + 1} = \frac{20}{1 + 20} = 0.952$$

(c) To find emitter current, I_E

$$I_E = I_B + I_C = 100 \times 10^{-6} + 2 \times 10^{-3}\ \text{A}$$
$$= (0.01 + 2) \times 10^{-3} = 2.01 \times 10^{-3}\ \text{A} = 2.01\ \text{mA}$$

(d) To find the new value of β when $\Delta I_B = 25\ \mu\text{A}$ and $\Delta I_C = 0.6$ mA

Therefore, $I_B = (100 + 25)\ \mu\text{A} = 125\ \mu\text{A}$

$I_C = (2 + 0.6)$ mA $= 2.6$ mA

The new value of β of the transistor,

$$\beta = \frac{I_C}{I_B} = \frac{2.6 \times 10^{-3}}{125 \times 10^{-6}} = 20.8$$

EXAMPLE 2.12

For a transistor circuit having $\alpha = 0.98$, $I_{CBO} = I_{CO} = 5$ μA and $I_B = 100$ μA, find I_C and I_E.

Solution Given $\alpha = 0.98$, $I_{CBO} = I_{CO} = 5\ \mu\text{A}$ and $I_B = 100\ \mu\text{A}$

The collector current is

$$I_C = \frac{\alpha \cdot I_B}{1-\alpha} + \frac{I_{CO}}{1-\alpha} = \frac{0.98 \times 100 \times 10^{-6}}{1-0.98} + \frac{5 \times 10^{-6}}{1-0.98} = 5.15 \text{ mA}$$

The emitter current is

$$I_E = I_B + I_C = 100 \times 10^{-6} + 5.15 \times 10^{-3} = 5.25 \text{ mA}$$

EXAMPLE 2.13

A germanium transistor used in a complementary symmetry amplifier has I_{CBO} = 10 μA at 27°C and h_{FE} = 50. (a) Find I_C when I_B = 0.25 mA and (b) assuming h_{FE} does not increase with temperature, find the value of new collector current, if the transistor's temperature rises to 50°C.

Solution Given $I_{CBO} = 10$ μA and $h_{FE} (= \beta) = 50$

(a) To find the value of collector current when I_B = 0.25 mA

$$I_C = \beta I_B + (1 + \beta) I_{CBO}$$
$$= 50 \times (0.25 \times 10^{-3}) + (1 + 50) \times (10 \times 10^{-6}) = 13.01 \text{ mA}$$

(b) To find the value of new collector current if temperature rises to 50°C

We know that I_{CBO} doubles for every 10°C rise in temperature. Therefore,

$$I'_{CBO} (\beta = 50) = I_{CBO} \times 2^{(T_2 - T_1)/10} = 10 \times 2^{(50-27)/10} \text{ μA}$$
$$= 10 \times 2^{2.3} \text{ μA} = 49.2 \text{ μA}$$

Therefore, the collector current at 50°C is

$$I_C = \beta \cdot I_B + (1 + \beta) I'_{CBO}$$
$$= 50 \times (0.25 \times 10^{-3}) + (1 + 50) \times 49.2 \times 10^{-6} \text{ A} = 15.01 \text{ mA}$$

EXAMPLE 2.14

When the emitter current of a transistor is changed by 1 mA, there is a change in collector current by 0.99 mA. Find the current gain of the transistor.

Solution The current gain of the transistor is $\alpha = \frac{\Delta I_C}{\Delta I_E} = \frac{0.99 \times 10^{-3}}{1 \times 10^{-3}} = 0.99$

EXAMPLE 2.15

The DC current gain of a transistor in CE mode is 100. Determine its DC current gain in CB mode.

Solution The DC current gain of the transistor in CB mode is $\alpha_{DC} = \frac{\beta_{DC}}{1 + \beta_{DC}}$

$$= \frac{100}{1 + 100} = 0.99$$

EXAMPLE 2.16

When I_E of a transistor is changed by 1 mA, its I_C changes by 0.995 mA. Find its common base current gain α, and common-emitter current gain β.

Solution Common-base current gain is $\alpha = \dfrac{\Delta I_C}{\Delta I_E} = \dfrac{0.995 \times 10^{-3}}{1 \times 10^{-3}} = 0.995$

Common-emitter current gain is $\beta = \dfrac{\alpha}{1 - \alpha} = \dfrac{0.995}{1 - 0.995} = 199$

EXAMPLE 2.17

The current gain of a transistor in CE mode is 49. Calculate its common-base current gain. Find the base current when the emitter current is 3 mA.

Solution Given $\beta = 49$

We know that $\alpha = \dfrac{\beta}{1 + \beta}$

Therefore, the common-base current gain is $\alpha = \dfrac{49}{1 + 49} = 0.98$

We also know that $\alpha = \dfrac{I_C}{I_E}$

Therefore, $I_C = \alpha I_E = 0.98 \times 3 \times 10^{-3} = 2.94$ mA

EXAMPLE 2.18

Determine I_C, I_E and α for a transistor circuit having I_B = 15 μA and β = 150.

Solution The collector current, $I_C = \beta I_B = 150 \times 15 \times 10^{-6} = 2.25$ mA

The emitter current, $I_E = I_C + I_B = 2.25 \times 10^{-3} + 15 \times 10^{-6} = 2.265$ mA

Common-base current gain, $\alpha = \dfrac{\beta}{1 + \beta} = \dfrac{150}{151} = 0.9934$

EXAMPLE 2.19

Determine the base, collector and emitter currents and V_{CE} for a CE circuit shown in Fig. 2.18. For V_{CC} = 10 V, V_{BB} = 4 V, R_B = 200 kΩ, R_C = 2 kΩ, V_{BE} (on) = 0.7 V, β = 200.

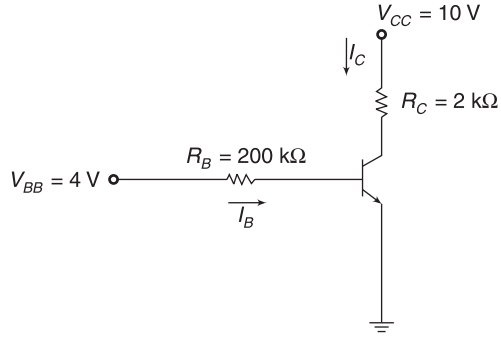


Fig. 2.18

Solution Referring to Fig. 2.18, the base current is

$$I_B = \frac{V_{BB} - V_{BE}(on)}{R_B} = \frac{4 - 0.7}{200 \times 10^3} = 16.5\ \mu A$$

The collector current is

$$I_C = \beta I_B = 200 \times 16.5 \times 10^{-6} = 3.3\ mA$$

The emitter current is

$$I_E = I_C + I_B = 3.3 \times 10^{-3} + 16.5 \times 10^{-6} = 3.3165\ mA$$

Therefore, $V_{CE} = V_{CC} - I_C R_C = 10 - 3.3 \times 10^{-3} \times 2 \times 10^3 = 3.4\ V$

EXAMPLE 2.20

Calculate the values of I_C and I_E for a transistor with $\alpha_{DC} = 0.99$ and $I_{CBO} = 5\ \mu A$. I_B is measured as 20 μA.

Solution Given $\alpha_{DC} = 0.99$, $I_{CBO} = 5\ \mu A$ and $I_B = 20\ \mu A$

$$I_C = \frac{\alpha_{DC} I_B}{1 - \alpha_{DC}} + \frac{I_{CBO}}{1 - \alpha_{DC}} = \frac{0.99 \times 20 \times 10^{-6}}{1 - 0.99} + \frac{5 \times 10^{-6}}{1 - 0.99} = 2.48\ mA$$

Therefore, $I_E = I_B + I_C = 20 \times 10^{-6} + 2.48 \times 10^{-3} = 2.5\ mA$

EXAMPLE 2.21

The reverse leakage current of the transistor when connected in CB configuration is 0.2 μA and it is 18 μA when the same transistor is connected in CE configuration. Calculate α_{DC} and β_{DC} of the transistor.
(AU May/June 2016)

Solution The leakage current $I_{CBO} = 0.2\ \mu A$

$$I_{CEO} = 18\ \mu A$$

Assume that $I_B = 30\ \mu A$

$$I_E = I_B + I_C$$

$$I_C = I_E - I_B = \beta I_B + (1 + \beta)\, I_{CBO}$$

We know that $I_{CEO} = \frac{I_{CBO}}{1 - \alpha} = (1 + \beta) I_{CBO}$

$$\beta = \frac{I_{CEO}}{I_{CBO}} - 1 = \frac{18}{0.2} - 1 = 89$$

$$I_c = \beta I_B + (1 + \beta) I_{CBO}$$

$$= 89\ (30 \times 10^{-6}) + (1 + 89)\ (0.2 \times 10^{-6})$$

$$= 2.418\ mA$$

$$\alpha_{DC} = 1 - \frac{I_{CBO}}{I_{CEO}} = 1 - \frac{0.2 \times 10^{-6}}{18 \times 10^{-6}} = 0.988$$

$$\beta_{DC} = \frac{I_C - I_{CBO}}{I_B - I_{CEO}}$$

$$= \frac{2.418 \times 10^{-3} - 0.2 \times 10^{-6}}{30 \times 10^{-3} - 18 \times 10^{-6}} = 201$$

EXAMPLE 2.22

If $\alpha_{DC} = 0.99$ and $I_{CBO} = 50\ \mu A$, find emitter current.

Solution Given $\alpha_{DC} = 0.99$ and $I_{CBO} = 50\ \mu A$,

Assume that $I_B = 1$ mA

$$I_C = \frac{\alpha_{DC} I_B}{1 - \alpha_{DC}} + \frac{I_{CBO}}{1 - \alpha_{DC}} = \frac{0.99\ (1 \times 10^{-3})}{1 - 0.99} + \frac{50 \times 10^{-6}}{1 - 0.99}$$

$$= \frac{0.99 \times 10^{-3}}{0.01} + \frac{50 \times 10^{-6}}{0.01} = 99\ \text{mA} + 5\ \text{mA} = 104\ \text{mA}$$

$$I_E = I_C + I_B = 104\ \text{mA} + 1\ \text{mA} = 105\ \text{mA}$$

EXAMPLE 2.23

For the CE amplifier circuit shown in Fig. 2.19, find the percentage change in collector current if the transistor with $h_{fc} = 50$ is replaced by another transistor with $h_{fc} = 150$. Assume $V_{BE} = 0.6$ V.

Solution

$$V_B = \frac{R_2}{R_1 + R_2} \times V_{CC} = \frac{5 \times 10^3}{5 \times 10^3 + 25 \times 10^3} \times 12 = 2\ \text{V}$$

$$V_E = V_B - V_{BE} = 2 - 0.6 = 1.4\ \text{V}$$

Here, $I_E = \frac{V_E}{R_E} = \frac{1.4}{100} = 14\ \text{mA}$

For $\beta = 50$, $I_B = \frac{I_E}{1 + \beta} = \frac{14 \times 10^{-3}}{51} = 274.5\ \mu A$

Therefore, $I_{C1} = \beta I_B = 50 \times 274.5 \times 10^{-6} = 13.725\ \text{mA}$

For $\beta = 150$, $I_B = \frac{I_E}{1 + \beta} = \frac{14 \times 10^{-3}}{151} = 92.715\ \text{mA}$

Therefore, $I_{C2} = \beta I_B = 150 \times 92.715 \times 10^{-6} = 13.907\ \text{mA}$

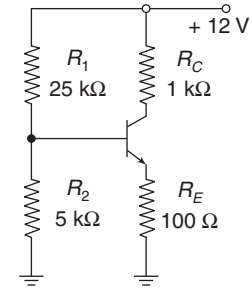


Fig. 2.19

Hence, the percentage change in collector current is calculated as

$$\frac{I_{C2} - I_{CE}}{I_{C1}} \times 100 = \frac{13.907 \times 10^{-3} - 13.725 \times 10 \times 10^{-3}}{13.725 \times 10^{-3}} \times 100 = 1.326\%$$

EXAMPLE 2.24

Given an *NPN* transistor for which $\alpha = 0.98$, $I_{CO} = 2\ \mu A$ and $I_{CEO} = 16\ \mu A$. A common emitter connection is used as shown in Fig. 2.20 with $V_{CC} = 12$ V and $R_C = 4\ k\Omega$. What is the minimum base current required in order that transistor enter into saturation region.

Solution Given $\alpha = 0.98$, $I_{CO} = 2\ \mu A$, $I_{CEO} = 1.6\ \mu A$, $V_{CC} = 12$ V and $R_C = 4\ k\Omega$.

$$I_{C(sat)} = \frac{V_{CC}}{R_C} = \frac{12}{4 \times 10^3} = 3\ mA$$

$$\beta = \frac{\alpha}{1-\alpha} = \frac{0.98}{1-0.98} = 49$$

$$I_{B(min)} = \frac{I_{C(sat)}}{\beta}$$

$$= \frac{3 \times 10^{-3}}{49} = 61.224 \times 10^{-6} = 61.224\ \mu A$$

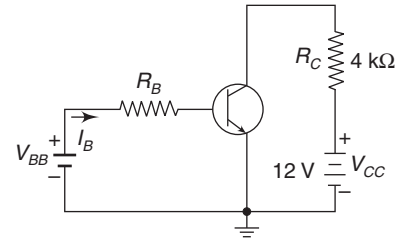


Fig. 2.20

EXAMPLE 2.25

A transistor operating in *CB* configuration has $I_C = 2.98$ mA, $I_E = 3$ mA and $I_{CO} = 0.01$ mA . What current will flow in the collector circuit of this transistor when connected in *CE* configuration with a base current of 30 μA?

Solution

Given $I_C = 2.98$ mA, $I_E = 3$ mA, $I_{CO} = 0.01$ mA and $I_B = 30\ \mu A$.

For CB configuration, $I_C = \alpha I_E + I_{CO}$

Therefore, $$\alpha = \frac{I_C - I_{CO}}{I_E} = \frac{(2.98 - 0.01) \times 10^{-3}}{3 \times 10^{-3}} = 0.99$$

$$\beta = \frac{\alpha}{1-\alpha} = \frac{0.99}{1-0.99} = 99$$

For CE configuration, $I_C = \beta I_B + (1 + \beta)\, I_{CO}$

$$= 99 \times 30 \times 10^{-6} + (1 + 99) \times 0.01 \times 10^{-3} = 3.97\ mA.$$

2.9 VOLTAGE-DIVIDER BIAS OR SELF-BIAS [AU May 2014, 16 marks]

A simple circuit used to establish a stable operating point is the self-biasing configuration. The self-bias, also called emitter bias, or emitter resistor, and potential divider circuit that can be used for low collector resistance is shown in Fig. 2.21. The current in the emitter resistor R_E causes a voltage drop

which is in the direction to reverse bias the emitter junction. For the transistor to remain in the active region, the base-emitter junction has to be forward biased. The required base bias is obtained from the power supply through the potential divider network of the resistances R_1 and R_2.

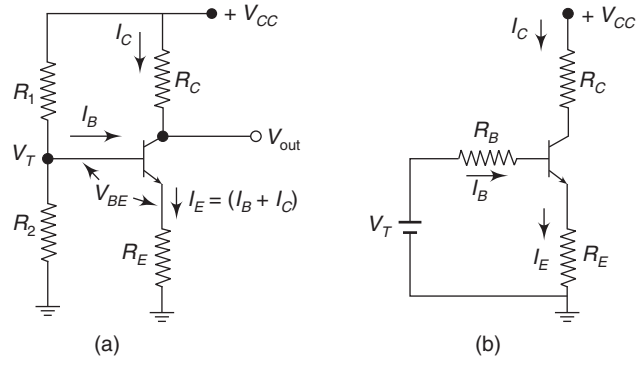


Fig. 2.21 *(a) Self-bias circuit (b) Thevenin's equivalent circuit*

To determine stability factor, S

The extent to which the collector current I_C is stabilized with varying I_{CO} is measured by a stability factor S. It is defined as the rate of change of collector current I_C with respect to the collector-base leakage current I_{CO}, keeping both the current I_B and the current gain β constant.

$$S = \frac{\partial I_C}{\partial I_{CO}} \approx \frac{dI_C}{dI_{CO}} = \frac{\Delta I_C}{\Delta I_{CO}}, \text{ where } \beta \text{ and } I_B \text{ constant}$$

The collector current for a CE amplifier is given by

$$I_C = \beta I_B + (\beta + 1) I_{CO}$$

Differentiating the above equation with respect to I_C, we get

$$1 = \beta \frac{dI_B}{dI_C} + (\beta + 1) \frac{dI_{CO}}{dI_C}$$

i.e., $$\left(1 - \beta \frac{dI_B}{dI_C}\right) = \frac{(\beta + 1)}{S}$$

Therefore, $$S = \frac{(1 + \beta)}{\left(1 - \beta \frac{dI_B}{dI_C}\right)}$$

Applying Thevenin's theorem to the circuit of Fig. 2.21, for finding the base current, we have,

$$V_T = \frac{R_2 V_{CC}}{R_1 + R_2} \quad \text{and} \quad R_B = \frac{R_1 R_2}{R_1 + R_2}$$

The loop equation around the base circuit can be written as

$$V_T = I_B R_B + V_{BE} + (I_B + I_C) R_E$$

Differentiating this equation with respect to I_C, we get

$$\frac{dI_B}{dI_C} = -\frac{R_E}{R_E + R_B}$$

Hence, the stability factor is

$$S = \frac{(1+\beta)}{\left(1-\beta\dfrac{dI_B}{dI_C}\right)} = \frac{(1+\beta)}{\left(1+\beta\dfrac{R_E}{R_E+R_B}\right)}$$

$$= (1+\beta)\frac{1+\dfrac{R_B}{R_E}}{\left(1+\beta\dfrac{R_B}{R_E}\right)}$$

As can be seen, the value of S is equal to one if the ratio R_B/R_E is very small as compared to 1. As this ratio becomes comparable to unity, and beyond towards infinity, the value of the stability factor goes on increasing till $S = 1 + \beta$.

This improvement in the stability up to a factor equal to 1 is achieved at the cost of power dissipation. To improve the stability, the equivalent resistance R_B must be decreased, forcing more current in the voltage divider network of R_1 and R_2.

Often, to prevent the loss of gain due to the negative feedback, R_E is shunted by a capacitor C_E. The capacitive reactance X_{CE} must be equal to about one-tenth of the value of the resistance R_E at the lowest operating frequency.

2.10 BREAKDOWN IN TRANSISTORS

There is a possibility of voltage breakdown in the transistor at high voltages even though the rated dissipation of the transistor is not exceeded. Therefore, there is an upper limit to the maximum allowable collector junction voltage. There are two types of breakdown, namely, *Avalanche multiplication* or *Avalanche breakdown* and *reach-through* or *punch-through*.

2.10.1 Avalanche Breakdown and Multiplication

When a diode is reverse biased, there is a limit on the voltage that can be applied which is the Avalanche voltage. Similarly, in the transistor, the maximum reverse biasing voltage which may be applied before breakdown between the collector and base terminals with the emitter open is called breakdown voltage BV_{CBO}. Therefore, an upper limit is set on the collector voltage V_{CB} by avalanche breakdown in the reverse biased collector–base junction. Fields of order 10^6 V/m are required in the depletion layer for breakdown to occur, which usually limits V_{CB} to a maximum of several tens of volts.

Breakdown may occur because of Avalanche multiplication of the current I_{CO} that crosses the collector junction. As a result of this multiplication, the current becomes MI_{CO}, where M is the Avalanche multiplication factor. At the breakdown voltage BV_{CBO}, multiplication factor M becomes infinite and the current rises abruptly in the breakdown region, as shown in Fig. 2.22, there will be large changes in current with small changes in applied voltage.

The Avalanche multiplication factor depends on the voltage V_{CB} between collector and base, which has been found to be given empirically by

$$M = \frac{1}{1 - \left(\frac{V_{CB}}{BV_{CBO}}\right)^n} \tag{2.27}$$

where the empirical constant, n, depends on the lattice material and the carrier type, which is usually in the range of about 2 to 10; for N-type silicon $n \approx 4$ and for P-type $n \approx 2$, and controls the sharpness of the onset breakdown. Taking Avalanche multiplication into account, I_C has the magnitude $M \alpha I_E$, where α is the common base current gain. As a result, in the presence of Avalanche multiplication, the current gain of CB transistor has become $M\alpha$. As the effective current gain exceeds unity, the emitter circuit may then display negative resistance effects, which can lead to undesirable instabilities.

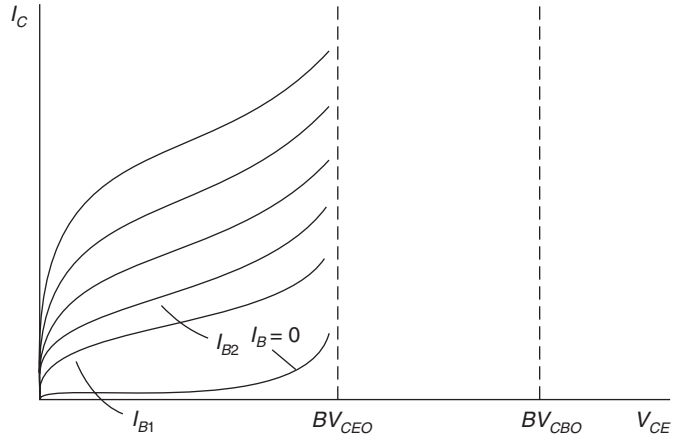


Fig. 2.22 *CE characteristics in the breakdown region*

For the CE configuration, the collector to emitter breakdown voltage BV_{CEO} with base open is

$$BV_{CEO} = BV_{CBO} \sqrt[n]{\frac{1}{h_{FE}}} \tag{2.28}$$

In general, BV_{CEO} is 40–50% of BV_{CBO}. This is the upper limit of V_{CE} that can be placed across the transistor without damaging it.

2.10.2 Reach-Through or Punch-Through

According to Early effect, the width of the collector-junction transition region increases with increased collector-junction voltage. As the voltage applied across the junction V_{CB} increases the transition region penetrates deeper into the base and will have spread completely across the base to reach the emitter junction, as the base is very thin. Thus, the collector voltage has reached through the base region. This effect, known as reach-through, also affects the output characteristics of a transistor since I_C versus V_{CB} curves are no longer horizontal but take on a positive slope indicating that the device has a finite output impedance that is voltage dependent. Since the input characteristics are also affected, the input impedance is also influenced by V_{CB}.

It is possible to raise the punch-through voltage by increasing the doping concentration in the base, but this automatically reduces the emitter efficiency.

Punch-through takes place at a fixed voltage between collector and base and is not dependent on circuit configuration, whereas Avalanche multiplication takes place at different voltages depending upon the circuit configuration. Therefore, the voltage limit of a particular transistor is determined by either of the two types of breakdown, whichever occurs at lower voltage.

2.11 EBERS–MOLL MODEL [AU May 2017, Dec 2015, Nov 2014, 8 marks]

The general expression for collector current I_C of a transistor for any voltage across collector junction V_C and emitter current I_E is

$$I_C = -\alpha_N I_E - I_{CO}\left(e^{\frac{V_C}{V_T}} - 1\right) \tag{2.29}$$

where α_N is the current gain in *normal* operation and I_{CO} is the collector junction reverse saturation current.

In inverted mode of operation, the above equation can be written as

$$I_E = -\alpha_I I_C - I_{EO}\left(e^{\frac{V_E}{V_T}} - 1\right) \tag{2.30}$$

where α_I is the inverted common-base current gain and I_{EO} is the emitter junction reverse saturation current.

The above four parameters are related by the condition

$$\alpha_I I_{CO} = \alpha_N I_{EO} \tag{2.31}$$

For many transistors I_{EO} lies in the range 0.5 I_{CO} to I_{CO}.

Figure 2.23 shows the Ebers–Moll model for a *PNP* transistor. Here, two separate ideal diodes are connected back to back with saturation currents $-I_{EO}$ and $-I_{CO}$ and there are two dependent current-controlled current sources shunting the ideal diodes. The current sources account for the minority carrier transport across the base. An application of Kirchhoff's current law to the collector node of Fig. 2.23 gives

$$I_C = -\alpha_N I_E + I = -\alpha_N I_E + I_O\left(e^{\frac{V_C}{V_T}} - 1\right) \tag{2.32}$$

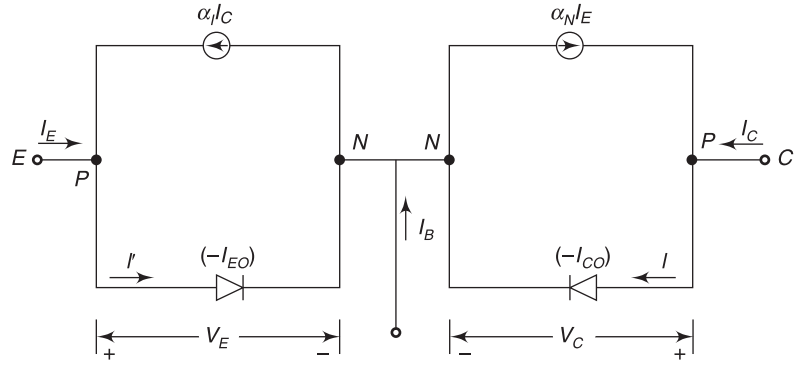


Fig. 2.23 *Ebers–Moll model for a PNP transistor*

where I is the diode current. As I_O is the magnitude of reverse saturation current, then $I_O = -I_{CO}$. Substituting this value of I_O in Eq. (2.32), we get

$$I_C = -\alpha_N I_E - I_{CO}\left(\frac{V_C}{e^{V_T}} - 1\right)$$

which is nothing but the general expression for collector current of a transistor given in Eq. (2.29). Hence, this model is valid for both forward and reverse static voltages applied across the transistor junctions. Here, base spreading resistance has been omitted and the difference between I_{CBO} and I_{CO} have been neglected.

The dependent current sources may be removed from the Fig. 2.23, provided $\alpha_N = \alpha_I = 0$. If the base-width is made much larger than the diffusion length of minority carriers in the base, all minority carriers will recombine in the base and no minority carrier will be available to reach the collector. Therefore, the transistor amplification factor α will become zero. As a result, transistor action ceases. Hence, it is not possible to construct a transistor by simply placing two isolated diodes back to back.

2.12 *h* -PARAMETER MODEL

[AU April 2015 and Nov 2014, 16 marks]

A transistor can be treated as a two-port network. The terminal behaviour of any two-port network can be specified by the terminal voltages V_1 and V_2 at ports 1 and 2, respectively, and currents I_1 and I_2, entering ports 1 and 2, respectively, as shown in Fig. 2.24. Of these four variables, V_1, V_2, I_1 and I_2, two can be selected as independent variables and the remaining two can be expressed in terms of these independent variables. This leads to various two-port parameters out of which the following are very important.

(i) *h*-parameters or *hybrid parameters*
(ii) *Z*-parameters or *impedance parameters*
(iii) *Y*-parameters or *admittance parameters*

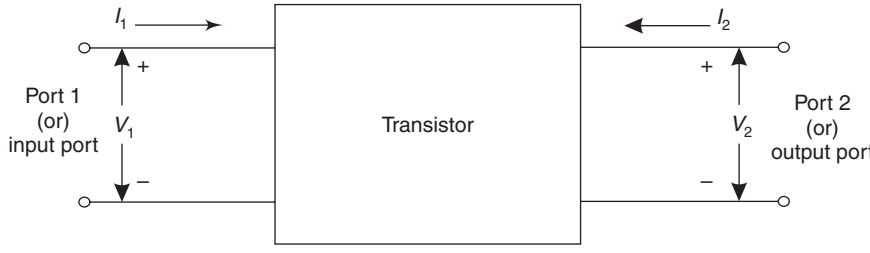


Fig. 2.24 *Two-port network*

The *h*-parameter representation is widely utilised in modelling of transistors. Hence, both short-circuit and open-circuit terminal conditions are used. If the input current I_1 and the output voltage V_2 are taken as independent variables, the input voltage V_1 and output current I_2 can be written as

$$\begin{bmatrix} V_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ V_2 \end{bmatrix}$$

$$V_1 = h_{11}I_1 + h_{12}V_2$$
$$I_2 = h_{21}I_1 + h_{22}V_2$$

The four hybrid parameters h_{11}, h_{12}, h_{21} and h_{22} are defined as follows:

When $V_1 = 0$, i.e., with output port short-circuited,

$$h_{11} = \left[\frac{V_1}{I_1}\right] \text{ with } V_2 = 0$$
$$= \text{input impedance}$$

and
$$h_{21} = \left[\frac{I_2}{I_1}\right] \text{ with } V_2 = 0$$
$$= \text{forward current gain or forward transfer ratio}$$

When $I_1 = 0$, i.e., with input port open-circuited,

$$h_{22} = \left[\frac{I_2}{V_2}\right] \text{ with } I_2 = 0$$
$$= \text{output admittance}$$

and
$$h_{12} = \left[\frac{V_1}{V_2}\right] \text{ with } I_1 = 0$$
$$= \text{reverse voltage gain or reverse transfer ratio}$$

The equivalent circuit of the *h*-parameter representation is shown in Fig. 2.25. Here, $h_{12}V_2$ is the controlled voltage source and $h_{21}I_2$ is the controlled current source.

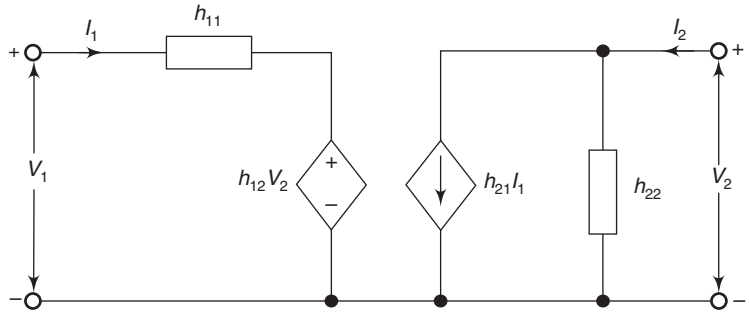


Fig. 2.25 *Equivalent circuit of h-parameter model*

Note: The dimensions of *h*-parameters are as follows:

$h_{11} \rightarrow \Omega$

$h_{22} \rightarrow \mho$

$h_{12}, h_{21} \rightarrow$ dimensionless

As the dimensions are not alike, i.e., they are hybrid in nature, these parameters are called *hybrid parameters.*

The alternative subscript notations recommended by IEEE commonly used are

i = 11 = input, o = 22 = output

f = 21 = forward transfer, r = 12 = reverse transfer

Based on the definition of hybrid parameters, the mathematical model for two-port networks known as *h*-parameter model can be developed as shown in Fig. 2.25.

When *h*-parameters are applied to transistors, it is a common practice to add a second subscript to designate the type of configuration considered, *e* for common emitter, *b* for common base and *c* for common collector. Thus, for a common emitter (CE) configuration,

$h_{ie} = h_{11e}$ = short-circuit input impedance

$h_{oe} = h_{22e}$ = open-circuit output admittance

$h_{re} = h_{12e}$ = open-circuit reverse-voltage gain

$h_{fe} = h_{21e}$ = short-circuit forward-current gain

The proposed model shown in Fig. 2.26 should satisfy the *h*-parameter equations and it can be readily verified by writing Kirchhoff's voltage-law equation in the input loop and Kirchhoff's current-law equation for the output node. It is to be noted that the input circuit has a dependent voltage generator and the output circuit contains a dependent current generator.

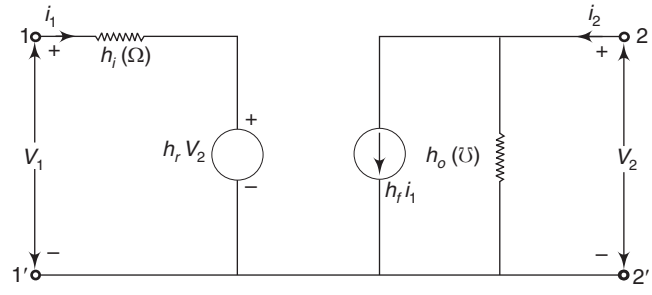
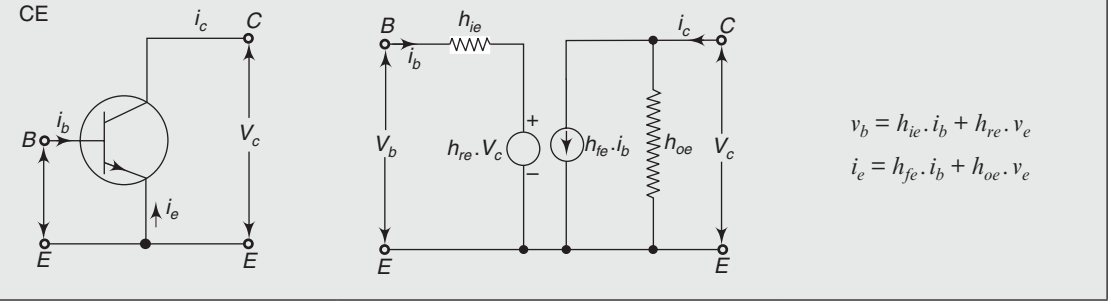


Fig. 2.26 *h-parameter model for two-port network*

The *h*-parameter models and equations given in Table 2.2 are valid for *NPN* as well as *PNP* transistors and hold good for all types of loads and methods of biasing. Table 2.3 gives the typical *h*-parameter values for a transistor and Table 2.4 gives the conversion formulae to find the *h*-parameters for CC and CB configurations given the *h*-parameters for CE configuration.

Table 2.2 *h-parameter model for the transistor in three different configurations*



(*Contd.*)

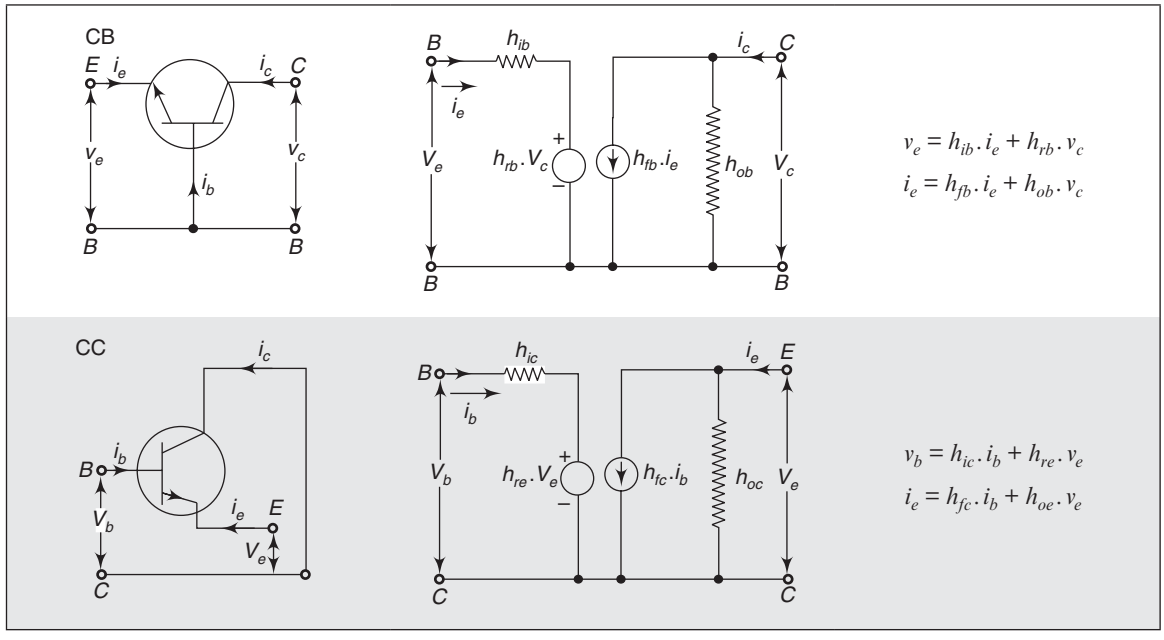


Table 2.3 *Typical h-parameter values for a transistor*

Parameter	*CE*	*CC*	*CB*
h_i	1,100 Ω	1,100 Ω	22 Ω
h_r	2.5×10^{-4}	1	2.5×10^{-4}
h_f	50	–51	–0.98
h_o	25 μA/V	25 μA/V	0.49 μA/V

Table 2.4 *Conversion formulae for h-parameters*

CC	*CB*
$h_{ic} = h_{ie}$	$h_{ib} = \dfrac{h_{ie}}{1 + h_{fe}}$
$h_{rc} = 1$	$h_{rb} = \dfrac{h_{ie}h_{oe}}{1 + h_{fe}} - h_{re} - h_{\mathrm{re}}$
$h_{fc} = -(1 + h_{fe})$	$h_{fb} = \dfrac{-h_{fe}}{1 + h_{fe}}$
$h_{oc} = h_{oe}$	$h_{ob} = \dfrac{h_{oe}}{1 + h_{fe}}$

Simplified h-parameter model

As the *h*-parameters themselves vary widely for the same type of transistor, it is justified to make approximations for the small-signal analysis of transistor amplifier in terms of current gain (A_I), voltage gain (A_V), power gain (A_P), input resistance (R_i) and output resistance (R_o). In addition, a better understanding of the behaviour of the transistor circuit can be obtained by using the simplified *h*-parameter model. Since CE configuration is more useful and general, it is taken for consideration.

The *h*-parameter equivalent circuit of the transistor in the CE configuration is shown in Fig. 2.27. Here, $1/h_{oe}$ is in parallel with the load resistor, R_L. The parallel combination of two unequal impedances, i.e., $1/h_{oe}$ and R_L, is approximately equal to the lower value, i.e., R_L. Hence, if $1/h_{oe} >> R_L$, then the term h_{oe} may be neglected. Further, if h_{oe} is omitted, the collector current I_C is given by $I_C = h_{fe} I_b$.

Under this condition, the magnitude of voltage of generator in the emitter circuit is

$$h_{re} |V_C| = h_{re} I_C R_L = h_{re} h_{fe} I_b R_L$$

Since $h_{re}\ h_{fe} = 0.01$, this voltage may be neglected in comparison with the voltage drop across $h_{ie} = h_{ie} I_b$ provided that R_L is not too large.

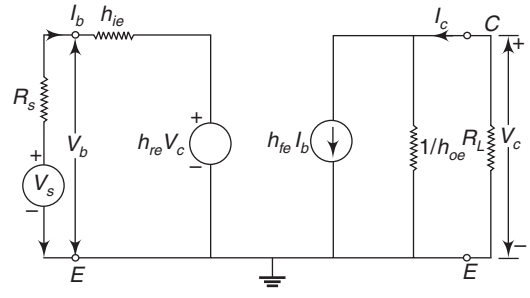


Fig. 2.27 *Exact CE h-parameter model*

To conclude, if the load resistance R_L is small, it is possible to neglect the parameter h_{re} and h_{oe} and obtain the approximate equivalent circuit as shown in Fig. 2.28. It can be shown that if $h_{oe} R_L \leq 0.1$, the error in calculating A_I, A_V, R_i and R_o for *CE* configuration is less than 10%.

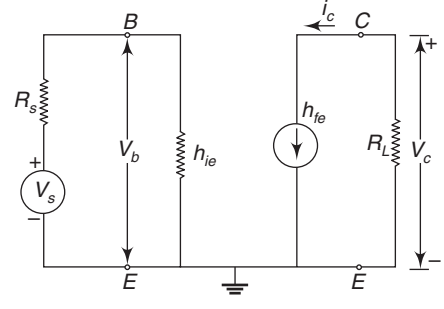


Fig. 2.28 *Approximate CE h-parameter model*

Generalised approximate model

Figure 2.29 shows the simplified *h*-parameter circuit that can be used for any configuration by simply grounding the appropriate terminal. The source is connected between input and ground and the load is connected between output and ground. The errors introduced in calculating the various parameters using the approximate hybrid model are to be found out now. This is done for CE configuration.

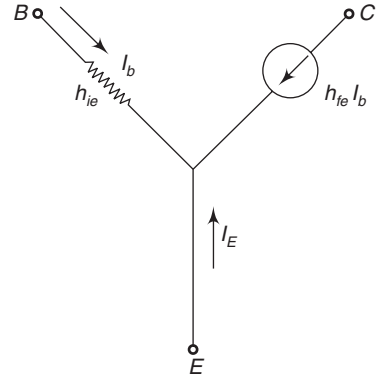


Fig. 2.29 *Approximate h-parameter model valid for all configurations*

2.13 HYBRID-π MODEL [AU May 2016, Dec 2015, 8 marks]

At high frequencies, the capacitive effects of the transistor junctions and the delay in response of the transistor caused by the process of diffusion of carriers should be taken into account in determining the high-frequency model of a transistor.

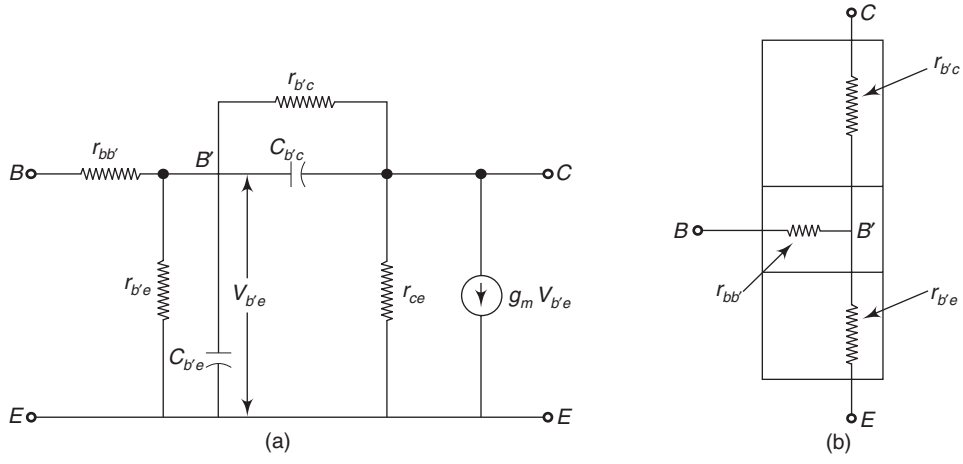


Fig. 2.30 *(a) Hybrid π model for a transistor in the CE configuration, (b) Diagram showing virtual base B′ and ohmic base-spreading resistance $r_{bb'}$*

A high-frequency-π or Giacoletto model for a transistor is shown in Fig. 2.30, where r_{bb}- base spreading resistance between the actual base *B* and virtual base *B′*. It represents the bulk resistance of the base. Its typical value is 100 Ω.

$r_{b'e}$—resistance between the virtual base B' and the emitter terminal E whose Typical value is 1 kΩ.

Input resistance from base to emitter with the output shorted is simply $r_{bb'} + r_{b'e}$ and this is the same as h_{ie}. Hence, $h_{ie} = r_{bb'} + r_{b'e}$.

$r_{b'c}$—resistance between the virtual base B' and the collector terminal C whose typical value is 4 MΩ

$C_{b'e}$—diffusion capacitance of the normally forward biased base-emitter junction. It has a typical value of 100 pF.

$C_{b'c}$—transistor capacitance of the normally reverse biased collector-base junction. It has a typical value of 3 pF.

r_{ce}—output resistance with a typical value of 80 kΩ. Since $r_{ce} >> R_L$, if a load R_L is connected r_{ce} can be neglected.

$g_m V_{b'e}$—output current generator value where g_m is the transconductance of the transistor.

2.13.1 Hybrid-π Conductances

The hybrid-π model for the CE transistor at low frequencies is shown in Fig. 2.31(a). The *h*-parameter model for the same is shown in Fig. 2.31(b). As the hybrid-π model is drawn for low frequencies, the capacitive elements are considered as open circuit.

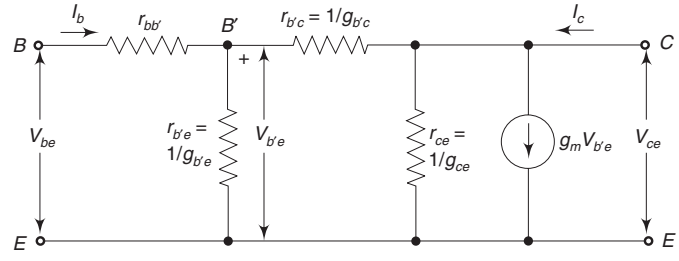


Fig. 2.31 (a) *Hybrid-π model for a common-emitter transistor at low frequency*

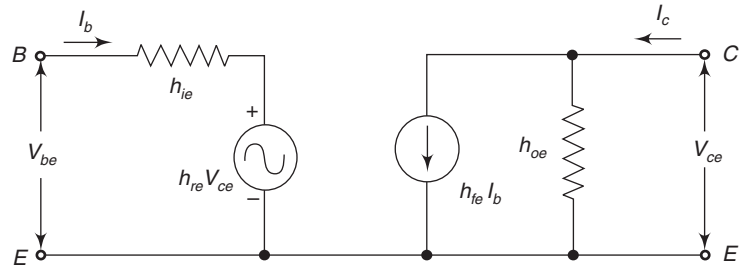


Fig. 2.31 *(b) h-parameter model for a common-emitter transistor at low frequency*

Base Spreading Resistance ($r_{bb'}$)

In the circuit shown in Fig. 2.31(b), the value of input resistance is equal to h_{ie} when the output terminals are short circuited, i.e., $V_{ce} = 0$. Under these conditions, for the circuit in Fig. 2.31(a), the input impedance is given by

$$Z_i\big|_{V_{ce}=0} = r_{bb'} + r_{b'e} \parallel r_{b'c}$$

Therefore, $$h_{ie} = r_{bb'} + r_{b'e} \parallel r_{b'c}$$

As $r_{b'c} >> r_{b'e}$ the above equation can be written as

$$h_{ie} = r_{bb'} + r_{b'e}$$

Conductance between terminals B′ and C or the feedback conductance ($g_{b'c}$)

In the circuit shown in Fig. 2.31 (b), if the input terminals are open-circuited, then the reverse voltage gain h_{re} can be written for the circuit shown in Fig. 2.31(a), and it is given by

$$h_{re} = \frac{V_{b'e}}{V_{ce}} = \frac{r_{b'e}}{r_{b'e} + r_{b'c}}$$

Rearranging the above equation, we get

$$r_{b'e}(1 - h_{re}) = h_{re} r_{b'c}$$

As the value of h_{re} is in the range of 10^{-4}, the above equation can be approximated by

$$r_{b'e} = h_{re} r_{b'c} \text{ or } g_{b'c} = h_{re} g_{b'e}$$

The equation also shows that the value of resistance $r_{b'c}$ is much larger than resistance ($r_{b'e}$), i.e., $r_{b'c} >> r_{b'e}$.

Conductance between terminals C and E (g_{ce})

In the circuit shown in Fig. 2.31(b), if the input terminals are open circuited, then

$$V_{b'e} = h_{re} V_{ce}$$

For the circuit in Fig. 2.31(a), with the input terminals open, i.e., $I_b = 0$, then the collector current I_c is given by

$$I_c = \frac{V_{ce}}{r_{ce}} + \frac{V_{ce}}{r_{b'e} + r_{b'c}} + g_m V_{b'e}$$

The value of output admittance h_{oe} is given by

$$h_{oe} = \left.\frac{I_c}{V_{ce}}\right| \quad = \frac{1}{r_{ce}} + \frac{1}{r_{b'e} \quad r_{b'c}} + \frac{g_m V_{b'e}}{V_{ce}}$$

Substituting the value of $V_{b'e}$ in the above equation, we get

$$h_{oe} = \frac{1}{r_{ce}} + \frac{1}{r_{b'e} + r_{b'c}} + g_m h_{re}$$

Assuming that $r_{b'c} >> r_{b'e}$, the above equation can be rewritten in terms of conductance as

$$h_{oe} = g_{ce} + g_{b'c} + g_m \frac{g_{b'c}}{g_{b'e}}$$

Substituting $g_m = h_{fe} g_{b'e}$ in the equation, we get

$$h_{oe} = g_{ce} + g_{b'c} + g_{b'c} h_{fe}$$

Rearranging the terms in the above equation, we get

$$g_{ce} = h_{oe} - (1 + h_{fe}) g_{b'c}$$

Since $h_{fe} >> 1$, the above equation can be written as

$$g_{ce} \cong h_{oe} - h_{fe} g_{b'c} \cong h_{oe} - g_m h_{fe}$$

Conductance between terminals B′ and E or the input conductance ($g_{b'e}$)

As the value of $r_{b'c}$ is much greater than $r_{b'e}$, most of the current I_b flows through $r_{b'e}$ in the circuit shown in Fig. 2.31(b) and the value of $V_{b'e}$ is given by

$$V_{b'e} \cong I_b r_{b'e}$$

The short-circuit collector current, I_c, is given by

$$I_c = g_m V_{b'e} \cong I_b r_{b'e}$$

The short-circuit current gain, h_{fe}, is defined as

$$h_{fe} = \left.\frac{I_c}{I_b}\right|_{V_{ce}} \cong g_m r_{b'e}$$

Rearranging the above equation, we get

$$r_{b'e} = \frac{h_{fe}}{g_m}$$

or

$$g_{b'e} = \frac{g_m}{h_{fe}}$$

Transistor's transconductance (g_m)

The transconductance of a transistor (g_m) is defined as the ratio of change in I_c to change in $V_{b'e}$ for constant value of collector-emitter voltage. For common-emitter transistor configuration, the expression for collector current is given by

$$I_c = I_{CO} + \alpha I_e$$

The value of g_m is given by

$$g_m = \left.\frac{\partial I_c}{\partial V_{b'e}}\right|_{V_{CE} = \text{constant}} = \alpha \frac{\partial I_e}{\partial V_{b'e}} = \alpha \frac{\partial I_e}{\partial V_e}$$

The partial derivative emitter voltage with respect to the emitter current (i.e., $\partial V_e / \partial I_e$) can be represented as the emitter resistance (r_e) and the dynamic resistance of a forward-biased diode (r_d) is given as

$$r_d = \frac{V_T}{I_D}$$

where V_T is the volt equivalent of temperature and I_D is the diode current. Therefore, the value of g_m can be generalised as

$$g_m = \frac{\alpha I_e}{V_T} = \frac{I_c - I_{co}}{V_T}$$

As $I_c >> I_{CO}$, the value of g_m for an *NPN* transistor is positive. For a *PNP* transistor, the analysis can be carried out on similar lines and the value of g_m in the case of a *PNP* transistor is also positive. Therefore, the above expression for g_m is written as

$$g_m = \frac{|I_c|}{V_T}$$

2.13.2 Hybrid-π Capacitances

In the hybrid-π model shown in Fig. 2.31 (b), there are two capacitances, namely the collector-junction barrier capacitance (C_c) and the emitter-junction diffusion capacitance (C_e).

Collector-junction capacitance (C_c)

The capacitance C_c is the output capacitance of the common-base transistor configuration with the input open ($I_e = 0$). As the collector-base junction is reverse-biased, C_c is the transition capacitance and it varies as $(V_{CB})^{-n}$, where n is 1/2 for abrupt junction and 1/3 for a graded junction.

Emitter-junction capacitance (C_e)

The capacitance C_e is the diffusion capacitance of the forward-biased emitter junction and is proportional to the emitter current I_e and is almost independent of temperature.

2.14 GUMMEL–POON MODEL [AU Nov 2016, 8 marks]

The DC and dynamic currents of the transistor in response to V_{BE} and V_{CE} can be represented accurately by the BJT model used in circuit simulators such as SPICE. A typical circuit-simulation model or compact model is made of the Ebers–Moll model (having V_{BE} and V_{CE} as the two driving forces for I_C and I_B) along with additional enhancements for high-level injection, parasitic resistances and voltage-dependent capacitances that accurately represent the charge storage in the transistor. This BJT model is an extension of Ebers–Moll model and is known as the *Gummel–Poon model.*

The Gummel-Poon model is valid for both positive and negative values of V_{BE} and V_{CE} whereas the Ebers–Moll model is limited to positive values of V_{BE} and V_{CE}. Gummel–Poon model was an improvement over the Ebers–Moll model by taking into account the early effect and high-level injection effects. From the Ebers–Moll model, we know that for an *NPN* transistor, the emitter current is composed of hole current (I_{pe}) flowing in *N*-type emitter and electron current (I_{nc}) flowing from emitter to collector through *P*-type base and it is given by

$$I_E = -(I_{pe} + I_{nc}) \tag{2.33}$$

The collector current is due to electron current (I_{nc}) and hole current (I_{pc}) flowing through the base-collector depletion region and it is given by

$$I_C = I_{nc} - I_{pc} \tag{2.34}$$

in which I_{nc} is the collector current component due to electrons. These electrons are injected from the emitter and diffuse across the base to the collector. They contribute to both the collector and emitter currents. I_{pc} is the collector current component due to holes injected from the base into the collector. I_{pe} is the emitter current component due to holes injected from the base to the emitter. As the recombination in the base is considered to be small, I_{pe} is equal to the base current when the transistor is biased in normal forward operation with the base-emitter junction forward biased and the base-collector junction reversed biased. For an *NPN* transistor, the collector current due to electrons is given by

$$I_{nc} = I_{be1} - I_{bc1} \tag{2.35}$$

where $$I_{be1} = I_o\left(e^{\frac{V_{BE}}{V_T}} - 1\right) \text{ and } I_{bc1} = I_o\left(e^{\frac{V_{CE}}{V_T}} - 1\right)$$

Here, I_o is the diode reverse saturation current at room temperature.

Therefore,

$$I_{nc} = I_o\left(e^{\frac{V_{BE}}{V_T}} - e^{\frac{V_{CE}}{V_T}}\right) \qquad (2.36)$$

Also, in the normal forward operating region, I_{be1} is the collector current and I_{pe} is the base current. Therefore,

$$I_{be1} = \beta_F I_{pe} \qquad (2.37)$$

The equations are symmetric so that in the reverse condition

$$I_{be1} = \beta_R I_{pc} \qquad (2.38)$$

Substituting Eq. (2.35) and Eq. (2.37) in Eq. (2.33) results in

$$I_E = (I_{be1} - I_{bc1}) - \frac{I_{be1}}{\beta_F} \qquad (2.39)$$

Substituting Eq. (2.35) and Eq. (2.38) in Eq. (2.34) results in

$$I_C = I_{be1} - I_{bc1} - \frac{I_{bc1}}{\beta_R} \qquad (2.40)$$

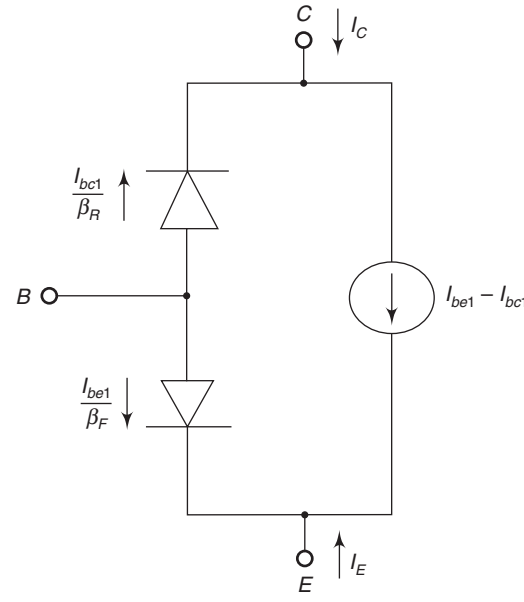


Fig. 2.32 *Gummel–Poon NPN model without early effect and high-level injection effects*

Equations (2.39) and (2.40) can be represented schematically in Fig. 2.32 and these two equations are the Ebers–Moll model formulated in a way that allows the charge control concept used by Gummel–Poon to be included. A base charge factor (K_{qb}), is added to these equations and it is a normalized integer representing positive mobile charge in the base. When the collector voltage increases, base charge factor becomes smaller because the collector base depletion region increases, reducing the base width and, therefore, the charge in the base. This is known as the early effect which causes the collector current to increase.

When holes are injected at a higher level from the emitter, the extra electrons attract extra holes, increasing the positive mobile charge in the base. If the density of these added charges approaches the doping level in the base, the voltage necessary to maintain the positive charge becomes important. A portion of the applied base-emitter voltage appears across the positive charge in the base and is not available to the base-emitter junction. This is known as the high-level injection effect which causes collector current to decrease. High-level injection is modeled as an increase in K_{qb}. Adding the base charge factor K_{qb} to Eq. (2.39) and Eq. (2.40) results in

$$I_E = \frac{I_{be1} - I_{bc1}}{K_{qb}} - \frac{I_{be1}}{\beta_F} \qquad (2.41)$$

$$I_C = \frac{I_{be1} - I_{bc1}}{K_{qb}} - \frac{I_{ce1}}{\beta_R} \qquad (2.42)$$

Equations (2.42) and (2.43) can be represented schematically in Fig. 2.33. I_{be1} and I_{bc1} are the currents flowing in two additional diodes in Fig. 2.33 as compared to Fig. 2.32. These currents model the base current due to recombination in the depletion regions. Base-to-emitter and base-to-collector capacitances are also shown in Fig. 2.33. These capacitances are the sum of the junction and diffusion capacitances for the junctions.

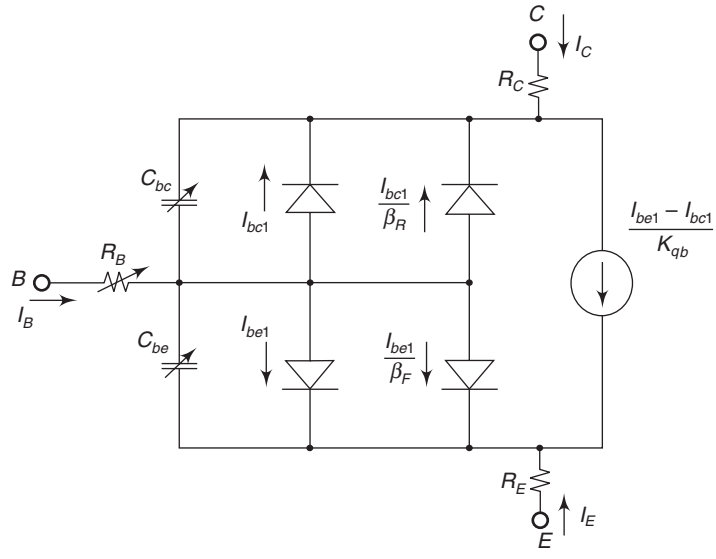


Fig. 2.33 *Gummel–Poon model with high-level injection effect and early effect*

Gummel plot is shown in Fig. 2.34 which represents the plot of the log of the collector and base currents as a function of the base-emitter voltage. The logarithmic vertical axis results in a straight line plot, with a slope of $1/V_T$ over a wide range. This is valid since

$$I_C = I_o e^{\frac{V_{BE}}{V_T}} \tag{2.43}$$

$$\ln(I_C) = \ln(I_o) + \frac{V_{BE}}{V_T} \tag{2.44}$$

Since the logarithms of the collector and base currents are plotted in Fig. 2.34, the log of β, the ratio of I_C to I_B, is the distance between the curves $\ln(I_C) - \ln(I_B)$. The current gain, β decreases at both high and low values of collector current. At high levels, collector current is reduced by high-level injection effects. The plot of collector current is a straight line up to about the forward knee current, I_{KF}. At larger current values, high-level injection effects reduce the slope of the current plot to a value close to $1/2V_T$. At low current levels, base current is larger than expected due to recombination current and generation current. This current is represented by I_{be1} flowing in the diode as shown in Fig. 2.33 and it does not contribute to the collector current.

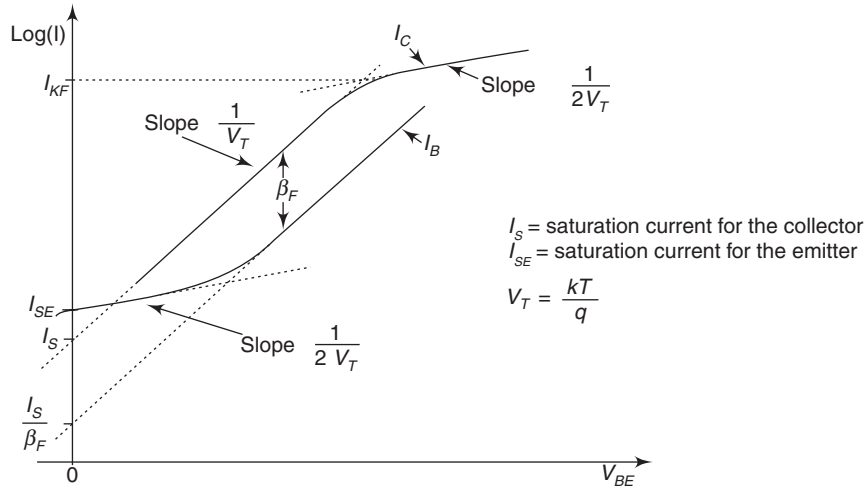


Fig. 2.34 *Gummel plot*

2.15 MULTI-EMITTER TRANSISTOR

Multi-emitter transistors are mainly used in applications such as Transistor-Transistor Logic (TTL). TTL is fast, inexpensive and easy to use. This switch uses a multiple-emitter transistor which can be easily and economically fabricated. Figure 2.35 shows the cross-sectional view of two N^+ emitter regions diffused in three places inside the *P*-type base and its symbol. This arrangement saves the chip area and enhances the component density of the IC.

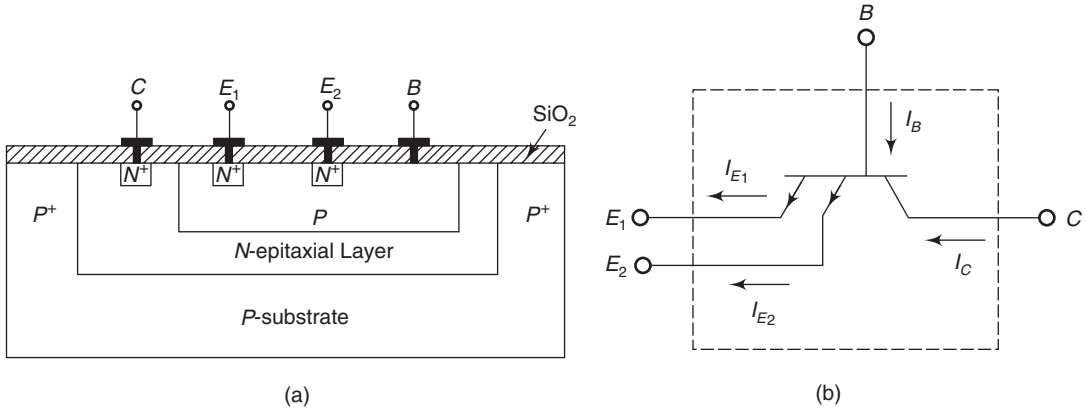


Fig. 2.35 *Multi-emitter transistor: (a) Cross-sectional view, (b) Symbol*

The fastest saturating logic circuit is the transistor–transistor logic two-input NAND gate shown in Fig. 2.36. The input voltages *A* and *B* are either low (ground) or high (+ 5 V). If *A* or *B* are low, the base of Q_1 is pulled down to approximately 0.7 V. This reduces the base voltage of Q_2 to almost zero. Therefore, Q_2 is cut off. With Q_2 open, the output *Y* is pulled up to a high voltage.

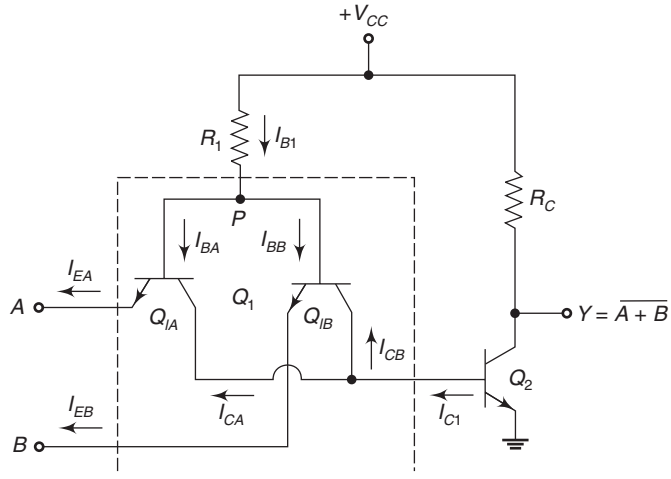


Fig. 2.36 *Basic two-input TTL NAND gate*

On the other hand, when A and B are both high voltages, the emitter diodes of Q_1 stop conducting, and the collector diode goes into forward conduction. This forces Q_2 base to go high. In turn, Q_2 goes into saturation, producing a low output.

Two Mark Questions and Answers

1. **If a transistor has an α of 0.97, find the value of β.** **(April/May 2017)**
   Refer to Example 2.5.

2. **The common-base DC current gain of a transistor is 0.967. If the emitter current is 10 mA, what is the value of base current?** **(Dec 2015/Jan 2016)**
   Refer to Example 2.3.

3. **A transistor has $\beta = 150$. Find the collector and base currents, if $I_E = 10$ mA.** **(May/June 2016)**
   Refer to Example 2.7

4. **Sketch the Ebers-Moll model.** **(April/May 2015)**
   Refer to Fig. 2.23.

5. **Calculate the collector and emitter current levels for a BJT with $\alpha_{dc} = 0.99$ and $I_B = 20$ μA.** **(Nov/Dec 2014)**

   Solution:
   Given $\alpha_{dc} = 0.99$ and $I_B = 20$ μA.

   We know that, $$\alpha_{dc} = \frac{I_C}{I_E} = \frac{I_C}{I_B + I_C}$$

   i.e., $$0.99\,(20 \times 10^{-6} + I_C) = I_C$$

Hence, $$I_C = \frac{19.8\times10^{-6}}{0.01} = 1.98\times10^{-3} = 1.98\,\text{mA}$$

and $$I_E = I_B + I_C = (0.02+1.98)\times10^{-3} = 2\,\text{mA}$$

6. What is the need for biasing in the transistor? (May/June 2014)

To operate the transistor in active mode, it is required to forward bias the emitter-base junction and reverse bias the collector-base junction. By providing proper bias voltage, the transistor can be made to work as an amplifier.

7. Draw the *h*-parameter model for CE transistor. (May/June 2014)

Refer to Fig. 2.26.

8. Draw input characteristics of CB transistor. (Nov/Dec 2013)

Refer to Fig. 2.8.

9. Define and explain the term 'early effect'. (May/June 2012)

As the collector voltage V_{CC} is made to increase the reverse bias, the space charge width between collector and base tends to increase, with the result that the effective width of the base decreases. This dependency of base-width on collector to emitter voltage is known as the *early effect*.

10. What is "early effect" in CB configuration and give its consequences? (June 2010)

For Early effect: Refer to Answer of Q. No. 9.

The decrease in effective base-width due to early effect has three consequences:

(i) There is less chance for recombination within the base region. Hence, the current gain α increases with increasing $|V_{CB}|$.

(ii) The charge gradient is increased within the base, and consequently, the current of minority carriers injected across the emitter junction increases.

(iii) For extremely large voltages, the effective base-width may be reduced to zero, causing voltage breakdown in the transistor. This phenomenon is called the *punch through*.

11. Given the biasing arrangement for an NPN transistor to operate in the active region. (May/June 2013)

Refer to Fig. 2.3.

12. Define α_{dc} and β_{dc} of a transistor. (May/June 2014)

The dc current gain in a common base transistor α_{dc} is defined as the ratio of the collector current I_C to the emitter current I_E. It is given by

$$\alpha_{dc} = \frac{I_C}{I_E}$$

The dc current gain in a common emitter transistor β_{dc} is defined as the ratio of the collector current I_C to the base current I_B. It is given by

$$\beta_{dc} = \frac{I_C}{I_B}$$

13. Define base-width modulation and bring out its consequences. **(Nov/Dec 2010)**

Refer to Answers of Q. No. 9 and 10.

14. What is thermal runaway in semiconductor devices? **(Nov/Dec 2011)**

When the reverse bias across collector-base junction in BJT is increased, there will be an increase in reverse leakage current. This increase in reverse leakage current will further increase the current flowing through a transistor and thus the power dissipation, causing a further increase in collector-to-emitter leakage current. This process is cumulative and is termed as thermal runaway.

15. What is the major difference between a bipolar and unipolar device? **(Nov/Dec 2014)**

In Bipolar devices like BJT, the current flowing is controlled by both minority and majority charge carriers whereas in unipolar devices like FET, only the majority charge carriers are responsible for the current flow.

Review Questions

1. What is a Bipolar junction transistor? How are its terminals named?
2. Explain the operation of *NPN* and *PNP* transistors.
3. What are the different configurations of BJT?
4. Explain the input and output characteristics of a transistor in CB configuration.
5. Explain the early effect and its consequences.
6. Derive the relationship between α and β.
7. Why does the CE configuration provide large current amplification while the CB configuration does not?
8. Draw the circuit diagram of an *NPN* junction transistor CE configuration and describe the static input and output characteristics. Also, define active, saturation and cut-off regions, and saturation resistance of a CE transistor.
9. How will you determine *h*-parameters from the characteristics of CE configuration?
10. Determine the *h*-parameters from the characteristics of CB configuration.
11. What is the relation between I_B, I_E and I_C in CB configuration?
12. Explain the laboratory setup for obtaining the CC characteristics.
13. Compare the performance of a transistor in different configurations.
14. Define α, β and γ of a transistor. Show how are they related to each other?
15. Explain how a transistor is used as an amplifier.
16. From the characteristics of CE configuration, explain the large signal, DC, and small signal CE values of current gain.
17. Calculate the values of I_C and I_E for a transistor with $\alpha_{DC} = 0.99$ and $I_{CBO} = 5\ \mu A$. I_B is measured as 20 μA. [*Ans.* $I_C = 2.48$ mA, $I_E = 2.5$ mA]
18. If $\alpha_{DC} = 0.99$ and $I_{CBO} = 50\ \mu A$, find emitter current. [*Ans.* $I_C = 104$ mA, $I_E = 105$ mA]
19. If I_C is 100 times larger than I_B, find the value of β_{DC}. [*Ans.* 100]
20. Find the value of α_{DC}, if β_{DC} is equal to 100. [*Ans.* 0.99]
21. Find the voltage gain of a transistor amplifier if its output is 5 V r.m.s. and the input is 100 mV r.m.s. [*Ans.* 50]
22. Find the value of α_{DC}, when $I_C = 8.2$ mA and $I_E = 8.7$ mA. [*Ans.* 0.943]
23. If α_{DC} is 0.96 and $I_E = 9.35$ mA, determine I_C. [*Ans.* 8.98 mA]

24. Describe the two types of breakdown in transistors.
25. Draw the Ebers–Moll model for a *PNP* transistor and give the equations for emitter current and collector current.
26. Draw the high frequency π model of a transistor and explain it.
27. Derive the relationship between low frequency *h*-parameters and high frequency hybrid π parameters for a transistor in CE configuration.
28. Explain the Gummel–Poon Model with equivalent circuit.
29. Describe briefly multi-emitter transistor and its applications.

FIELD EFFECT TRANSISTORS

3.1 INTRODUCTION

The FET is a device in which the flow of current through the conducting region is controlled by an electric field. Hence, the name Field Effect Transistor (FET). As current conduction is only by majority carriers, FET is said to be a unipolar device.

Based on the construction, the FET can be classified into two types as *Junction FET(JFET)* and *Metal Oxide Semiconductor FET(MOSFET)* or *Insulated Gate FET (IGFET)* or *Metal Oxide Silicon Transistor(MOST)*.

Depending upon the majority carriers, JFET has been classified into two types, namely, (1) *N-channel JFET* with electrons as the majority carriers, and (2) *P-Channel JFET* with holes as the majority carriers.

3.2 CONSTRUCTION OF *N*-CHANNEL JFET

[AU May 2017, June 2013, Nov 2012 and June 2011, 16 marks]

It consists of a *N*-type bar which is made of silicon. Ohmic contacts (terminals), made at the two ends of the bar, are called Source and Drain.

Source (S) This terminal is connected to the negative pole of the battery. Electrons which are the majority carriers in the *N*-type bar enter the bar through this terminal.

Drain (D) This terminal is connected to the positive pole of the battery. The majority carriers leave the bar through this terminal.

Gate (G) Heavily doped *P*-type silicon is diffused on both sides of the *N*-type silicon bar by which *PN* junctions are formed. These layers are joined together and called Gate *G*.

Channel The region *BC* of the *N*-type bar between the depletion region is called the channel. Majority carriers move from the source to drain when a potential difference V_{DS} is applied between the source and drain.

3.3 OPERATION OF *N*-CHANNEL JFET

[AU Dec 2015, May 2015, May 2014, Nov 2014, Nov 2012, June 2012 and Nov 2010, 10 marks]

When V_{GS} = 0 and V_{DS} = 0 When no voltage is applied between drain and source, and gate and source, the thickness of the depletion regions round the *PN* junction is uniform as shown in Fig. 3.1.

When V_{DS} = 0 and V_{GS} is decreased from zero In this case, the *PN* junctions are reverse biased and hence the thickness of the depletion region increases. As V_{GS} is decreased from zero, the reverse bias voltage across the *PN* junction is increased and hence, the thickness of the depletion region in the channel also increases until the two depletion regions make contact with each other. In this condition, the channel is said to be cut-off. The value of V_{GS} which is required to cut-off the channel is called the cut-off voltage V_C.

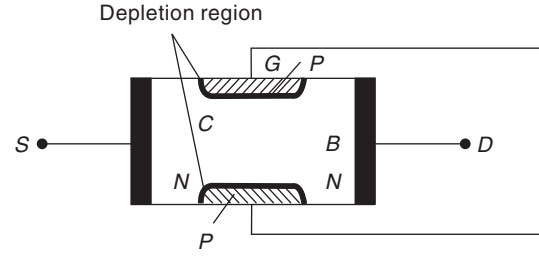


Fig. 3.1 *JFET construction*

When V_{GS} = 0 and V_{DS} is increased from zero Drain is positive with respect to the source with V_{GS} = 0. Now the majority carriers (electrons) flow through the *N*-channel from source to drain. Therefore the conventional current I_D flows from drain to source. The magnitude of the current will depend upon the following factors:

1. The number of majority carriers (electrons) available in the channel, i.e., the conductivity of the channel.
2. The length L of the channel.
3. The cross-sectional area A of the channel at B.
4. The magnitude of the applied voltage V_{DS}. Thus the channel acts as a resistor of resistance R given by

$$R = \frac{\rho L}{A} \tag{3.1}$$

$$I_D = \frac{V_{DS}}{R} = \frac{A V_{DS}}{\rho L} \tag{3.2}$$

where ρ is the resistivity of the channel. Because of the resistance of the channel and the applied voltage V_{DS}, there is a gradual increase of positive potential along the channel from source to drain. Thus the reverse voltage across the *PN* junctions increases and hence the thickness of the depletion regions also increases. Therefore, the channel is wedge shaped as shown in Fig. 3.2.

As V_{DS} is increased, the cross-sectional area of the channel will be reduced. At a certain value V_P of V_{DS}, the cross-sectional area at B becomes minimum. At this voltage, the channel is said to be pinched off and the drain voltage V_P is called the pinch-off voltage.

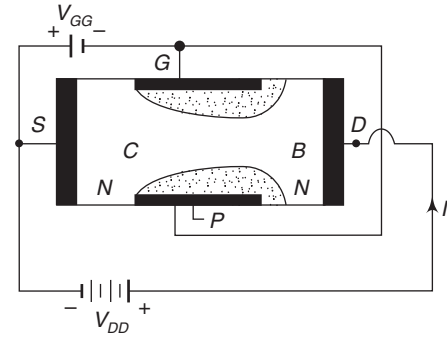


Fig. 3.2 *JFET under applied bias*

As a result of the decreasing cross-section of the channel with the increase of V_{DS}, the following results are obtained.

(i) As V_{DS} is increased from zero, I_D increases along OP, and the rate of increase of I_D with V_{DS} decreases as shown in Fig. 3.3. The region from V_{DS} = 0V to $V_{DS} = V_P$ is called the ohmic region. In the channel ohmic region, the drain to source resistance $\frac{V_{DS}}{I_D}$ is related to the gate voltage V_{GS}, in an almost linear manner. This is useful as a voltage variable resistor (VVR) or voltage dependent resistor (VDR).

(ii) When $V_{DS} = V_P$, I_D becomes maximum. When V_{DS} is increased beyond V_P, the length of the pinch-off or saturation region increases. Hence, there is no further increase of I_D.

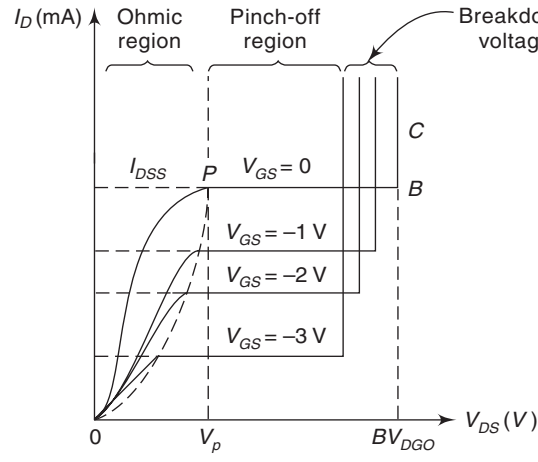


Fig. 3.3 *Drain characteristics*

(iii) At a certain voltage corresponding to the point B, I_D suddenly increases. This effect is due to the Avalanche multiplication of electrons caused by breaking of covalent bonds of silicon atoms in the depletion region between the gate and the drain. The drain voltage at which the breakdown occurs is denoted by BV_{DGO}. The variation of I_D with V_{DS} when V_{GS} = 0 is shown in Fig. 3.3 by the curve $OPBC$.

When V_{GS} is negative and V_{DS} is increased When the gate is maintained at a negative voltage less than the negative cut-off voltage, the reverse voltage across the junction is further increased. Hence for a negative value of V_{GS}, the curve of I_D versus V_{DS} is similar to that for V_{GS} = 0, but the values of V_P and BV_{DGO} are lower, as shown in Fig. 3.3.

From the curves, it is seen that above the pinch-off voltage, at a constant value of V_{DS}, I_D increases with an increase of V_{GS}. Hence, a JFET is suitable for use as a voltage amplifier, similar to a transistor amplifier.

It can be seen from the curve that for voltage $V_{DS} = V_P$, the drain current is not reduced to zero. If the drain current is to be reduced to zero, the ohmic voltage drop along the channel should also be reduced to zero. Further, the reverse biasing to the gate-source PN junction essential for pinching off the channel would also be absent.

The drain current I_D is controlled by the electric field that extends into the channel due to reverse biased voltage applied to the gate; hence, this device has been given the name *Field Effect Transistor*.

In a bar of P-type semiconductor, the gate is formed due to N-type semiconductor. The working of the P-channel JFET will be similar to that of N-channel JFET with proper alterations in the biasing circuits; in this case holes will be the current carriers instead of electrons. The circuit symbols for N-channel and P-channel JFETs are shown in Fig. 3.4. It should be noted that the direction of the arrow points in the direction of conventional current which would flow into the gate if the PN junction was forward biased.

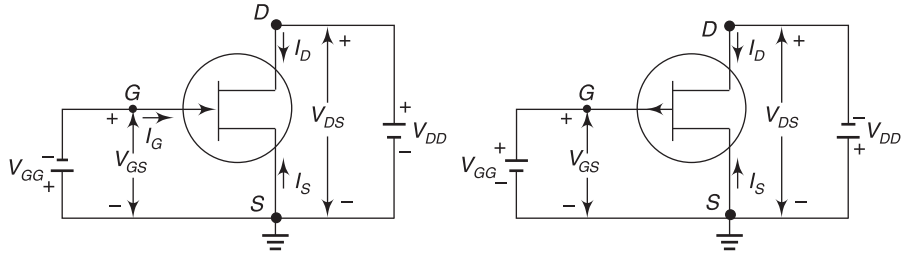


Fig. 3.4 *Circuit symbols for N- and P-channel JFET*

3.4 CHARACTERISTIC PARAMETERS OF THE JFET

[AU Nov 2015, Nov 2010, 6 marks]

In a JFET, the drain current I_D depends upon the drain voltage V_{DS} and the gate voltage V_{GS}. Any one of these variables may be fixed and the relation between the other two are determined. These relations are determined by the three parameters which are defined below.

(1) Mutual conductance or transconductance, g_m It is the slope of the transfer characteristic curves, and is defined by

$$g_m = \left(\frac{\partial I_D}{\partial V_{GS}}\right)_{V_{DS}} = \frac{\Delta I_D}{\Delta V_{GS}},\ V_{DS} \text{ held constant.}$$

It is the ratio of a small change in the drain current to the corresponding small change in the gate voltage at a constant drain voltage. The change in I_D and V_{GS} should be taken on the straight part of the transfer characteristics. It has the unit of conductance in mho.

(2) Drain resistance, r_d It is the reciprocal of the slope of the drain characteristics and is defined by

$$r_d = \left(\frac{\partial V_{DS}}{\partial I_D}\right)_{V_{GS}} = \frac{\Delta V_{DS}}{\Delta I_D},\ V_{GS} \text{ held constant.}$$

It is the ratio of a small change in the drain voltage to the corresponding small change in the drain current at a constant gate voltage. It has the unit of resistance in ohms.

The drain resistance at $V_{GS} = 0$ V, i.e., when the depletion regions of the channel are absent, is called as *drain-source ON resistance,* represented as R_{DS} or $R_{DS(ON)}$.

The reciprocal of r_d is called the drain conductance. It is denoted by g_d or g_{ds}.

(3) Amplification factor, μ It is defined by

$$\mu = \left(\frac{\partial V_{DS}}{\partial V_{GS}}\right)_{I_D} = -\frac{\Delta V_{DS}}{\Delta V_{GS}},\ I_D \text{ held constant.}$$

It is the ratio of a small change in the drain voltage to the corresponding small change in the gate voltage at a constant drain current. Here, the negative sign shows that when V_{GS} is increased, V_{DS} must be decreased for I_D to remain constant.

(4) Relationship among FET parameters As I_D depends on V_{DS} and V_{GS}, the functional equation can be expressed as

$$I_D = f(V_{DS}, V_{GS})$$

If the drain voltage is changed by a small amount from V_{DS} to $(V_{DS} + \Delta V_{DS})$ and the gate voltage is changed by a small amount from V_{GS} to $(V_{GS} + \Delta V_{GS})$, then the corresponding small change in I_D may be obtained by applying Taylor's theorem with neglecting higher order terms. Thus the small change ΔI_D is given by

$$\Delta I_D = \left(\frac{\partial I_D}{\partial V_{DS}}\right)_{V_{GS}} \Delta V_{DS} + \left(\frac{\partial I_D}{\partial V_{GS}}\right)_{V_{DS}} \Delta V_{GS}$$

Dividing both the sides of this equation by ΔV_{GS}, we obtain

$$\frac{\Delta I_D}{\Delta V_{GS}} = \left(\frac{\partial I_D}{\partial V_{DS}}\right)_{V_{GS}} \left(\frac{\Delta V_{DS}}{\Delta V_{GS}}\right) + \left(\frac{\partial I_D}{\partial V_{GS}}\right)_{V_{DS}}$$

If I_D is constant, then $\dfrac{\Delta I_D}{\Delta V_{GS}} = 0$

Therefore, we have

$$0 = \left(\frac{\partial I_D}{\partial V_{DS}}\right)_{V_{GS}} \left(\frac{\Delta V_{DS}}{\Delta V_{GS}}\right)_{I_D} + \left(\frac{\partial I_D}{\partial V_{GS}}\right)_{V_{DS}}$$

Substituting the values of the partial differential coefficients, we get

$$0 = \left(\frac{1}{r_d}\right)(-\mu) + g_m$$

Hence,

$$\mu = g_m r_d$$

Therefore, amplification factor (μ) is the product of transconductance (g_m) and drain resistance (r_d).

(5) Power dissipation, P_D The FET's continuous power dissipation, P_D, is the product of I_D and V_{DS}.

(6) Pinch-off Voltage, V_P A single-ended-geometry junction FET is shown in Fig. 3.5 in which the diffusion is done from one side only. The substrate is of *P*-type material which is epitaxially grown on an *N*-type channel. A *P*-type gate is then diffused into the *N*-type channel. The substrate functions as a second gate which is of relatively low resistivity material. The diffused gate is also of very low resistivity material, allowing the depletion region to spread mostly into the *N*-type channel. A slab of *N*-type semiconductor is sandwiched between two layers of *P*-type material forming two *PN* junction in this device.

The gate reverse voltage that removes all the free charge from the channel is called the pinch-off voltage V_P. We consider that the *P*-type region is doped with N_A acceptor atoms, the *N*-type region is doped with N_D donor atoms and the junction formed is abrupt.

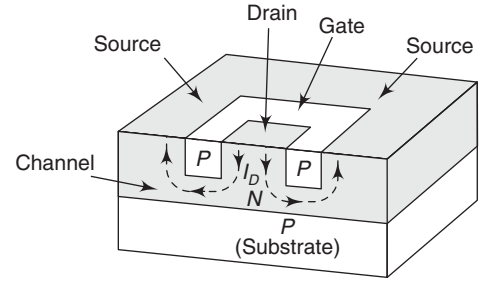


Fig. 3.5 *Single-ended-geometry junction FET*

Moreover, if the acceptor impurity density is assumed to be much larger than the donor density, then the depletion region width in the *P*-region will be much smaller than the depletion width of the *N*-region, i.e., $N_A >> N_D$, then $W_P << W_N$ and $W_P = W$. We know that, the relationship between potential and charge density is given by

$$\frac{d^2V}{dx^2} = \frac{-qN_D}{\varepsilon}$$

Integrating the above equation subject to boundary conditions, we get

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

Integrating again, we get

$$V = \frac{-qN_D}{2\varepsilon}(x^2 - 2Wx)$$

At $x = W$, $V = V_B$ which is the junction or barrier potential. Thus,

$$V_B = \frac{qN_D W^2}{2\varepsilon}$$

As the barrier potential represents a reverse voltage, it is lowered by an applied forward voltage $V(x)$ at x and it is expressed as $V_B = V_P - V(x)$. Now, the space-charge width, $W_n(x) = W(x)$ at a distance x along the channel in Fig. 3.6 becomes

$$W(x) = a - b(x) = \left\{\frac{2\varepsilon}{qN_D}[V_0 - V(x)]\right\}^{\frac{1}{2}}$$

where ε is the dielectric constant of channel material, q is the magnitude of electronic charge, V_0 is the junction contact potential at x, $V(x)$ is the applied potential across space-charge region at x and is a negative value for an applied reverse bias, a is the metallurgical channel thickness between the substrate and P^+ gate region, $a - b(x)$ is the penetration $W(x)$ of depletion region into channel at a point x along channel as shown in Fig. 3.6 and W_{PO} is the depletion region width at pinch-off.

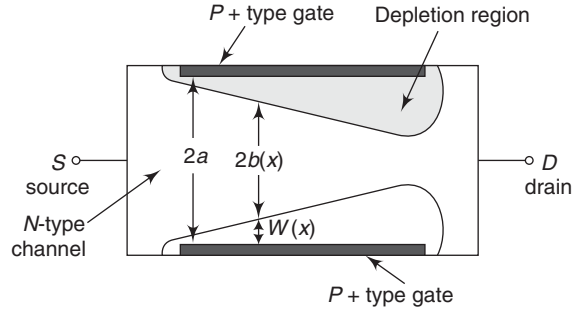


Fig. 3.6 *Detailed structure of N-channel JFET*

If the drain current is zero, $b(x)$ and $V(x)$ are independent of x and hence $b(x) = b$. If we substitute $b(x) = b = 0$ in the previous equation and solve for V with the assumption that $|V_0| << |V|$, the pinch-off voltage V_P can be obtained as

$$|V_P| = \frac{qN_D}{2\varepsilon}a^2$$

Here, at pinch-off, the depletion width $W_{PO} = a$ when $|V_P| = |V_{GS}|$ and hence

$$W_{PO} = \left(\frac{2\varepsilon}{qN_D}V_{GS}\right)^{1/2}$$

To study the effect of V_{GS}, given $V_{GS} = V_0 - V(x)$ in the space charge width equation, we get

$$a - b = \left\{\frac{2\varepsilon}{qN_D}V_{GS}\right\}^{1/2}$$

Upon solving, we get

$$a - b = \left[\frac{a^2}{V_P} V_{GS}\right]^{1/2}$$

Further simplifying,

$$\frac{a-b}{a} = \left(\frac{V_{GS}}{V_P}\right)^{1/2}$$

Therefore,

$$\left(1 - \frac{b}{a}\right)^2 = \frac{V_{GS}}{V_P}$$

Hence, the gate source voltage is

$$V_{GS} = \left(1 - \frac{b}{a}\right)^2 V_P$$

For a reverse biased $P^+ N$ junction, V_{GS} must be a negative voltage across the gate junction and is independent of distance along the channel if $I_D = 0$.

EXAMPLE 3.1

When a reverse gate voltage of 12 V is applied to JFET, the gate current is 1 nA. Determine the resistance between gate and source.

Solution Given $V_{GS} = 12$ V and $I_G = 10^{-9}$ A.

Therefore, gate-to-source resistance $= \dfrac{V_{GS}}{I_G} = \dfrac{12}{10^{-9}} = 12{,}000\ \text{M}\Omega$

EXAMPLE 3.2

When the reverse gate voltage of JFET changes from 4.0 to 3.9 V, the drain current changes from 1.3 to 1.6 mA. Find the value of transconductance.

Solution $\Delta V_{GS} = 4.0 - 3.9 = 0.1$ V

$\Delta I_D = 1.6 - 1.3 = 0.3$ mA

Therefore, transconductance, $g_m = \dfrac{\Delta I_D}{\Delta V_{GS}} = \dfrac{0.3 \times 10^{-3}}{0.1} = 3\ \text{m}\ \Omega^{-1}$

3.5 EXPRESSION FOR SATURATION DRAIN CURRENT

[AU May 2015, 10 marks]

For the transfer characteristics, V_{DS} is maintained constant at a suitable value greater than the pinch-off voltage V_P. The gate voltage V_{GS} is decreased from zero till I_D is reduced to zero. The transfer characteristics I_D versus V_{GS} is shown in Fig. 3.7. The shape of the transfer characteristic is very nearly a parabola. It is found that the characteristic is approximately represented by the parabola,

$$I_{DS} = I_{DSS}\left(1 - \frac{V_{GS}}{V_P}\right)^2 \tag{3.3}$$

where I_{DS} is the saturation drain current, I_{DSS} is the value of I_{DS} when $V_{GS} = 0$, and V_P is the pinch-off voltage. The expression for the drain current in saturation region is derived in Section 3.18.

Differentiating Eq. (3.3) with respect to V_{GS} we can obtain an expression for g_m.

$$\frac{\partial I_{DS}}{\partial V_{GS}} = I_{DSS} \times 2\left(1 - \frac{V_{GS}}{V_P}\right)\left(-\frac{1}{V_P}\right)$$

We know that $g_m = \dfrac{\delta I_{DS}}{\delta V_{GS}}$, V_{DS} is constant.

Therefore, $$g_m = \frac{-2\, I_{DSS}}{V_P}\left(1 - \frac{V_{GS}}{V_P}\right) \quad (3.4)$$

Now from Eq. (3.3), we have

$$\left(1 - \frac{V_{GS}}{V_P}\right) = \sqrt{\frac{I_{DS}}{I_{DSS}}}$$

Substituting this value in Eq. (3.4), we get

$$g_m = \frac{-2\sqrt{I_{DS}\, I_{DSS}}}{V_P} \quad (3.5)$$

Suppose $g_m = g_{mo}$, when $V_{GS} = 0$, then from Eq. (3.4)

$$g_{mo} = -\frac{2I_{DSS}}{V_P} \quad (3.6)$$

Therefore, from Eqs (3.4) and (3.6), we get

$$g_m = g_{mo}\left(1 - \frac{V_{GS}}{V_P}\right) \quad (3.7)$$

Equation (3.5) shows that g_m varies as the square root of the saturation drain current I_{DS}, and Eq. (3.7) shows that g_m decreases linearly with increase of V_{GS}.

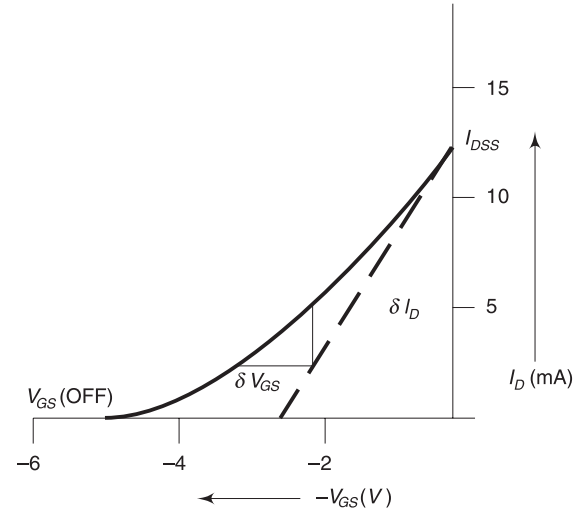


Fig. 3.7 *Transfer characteristics of JFET*

3.6 SLOPE OF THE TRANSFER CHARACTERISTIC AT I_{DSS}

From Eq. (3.5), we have

$$g_m = \frac{-2\sqrt{I_{DS}\, I_{DSS}}}{V_P}$$

or $$\frac{\partial I_{DS}}{\partial V_{GS}} = \frac{-2\sqrt{I_{DS}\, I_{DSS}}}{V_P}$$

Substituting $I_{DS} = I_{DSS}$,

$$\frac{\partial I_{DS}}{\partial V_{GS}} = -\frac{2I_{DSS}}{V_P} = \frac{I_{DSS}}{\frac{-V_P}{2}}$$

This equation shows that the tangent to the curve at $I_{DS} = I_{DSS}$, $V_{GS} = 0$, will have an intercept at $\frac{-V_P}{2}$ on the axis of V_{GS} as shown in Fig. 3.7. Therefore, the value of V_P can be found by drawing the tangent at $I_{DS} = I_{DSS}$, $V_{GS} = 0$.

The gate-source cut off voltage, $V_{GS(off)}$, on the transfer characteristic is equal to the pinch off voltage, V_P, on the drain characteristics, i.e., $V_P = |V_{GS(off)}|$.

Therefore,
$$I_D = I_{DSS}\left[1 - \frac{V_{GS}}{V_{GS(off)}}\right]^2$$

3.7 BIASING FOR ZERO CURRENT DRIFT

Figure 3.8 shows the transfer characteristics of *N* channel JFET for different values of temperature. Here, the drain current varies due to change in temperature. The two factors which make drain current to change with temperature are the mobility of majority carriers and the depletion region width.

The mobility of majority carriers decreases with increase in temperature. As the temperature increases, the lattice ions vibrate more vigorously and hence, the carriers cannot move freely in the crystalline structure. For the given gate-source voltage V_{GS}, their velocity is decreased and this reduces the drain current. The reduction in I_D is 0.7% for 1°C increase in the temperature.

The width of the depletion region (increase in channel width) decreases with increase in temperature. This allows I_D to increase and the increase in I_D is equivalent to a change of 2.2 mV/1°C in $|V_{GS}|$. This is a similar phenomenon to the change of $|V_{BE}|$ of 2.5 mV/°C in bipolar transistor.

Hence, it is required to design a biasing circuit which compensates for these two factors, so that there is no change of drain current with temperature. Such a biasing is called Biasing for zero current drift.

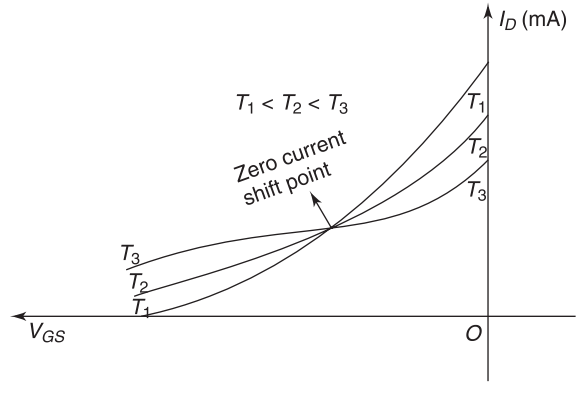


Fig. 3.8 *Transfer characteristics for an N channel FET as a function of temperature T*

Due to the change in I_D and change in V_{GS} for 1°C, the condition for zero current drift can be obtained as

$$0.007|I_D| = 0.0022\, g_m$$

i.e., $$\frac{|I_D|}{g_m} = 0.314 \text{ V}$$

We know that, $I_D = I_{DSS}\left[1-\frac{V_{GS}}{V_P}\right]^2$ and $g_m = \frac{-2I_{DSS}}{V_P}\left[1-\frac{V_{GS}}{V_P}\right]$.

Hence, $$\frac{|I_D|}{g_m} = \frac{I_{DSS}\left(1-\frac{V_{GS}}{V_P}\right)^2}{-\frac{2I_{DSS}}{V_P}\left[1-\frac{V_{GS}}{V_P}\right]} = \frac{|V_P|-|V_{GS}|}{-2} = 0.314 \text{ V}$$

Therefore, $$|V_P|-|V_{GS}| = -0.628\text{V} \approx -0.63\text{ V}$$

From the above equation, it is seen that, if the value of V_P is known, the value of V_{GS} for zero drift current can be obtained.

When V_{GS} is adjusted for zero drift current, the drain current I_D and the transconductance g_m are given by

$$I_D = I_{DSS}\left[1-\frac{V_P+0.63}{V_P}\right]^2 = I_{DSS}\left[\frac{0.63}{V_P}\right]^2$$

and $$g_m = g_{mo}\left[1-\frac{V_P+0.63}{V_P}\right] = g_{mo}\left(\frac{0.63}{|V_P|}\right)$$

EXAMPLE 3.3

A FET has a drain current of 4 mA. If I_{DSS} = 8 mA and $V_{GS(off)}$ = –6 V. Find the values of V_{GS} and V_P.

Solution The drain current, $I_D = I_{DSS}\left[1 - \frac{V_{GS}}{V_{GS(off)}}\right]^2$

$$4 = 8\left[1 + \frac{V_{GS}}{6}\right]^2$$

$$1 + \frac{V_{GS}}{6} = \sqrt{\frac{4}{8}} = \frac{1}{\sqrt{2}} = 0.707$$

Therefore, $$V_{GS} = -1.76 \text{ V}$$

$$V_P = |V_{GS(off)}| = 6 \text{ V}$$

EXAMPLE 3.4

An *N*-channel JFET has I_{DSS} = 8 mA and V_P = – 5 V. Determine the minimum value of V_{DS} for pinch-off region and the drain current I_{DS}, for V_{GS} = – 2 V in the pinch-off region.

Solution The minimum value of V_{DS} for pinch-off to occur for $V_{GS} = -2$ V is

$$V_{DS\,\min} = V_{GS} - V_P = -2 - (-5) = 3 \text{ V}$$

$$I_{DS} = I_{DSS}\left[1 - \frac{V_{GS}}{V_P}\right]^2$$

$$= 8 \times 10^{-3}\,[1 - (-2)/(-5)]^2 = 2.88 \text{ mA}$$

EXAMPLE 3.5

Determine the pinch-off voltage for an *N*-channel silicon FET with a channel width of 5.6×10^{-4} cm and a donor concentration of 10^{15} /cm^3. Given, the dielectric constant of Si is 12.

Solution Given $N_D = 10^{15}/\text{cm}^3 = 10^{21}/\text{m}^3$, $\varepsilon = \varepsilon_r\,\varepsilon_o = 12\,\varepsilon_0$ and

$a = 5.6 \times 10^{-4}$ cm $= 5.6 \times 10^{-6}$m

The pinch-off voltage, $|V_P| = \dfrac{qN_D}{2\varepsilon}\,a^2$

where $\varepsilon =$ dielectric constant of channel material (Si) $= \varepsilon_r\,\varepsilon_0 = 12\,\varepsilon_0$

$q =$ magnitude of electronic charge

a is in metres and N_D is in electrons/m^3

$$|V_P| = \frac{1.602 \times 10^{-19} \times 10^{21}}{2 \times 12 \times 8.854 \times 10^{-12}} \times (5.6 \times 10^{-6})^2 = 23.6 \text{ V}$$

EXAMPLE 3.6

For a *P*-channel silicon FET, with an effective channel width, $a = 2 \times 10^{-4}$ cm and channel resistively ρ = 20 Ω-cm, find the pinch-off voltage.

Solution We know that the pinch-off voltage for *P*-channel FET, $V_P = \dfrac{qN_A}{2\varepsilon} \times a^2$

For silicon, $\varepsilon = 12\,\varepsilon_0$ and $\mu_p = 500$ cm^2/V-sec

$$\sigma = \frac{1}{\rho} = qN_A\mu_p$$

i.e., $$qN_A = \frac{1}{\rho\mu_p} = \frac{1}{20 \times 500} = 1 \times 10^{-4}$$

Therefore, $$V_P = \frac{qN_A}{2\varepsilon} \times a^2 = \frac{1 \times 10^{-4}}{2 \times 12 \times 8.854 \times 10^{-12}} \times (2 \times 10^{-4})^2 = 1.89 \text{ V}$$

3.8 COMPARISON OF JFET AND BJT [AU May 2014 and Nov 2011, 6 marks]

1. FET operation depends only on the flow of majority carriers-holes for *P*-channel FETs and electrons for *N*-channel FETs. Therefore, they are called unipolar devices. Bipolar transistor (BJT) operation depends on both minority and majority current carriers.
2. As FET has no junctions and the conduction is through an *N*-type or *P*-type semiconductor material, FET is less noisy than BJT.
3. As the input circuit of FET is reverse biased, FET exhibits a much higher input impedance (in the order of 100 MΩ) and lower output impedance and there will be a high degree of isolation between input and output. So, FET can act as an excellent buffer amplifier but the BJT has low input impedance because its input circuit is forward biased.
4. FET is a voltage controlled device, i.e., voltage at the input terminal controls the output current, whereas BJT is a current controlled device, i.e., the input current controls the output current.
5. FETs are much easier to fabricate and are particularly suitable for ICs because they occupy less space than BJTs.
6. The performance of BJT is degraded by neutron radiation because of the reduction in minority-carrier lifetime, whereas FET can tolerate a much higher level of radiation since they do not rely on minority carriers for their operation.
7. The performance of FET is relatively unaffected by ambient temperature changes. As it has a negative temperature coefficient at high current levels, it prevents the FET from thermal breakdown. The BJT has a positive temperature co-efficient at high current levels which leads to thermal breakdown.
8. Since FET does not suffer from minority carrier storage effects, it has higher switching speeds and cut-off frequencies. BJT suffers from minority carrier storage effects and therefore has lower switching speed and cut-off frequencies.
9. FET amplifiers have low gain bandwidth product due to the junction capacitive effects and produce more signal distortion except for small signal operation.
10. BJTs are cheaper to produce than FETs.

3.9 APPLICATIONS OF JFET [AU May 2014, 6 marks]

1. FET is used as a buffer in measuring instruments, receivers since it has high input impedance and low output impedance.
2. FETs are used in RF amplifiers in FM tuners and in communication equipment for its low noise level.
3. Since the input capacitance is low, FETs are used in cascade amplifiers in measuring and test equipments.
4. Since the device is voltage controlled, it is used as a voltage variable resistor in operational amplifiers and tone controls.
5. FETs are used in mixer circuits in FM and TV receivers, and in communication equipment because inter modulation distortion is low.
6. It is used in oscillator circuits because frequency drift is low.

7. As the coupling capacitor is small, FETs are used in low frequency amplifiers in hearing aids and in inductive transducers.
8. FETs are used in digital circuits in computers, LSD and memory circuits because of its small size.

3.10 METAL OXIDE SEMICONDUCTOR FIELD EFFECT TRANSISTOR (MOSFET)

[AU Dec 2016, May 2015, May 2014, June 2011 and June 2010, 10 marks]

MOSFET is the common term for the Insulated Gate Field Effect Transistor (IGFET). There are two basic forms of MOSFET: (i) Enhancement MOSFET and (ii) Depletion MOSFET.

Principle By applying a transverse electric field across an insulator deposited on the semiconducting material, the thickness and hence the resistance of a conducting channel of a semiconducting material can be controlled.

In a depletion MOSFET, the controlling electric field reduces the number of majority carriers available for conduction, whereas in the enhancement MOSFET, application of electric field causes an increase in the majority carrier density in the conducting regions of the transistor.

3.11 ENHANCEMENT MOSFET

[AU May 2017, Dec 2015, May 2015, Nov 2014, May 2014 and Nov 2011, 10 marks]

Construction The construction of an *N*-channel enhancement MOSFET is shown in Fig. 3.9(a) and the circuit symbols for an *N*-channel and a *P*-channel enhancement MOSFET are shown in Fig. 3.9(b) and (c), respectively. As there is no continuous channel in an enhancement MOSFET, this condition is represented by the broken line in the symbols.

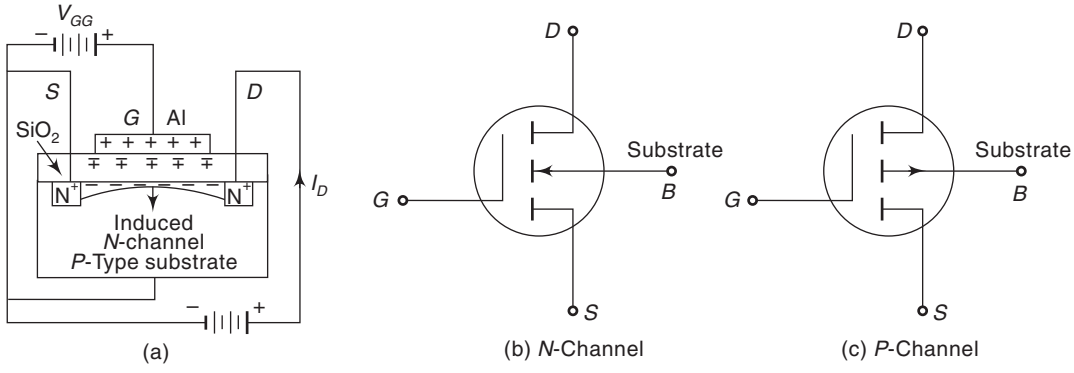


Fig. 3.9 *(a) N-channel enhancement MOSFET, (b) and (c) Circuit symbols for enhancement MOSFET*

Two highly doped N^+ regions are diffused in a lightly doped substrate of *P*-type silicon substrate. One N^+ region is called the source *S* and the other one is called the drain *D*. They are separated by 1 mil (10^{-3} inch). A thin insulating layer of SiO_2 is grown over the surface of the structure and holes are cut into the oxide layer, allowing contact with source and drain. Then a thin layer of metal aluminium is formed over the layer of SiO_2. This metal layer covers the entire channel region and it forms the gate *G*.

The metal area of the gate, in conjunction with the insulating oxide layer of SiO_2 and the semiconductor channel forms a parallel plate capacitor. This device is called the insulated gate FET because of the insulating layer of SiO_2. This layer gives an extremely high input impedance for the MOSFET.

Operation If the substrate is grounded and a positive voltage is applied at the gate, the positive charge on *G* induces an equal negative charge on the substrate side between the source and drain regions. Thus, an electric field is produced between the source and drain regions. The direction of the electric field is perpendicular to the plates of the capacitor through the oxide. The negative charge of electrons which are minority carriers in the *P*-type substrate forms an inversion layer. As the positive voltage on the gate increases, the induced negative charge in the semiconductor increases. Hence, the conductivity increases and current flows from source to drain through the induced channel. Thus the drain current is enhanced by the positive gate voltage as shown in Fig. 3.10.

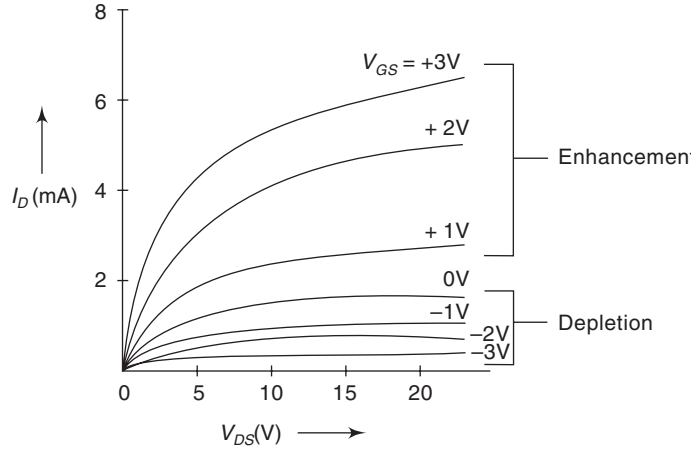


Fig. 3.10 *Volt-ampere characteristics of MOSFET*

3.12 DEPLETION MOSFET

[AU May 2016, Dec 2015, Nov 2013, June 2013 and Nov 2010, 16 marks]

The construction of an *N*-channel depletion MOSFET is shown in Fig. 3.11(a) where an *N*-channel is diffused between the source and drain to the basic structure of MOSFET. The circuit symbols for an *N*-channel and a *P*-channel depletion MOSFET are shown in Figs 3.11(b) and (c), respectively.

With $V_{gs} = 0$ and the drain *D* at a positive potential with respect to the source, the electrons (majority carriers) flow through the *N*-channel from *S* to *D*. Therefore, the conventional current I_D flows through the channel *D* to *S*. If the gate voltage is made negative, positive charge consisting of holes is induced in the channel through SiO_2 of the gate-channel capacitor. The introduction of the positive charge causes depletion of mobile electrons in the channel. Thus a depletion region is produced in the channel. The shape of the depletion region depends on V_{GS} and V_{DS}. Hence the channel will be wedge shaped as shown in Fig. 3.11. When V_{DS} is increased, I_D increases and it becomes practically constant at a certain value of V_{DS}, called the pinch-off voltage. The drain current I_D almost gets saturated beyond the pinch-off voltage.

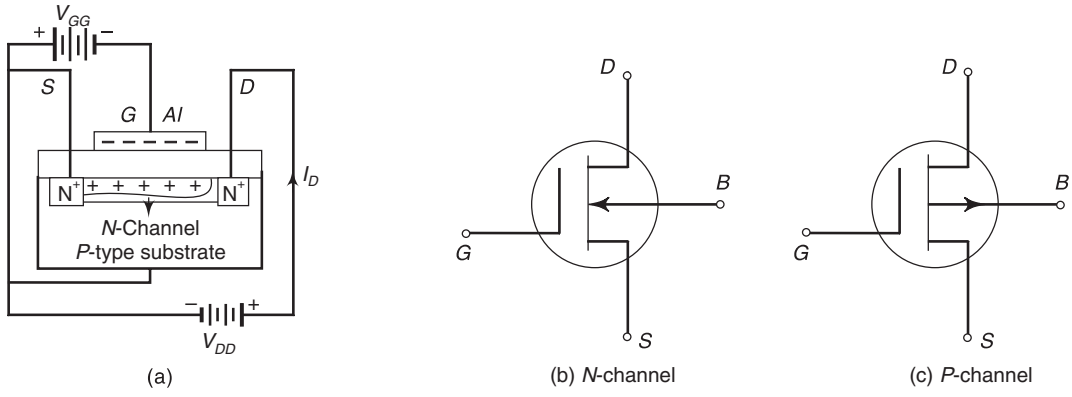


Fig. 3.11 *(a) N-channel depletion MOSFET, (b) and (c) Circuit symbols for depletion MOSFETs*

Since the current in an FET is due to majority carriers (electrons for an *N*-type material), the induced positive charges make the channel less conductive, and I_D drops as V_{GS} is made negative.

The depletion MOSFET may also be operated in an enhancement mode. It is only necessary to apply a positive gate voltage so that negative charges are induced into the *N*-type channel. Hence, the conductivity of the channel increases and I_D increases. As the depletion MOSFET can be operated with bipolar input signals irrespective of doping of the channel, it is also called as *dual mode MOSFET*. The volt-ampere characteristics are indicated in Fig. 3.10.

The curve of I_D versus V_{GS} for constant V_{DS}, as shown in Fig. 3.12, is called the transconductance curve for D-MOSFET which is very similar to the curve for JFET. When V_{GS} is negative, $I_D < I_{DSS}$. When $V_{GS} = V_{GS(off)}$, I_D is decreased to approximately 0 mA. When V_{GS} is positive, $I_D > I_{DSS}$. Hence, I_{DSS} is not the maximum possible value for MOSFET.

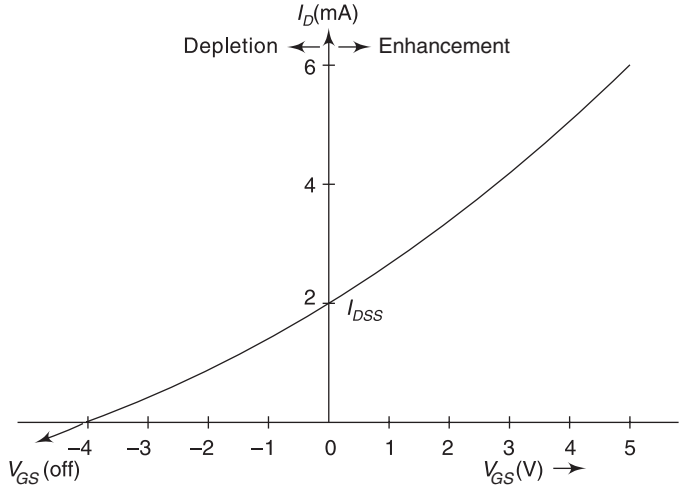


Fig. 3.12 *Transfer characteristics of MOSFET*

3.12.1 Effect of Channel Length Modulation [AU May 2017, Dec 2015, 6 marks]

In actual MOSFET characteristic as shown in Fig. 3.13, a non-zero slope exists beyond the saturation point. For the saturation region, i.e., ($V_{DS} > V_{DS}$ (sat)), the effective channel length decreases and this phenomenon is called channel length modulation. For an *N*-channel device, the slope of the curve in the saturation region can be expressed by using the drain current I_D given by

$$I_D = K_N(V_{GS} - V_{TN})^2 (1 + \lambda V_{DS})$$

where λ is a positive quantity called the channel length modulation parameter or λ^{-1} is analogous to the early voltage in bipolar transistors, K_N is the conduction parameter and V_{TN} is the threshold voltage. The curves are extended so that the intercept at a point $I_D = 0$, $V_{DS} = -V_E$ which means that $V_E = 1/\lambda$.

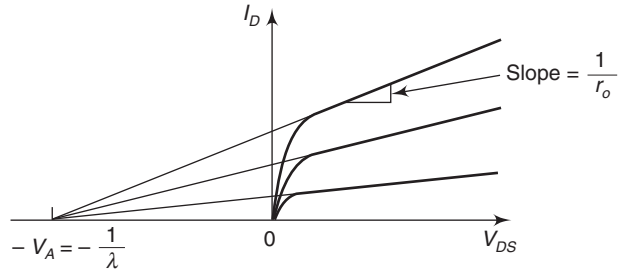


Fig. 3.13 *Effect of channel length modulation due to the non-zero slope in the saturation region, resulting in a finite output resistance*

The output resistance due to the channel length modulation is expressed by

$$r_0 = \left(\frac{\partial I_D}{\partial V_{DS}}\right)\Bigg|_{V_{GS}} = \text{const.}$$

The output resistance can be determined at the *Q*-point by

$$r_0 = [\lambda K_N(V_{GSQ} - V_{TN})^2]^{-1}$$

$$= \left[\lambda I_{DSS}\left(1 - \frac{V_{GSQ}}{V_P}\right)^2\right]^{-1}$$

i.e.,

$$r_0 \cong [\lambda I_{DQ}]^{-1} = \frac{1}{\lambda I_{DQ}} = \frac{V_E}{I_{DQ}}$$

The output resistance is an important factor in the analysis of small signal equivalent circuit of MOSFET.

3.12.2 Threshold Voltage in MOSFET [AU May 2015, 6 marks]

The threshold voltage is the applied gate voltage needed to achieve the threshold inversion point. The condition for this threshold inversion point is that the surface potential, $\Phi_s = 2\Phi_{fp}$ for *P*-type semiconductor and $\Phi_s = 2\Phi_{fn}$ for *N*-type semiconductor. Here, Φ_{fp} and Φ_{fn} are the difference in potential

between the Fermi energy levels for *P*-type and *N*-type semiconductors. The potential Φ_{fp} for *P*-type semiconductor is given by

$$\Phi_{fp} = V_T \ln\left(\frac{N_A}{n_i}\right)$$

where $V_T = \dfrac{kT}{q} = 26\,\text{mV}$ at room temperature, N_A is the acceptor doping concentration and n_i is the intrinsic carrier concentration. The threshold voltage can be derived in terms of the electrical and geometrical properties of the MOS capacitor.

Figure 3.14 shows the charge distribution in the MOS capacitor at the threshold inversion point for a *P*-type semiconductor substrate. Consider that the space charge width reaches its maximum value. Let us assume that there are equivalent oxide charge, Q'_{ss}, and the positive charge on the metal gate at threshold, Q'_{mT}. Here, the prime on Q'_{ss} and Q'_{mT} indicates the charge per unit area and the inversion layer charge is neglected at this threshold inversion point.

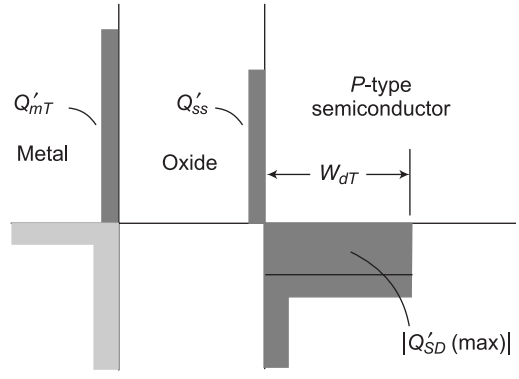


Fig. 3.14 *Charge distribution through the MOS capacitor for a P-type substrate*

The space charge width is given by

$$W_d = \left(\frac{2\varepsilon_s \Phi_s}{qN_A}\right)^{1/2}$$

where ε_s is the permittivity of the semiconductor and q is the charge of an electron. From conservation of charge, we can write

$$Q'_{mT} + Q'_{ss} = |Q'_{SD}(\text{max})|$$

where $|Q'_{SD}(\text{max})| = qN_A W_{dT}$ is the magnitude of the maximum space charge density per unit area of the depletion region and W_{dT} is the maximum space charge width.

Figure 3.15 shows the energy-band diagram of the MOS device with an applied positive gate voltage. If a gate voltage is applied, the voltage across the oxide, V_{ox}, and the surface potential, Φ_s, will change. Therefore,

$$V_G = \Delta V_{ox} + \Delta \Phi_s = V_{ox} + \Phi_s + \Phi_{ms}$$

where Φ_{ms} is metal-semiconductor work function potential difference. We know that, at threshold inversion point, $V_G = V_{TN}$. Since the surface potential for a *P*-type semiconductor is $\Phi_s = 2\Phi_{fp}$, the threshold voltage that creates the electron inversion layer charge is

$$V_{TN} = V_{oxT} + 2\Phi_{fp} + \Phi_{ms}$$

where V_{oxT} is the voltage across the oxide at the threshold inversion point.

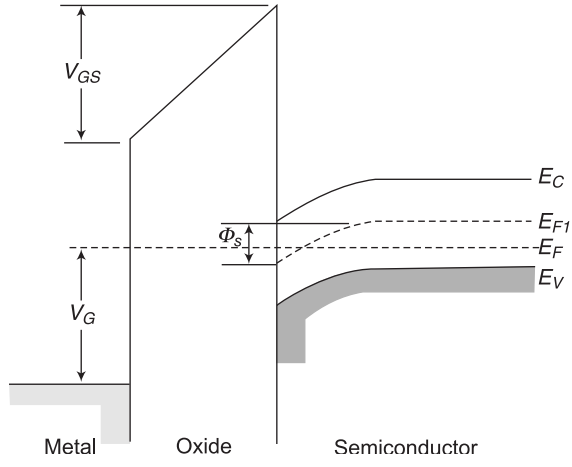


Fig. 3.15 *Energy band diagram for MOSFET with a +ve gate bias, V_G*

This voltage V_{oxT} is related to the charge on the metal and oxide capacitance as given by

$$V_{oxT} = \frac{Q'_{mT}}{C_{ox}} = (|\,Q'_{SD}(\max)\,| - Q'_{ss})\frac{1}{C_{ox}}$$

where $Q'_{mT} = (|\,Q'_{SD}(\max)\,| - Q'_{ss})$ and C_{ox} is the oxide capacitance per unit area. Therefore, the threshold voltage is

$$\begin{aligned} V_{TN} &= V_{oxT} + 2\Phi_{fp} + \Phi_{ms} \\ &= (|\,Q'_{SD}(\max)\,| - Q'_{ss})\frac{1}{C_{ox}} + 2\Phi_{fp} + \Phi_{ms} \\ &= \left(|\,Q'_{SD}(\max)\,| - Q'_{ss}\right)\left(\frac{t_{ox}}{\varepsilon_{ox}}\right) + 2\Phi_{fp} + \Phi_{ms} \end{aligned}$$

where $C_{ox} = \dfrac{\varepsilon_{ox}}{t_{ox}}$. Hence, the threshold voltage is a function of semiconductor doping, oxide charge Q'_{ss} and oxide thickness t_{ox} for a given semiconductor material, oxide material, and gate metal.

The threshold voltage must be within the voltage range of a circuit design. If the threshold voltage of a MOSFET is beyond the operating voltage range, then the circuit cannot be turned ON and OFF.

EXAMPLE 3.7

Consider an N^+ polysilicon gate and a *P*-type silicon substrate doped to $N_A = 3 \times 10^{22}$ m^{-3}. Assume $n_i = 1.5 \times 10^{16}$ m^{-3}, $\Phi_{ms} = -1.13$ V, $\varepsilon_s = 11.7$ for Si, $\varepsilon_{ox} = 3.9$ for SiO_2 and $Q'_{ss} = 10^{15}$ m^{-2}. Determine the oxide thickness such that $V_{TN} = 0.75$ V.

Solution The potential Φ_{fp} for *P*-type semiconductor is

$$\Phi_{fp} = V_T \ln\left(\frac{N_A}{n_i}\right)$$

$$= 26\times10^{-3} \times \ln\left(\frac{3\times10^{22}}{1.5\times10^{16}}\right) = 0.376\,\text{V}$$

The space charge width is

$$W_d = \left(\frac{2\varepsilon_s \Phi_s}{qN_A}\right)^{1/2}$$

The space charge width is maximum when $\Phi_s = 2\Phi_{fp}$ and it is given by

$$W_{dT} = \left(\frac{4\varepsilon_s \Phi_{fp}}{qN_A}\right)^{1/2}$$

$$= \left\{\frac{4\times11.7\times8.854\times10^{-12}\times0.376}{1.602\times10^{-19}\times3\times10^{22}}\right\}^{1/2} = 0.18\,\mu\text{m}$$

Therefore, the magnitude of the maximum space charge density is

$$|Q'_{SD}(\max)| = qN_A W_{dT}$$
$$= 1.602\times10^{-19}\times3\times10^{22}\times0.18\times10^{-6}$$
$$= 8.65\times10^{-4}\,\text{C/m}^2$$

The threshold voltage equation is

$$V_{TN} = \left(|Q'_{SD}(\max)| - Q'_{ss}\right)\left(\frac{t_{ox}}{\varepsilon_{ox}}\right) + 2\Phi_{fp} + \Phi_{ms}$$

$$0.75 = \frac{[(8.65\times10^{-4}) - 10^{15}\times1.602\times10^{-19}]}{3.9\times8.854\times10^{-12}} \times t_{ox} + 2(0.376) - 1.13$$

$$= 2.04\times10^{7}\times t_{ox} - 0.248$$

Hence,

$$t_{ox} = \frac{0.75 + 0.248}{2.04 \times 10^{7}} = 0.4892 \times 10^{-7}\ \text{m} = 48.92\ \text{nm}$$

EXAMPLE 3.8

Calculate the threshold voltage of a MOSFET using the aluminum gate. Consider a *P*-type silicon substrate at $T = 300$ K doped to $N_A = 10^{20}$ m^{-3}. Here, $Q'_{ss} = 10^{14}$ m^{-2} and $t_{ox} = 5 \times 10^{-8}$ m. Assume the oxide is silicon dioxide, $n_i = 1.5 \times 10^{16}$ m^{-3}, $\varepsilon_s = 11.7$ and $\Phi_{ms} = -1$ V.

Solution We know that,

$$\Phi_{fp} = V_T \ln\left(\frac{N_A}{n_i}\right)$$

$$= 26 \times 10^{-3} \times \ln\left(\frac{10^{20}}{1.5 \times 10^{16}}\right) = 0.228\ \text{V}$$

The maximum space charge width is

$$W_{dT} = \left(\frac{4\varepsilon_s \Phi_{fp}}{qN_A}\right)^{1/2}$$

$$= \left\{\frac{4 \times 11.7 \times 8.854 \times 10^{-12} \times 0.228}{1.602 \times 10^{-19} \times 10^{20}}\right\}^{1/2} = 2.43\,\mu\text{m}$$

Therefore, the magnitude of the maximum space charge density is

$$|Q'_{SD}(\text{max})| = qN_A W_{dT}$$

$$= 1.602 \times 10^{-19} \times 10^{20} \times 2.43 \times 10^{-6}$$

$$= 3.89 \times 10^{-5}\ \text{C/m}^2$$

The threshold voltage can be calculated as

$$V_{TN} = \left(|Q'_{SD}(\text{max})| - Q'_{ss}\right)\left(\frac{t_{ox}}{\varepsilon_{ox}}\right) + 2\Phi_{fp} + \Phi_{ms}$$

$$= \left[(3.89 \times 10^{-5}) - 10^{14} \times 1.602 \times 10^{-19}\right]\left[\frac{5 \times 10^{-8}}{3.9 \times 8.854 \times 10^{-12}}\right] + (2 \times 0.228) - 1$$

$$= -0.511\ \text{V}$$

Since V_{TN} is negative for a *P*-type substrate, it implies that the device MOSFET is in depletion mode. The *P*-type substrate should be heavily doped to obtain an enhancement mode MOSFET.

3.12.3 Temperature Effects

The threshold voltage V_{TN} or V_{TP} and conduction parameter K_N or K_P are functions of temperature. The magnitude of threshold voltage decreases with temperature and hence the drain current I_D increases with temperature at a given V_{GS}. The conduction parameter is directly proportional to mobility μ_n or μ_p of the carrier, which increases as the temperature decreases. Here the temperature dependent of mobility is larger than that of the threshold voltage. Hence the net effect of decreasing drain current at a given V_{GS} due to increase in temperature provides a negative feedback condition and hence the stability for a power MOSFET.

3.13 COMPARISON OF MOSFET WITH JFET [AU May 2014, 6 marks]

1. In enhancement and depletion types of MOSFET, the transverse electric field induced across an insulating layer deposited on the semiconductor material controls the conductivity of the channel. In the JFET, the transverse electric field across the reverse biased *PN* junction controls the conductivity of the channel.
2. The gate leakage current in a MOSFET is of the order of 10^{-12} A. Hence the input resistance of a MOSFET is very high in the order of 10^{10} to 10^{15} Ω. The gate leakage current of a JFET is of the order of 10^{-9} A and its input resistance is of the order of 10^{8} Ω.
3. The output characteristics of the JFET are flatter than those of the MOSFET and hence, the drain resistance of a JFET (0.1 to 1 MΩ) is much higher than that of a MOSFET (1 to 50 kΩ).
4. JFETs are operated only in the depletion mode. The depletion type MOSFET may be operated in both depletion and enhancement mode.
5. Comparing to JFET, MOSFETs are easier to fabricate.
6. MOSFET is very susceptible to overload voltage and needs special handling during installation. It gets damaged easily if it is not properly handled.
7. MOSFET has zero offset voltage. As it is a symmetrical device, the source and drain can be interchanged. These two properties are very useful in analog signal switching.
8. Special digital CMOS circuits are available which involve near-zero power dissipation and very low voltage and current requirements. This makes them most suitable for portable systems.

MOSFETs are widely used in digital VLSI circuits than JFETs because of their advantages.

3.14 COMPARISON OF *N*- WITH *P*-CHANNEL MOSFETs

1. The *P*-channel enhancement MOSFET is very popular because it is much easier and cheaper to produce than the *N*-channel device.
2. The hole mobility is nearly 2.5 times lower than the electron mobility. Thus, a *P*-channel MOSFET occupies a larger area than an *N*-channel MOSFET having the same I_D rating.
3. The drain resistance of *P*-channel MOSFET is three times higher than that for an identical *N*-channel MOSFET.
4. The *N*-channel MOSFET has the higher packing density which makes it faster in switching applications due to the smaller junction areas and lower inherent capacitances.

5. The *N*-channel MOSFET is smaller for the same complexity than *P*-channel device.
6. Due to the positively charged contaminants, the *N*-channel MOSFET may turn ON prematurely, whereas the *P*-channel device will not be affected.

3.15 COMPARISON OF *N*- WITH *P*-CHANNEL FETs [AU May 2012, 4 marks]

1. In an *N*-channel JFET the current carriers are electrons, whereas the current carriers are holes in a *P*-channel JFET.
2. Mobility of electrons is large in *N*-channel JFET; mobility of holes is poor in *P*-channel JFET.
3. The input noise in less in *N*-channel JFET than that of *P*-channel JFET.
4. The transconductance is larger in *N*-channel JFET than that of *P*-channel JFET.

3.16 USE OF JFET AS VOLTAGE-VARIABLE RESISTOR

FET is operated in the constant-current portion of its output characteristics for the linear applications. In the region before pinch-off, where V_{DS} is small, the drain to source resistance r_d can be controlled by the bias voltage V_{GS}. The FET is useful as a *Voltage Variable Resistor* (*VVR*) or *Voltage Dependent Resistor* (*VDR*).

In JFET, the drain-to-source conductance, $g_d = \dfrac{I_D}{V_{DS}}$ for small values of V_{DS}, which may also be expressed as,

$$g_d = g_{do}\left[1 - \left(\frac{V_{GS}}{V_P}\right)^{\frac{1}{2}}\right]$$

where g_{do} is the value of drain conductance when the bias voltage V_{GS} is zero.

The variation of the r_d with V_{GS} can be closely approximated by the empirical expression,

$$r_d = \frac{r_0}{1 - KV_{GS}}$$

where r_0 = drain resistance at zero gate bias, and K = a constant, dependent upon FET type.

Thus, small signal FET drain resistance r_d varies with applied gate voltage V_{GS} and FET acts like a variable passive resistor.

FET finds wide applications where VVR property is useful. For example, the VVR can be used in Automatic Gain Control (AGC) circuit of a multistage amplifier.

3.17 ADVANTAGES OF BJT OVER MOSFET

BJTs have some advantages over MOSFETs for at least two digital applications. Firstly, in high speed switching, they do not have the *larger* capacitance factor imposed by the gate, which when multiplied by the resistance of the channel gives the intrinsic time constant of the process. The intrinsic time constant places a limit on the speed performance at which a MOSFET can operate, since the higher frequency signals may be filtered out. Widening the channel reduces the resistance of the channel. However, it increases the capacitance by the same amount. On the other hand, reducing the width of the channel increases the resistance, but reduces the capacitance by the same amount. For example, if $R \times C = T_{c1}$,

the value of $0.5\ R \times 2C = T_{c1}$ and $2R \times 0.5C = T_{c1}$. There is no other way to minimize the intrinsic time constant for a certain process. The various processes using different channel lengths, channel heights, gate thicknesses and materials will also have varying intrinsic time constants. This problem is mostly avoided while using a BJT, since it does not use a gate.

The second application where the BJTs prove more advantageous over the MOSFETs stems from the first reasoning. While driving several fan-out gates, the resistance of the MOSFET is in series with the gate capacitances of the other FETs, creating a secondary time constant. Delay circuits use this fact to create a fixed signal delay by using a small CMOS device to send a signal to many other, larger CMOS fan-out devices. The secondary time constant can be minimized by increasing the driving FET's channel width to decrease its resistance and decreasing the channel widths of the FETs being driven, for reducing the fan-out or driven capacitance. However, its drawback the is that, this increases the capacitance of the driving FET and increases the resistance of the FETs being driven. Usually this drawback is a negligible factor, while comparing with the timing problem. The BJTs are better able to drive the other gates, since they can provide more output current than the MOSFETs. This allows the FETs being driven to get charged faster. Therefore, many integrated circuits employ the MOSFET driven inputs and BiCMOS outputs.

3.18 CURRENT EQUATION FOR JFET

[AU May 2016, 16 marks]

The derivation of the ideal drain current equation for the JFET in depletion mode is more complex and a good approximation is made to the *I-V* characteristics for deriving the drain current when the JFET is biased in the saturation region. Unlike in BJT, there is no linear relationship between the input and output quantities for JFET. The relationship between I_D and V_{GS} is defined by Shockley's equation as given by

$$I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_p}\right)^2 \tag{3.8}$$

where I_{DSS} is the saturation current when $V_{GS} = 0$. The approximate drain current equation is given by Eq. (3.8) and the ideal drain current equation or the current-voltage relationship of JFET is derived by starting with Ohm's law.

An *N*-channel JFET is considered with the geometry shown in Fig. 3.16. The differential resistance of the channel at a point x in the channel, by considering half of the two-sided symmetrical geometry, is given by

$$dR = \frac{\rho dx}{A(x)} \tag{3.9}$$

where ρ is the resistivity and A is the cross-sectional area. By neglecting the minority carrier holes in the *N*-channel, the channel resistivity is given by

$$\rho = \frac{1}{q\mu_n N_D} \tag{3.10}$$

where N_D is the doping concentration.

The cross-sectional area is given by

$$A(x) = (a - h(x))W \tag{3.11}$$

where W is the width of the channel and a is the channel thickness.

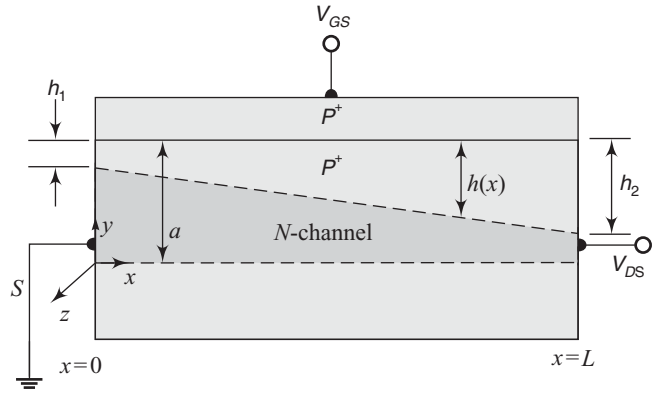


Fig. 3.16 *N-channel JFET geometry*

Substituting Eq. (3.10) and Eq. (3.11) in Eq. (3.9), the differential resistance of the channel is given by

$$dR = \frac{dx}{q\mu_n N_D (a - h(x))W} \tag{3.12}$$

The differential voltage across a differential length dx can be written as

$$dV(x) = I_{D1}\, dR \tag{3.13}$$

where the drain current I_{D1} is constant through the channel.

Substituting Eq. (3.12) in Eq. (3.13), we get

$$dV(x) = \frac{I_{D1} dx}{q\mu_n N_D W(a - h(x))} \tag{3.14}$$

Therefore,

$$I_{D1}\, dx = q\mu_n N_D W(a - h(x))\, dV(x) \tag{3.15}$$

The depletion width $h(x)$ is given by

$$h(x) = \left\{\frac{2\varepsilon(V(x) + V_{bi} - V_{GS})}{qN_D}\right\}^{\frac{1}{2}} \tag{3.16}$$

where $V(x)$ is the potential in the channel due to the drain-source voltage, V_{bi} is the built-in potential barrier and ε is the permittivity of the substrate. Solving for $V(x)$ in Eq. (3.16) and differentiating $V(x)$ w.r.t $h(x)$, we get

$$dV(x) = \frac{qN_D h(x)\, dh(x)}{\varepsilon} \tag{3.17}$$

Substituting Eq. (3.17) in Eq. (3.15), the drain current through dx can be written as

$$I_{D1} dx = \frac{\mu_n (qN_D)^2 W}{\varepsilon} [ah(x)\, dh(x) - h(x)^2\, dh(x)] \tag{3.18}$$

The drain current I_{D1} can then be obtained by integrating Eq. (3.18) along the channel length. Assuming the current and mobility are constant through the channel, the drain current is given by

$$I_{D1} = \frac{\mu_n (qN_D)^2 W}{\varepsilon L}\left\{\int_{h_1}^{h_2} ah\,dh - \int_{h_1}^{h_2} h^2 dh\right\} \tag{3.19}$$

$$I_{D1} = \frac{\mu_n \left(qN_D\right)^2 W}{\varepsilon L}\left\{\frac{a}{2}\left(h_2^2 - h_1^2\right) - \frac{1}{3}\left(h_2^3 - h_1^3\right)\right\} \tag{3.20}$$

For JFET, h_2 is the depletion width at the drain terminal which is given by

$$h_2 = \left[\frac{2\varepsilon\left(V_{DS} + V_{bi} - V_{GS}\right)}{qN_D}\right]^{\frac{1}{2}} \tag{3.21}$$

and h_1 is the depletion width at the source terminal which is given by

$$h_1 = \left[\frac{2\varepsilon\left(V_{bi} - V_{GS}\right)}{qN_D}\right]^{\frac{1}{2}} \tag{3.22}$$

and the internal pinch-off voltage is given by

$$V_{po} = \frac{qa^2 N_D}{2\varepsilon} \tag{3.23}$$

Using Eq. (3.21) to Eq. (3.23), Eq. (3.20) can be written as

$$I_{D1} = \frac{\mu_n \left(qN_D\right)^2 Wa^3}{2\varepsilon L}\left\{\frac{V_{DS}}{V_{po}} - \frac{2}{3}\left[\frac{V_{DS} + V_{bi} - V_{GS}}{V_{po}}\right]^{\frac{3}{2}} + \frac{2}{3}\left[\frac{V_{bi} - V_{GS}}{V_{po}}\right]^{\frac{3}{2}}\right\} \tag{3.24}$$

The pinch-off current I_{p1} for JFET is given by

$$I_{p1} \equiv \frac{\mu_n \left(qN_D\right)^2 Wa^3}{6\varepsilon L}$$

Using the above pinch-off current equation, the drain current in Eq. (3.24) becomes

$$I_{D1} = I_{p1}\left\{3\left(\frac{V_{DS}}{V_{po}}\right) - 2\left(\frac{V_{DS} + V_{bi} - V_{GS}}{V_{po}}\right)^{\frac{3}{2}} + 2\left(\frac{V_{bi} - V_{GS}}{V_{po}}\right)^{\frac{3}{2}}\right\} \tag{3.25}$$

The above drain current equation is valid in the non-saturation region for $0 \le |V_{GS}| \le |V_P|$ and $0 \le |V_{DS}| \le |V_{DS(\text{sat})}|$. The pinch-off current I_{p1} will be the maximum drain current in the JFET if the zero-biased depletion region was ignored or if V_{GS} and V_{bi} were both zero. Equation (3.25) is the ideal current-voltage relationship for the one-sided *N*-channel JFET in the non-saturation region.

For the two-sided symmetrical JFET shown in Fig. 3.17, the total drain current will be $I_{D2} = 2I_{D1}$. Equation (3.24) can also be written as

$$I_{D1} = G_{01}\left\{V_{DS} - \frac{2}{3}\sqrt{\frac{1}{V_{po}}}\left[\left(V_{DS} + V_{bi} - V_{GS}\right)^{\frac{3}{2}} - \left(V_{bi} - V_{GS}\right)^{\frac{3}{2}}\right]\right\} \tag{3.26}$$

where $G_{01} = \dfrac{\mu_n (qN_D)^2 Wa^3}{2\varepsilon L V_{po}} = \dfrac{q\mu_n N_D Wa}{L} = \dfrac{3I_{P1}}{V_{po}}$

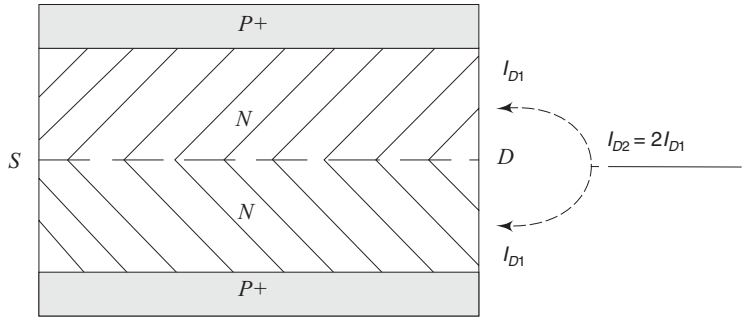


Fig. 3.17 *Drain current of a two-sided symmetrical JFET*

The channel conductance is defined as

$$g_d = \left.\frac{\partial I_{D1}}{\partial V_{DS}}\right|_{V_{DS}=0}$$

Taking the derivative of Eq. (3.26) with respect to V_{DS}, we get

$$g_d = \left.\frac{\partial I_{D1}}{\partial V_{DS}}\right|_{V_{DS}=0} = G_{01}\left\{1 - \left(\frac{V_{bi} - V_{GS}}{V_{po}}\right)^{\frac{1}{2}}\right\} \tag{3.27}$$

From the above equation, G_{01} is the conductance of the channel if both V_{bi} and V_{GS} are zero. This condition will exist if no space charge regions existed in the channel. Equation (3.27) also shows that the channel conductance is modulated or controlled by the gate voltage. This channel conductance modulation is the basis of the field effect phenomenon.

Figure 3.18 shows the ideal current-voltage characteristics or drain characteristics of an *N*-channel JFET. From Fig. 3.18, the drain gets pinched off, when $V_{DS} = V_{DS(\text{sat})} = V_{po} - (V_{bi} - V_{GS})$. By setting $V_{DS} = V_{DS(\text{sat})}$ in Eq. (3.25), the drain current equation in the saturation region is given by

$$I_{D1} = I_{D1(\text{sat})} = I_{p1}\left\{1 - 3\left(\frac{V_{bi} - V_{GS}}{V_{po}}\right)\left[1 - \frac{2}{3}\sqrt{\frac{V_{bi} - V_{GS}}{V_{po}}}\right]\right\} \tag{3.28}$$

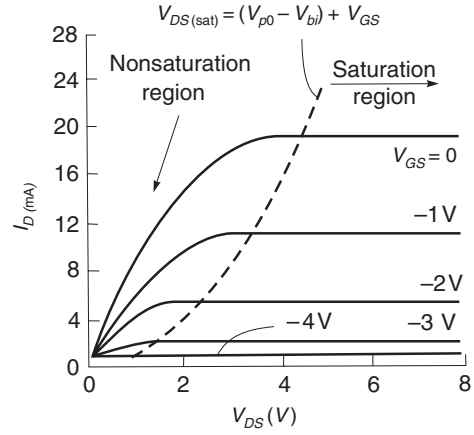


Fig. 3.18 *Ideal current-voltage characteristics of N-channel JFET*

The ideal drain saturation current is now independent of the drain-to-source voltage. Equations (3.25) and (3.28) show the exact drain current equation in the non-saturation region and in the saturation region. Both these equations are rather complex to use in any calculation. Hence, a good approximation is made in the saturation region and the drain current equation is written as

$$I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_P}\right)^2 \tag{3.29}$$

where V_p is the pinch-off voltage, V_{GS} is the gate-source voltage and I_{DSS} is the maximum drain saturation current when $V_{GS} = 0$.

Figure 3.19 shows the transfer characteristics for an *N*-channel depletion mode JFET in which both V_{GS} and V_p are negative whereas for *P*-channel depletion mode JFET, both are positive. The comparison between the ideal curve and approximation curve for the drain current of JFET is also shown in Fig. 3.19. The approximate drain current equation mentioned above in saturation region applies to an *N*-channel depletion mode MOSFET also.

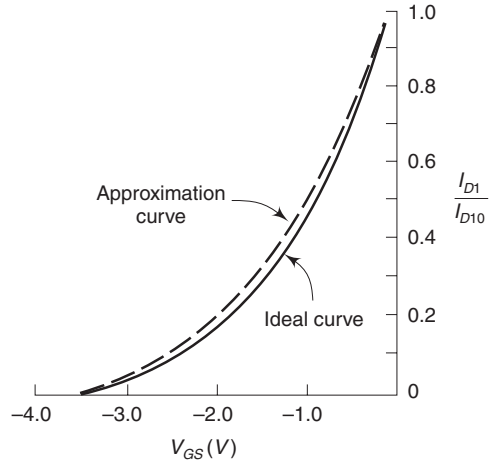


Fig. 3.19 *Transfer characteristics of N-channel JFET in saturation region*

The ideal current-voltage characteristics of the enhancement-mode JFET are the same as the depletion mode and the only difference is the relative values of the internal pinch-off voltage. We know that the drain current in the saturation region is given by

$$I_{D1} = I_{D1(\text{sat})} = I_{p1}\left\{1 - 3\left(\frac{V_{bi} - V_{GS}}{V_{po}}\right)\left[1 - \frac{2}{3}\sqrt{\frac{V_{bi} - V_{GS}}{V_{po}}}\right]\right\}$$

The threshold voltage for the *N*-channel device is defined as $V_T = V_{bi} - V_{po}$. For an *N*-channel depletion-mode JFET, $V_T < 0$, and for the enhancement-mode device, $V_T > 0$. Hence, for the enhancement-mode device, $V_{bi} > V_{po}$. The design of enhancement-mode JFET implies the use of narrow channel thicknesses and low channel doping concentrations to achieve this condition. Therefore, the built-in potential barrier can be written as,

$$V_{bi} = V_T + V_{po}$$

Substituting the above expression in Eq. (3.28), we get

$$I_{D1(\text{sat})} = I_{p1}\left\{1 - 3\left[1 - \left(\frac{V_{GS} - V_T}{V_{po}}\right)\right] + 2\left[1 - \left(\frac{V_{GS} - V_T}{V_{po}}\right)\right]^{\frac{3}{2}}\right\} \tag{3.30}$$

The above equation is valid for $V_{GS} \geq V_T$. So, when the transistor switches ON, we have

$$(V_{GS} - V_T) << V_{po}.$$

Expanding Eq. (3.30) into a Taylor series, we get

$$I_{D1(\text{sat})} \approx I_{p1}\left[\frac{3}{4}\left(\frac{V_{GS} - V_T}{V_{po}}\right)\right]^2 \tag{3.31}$$

Substituting the expressions for I_{p1} and V_{po}, Eq. (3.31) can be written as

$$I_{D1(\text{sat})} = \frac{\mu_n \varepsilon W}{2aL} \quad \text{for } V_{GS} \geq V_T$$

The above equation can now be written as

$$I_{D1(\text{sat})} = K_N(V_{GS} - V_T)^2 \tag{3.32}$$

where $K_N = \dfrac{\mu_n \varepsilon W}{2aL}$ and this factor K_N is called a conduction parameter. Equation (3.32) is the drain current equation for *N*-channel enhancement-mode JFET and MOSFET.

3.19 EQUIVALENT CIRCUIT MODEL OF JFET AND MOSFET AND THEIR PARAMETERS

For the analysis of any transistor circuits like JFET and MOSFET, a mathematical model or equivalent circuit of the transistor is needed. One of the most useful models is the small-signal equivalent circuit, which applies to transistors used in linear amplifier circuits. This equivalent circuit also introduces frequency effects in the transistor through the equivalent capacitor and resistor circuits. The physical

factor affecting the frequency limitations is the transistor cut-off frequency, which is a figure of merit for both JFET and MOSFET.

3.19.1 JFET Small-Signal Equivalent Circuit

The cross section of an *N*-channel JFET is shown in Fig. 3.20, which includes the series resistances of source and drain terminal. The substrate used is a P^+ type Si substrate. The corresponding small-signal equivalent circuit for the JFET is shown in Fig. 3.21. The internal gate-source voltage is denoted by $V_{g's'}$ that controls the drain current. The gate-source diffusion resistance and junction capacitance are represented by parameters r_{gs} and C_{gs} respectively. The gate-source junction is reverse biased for depletion-mode devices and forward biased for enhancement-mode devices, so that normally r_{gs} is large. The parameters r_{gd} and C_{gd} are the gate-drain resistance and capacitance respectively. The resistance r_{ds} is the finite drain source resistance, which is a function of the channel-length modulation effect. The capacitance C_{ds} mainly a drain-source parasitic capacitance and C_s is the drain-substrate capacitance.

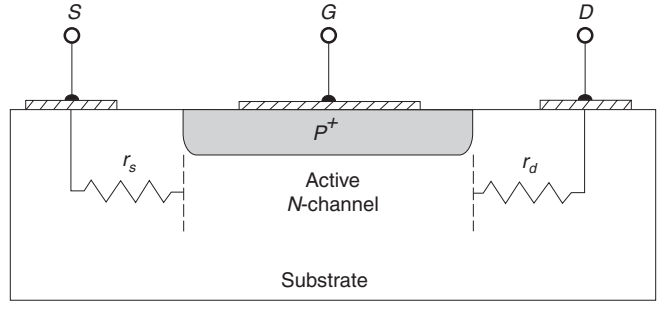


Fig. 3.20 *Cross section of JFET with source and drain series resistance*

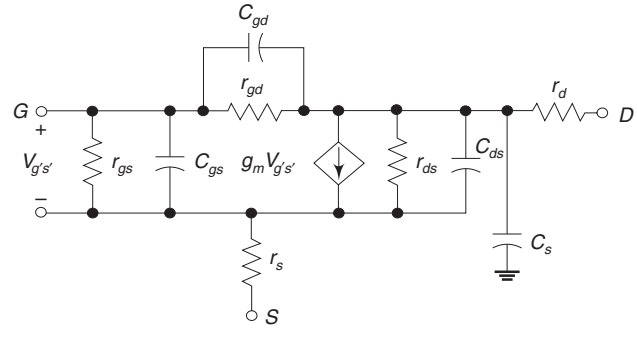


Fig. 3.21 *Small-signal equivalent circuit of JFET*

The ideal small-signal equivalent circuit at low frequency is shown in Fig. 3.22(a). All diffusion resistances are infinite, the series resistances are zero and the capacitances become open circuits. The small-signal drain current is given by

$$I_{ds} = g_m V_{gs}$$

which is a function of transconductance and input signal voltage. Figure 3.22(b) shows the equivalent circuit of JFET which includes the series source resistance. From Fig. 3.22(b), we know that,

$$I_{ds} = g_m V_{g's'}$$

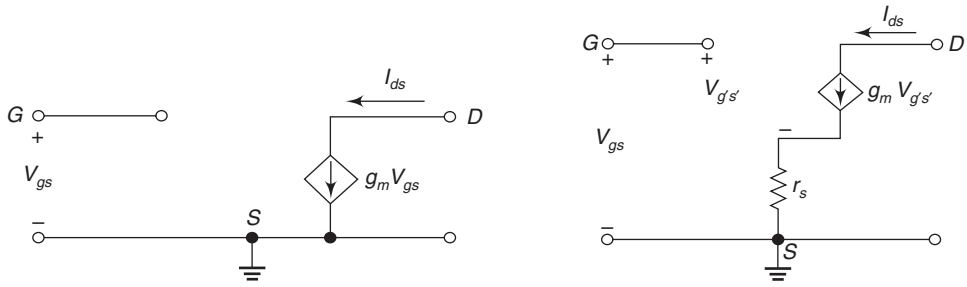


Fig. 3.22 *Small-signal equivalent circuit (a) without r_s, (b) with r_s*

The relation between V_{gs} and $V_{g's'}$ can be obtained from

$$V_{gs} = V_{g's'} + (g_m V_{g's'}) r_s = (1 + g_m r_s) V_{g's'}$$

Now, the drain current can then be written as

$$I_{ds} = \left(\frac{g_m}{1 + g_m r_s} \right) V_{gs} = g'_m V_{gs}$$

where g_m is the transconductance without r_s and g'_m is the transconductance with r_s. The effect of the source resistance is to reduce the effective transconductance or transistor gain. As g_m is a function of V_{gs}, g'_m will also be a function of V_{gs} as shown in Fig. 3.23.

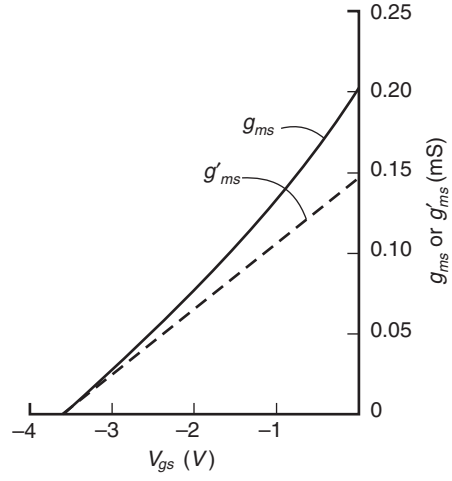


Fig. 3.23 *JFET transconductance versus V_{gs} (without r_s and with r_s)*

There are two frequency limitation factors in a JFET. The first is the channel transit time τ_t. For a channel length L of 1 μm and assuming the carriers are travelling at their saturation velocity cm/s then the transit time is of the order of

$$\tau_t = \frac{L}{v_s} = \frac{1 \times 10^{-4}}{1 \times 10^{7}} = 10\ ps$$

The channel transit time is normally not the limiting factor except in very high frequency devices.

The second frequency limitation factor is the capacitance charging time. Figure 3.24 is a simplified equivalent circuit at high frequency that includes the primary capacitances and ignores the diffusion resistances. The output current will be the short-circuit current. As the frequency of the input signal voltage V_{gs} increases, the impedance of C_{gd} and C_{gs} decreases, so the current through C_{gd} will increase. For a constant $g_m V_{gs}$, the I_{ds} current will then decrease. The output current then becomes a function of frequency.

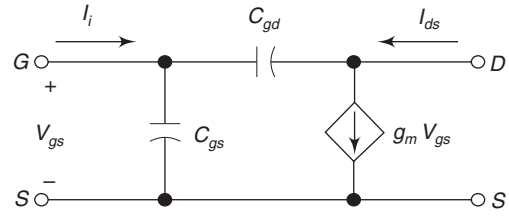


Fig. 3.24 *Small-signal equivalent circuit at high frequency including C_{gd} and C_{gs}*

3.19.2 MOSFET Small-Signal Equivalent Circuit

The small-signal equivalent circuit of the MOSFET is constructed from the basic MOSEFT geometry. A model based on the inherent capacitances and resistances within the transistor structure, along with elements that represent the basic device equations, is shown in Fig. 3.25. Here the source and substrate are both tied to ground. Two of the capacitances connected to the gate are inherent in the device. These capacitances are C_{gs} and C_{gd}, which represent the interaction between the gate and the channel charges near the source and drain terminals, respectively. The remaining two gate capacitances, C_{gsp} and C_{gdp}, are parasitic or overlap capacitances.

In real devices, the gate oxide will overlap the source and drain contacts because of tolerance or fabrication factors. The drain overlap capacitance, C_{gdp}, lowers the frequency response of the device. The parameter C_{ds} is the drain-to-substrate *PN*-junction capacitance. The series resistances, r_s and r_d, are the resistances associated with the source and drain terminals. The small-signal channel current is controlled by the internal gate-source voltage through the transconductance.

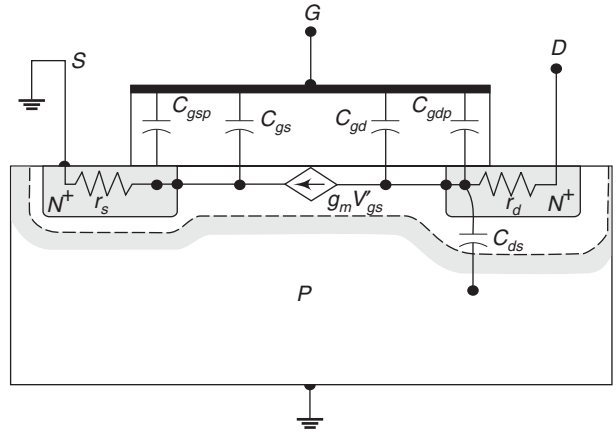


Fig. 3.25 *N-channel MOSFET structure with inherent resistances and capacitances*

The small-signal equivalent circuit for the *N*-channel common source MOSFET is shown in Fig. 3.26. The voltage V'_{gs} is the internal gate-source voltage that controls the channel current. The parameters C_{gsT} and C_{gdT} are the total gate-source and total gate-drain capacitances. The drain-source resistance r_{ds}, shown in Fig. 3.26, is not shown in Fig. 3.25 and this resistance is associated with the slope of I_D versus V_{DS}. In the ideal MOSFET biased in the saturation region, I_D is independent of V_{DS} so that r_{ds} would be infinite. In short-channel-length devices, in particular, r_{ds} is finite because of channel-length modulation.

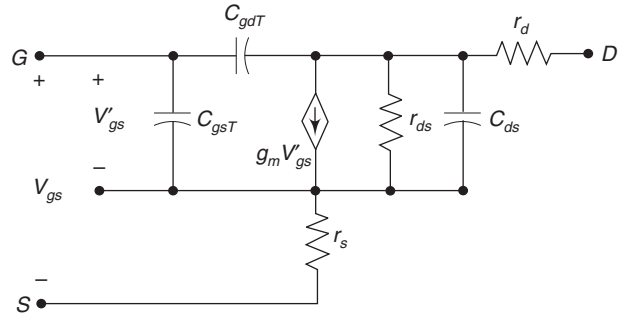


Fig. 3.26 *Small-signal equivalent circuit of a common source N-channel MOSFET*

A simplified small-signal equivalent circuit valid at low frequency for the *N*-channel common-source MOSFET is shown in Fig. 3.27. Neglecting the series resistances, r_s and r_d, the drain current is mainly a function of the gate-source voltage through the transconductance. The input gate impedance is infinite in this simplified model.

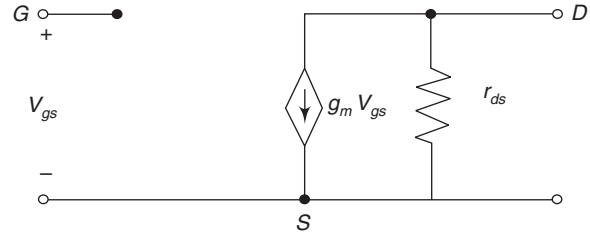


Fig. 3.27 *Simplified small-signal equivalent circuit of MOSFET at low frequency*

The source resistance, r_s, can have a significant effect on the MOSFET characteristics as in JFET. Figure 3.28 shows a simplified low-frequency equivalent circuit including r_s but neglecting r_{ds}. From Fig. 3.28, the drain current is given by

$$I_d = g_m V'_{gs}$$

The relation between V_{gs} and V'_{gs} can be obtained from

$$V_{gs} = V'_{gs} + (g_m V'_{gs}) r_s = (1 + g_m r_s) V'_{gs}$$

The drain-current equation can now be written as

$$I_d = \left(\frac{g_m}{1 + g_m r_s} \right) V_{gs} = g'_m V_{gs}$$

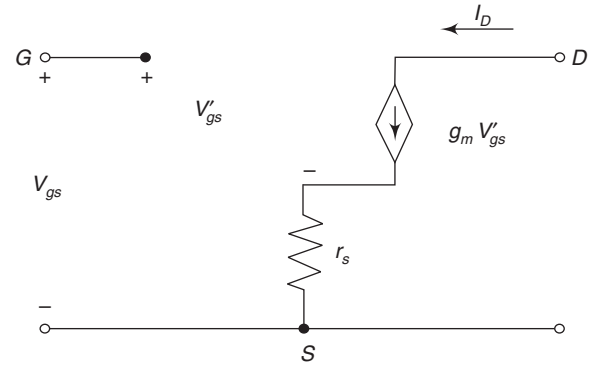


Fig. 3.28 *Small-signal equivalent circuit of MOSFET including r_s*

The presence of source resistance in Fig. 3.28 reduces the effective transconductance or transistor gain. The equivalent circuit of the *P*-channel MOSFET is exactly the same as that of the *N*-channel except that all voltage polarities and current directions are reversed. All the capacitances and resistances in the *N*-channel equivalent circuit model apply to the *P*-channel model also.

The two basic frequency limitation factors in the MOSFET are the channel transit time and the gate or capacitance charging time, just like JFET. Assume that the carriers are travelling at their saturation drift velocity v_s, and then the transit time is $\tau_t = L/v_s$, where L is the channel length. If $v_s = 10^7$ cm/s and L = 1 μm then $\tau_t = 10$ *ps*, which translates into a maximum frequency of 100 GHz. This frequency is much larger than the typical maximum frequency response of a MOSFET. For high-frequency analysis of MOSFET, the parameters r_s, r_d, r_{ds}, and C_{ds} are neglected and the resulting small-signal equivalent circuit of common source *N*-channel MOSFET is shown in Fig. 3.29 where R_L is the load resistance.

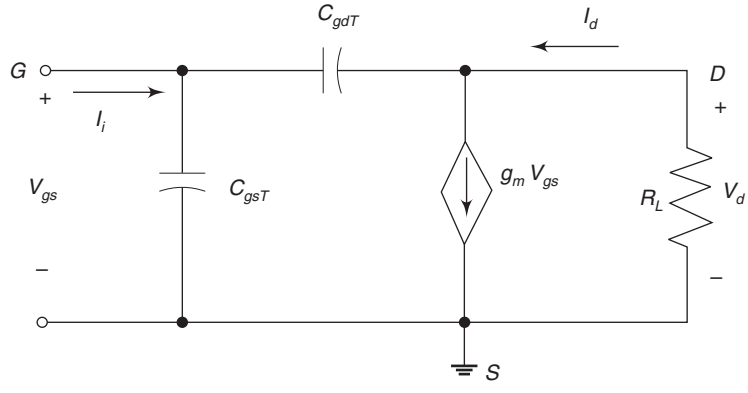


Fig. 3.29 *Small-signal equivalent of MOSFET at high frequency*

Two Mark Questions and Answers

1. **Give the current-voltage relationship of the D-MOSFET and E-MOSFET.** **(April/May 2017)**
   The current-voltage relationship of D-MOSFET is given by

$$I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_P}\right)^2$$

The current-voltage relationship of E-MOSFET is given by

$$I_D = K_N(V_{GS} - V_T)^2$$

2. Give some applications of JFET. **(May/June 2016)**

Refer to Section 3.9 Applications of JFET

3. What is JFET? Give its different modes of operation. **(Nov/Dec 2016)**

The JFET is an electronic device in which the flow of current through the conducting region is controlled by an electric field. Hence, the name Junction Field Effect Transistor (JFET). As current conduction is only by majority carriers, JFET is said to be a unipolar device. JFET works under depletion mode and enhancement mode.

4. What is the major difference between a bipolar and unipolar device? **(Nov/Dec 2014)**

Bipolar devices can have both minority and majority carriers flowing, whereas unipolar devices only have majority carriers flowing. The fact that bipolar devices have two types of charge carriers flowing simultaneously results in the name *Bi-polar*. The flow of minority carriers is responsible for collector conduction modulation and transistor storage time.

5. Draw the structure and symbol for an *N*-channel JFET. **(May/June 2014)**

Refer to Fig. 3.1 for structure and Fig. 3.4 (a) for symbol of an *N*-channel JFET.

6. In which region JFET acts as a resistor and why? **(May/June 2014)**

When a FET acts as a voltage variable resistor? **(Nov/Dec 2012)**

Ohmic region: When gate-source voltage, $V_{GS} = 0$, the depletion layer of the channel is very small and the JFET acts like a voltage controlled resistor since the current increases with increase in voltage.

7. Define drain to source resistance of JFET. **(Nov/Dec 2013)**

The drain resistance is given by

$$r_d = \frac{\Delta V_{DS}}{\Delta I_D},\ V_{GS} \text{ held constant}$$

Here, the drain resistance at $V_{GS} = 0$, i.e., when the depletion regions of the channel are absent, is called *drain-source ON resistance* and it is represented as R_{DS} or $R_{DS(ON)}$.

8. Compare BJT and FET. **(Nov/Dec 2012 and May/June 2014)**

Refer to Section 3.8 Comparison of JFET and BJT.

9. Write the equation for drain current of JFET. **(May/June 2013)**

The drain current of JFET is given by

$$I_{DS} = I_{DSS}\left(1 - \frac{V_{GS}}{V_P}\right)^2$$

10. What is channel length modulation? **(April/May 2015)**

The channel length modulation is a shortening of length of the inverted channel region with increase in drain bias for large drain biases. The channel length modulation results in an increase in current with drain bias and reduction in output resistance.

11. Define pinch-off voltage. **(Nov/Dec 2014)**

In JFET, the electrons flow through a semiconducting channel between source and drain terminals. By applying a reverse bias voltage to a gate terminal, the channel is made free of charge carriers, i.e., pinched off, so that the electric current is impeded or switched off completely.

12. Compare *P*-channel and *N*-channel JFETs. **(June 2010 and Nov/Dec 2011)**

Refer to Section 3.15 Comparison of *N* - with *P*-channel FETs.

13. Sketch the graph symbol for *N*-channel and *P*-channel MOSFETs. **(June 2011 and Nov/Dec 2014)**

Refer to Fig. 3.9 (b) and (c) for Enhancement MOSFET and Fig. 3.11 (b) and (c) for Depletion MOSFET.

14. Mention the advantages of MOSFET over JFET. **(Nov/Dec 2010)**

(i) JFET can only be operated in the *depletion mode* whereas MOSFET can be operated in either *depletion* or in *enhancement mode*. In a JFET, if the gate is forward biased, excess-carrier injunction occurs and the gate current is substantial. Thus, channel conductance is enhanced to some degree due to excess carriers but the device is never operated with gate forward biased because gate current is undesirable.

(ii) MOSFET has input impedance much higher than that of JFET. This is due to negligibly small leakage current.

(iii) JFET has characteristic curves more flat than those of MOSFET indicating a higher drain resistance.

(iv) When JFET is operated with a reverse bias on the junction, the gate current I_G is larger than it would be in a comparable MOSFET. The current caused by minority carrier extraction across a reverse-biased junction is greater, per unit area, than the leakage current that is supported by the oxide layer in a MOSFET. Thus, MOSFET devices are more useful in electrometer applications than are the JFETs.

15. Differentiate between enhancement-type and depletion-type MOSFETs. **(May/June 2012)**

(i) Gate formation of *N*-channel E-MOSFET is similar to the *N*-channel D-MOSFET.

(ii) E-MOSFET can be worked only in enhancement mode and hence, this MOSFET is called enhancement MOSFET or E-MOSFET.

(iii) E-MOSFET has no conducting channel between the source and gate terminals whereas D-MOSFET has conducting channel between the source and gate terminals.

Review Questions

1. Why a Field Effect Transistor is called so?
2. Explain the construction of *N* channel JFET.
3. With the help of neat sketches and characteristic curves, explain the operation of the junction FET.
4. How does the FET behave for small and large values of $|V_{DS}|$?
5. Define the pinch-off voltage V_P. Sketch the depletion region before and after pinch-off.
6. Explain the four distinct regions of the output characteristics of a JFET.
7. Define and explain the parameters transconductance g_m, drain resistance r_d and amplification factor μ of a JFET. Establish the relation between them.
8. Assuming that the saturation drain current I_{DS} is given by the parabolic relation,

$$I_{DS} = I_{DSS}\left(1 - \frac{V_{GS}}{V_P}\right)^2$$

Prove that the transconductance g_m is given by

$$g_m = g_{mo}\left(1 - \frac{V_{GS}}{V_P}\right) = -\frac{2}{V_P}\sqrt{I_{DS} I_{DSS}}$$

where g_{mo} is the value of g_m for $V_{GS} = 0$.

9. Explain how the transconductance of a JFET varies with drain current and gate voltage.
10. What are the relative merits of an *N*-channel and a P-channel FET?
11. Compare JFET with BJT.
12. Explain why BJTs are called bipolar devices while FETs are called unipolar devices.
13. Explain why a low power FET is called as a square law device.
14. Briefly describe some applications of JFET.
15. A certain JFET operates in the linear region with a constant drain voltage of 1 V. When the gate voltage is 2 V, a drain current of 10 mA flows, but when the gate voltage is changed to 1 V, the drain current becomes 22.8 mA. Find (a) the pinch-off voltage of the device and (b) the channel resistance for zero gate voltage.

(*Ans.* 3 V, 18.6 Ω).

16. Show that if a JFET is operated at sufficiently low drain voltage, it behaves as a resistance *R* given approximately by

$$R = \frac{R_o}{\left[1 - \left(\frac{V_{GS}}{V_P}\right)^{1/2}\right]}$$

where R_o is the channel resistance for zero gate voltage.

17. What is a MOSFET? How many types of MOSFETs are there?
18. With the help of suitable diagrams explain the working of different types of MOSFET.
19. How does the constructional feature of a MOSFET differ from that of a JFET?
20. Explain qualitatively the shapes of the I_D versus V_{DS} and I_D versus V_{GS} characteristics for the three types of FETs.

21. What is channel length modulation in MOSFET? Obtain the output resistance at the *Q*-point.
22. Define threshold voltage.
23. Explain the threshold voltage concept in MOSFET.
24. Derive the expression for threshold voltage in MOSFET.
25. Consider an N^+ polysilicon gate and a *P*-type silicon substrate doped to $N_A = 2 \times 10^{22}$ m^{-3}. Assume $n_i = 1.5 \times 10^{16}$ m^{-3}, $\Phi_{ms} = -0.9$ V, $\varepsilon_s = 11.7$ for Si, $\varepsilon_{ox} = 3.9$ for SiO_2 and $Q'_{ss} = 10^{15}$ m^{-2}. Determine the oxide thickness such that $V_{TN} = 0.8$ V.
26. Calculate the threshold voltage of a MOSFET using the aluminum gate. Consider a *P*-type silicon substrate at $T = 300$ K doped to $N_A = 3 \times 10^{20}$ m^{-3}. Here, $Q'_{ss} = 5 \times 10^{14}$ m^{-2} and $t_{ox} = 5 \times 10^{-8}$ m . Assume the oxide is silicon dioxide, $n_i = 1.5 \times 10^{16}$ m^{-3}, $\varepsilon_s = 11.7$ and $\Phi_{ms} = -1.2$ V.
27. Briefly explain the temperature effects in MOSFET.
28. Give the current–voltage relationships of the *N*-MOSFET and *P*-MOSFET.
29. Why are *N*-channel MOSFETs preferred over *P*-channel MOSFETs?
30. Explain how an FET is used as a voltage variable resistor.
31. Derive the exact drain current equation for depletion-mode JFET and compare it with approximate equation.
32. Using depletion-mode drain current equation, derive the expression for JFET drain current in enhancement mode.
33. Describe briefly about the small signal equivalent circuit of JFET and its associated parameters with necessary diagrams.
34. With neat diagrams, explain the small-signal equivalent circuit of MOSFET and its associated parameters.
35. Why does the MOSFET have high input impedance?

SPECIAL SEMICONDUCTOR DEVICES

4

4.1 INTRODUCTION

In addition to the *PN*-junction diode, other types of diodes are also manufactured for specific applications. These special semiconductor diodes are two-terminal devices with their doping levels carefully selected to give the desired characteristics. This chapter deals with Metal-Semiconductor Junction, Schottky Barrier Diode, Gallium Arsenide Devices, Dual Gate MOSFET, FinFET, PIN-FET, CNTFET Zener Diode, Varactor Diode, Tunnel Diode, Laser Diode, and LDR.

4.2 METAL-SEMICONDUCTOR JUNCTIONS

[AU Dec 2015, May 2015, 12 marks]

Metal-semiconductor junctions are very common in all semiconductor devices and have very high importance. Depending upon the doping concentration, materials, and the characteristics of the interface, the metal-semiconductor junctions can act as either an ohmic contact or as a Schottky barrier. An analysis of metal-semiconductor junction is presented in this section.

4.2.1 Structure of Metal-Semiconductor Junction

A metal-semiconductor junction, as the name indicates, consists of a metal in contact with a piece of semiconductor. The structure of a typical metal-semiconductor junction is shown in Fig. 4.1. The active junction is the interface between the metal, which acts as an anode, and the semiconductor. The other interface between the semiconductor and the metal, which acts as a cathode, is an Ohmic contact and there is no potential drop at this junction.

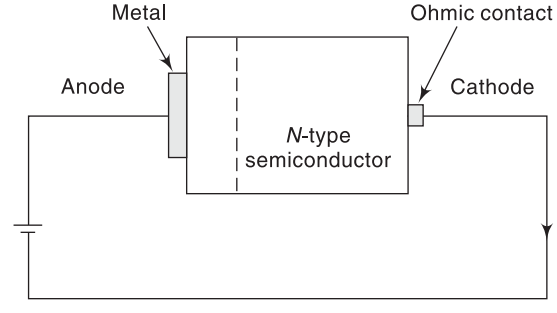


Fig. 4.1 *Structure of a metal-semiconductor junction*

4.2.2 Energy Band Diagram

The energy band diagram helps in identifying the barrier between the metal and the semiconductor. In order to understand the energy band structure at a metal–semiconductor junction, first let us consider the energy bands in metal and semiconductors separately, as shown in Fig. 4.2(a). The energy bands are

aligned at the same vacuum level. When the metal and semiconductor are brought together, the Fermi levels do align themselves at thermal equilibrium. The condition that exists just before the thermal equilibrium is reached is depicted in Fig. 4.2(b).

Let us define, Φ_B, the barrier height as the potential difference between the Fermi level of the metal and the band edge where the majority carriers exist. For an *N*-type semiconductor, the barrier height is given by the difference between the metal work function (Φ_M) and the electron affinity (χ).

$$\Phi_{BN} = \Phi_M - \chi \tag{4.1}$$

The work function, Φ_M varies depending upon surface preparation. For *P*-type semiconductor, the barrier height is given by the difference between the valence band level and the Fermi level in the metal,

$$\Phi_{BP} = \chi + \frac{E_g}{q} - \Phi_M \tag{4.2}$$

where E_g is the energy gap between the conduction and valence bands. The sum of the barrier heights on *N*-type and *P*-type substrate is expected to be equal to the energy gap, E_g, i.e., $(\Phi_{BN} + \Phi_{BP})\, q = E_g$.

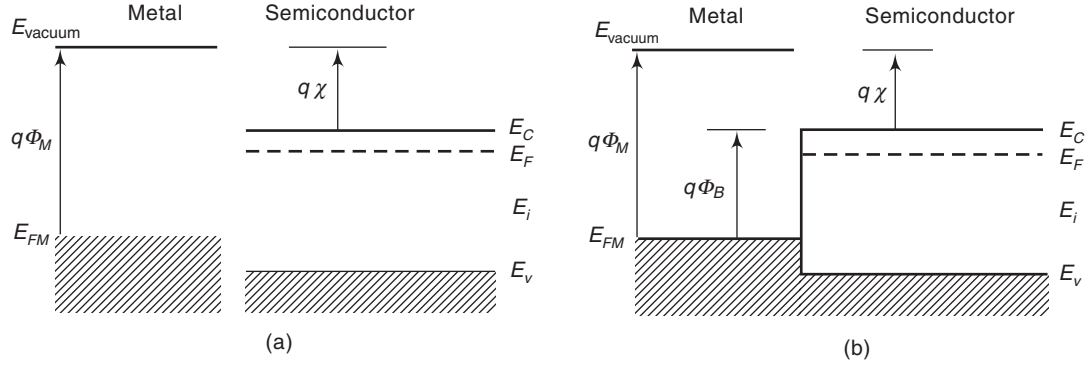


Fig. 4.2 *Energy band diagram of metal and semiconductor (a) before and (b) after contact is made*

In a metal-semiconductor junction, a barrier is formed if the Fermi level of the metal is somewhere between the valence and conduction band edges of the semiconductor, as shown in Fig. 4.2(b). Let us also define a built-in potential (Φ_1) as the difference between the Fermi level of the metal and the Fermi level of the semiconductor.

For an *N*-type semiconductor, the barrier height is given by

$$\Phi_{BN} = \Phi_M - \chi$$

$$\Phi_{IN} = \Phi_{BN} - \frac{E_C - E_F}{q} = \Phi_M - \chi - \frac{E_C - E_F}{q} \tag{4.3}$$

For a *P*-type semiconductor, the Fermi level is closer to the valence band and the built-in potential is given by

$$\Phi_{IP} = \chi + \frac{E_F - E_V}{q} - \Phi_M \tag{4.4}$$

The Fermi level in an N-type semiconductor is given by

$$E_F = E_C - kT \ln \frac{N_C}{N_D} \tag{4.5}$$

and the Fermi level in a P-type semiconductor is given by

$$E_F = E_V + kT \ln \frac{N_V}{N_A} \tag{4.6}$$

Substitution Eq. (4.5) and Eq. (4.6) in Eq. (4.3) and Eq. (4.4), respectively, would give expressions for built in potentials in terms of the barrier height and doping concentration, as follows.

$$\Phi_{IN} = \Phi_{BN} - \frac{E_C - E_F}{q} = \Phi_{BN} - \frac{kT}{q} \ln \frac{N_C}{N_D} \text{ for } N\text{-type semiconductor} \tag{4.7}$$

and

$$\Phi_{IP} = \Phi_{BP} - \frac{E_F - E_V}{q} = \Phi_{BP} - \frac{kT}{q} \ln \frac{N_V}{N_A} \text{ for } P\text{-type semiconductor} \tag{4.8}$$

4.2.3 Thermal Equilibrium

After the metal and semiconductor have been brought into contact, electrons start to flow from the semiconductor into the metal, and as a result, a depletion region of width x_d, with uncompensated donors (positive charge) is formed. Electrons continue to flow into the metal until the Fermi energy levels of metal and semiconductor align with each other. In metal, the electron current forms a negative surface charge layer. This results in an electric field and the band edges are lowered in the semiconductor as shown in Fig. 4.3.

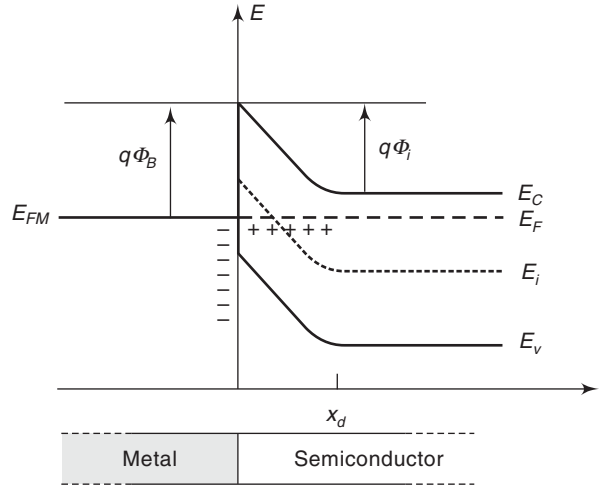


Fig. 4.3 *Energy band diagram in thermal equilibrium*

Table 4.1 gives the work function (F_M) and electron affinity (χ) of some commonly used metals and semiconductors.

Table 4.1 *Work function and electron affinity of some metals and semiconductor*

Element	*Work function, Φ_M (V)*
Aluminium, Al	4.28
Silver, Ag	4.26
Gold, Au	5.1
Chromium, Cr	4.5
Nickel, Ni	5.15
Platinum, Pt	5.65
Titanium, Ti	4.33
Tungsten, W	4.55
	Electron affinity, χ (Volt)
Silicon, Si	4.01
Germanium, Ge	4.13
Gallium Arsenide, GaAs	4.07

EXAMPLE 4.1

A metal–semiconductor junction is made of silver and silicon with $N_D = 4 \times 10^{17}$ cm^{-3}. Calculate the barrier height and the built-in potential.

Solution The work function (Φ_M) and electron affinity (χ) for silver and silicon are 4.26 V and 4.01 V, respectively (from Table 4.1). The barrier height for *N*-type material, from Eq. (4.1), is given by

$$\Phi_{BN} = \Phi_M - \chi = 4.26 - 4.01 = 0.25 \text{ V}$$

The built-in potential is given by Eq. (4.7), i.e.,

$$\Phi_{IN} = \Phi_{BN} - \frac{kT}{q} \ln \frac{N_C}{N_D} = 0.25 - \frac{1.38 \times 10^{-23} \times 300}{1.602 \times 10^{-19}} \ln \frac{2.8 \times 10^{25}}{4 \times 10^{17}} = 0.47 \text{ V}$$

(The density of state at conduction band edge for silicon, $N_C = 2.8 \times 10^{25}$ from Table 1.3**)**

4.2.4 Forward and Reverse Bias

When an external bias is applied, the metal-to-semiconductor barrier remains unchanged, whereas, the semiconductor-to-metal barrier is either decreased (forward bias) or increased (reverse bias).

When the metal is connected to a positive bias with respect to the semiconductor. The Fermi energy level of the metal is lowered from its equilibrium level. The depletion region is narrowed, and the potential barrier in the semiconductor is reduced. The number of electrons that diffuse from semiconductor to metal is now more than the number of electrons that drift from metal into the semiconductor. Thus, there will be a positive current through the device. Figure 4.4 illustrates a metal–semiconductor junction under forward-bias condition.

If the metal is connected to a negative bias with respect to the semiconductor, the metal is charged even more negatively than without any bias. The Fermi energy level of the metal is raised. The electrons in the semiconductor are repelled even more. The depletion region becomes wider and the potential barrier on semiconductor side is further increased, as shown in Fig. 4.5. However, the barrier on the

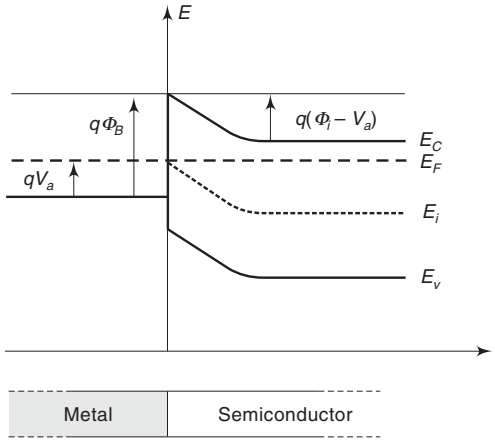


Fig. 4.4 *Energy band diagram under forward bias*

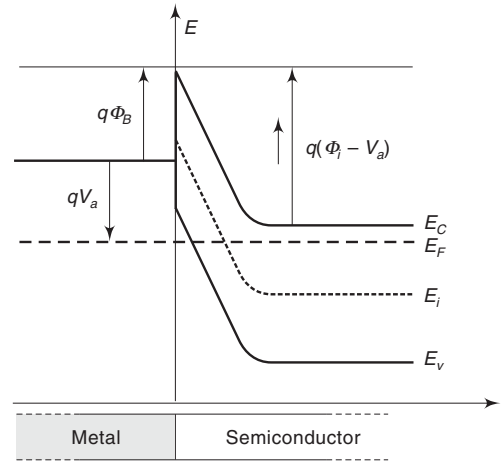


Fig. 4.5 *Energy band diagram under reverse bias*

metal side remains unchanged and limits the flow of electrons. A small current flows as a result of a few electrons in the metal acquiring enough thermal energy to overcome barrier.

4.2.5 Ohmic Contacts

An Ohmic contact is another type of metal–semiconductor junction. It is formed by applying a metal to a heavily doped semiconductor. Here the current is conducted equally in both directions and there will be a very little voltage drop across the junction. The usage of Ohmic contacts is to connect one semiconductor device to another on an IC, or to connect an IC to its external terminals.

Ohmic contacts are very common in semiconductor devices. Metal–semiconductor contacts cannot be considered to offer a resistance as low as that of two metals connected to each other. Metal–semiconductor junctions can act as either a rectifying junction or an Ohmic contact depending upon the Fermi energy levels of the metal and the semiconductor used. A proper choice of metal and semiconductor can offer a low resistance Ohmic contact. Alternatively, contacts that have a thin barrier can be created by heavily doping the semiconductor through which the carriers can tunnel. Both these types of contacts are presented in this section.

A metal–semiconductor junction can be an Ohmic contact if the Schottky barrier height, Φ_B, is zero or negative. This means, for an *N*-type semiconductor, that the metal work function, Φ_M, is either close to or smaller than the electron affinity (χ) of the semiconductor; and for a *P*-type semiconductor, the metal work function is either close to or greater than the sum of electron affinity and the bandgap energy.

That is, $\Phi_M \leq \chi$ for an *N*-type semiconductor

or, $\Phi_M \geq \chi + E_g$ for a *P*-type semiconductor

Figure 4.6 illustrates the energy band diagram when Q_M is less than the electron affinity, χ. In this case, the electrons flow from the metal into the semiconductor thus forming a positive surface charge layer in metal.

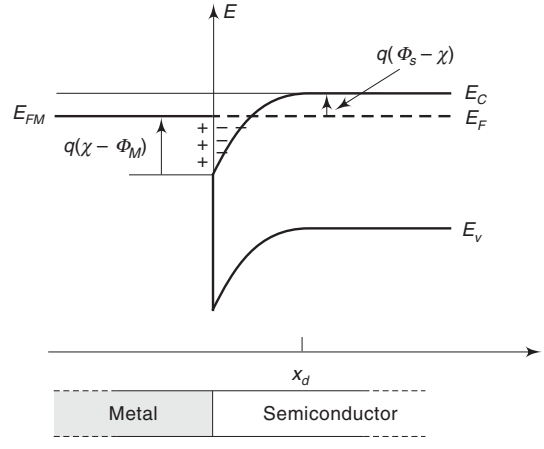


Fig. 4.6 *Energy band diagram in an Ohmic contact*

The resultant electric field sets up an electric potential and the energy bands of the semiconductor bend downward. There is no barrier for the flow of electrons in both directions. The current is directly proportional to the potential across the junction and is symmetric about the origin, as shown in Fig. 4.7.

A more practical method of providing contacts in semiconductor devices is to create a junction that consists of a thin barrier. Such contacts are also referred to as *tunnel contacts*. Such contacts have a positive barrier at the junction and a heavy doping in the semiconductor. This creates a very thin barrier separating the metal from the semiconductor, through which the carriers can easily tunnel. Figure 4.8 shows the energy band in a heavily doped Ohmic contact.

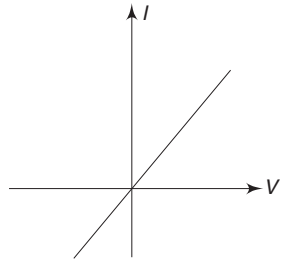


Fig. 4.7 *V-I characteristics of an Ohmic contact*

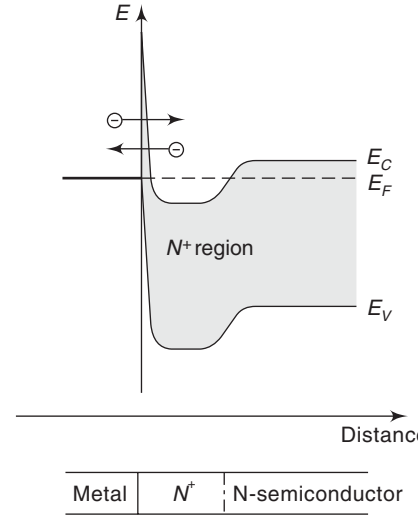


Fig. 4.8 *Energy band levels in an Ohmic contact*

4.3 GALLIUM ARSENIDE DEVICES

Gallium arsenide is a compound semiconductor made of a compound of two elements. Gallium having three valence electrons can be combined with arsenic, which has five valence electrons to form the compound GaAs. Figure 4.9 shows the arrangement of atoms in a gallium arsenide substrate. The gallium and arsenic atoms are alternatively positioned in exact crystallographic locations. Being a binary semiconductor, GaAs requires special care to avoid high temperatures during processing to prevent dissociation of the surface. This is one of the basic difficulties in the growth of bulk GaAs material.

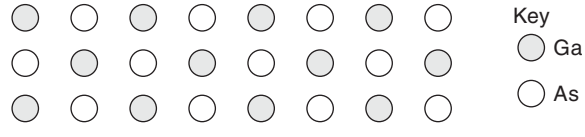


Fig. 4.9 *Arrangement of atoms in GaAs*

4.3.1 *N*-type Material

Group IV elements such as silicon can act as either donors (on Ga sites) or acceptors (on As sites). The covalent radius for Ga is 1.26 Armstrong units and As is 1.18 Armstrong units. As the silicon atoms are larger in size than arsenic, they tend to occupy gallium sites. Thus, silicon in GaAs is used as the dopant for the formation of *N*-type material. Figure 4.10 gives the structure of *N*-type GaAs material.

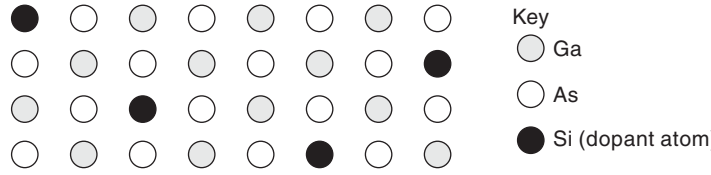


Fig. 4.10 *Arrangement of atoms in N-type GaAs*

4.3.2 *P*-type Material

Beryllium or magnesium (group II) can be used for the formation of *P*-type material. Beryllium being the lightest *P*-type dopant for GaAs, deep implantation of the dopant atoms can be accomplished. However, magnesium is also used as a suitable dopant in a number of processes.

4.3.3 Energy Band Structure

Gallium arsenide is a direct gap semiconductor, with its valence band maximum and conduction band minimum occurring at the same wave vectors. This means little momentum change is necessary for transition of an electron from the conduction band to the valence band. This property makes the GaAs an excellent light-emitting diode.

For the GaAs effective mass of the electrons travelling through the crystal is 0.067 times the mass of a free electron. This means electrons travel faster in gallium arsenide than in silicon. Gallium arsenide has a superior electron mobility characteristic because of the shapes of their conduction bands as shown in Fig. 4.11.

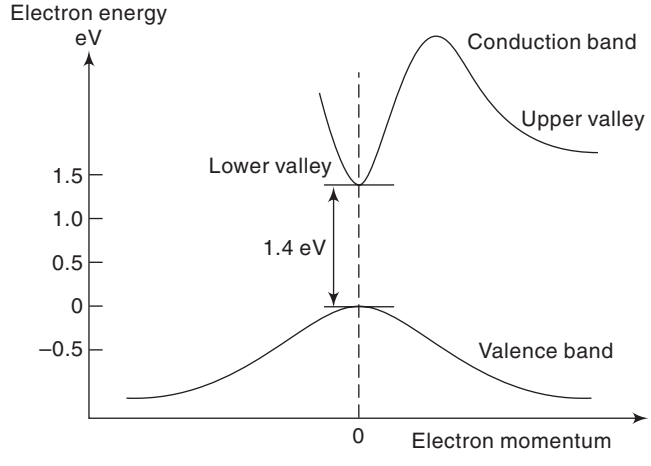


Fig. 4.11 *Energy band structure of GaAS*

4.3.4 GaAs Technology

As a result of considerable development in GaAs integrated circuitry and the technology, GaAs devices are widely gaining popularity. The GaAs devices provide the following typical characteristics:

- Less than one-micron gate geometry
- Less than two-micron metal pitch
- Up to four-layer metal
- ON and OFF devices
- Four-inch diameter wafers
- Suitability for clock rates in the range of 1 GHz–2 GHz.

The salient features of this technology include:

- Improved electron mobility over the silicon technology, resulting in very fast electron transit times.
- The saturation velocity for GaAs occurs at a lower threshold filled than for silicon.
- Large energy band gap offers bulk semi-insulating substrate with resistivities in the order of hundred to thousands of megaohms-cm.
- GaAs devices operate over a wider temperature range (–200°C to +200°C)
- Direct bandgap allows GaAs to be used as light emitters.
- Less power dissipation compared to silicon technology.

4.3.5 Types of GaAs Devices

Gallium arsenide (GaAs) devices have been used extensively in the development of Monolithic Microwave Integrated Circuits (MMIC) because of its suitability for both high-frequency transistors and low-loss passive components. GaAs has two key advantages over silicon for MMICs: (i) GaAs has higher saturated electron drift velocity and low-field mobility than silicon, resulting in faster devices, and (ii) GaAs can readily be made with high resistivity, making it a suitable substrate for microwave passive components. The GaAs device technology which is used for Radio Frequency Integrated Circuits (RFICs) and MMICs includes:

(a) GaAs MESFET
(b) GaAs HEMT
(c) GaAs HBT

The GaAs MEtal Semiconductor Field Effect Transistor (MESFET) was the first microwave transistor and remains an important building block for many MMIC designs. It is easily fabricated using ion-implantation for high-volume productions and has good noise figure and output power performance. MESFETs, having gate length of 0.5 μm, are useful for circuits operating up to 30 GHz. GaAs MESFETs are used for the fabrication of general-purpose, low DC power, high RF power switching devices, and also for the construction of enhancement mode/depletion-mode devices.

The GaAs based High Electron Mobility Transistor (HEMT) provides a considerable increase in transconductance over the MESFET and with short gate lengths of 0.1 μm results in the operation of circuits over 100 GHz. However, these high-performance devices are less suited to high volume manufacture than the MESFET, as the intricate material layers must be fabricated with either Molecular Beam Epitaxy (MBE) or Metal Organic Chemical Vapour Deposition (MOCVD), both of which are time-consuming processes. The gates must be defined with electron beam lithography, which is rather slow.

For good noise figures, the mushroom gate is needed, which must be produced using a precise multi-layer photoresist technique, PHEMT.

The GaAs Heterojunction Bipolar Transistor (HBT) has significant advantages over the HEMT because it is a vertical structure. The critical device dimensions are defined by the material growth and doping rather than by lithography. An HBT with an emitter width of several microns can offer good microwave performance. The HBT is intrinsically a very high-gain device, but tends to suffer from large parasitic resistances and capacitances. It can offer very high power density and efficiency. The high base resistance of HBTs means that their noise figure is considerably higher than that of HEMTs, but the device is continually being improved. HBT based products are particularly useful for manufacturing low-supply voltage components for mobile handsets.

4.3.6 GaAs Metal Semiconductor FET (MESFET)

The processing of a GaAs MESFET is relatively simple and requires only about 8 masking stages. The structure of the basic MESFET using GaAs technology consists of a thin *N*-type active region bridging two ohmic contacts with a narrow metal Schottky barrier *gate* that separates the more heavily doped *drain* and *source*. Fabrication involves photolithography, and ion implantation into semi-insulating GaAs structure. GaAs MESFETs are similar to silicon MOSFETs. The difference is the presence of a Schottky diode at the gate region.

4.3.7 Fabrication of GaAs Depletion-mode MESFET

A cross-sectional view of the structure of GaAs D-MESFET is shown in Fig. 4.12. A thin *N*-type region joins two ohmic contacts with a narrow metal Schottky barrier gate. In the process of fabricating the depletion-mode MESFET, the *N*-type dopants having a concentration density of approximately 10^{17} cm^{-3} are directly implanted into the semi-insulating GaAs substrate to form the *channel* and the more heavily doped *source* and the *drain.*

The *gate* and the first level interconnect metallizations are deposited by *E-beam evaporation* techniques. The transconductance of the device is determined by the gate length and its position relative to the drain and source contacts.

The fabrication process is illustrated in Fig. 4.12. Initially, the GaAs substrate is coated with the first level of insulator, a thin layer of silicon nitride. This thin film of insulator remains on the wafer throughout the processing steps and allows the annealing of GaAs at temperatures of up to 900 degrees centigrade. The second step is the formation of the *N*-type active layer, by direct ion implantation into the GaAs substrate. The Silicon nitride insulating layer is cut to make a window using photolithography processes for ion implantation of the *N* type active layer. Implantation of dopant ions takes place at about 200 to 230 keV to have an impurity concentration of up to 10^{12} cm^{-2}. There are only two main steps involved—formation of a shallow *N*-channel layer with high resistivity, and a deep N^+ layers for source and drain.

The wafer is then coated with the interleaved SiO_2 by chemical vapour deposition, and this layer provides protection against physical damage. This step is followed by encapsulation. Encapsulation is important as it prevents out diffusion of arsenic, brought about as the result of high vapour pressure of GaAs when exposed to high temperatures.

The next step in the process is formation of gates, the ohmic contacts and the first-level metal interconnects. The metals used must have a very low contact resistance. The ohmic contacts between the metal interconnect and the source and drain are deposited by evaporation using E-beam technology. A

thin layer of gold-germanium-platinum alloy is used for contacts. The following processes accomplish the first-level metallization, delineating photoresist patterns, plasma etching the underlying insulator, deposition of the metal on GaAs wafer by sputtering and photoresist lift-off.

Second-level metallization entails magnetron sputtered titanium/gold alloy, which is followed by filling the via between first-level and second-level metal. The final step in the fabrication is passivation and is used to protect the wafer from contamination. Passivation layer is deposited at low temperature, using plasma-enhanced chemical vapour deposition (CVD) process. The final structure of GaAs MESFET is shown in Fig. 4.13.

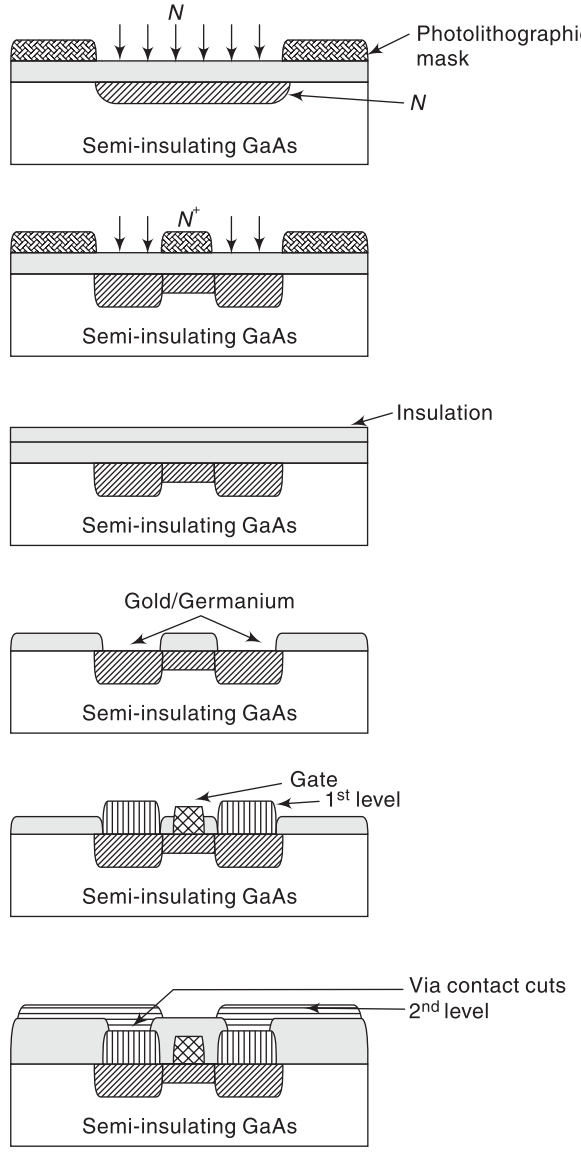


Fig. 4.12 *GaAs D-MESFET processing steps*

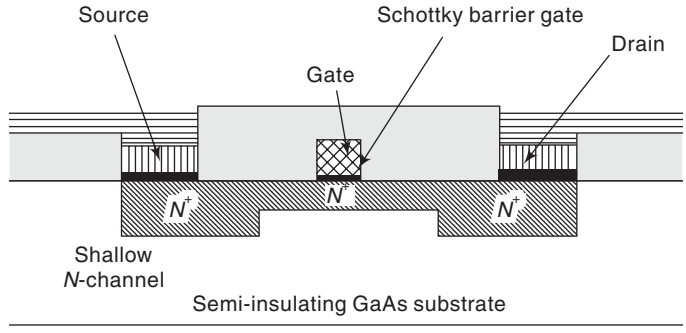


Fig. 4.13 *GaAs MESFET*

4.4 MESFET

MESFET is a modified form of MOSFET in which GaAs epitaxial layer was grown on semi-insulating GaAs substrate. Its operation is similar to *PN* junction JFET. The use of GaAs rather than silicon in MESFET provides two more significant advantages: first, the electron mobility at room temperature is more than 5 times larger, while the peak electron velocity is about twice that of silicon. Second, it is possible to fabricate semi-insulating (SI) GaAs substrates, which eliminates the problem of absorbing microwave power in the substrate due to free carrier absorption and provides better operation at high temperature. The MESFET has certain processing and performance advantages such as low temperature formation of the metal semiconductor barrier in comparison with *PN* made by diffusion or grown processes, low resistance and low voltage drop along the channel width and good heat dissipation for power devices.

4.4.1 Construction of MESFET

Metal-Semiconductor-Field-Effect-Transistor (MESFET) consists of a conducting channel positioned between a source and drain contact region as shown in Fig. 4.14(a). The carrier flow from source to drain is controlled by a Schottky metal gate. The control of the channel is obtained by varying the depletion layer width underneath the metal contact which modulates the thickness of the conducting channel and thereby the current between source and drain.

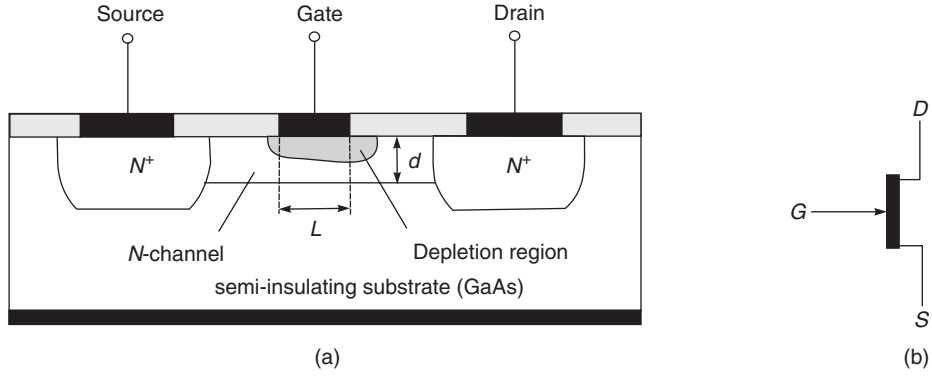


Fig. 4.14 *Depletion type MESFET: (a) Structure, (b) Symbol*

A thin epitaxial layer of GaAs is used as the active region as shown in Fig. 4.14(a). A very high-resistivity GaAs material referred to as semi-insulating substrate is used as the substrate. On this non-conducting substrate, a thin layer of lightly doped *N*-type GaAs is grown epitaxially to form the channel region of the FET. The photolithographic processing consists of defining patterns in the metal layers for source and drain ohmic contacts and Schottky barrier gate. The channel can be depleted to the semi-insulating substrate by reverse biasing the Schottky barrier gate. The channel conductance is modulated and a space charge region is induced under the metal gate by applying a reverse bias gate to source voltage V_{GS} as in the case of a JFET. The schematic symbol for *N*-channel depletion MESFET is shown in Fig. 4.14(b). Due to the presence of a metal-semiconductor junction, such FETs are called *Metal Semiconductor Field Effect Transistors* (*MESFETs*).

Operation

When a negative voltage is applied to the gate, it will attract free electrons (vacancies) in the channel to the metal surface. This reduces the number of carriers in the channel. Due to the reduction of carriers, drain current is also reduced, for the increasing values of negative voltage at the gate terminal. For positive voltages, additional electrons will be attracted into the channel and the current will rise as shown in the drain characteristics of Fig. 4.15. The drain and transfer characteristics of the depletion-type MESFET are very similar to those of depletion-type MOSFET.

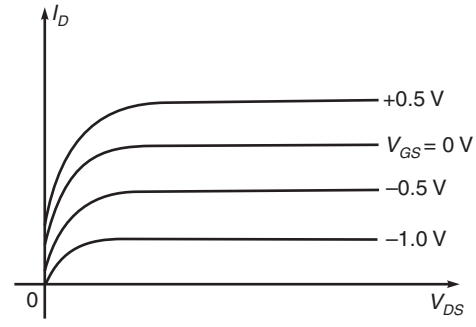


Fig. 4.15 *Drain characteristics of MESFET*

Enhancement-type MESFET is also possible with the same construction as depletion MESFET but without the initial channel between the source and drain. Figure 4.16(a) shows the structure of an *N*-channel enhancement MESFET. The response and characteristics for enhancement-type MESFET are essentially the same as depletion MESFET. But due to the Schottky barrier at the gate, the positive threshold voltage is limited to 0 V to about 0.4 V because the 'turn on' voltage for a Schottky barrier diode is about 0.7 V. When a positive voltage is applied to the gate, a channel is created between the source and drain, resulting in a large number of charge carriers (electrons). The schematic symbol for *N*-channel enhancement MESFET is shown in Fig. 4.16(b).

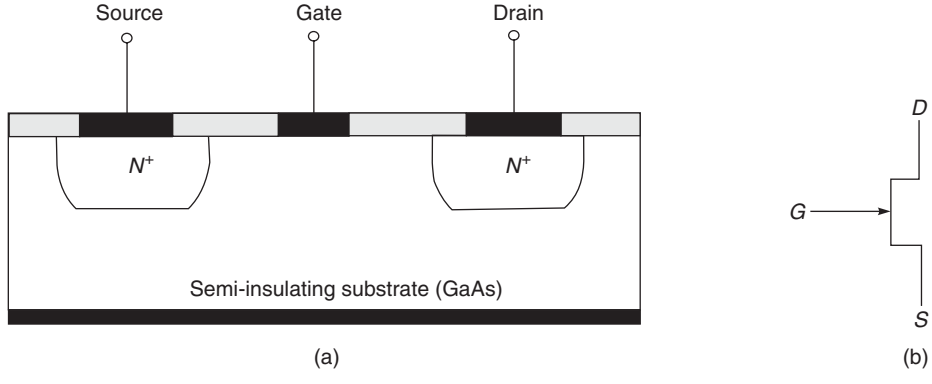


Fig. 4.16 *Enhancement type MESFET: (a) Structure, (b) Symbol*

As the mobility of holes in *P*-channel GaAs is relatively low compared to negatively charged carriers, it is not preferred for high-speed applications. Depletion-type and enhancement-type MESFETs are made with an *N*-channel between the source and the drain resulting in the commercial availability of *N*-type GaAs MESFETs. The length of the channel is made as small as possible (0.1 μm to 1 μm) for both types of MESFETs for high-speed applications.

Advantages

The key advantage of the MESFET is the higher mobility of the carriers in the channel as compared to the MOSFET. As the depletion region separates the carriers from the surface, their mobility is close to that of bulk material. The higher mobility leads to higher values for current, transconductance and transit frequency of the device.

Disadvantages

The disadvantage of the MESFET structure is the presence of the Schottky metal gate. It limits the forward bias voltage on the gate to the turn-on voltage of the Schottky diode. This turn-on voltage is typically 0.7 V for GaAs Schottky diodes. Therefore, the threshold voltage must be lower than this turn-on voltage. As a result, it is more difficult to fabricate circuits containing a large number of enhancement-mode MESFET.

Applications

The higher transit frequency of the MESFET makes it particularly of interest for microwave circuits. As MESFET provides a superior microwave amplifier or circuit, the limitation by the diode turn-on is easily tolerated. The use of GaAs substrates for MESFET offers some outstanding advantages over the traditional Si substrate, especially in the area of high-frequency and high-speed microwave circuits. Its application extends to low noise preamplifiers and linear amplifiers, oscillators and mixers in communication networks due to some of its outstanding physical properties, such as high cut-off frequency (more than 10 GHz for 1 μm gate length) and high electron mobility. Table 4.2 gives the comparison of MESFET with MOSFET.

Table 4.2 *Comparison of MESFET with MOSFET*

MOSFET	*MESFET*
Gate is controlled through a MOS barrier.	Gate control takes place through a Schottky barrier.
MOSFETs are usually fabricated in Si.	MESFETs are usually fabricated in GaAs.
MOSFET is normally in OFF state and there is a flow of current only when the channel is inverted by gate bias.	MESFET is normally in ON state and negative gate bias is needed to turn the current to OFF state (in case of *N*-channel MESFET).
Both *P*-channel and *N*-channel MOSFETs are possible giving rise to CMOS.	Typically *N*-channel MESFETs are feasible.
MOSFET is mainly used in integrated circuits, but microwave operation is not yet possible.	MESFET is mainly used in microwave devices but the integration level is not as high as the CMOS devices.

EXAMPLE 4.2

Consider an *N*-channel GaAs MESFET at $T = 300$ K with a gold Schottky barrier contact. Assume that the barrier height is $\phi_{Bn} = 0.89$ V. The *N*-channel doping is $N_D = 2 \times 10^{15}$ cm^{-3}. Determine the channel thickness such that $V_T = +0.25$ V, $N_C = 4.7 \times 10^{17}$cm^{-3} and ε_r of GaAs = 13.1. (AU April/May 2015)

Solution Given $T = 300$ K, $\phi_{Bn} = 0.89$ V, $N_D = 2 \times 10^{21}$ m^{-3}, $V_T = 0.25$ V and $N_C = 4.7 \times 10^{23}$ m^{-3}.

The threshold voltage for MESFET is given by

$$V_T = \phi_{Bn} - \frac{q \cdot N_D d^2}{2\varepsilon_o \varepsilon_r}$$

$$d^2 = \frac{2\varepsilon_o \varepsilon_r [\phi_{Bn} - V_T]}{q \cdot N_D}$$

$$d = \left[\frac{2\varepsilon_o \varepsilon_r [\phi_{Bn} - V_T]}{q \cdot N_D}\right]^{\frac{1}{2}}$$

Substituting the values, we get

$$d = \left[\frac{2 \times 8.854 \times 10^{-12} \times 12(0.89 - 0.25)}{1.602 \times 10^{-19} \times 2 \times 10^{21}}\right]^{\frac{1}{2}} = 0.6\ \mu\text{m}$$

4.5 SCHOTTKY BARRIER DIODE

[AU May 2017, Dec 2015, May 2014, 8 marks]

The metal contacts are required to be ohmic and no *PN* junctions to be formed between the metal and silicon layers. The N^+ diffusion region serves the purpose of generating ohmic contacts. On the other hand, if aluminium is deposited directly on the *N*-type silicon, then a metal-semiconductor diode can be said to be formed. Such a metal semiconductor diode junction exhibits the same type of *V-I* characteristics as that of an ordinary *PN* junction.

The cross-sectional view and symbol of a Schottky barrier diode are shown in Fig. 4.17(a) and (b) respectively. Contact 1 shown in Fig. 4.17(a) is a Schottky barrier and the contact 2 is an ohmic contact. The contact potential between the semiconductor and the metal generates a barrier for the flow of conducting electrons from semiconductor to metal. When the junction is forward biased, this barrier is lowered and the electron flow is allowed from semiconductor to metal, where the electrons are in large quantities.

The majority carriers carry the conduction current in the Schottky diode whereas in the *PN* junction diode, minority carriers carry the conduction current and it incurs an appreciable time delay from ON state to OFF state. This is due to the fact that the minority carriers stored in the junction have to be totally removed.

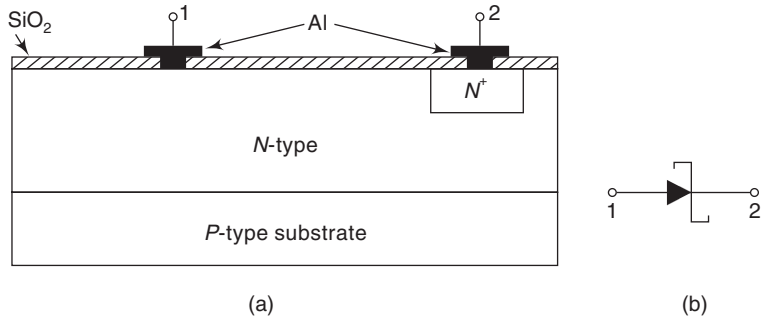


Fig. 4.17 *Schottky barrier diode: (a) Cross-sectional view, (b) Symbol*

This characteristic puts the Schottky barrier diode at an advantage since it exhibits negligible storage time to flow the electron from *N*-type silicon into aluminium almost right at the contact surface, where they mix with the free electrons. The other advantage of this type of diode is that it has less forward voltage (approximately 0.4 V). Thus, it can be used for clamping and detection in high frequency applications and microwave integrated circuits.

Schottky barrier diode is an extension of the oldest semiconductor device that is the point contact diode. Here, the metal–semiconductor interface is a surface, the Schottky barrier rather than a point contact. The Schottky diode is formed when a metal, such as Aluminum, is brought into contact with a moderately doped *N*-type semiconductor as shown in Fig. 4.18. It is a unipolar device because it has electrons as majority carriers on both sides of the junction. Hence, there is no depletion layer formed near the junction. It shares the advantage of point contact diode in that there is no significant current from the metal to the semiconductor with reverse bias. Thus, the delay present in the junction diodes due to hole–electron recombination time is absent here. Hence, because of the large contact area between the metal and semiconductor than in the point contact diode, the forward resistance is lower, and so is noise.

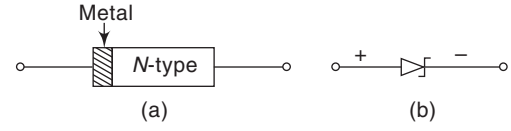


Fig. 4.18 *Schottky diode: (a) metal–semiconductor contact (b) circuit symbol*

The forward current is dominated by electron flow from semiconductor to metal, and the reverse current is mainly due to electron flow from metal to semiconductor. As there is very little minority carrier injection from semiconductor into metal, Schottky diodes are also said to be majority carrier devices.

The diode is also referred to as hot carrier diode because when it is forward biased, conduction of electrons on the *N* side gains sufficient energy to cross the junction and enter the metal. Since these electrons plunge into the metal with large energy, they are commonly called as hot carriers.

Figure 4.19 shows the *V-I* characteristics of a Schottky diode and a *PN* junction diode. The current in a *PN* junction diode is controlled by the diffusion of minority carriers whereas the current in the Schottky diode results from the flow of majority carriers over the potential barrier

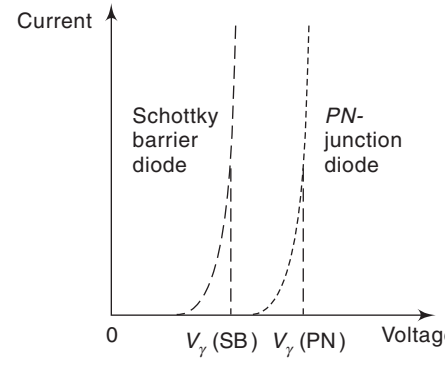


Fig. 4.19 *V–I characteristics of Schottky barrier diode and PN junction diode*

at the metal-semiconductor junction. The reverse saturation current for a Schottky diode is larger than that of a *PN* junction diode. The storage time for a Schottky diode is theoretically zero. The Schottky diode has a smaller turn-on voltage and shorter switching time than the *PN* junction diode.

Schottky diode can be used for rectification of signals of frequencies even exceeding 300 MHz. It is commonly used in switching power supplies at frequencies of 20 GHz. Its low noise figure finds application in sensitive communication receivers like radars. It is also used in clipping and clamping circuits and in computer gating.

4.6 DUAL-GATE MOSFETS [AU May 2016, May 2015 and May 2014, 6 marks]

The functions of integrated circuits have doubled per chip every two years since the invention of integrated circuits in 1960s. This continued progress in integrated-circuit development has been the driving factor for the information revolution. This trend known as *Moore's law* has paced the semiconductor industry. The sustained growth in VLSI technology is fuelled by the continued shrinking of transistors to very small dimensions. The benefits of miniaturization such as higher packing densities, higher circuit speeds, and lower power dissipation have been the key in the evolutionary progress leading to today's computers, wireless units and communication systems that offer superior performance, dramatically reduced cost per function and much reduced physical size, in comparison with their predecessors.

MOSFET forms the core component of integrated circuits. MOSFET has high switching speed and the circuit speed increases with miniaturization, which is the major advantage as compared to BJT. As the device dimension gets smaller and the circuit gets denser and more complex, problems are encountered in lithography, interconnects, and processing. At high frequencies, the gate-channel capacitance limits the switching speed, which can be changed via the channel resistance. For good high-frequency response, the ratio $g_m/2\pi C_g$ should be large, i.e., the transconductance, g_m, should be large and the total gate capacitance, C_g, should be small. Hence, the channel length should be short. This can be achieved by fabricating gates shorter with insulated gate technology.

As the dimension of the MOSFET gets scaled down, the close proximity between the source and the drain reduces the ability of the gate electrode to control the potential distribution and the flow of current in the channel region. These undesirable effects, called "short-channel effects" start plaguing MOSFETs.

Short-channel effects arise when control of the channel region by the gate is affected by electric field lines from source and drain. In the conventional bulk MOSFET shown in Fig. 4.20(a), the electric field lines propagate through the depletion region associated with the junctions. Their influence on the channel can be reduced by increasing the doping concentration in the channel region. In very small devices, the doping concentration becomes too high (10^{19} cm^{-3}) for proper device operation.

Most of the field lines propagate through the buried oxide (BOX) before reaching the channel region as shown in Fig. 4.20(b) and in a fully depleted SOI (FDSOI) device, short-channel effects can be reduced by using a thin buried oxide and an underlying ground plane. In FDSOI, most of the electric field lines from the source and drain terminate on the buried ground plane instead of the channel region as shown in Fig. 4.20(c). However, this approach has the inconvenience of increased junction capacitance.

An efficient device configuration is obtained by using the double-gate transistor structure. In a double-gate device, both gates are connected together. The electric field lines from source and drain underneath the device terminate on the bottom gate electrode and it cannot reach the channel region as shown in Fig. 4.20(d). Only the field lines that propagate through the silicon film itself can encroach on the channel region and degrade short-channel characteristics. This encroachment can be minimised by reducing the silicon film thickness.

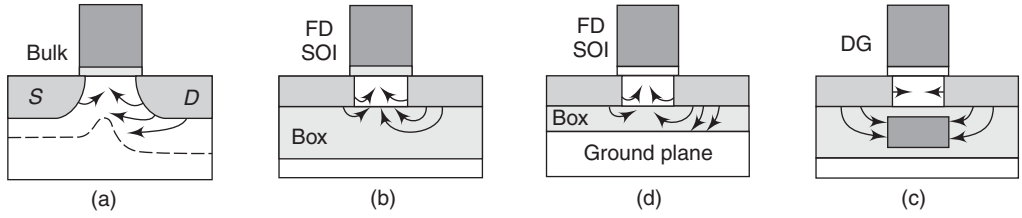


Fig. 4.20 *Encroachment of electric field lines from source and drain on the channel region in different types of MOSFETs (a) Bulk MOSFET, (b) Fully depleted SOI MOSFET, (c) Fully depleted SOI MOSFET with thin buried oxide and ground plane, (d) Double-gate MOSFET*

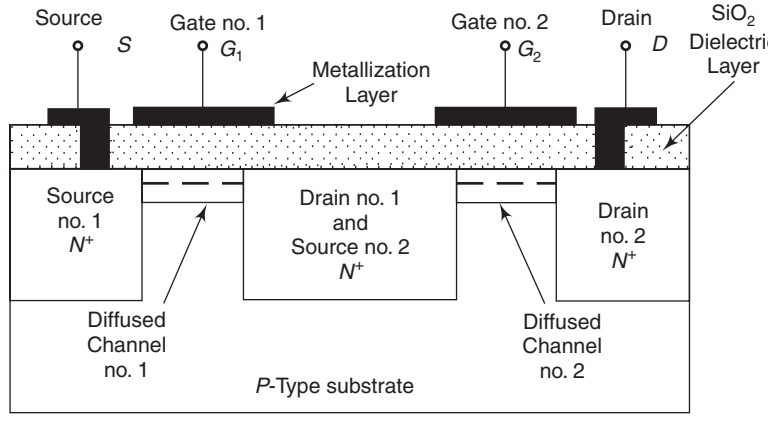
4.6.1 Construction of Dual-Gate MOSFET

The operation of conventional MOSFET is limited at high frequencies by its high gate-to-channel capacitance. The metal plate used for the gate is a conductor. The silicon dioxide between the gate and channel is a dielectric layer. Since the channel itself is considered a conductor, the combination of the three forms a capacitor.

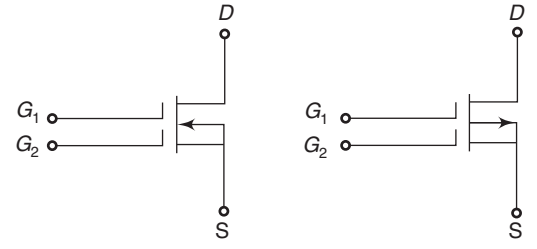
In a dual-gate MOSFET, two gate terminals are provided as compared to a conventional single-gate MOSFET. A *dual-gate* MOSFET uses two gates to reduce the overall high gate-to-channel capacitance at high frequencies. The voltages at both the gate terminals control the flow of current through the MOSFET. The dual-gate MOSFET can be considered the counterpart of a tetrode as the control is exerted by two gates similar to a tetrode. The physical construction of an *N*-channel dual-gate MOSFET is shown in Fig. 4.21(a) and the schematic symbols for both *N*-channel and *P*-channel are shown in Fig. 4.21(b). The device is normally wired so that the two gates are in *series*. Here, the N^+ region in the middle acts as drain for MOSFET-1 and source for MOSFET-2. When the dual-gate MOSFET is used as two series MOSFETs, the effect is similar to connecting two capacitors in series. As the total capacitance in a series connection is less than either individual value, by connecting the two gates in series, the overall capacitance is reduced.

Operation

In the *N*-channel dual-gate MOSFET, the voltages at both the gate terminals control the flow of current. When the voltages applied to the gate terminals such as gate-1 and gate-2 are greater than the threshold voltage, a channel is formed between the corresponding source and drain. When the gate voltage of either of the two gate terminals is made negative, the drain current decreases as shown in Fig. 4.22. The drain current gets enhanced when the gate voltage applied to both gate terminals (gate-1 and gate-2) are made positive. The transfer characteristics of the dual-gate *N*-channel MOSFET are shown in Fig. 4.22.



(a) *N*-channel dual gate MOSFET



(b) Schematic symbol for *N*-channel and *P*-channel MOSFET

Fig. 4.21

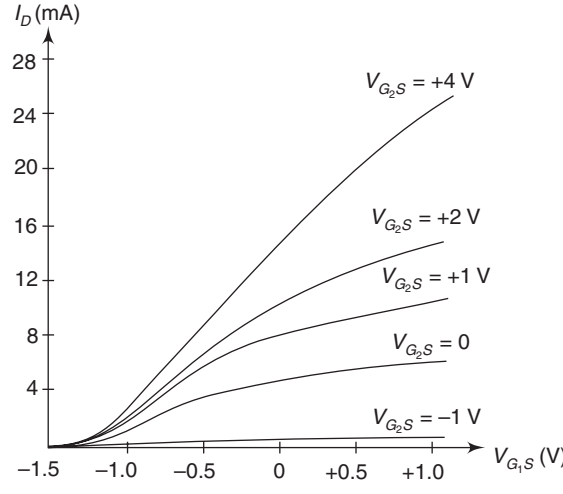


Fig. 4.22 *Transfer characteristics of N-channel dual gate MOSFET*

Applications

At high frequencies, the conventional MOSFET has high gate-to-channel capacitance. This capacitance is a result of the *metal-oxide-semiconductor* layers. A *dual-gate* MOSFET reduces the overall gate-to-channel capacitance because the total series capacitance is lower than either individual capacitance. Due to simultaneous control of two gate voltages, the dual-gate MOSFET has wide applications in mixers, demodulators, cascade amplifier, RF amplifier and AGC amplifiers. When the device is used in an AGC amplifier, the signal is connected to gate-1 to amplify the signal and the voltage is applied to gate-2 to control the gain.

4.7 FinFET

[AU Dec 2016, May 2014, 4 marks]

In many MOSFET devices, in addition to the usual gate formed on the top of the device, there can be another gate on the bottom of the device. The gate formed on top of the device is usually referred to as the *front gate,* and the one on the bottom is called the *back gate*. Both gates control the channel potential. A typical back-gated structure is shown in Fig. 4.23. Dual-gate devices have also been made as shown in Fig. 4.24, in which both the front and back gates control the channel. The two gates tend to shield the source region of the channel from the action of the drain, providing a high degree of isolation of the source and drain. As a result, short-channel effects are reduced in this structure.

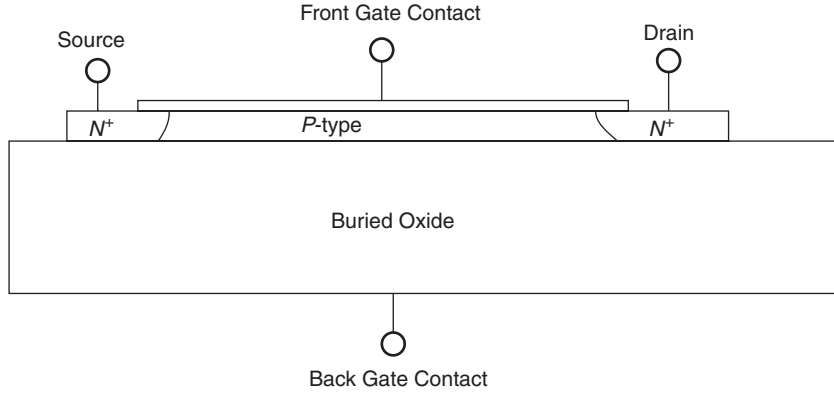


Fig. 4.23 *Back-contacted SOI MOSFET device*

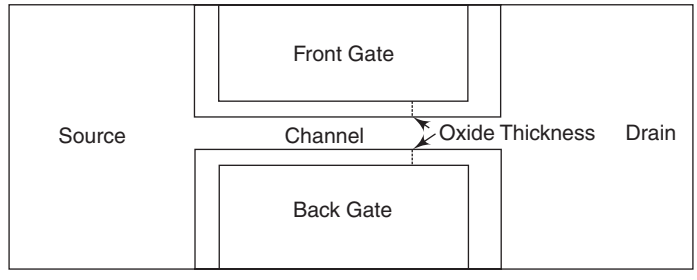


Fig. 4.24 *Cross section of a dual-gate MOSFET device*

The significant reduction of short-channel effects can be obtained by sandwiching a fully depleted SOI device between the two gate electrodes connected together. The first fabricated double-gate SOI MOSFET was the *fully DEpleted Lean-channel TrAnsistor* (*DELTA*) where the device is made in a tall and narrow silicon island called "finger", "leg" or "fin" as shown in Fig. 4.25.

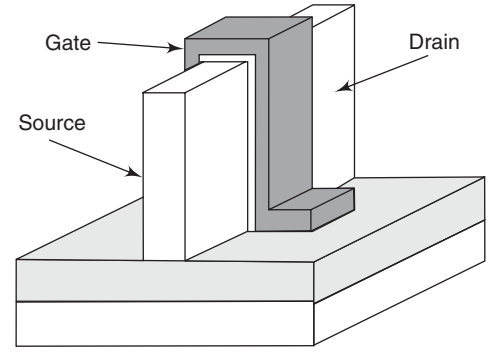


Fig. 4.25 *Double-gate MOS structure — DELTA MOSFET*

The FinFET structure is similar to DELTA, except for the presence of a dielectric layer called the hard mask on top of the silicon fin. The hard mask is used to prevent the formation of parasitic inversion channels at the top corners of the device. The channel is embodied in a fin sitting above the insulator in a wafer plane (SOI wafer). The silicon fin is either undoped or lightly doped. The gates of the FinFET are formed by wrapping the gate material around the sides of the silicon fin, resulting in front and back gates as shown in Fig. 4.26.

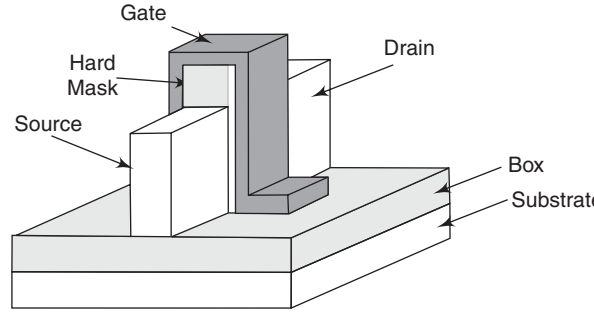


Fig. 4.26 *Double-gate MOS structure–FinFET*

4.7.1 Construction of FinFET

The FinFET transistor was initially named DELTA (fully DEpleted Lean-channel TrAnsistor) and has been renamed FinFET. The structure and symbol of a FinFET is shown in Fig. 4.27(a) and (b) respectively. Starting from an SOI wafer, very narrow silicon membranes are etched. Then, the gate stack is deposited and patterned on this topology. After spacer formation, selective epitaxial growth of silicon on the extension and source/drain areas is performed in order to reduce the series resistances as much as possible. These extension and source/drain electrode regions are then implanted and silicided. As for FDSOI and planar double-gate structures, the channel can be left undoped and the threshold voltage is adjusted by choosing the appropriate metal for the gate material. The height of this fin will be at the end of the electrical width of the device. Indeed, contrary to planar devices, the conduction takes place on the vertical sidewalls of the fin. The conduction width is thus twice that of the fin height (h_{fin}). As the fin height is limited to typically 50–100 nm, FinFETs are usually designed as multifinger transistors, with a conduction width quantified by $2h_{\text{fin}}$.

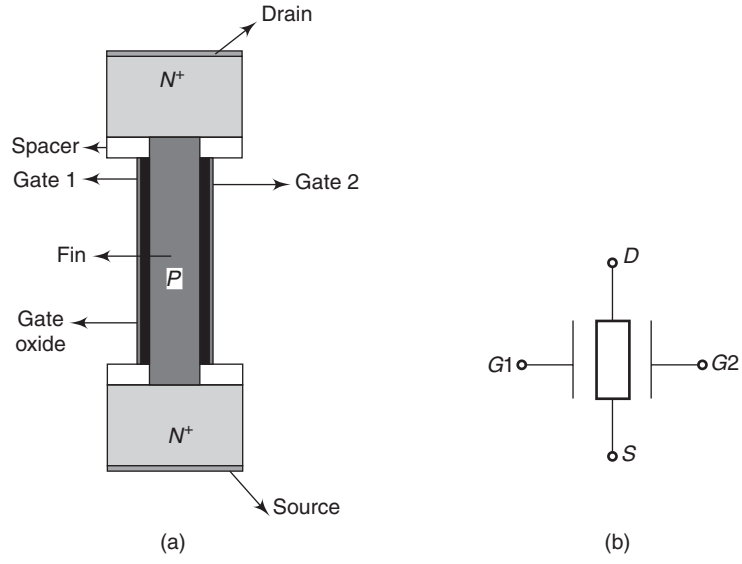


Fig. 4.27 *FinFET: (a) Structure, (b) Schematic symbol*

Operation

The independently driven FinFET is a double-gate device in which the two gate electrodes are not connected together and can be biased with different potentials. The I_D versus V_G characteristics of an independently driven double gate FinFET is shown in Fig. 4.28.

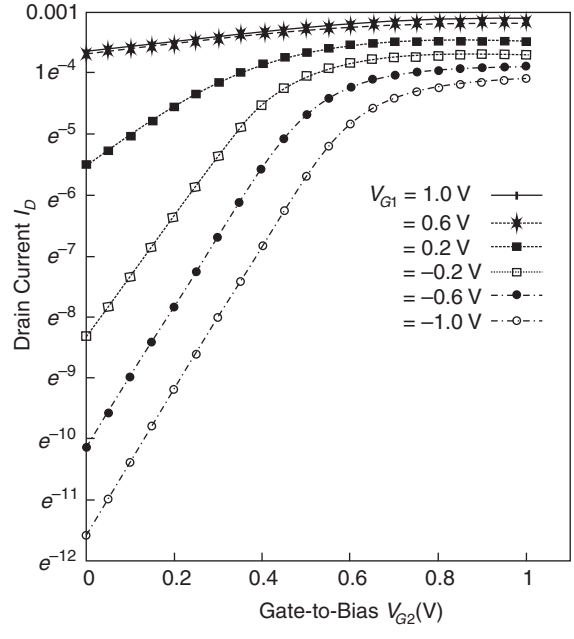


Fig. 4.28 *I_D versus V_G characteristics of a FinFET*

Applications

The double-gate FinFETs which are driven independently are used to construct low-power logic gates, single transistor mixers and SRAMs.

4.8 PIN-FET

A PIN-FET (P-Intrinsic-N, Field Effect Transistor) is an electronic device formed by the combination of a PIN photodiode and a Field Effect Transistor (FET), and packaged in a single housing. This device is mainly used as low-noise, high-speed photoreceiver in optical communication and data-processing systems. The photoreceiver fabricated by using PIN-FET has wide dynamic range and increased sensitivity. The term "PIN-FET" indicates the integration of a PIN photodiode and a discrete, high input impedance field effect transistor.

4.8.1 PIN Photodiode

A PIN photodiode is used for detection of light at the receiving end in optical communication. It is a three-region reverse-biased junction diode. A layer of intrinsic silicon is sandwiched between heavily doped *P* and *N* type semiconductor materials. As shown in Fig. 4.29, the depletion region extends almost to the entire intrinsic layer where most of the absorption of light photons takes place. The width of the intrinsic layer is large compared to the width of the other two layers. This ensures large absorption of light photons in the depletion region which also forms the absorption region. Light photons incident on the PIN photodiode are absorbed in the absorption region which leads to the generation of electron-hole pairs. These charge carriers present in the depletion region drift under the influence of the existing electric field that is set up due to the applied reverse bias. The reverse current flowing in the external circuit linearly increases with the level of illumination.

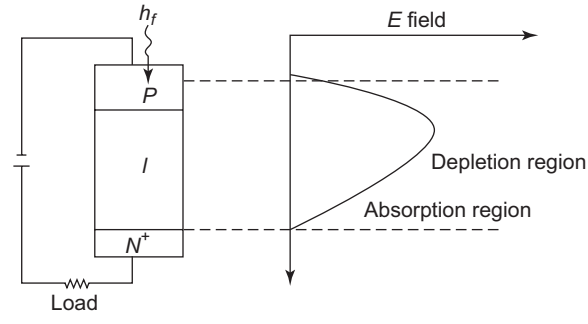


Fig. 4.29 *Structure of a PIN photodiode*

As the process of drifting is quicker than diffusion, the transit time of the charge carriers is small so that the response time is considerably reduced. The large width of the depletion region results in achieving high quantum efficiency.

4.8.2 Construction of PIN-FET Device

A monolithic PIN-FET optical receiver is fabricated on a semi-insulating InGaAs substrate by using Metal-Organic Chemical Vapour Deposition (MOCVD) technique grown multilayer structure and a latest interconnection technique.

The structures can be classified based on vertical or horizontal integration. For horizontal integration, the photodiode material is embedded in a semi-insulation substrate and the FET is fabricated on an adjacent portion of the substrate. In vertical integration, the epitaxial layers for the photodiode and FET are sequentially grown, separated by a high resistivity isolation layer.

Here, the PIN-FET is a horizontally integrated structure consisting of a PIN photodiode with a transimpedance preamplifier of depletion mode FET. Its response time can be varied from 1 ns to 40 ns by varying the parasitic capacitances.

The PIN-FET is an integration of an InGaAs PIN photodiode and an InGaAs FET. The schematic cross-section of PIN-FET device and its circuit are shown in Fig. 4.30. An integrated PIN-FET amplifier with high sensitivity is made by high performance ion-implanted InP FET. The material is grown on an N^+ InP substrate in the following two steps:

1. The InGaAs layer of low electron concentration is grown on the layer of undoped InP by chloride vapour phase epitaxy (VPE)
2. The second growth is done by atmospheric pressure metal organic chemical vapour deposition (MOCVD) technique.

The wafer with grown epilayers is then subjected to a series of ion implantations to form the FET. Now, the channel is formed by implanting Si, which is followed by the source-drain implant to form highly doped regions. Subsequently, As/Be co-implant is used to form the P^+ gate region. All implants are selective, non-self-aligned by using SiN_x and photoresist as an implantation mask. The implanted wafer is then activated in a Rapid Thermal Annealer (RTA) by using a dielectric capping layer.

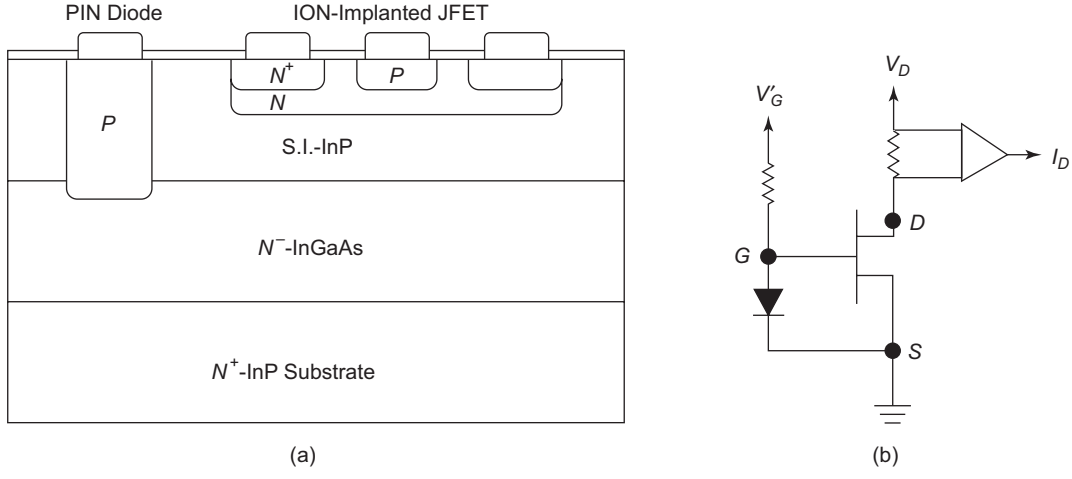


Fig. 4.30 *(a) Cross-sectional view of PIN-FET device (b) Circuit of PIN-FET device*

After the deposition of a SiN_x passivation layer, the diffusion of Zn is performed to form the PIN diode by using an evaporated film of zinc phosphide. The diffusion is performed in an RTA, and the resulting PN junction is located near the InP-InGaAs interface. Then, the gate, source and drain, and interconnections are performed by using metallization processing techniques.

4.8.3 Application of PIN-FET Device

The PIN-FET is used in line-terminating equipment (repeater) and in optical sensor system. Here, light is coupled into the detector stage through a multimode optical fiber pigtail, which allows excellent coupling with either single mode or multimode fiber systems. It offers long wavelength receiver system with high sensitivity and wide dynamic range. As an example, if existing high speed LED transmitter is used along with PIN-FET receiver, an optical data link length of up to 40 km is possible for data rate of over 100 Mbit/s.

4.9 CARBON NANOTUBE FIELD EFFECT TRANSISTOR (CNTFET)

4.9.1 Introduction to Carbon Nanotube

A carbon nanotube (CNT) is a cylindrical shaped nanostructure, made of allotropes of carbon, having a diameter measuring on the nanometer scale. A nanometer is one-billionth of a meter, or about 10,000 times smaller than a human hair. CNT is unique because the bonding between the atoms is very strong. Depending on the diameter of the nanotube, the intrinsic carrier concentration and bandgap will vary. The electrical properties of semiconducting CNT such as large mean free path, excellent carrier mobility, ballistic carrier transport for several hundred nanometers and improved electrostatics have made it a suitable replacement for silicon at the nanoscale dimensions. By using a single-wall CNT as the channel between two electrodes which work as the source and drain contacts of a FET, a CNTFET can be fabricated. Since the gate contact wraps all around the channel (CNT), it has a very good control on carrier, and allows for better electrostatics.

Carbon 60 (C60) Fullerene Carbon Nanostructure

There are a number of *carbon nanotube* production methods as well as fullerene synthesis techniques. Fullerenes were first observed after vaporizing graphite. A fullerene carbon nanostructure is shown in Fig. 4.31(a). Each carbon atom is at the intersection of two hexagons and one pentagon. Numbered ones are pentagons i.e. there are 12 pentagons and 20 hexagons in Fig. 4.31(b).

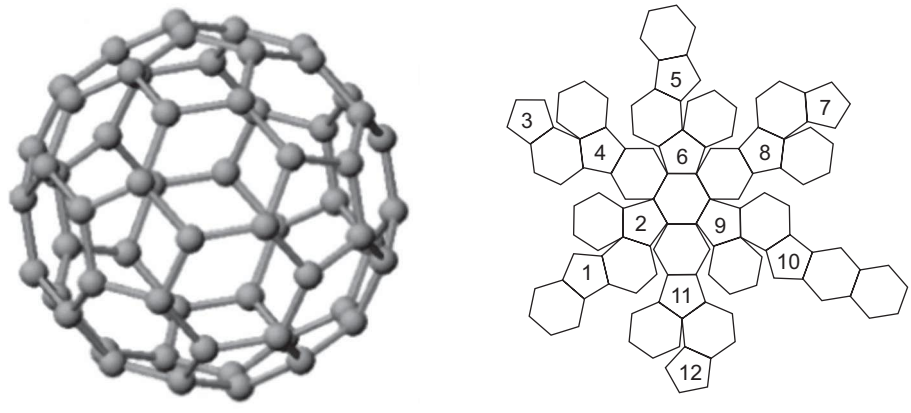


Fig. 4.31 *(a) Fullerene Carbon nanostructure (b) Carbon atom at intersection of two hexagons and one pentagon*

Graphene Carbon Nanotube

Carbon nanotube (CNT) can also be formed with seamless cylinders of one or more layers of graphene. The hexagonal structure of graphene carbon nanotube structure is shown in Fig. 4.32. In this structure, each carbon atom bonds with its neighbouring atoms, consisting of total three bonds.

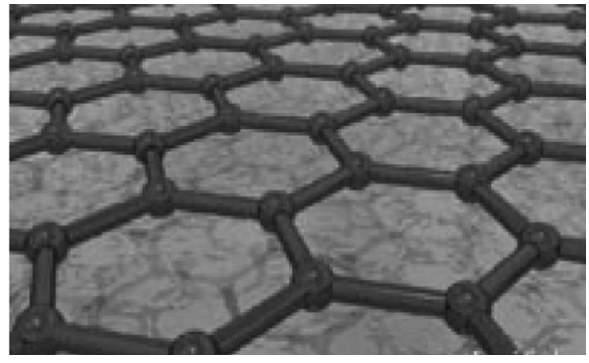


Fig. 4.32 *Structure of Graphene Carbon Nanotube*

4.9.2 Types of Carbon Nanotubes

There are two types of carbon nanotubes based on how they are constructed. They are:

1. Single Walled Carbon Nanotube (SWCNT)
2. Multi Walled Carbon Nanotube (MWCNT)

Single Walled Carbon Nanotube

A Single Walled Carbon Nanotube (SWCNT) capped with half of a C_{60} fullerene molecule at the ends is shown in the Fig. 4.33. The metallic and semiconducting properties of SWCNT can carry high current in the device.

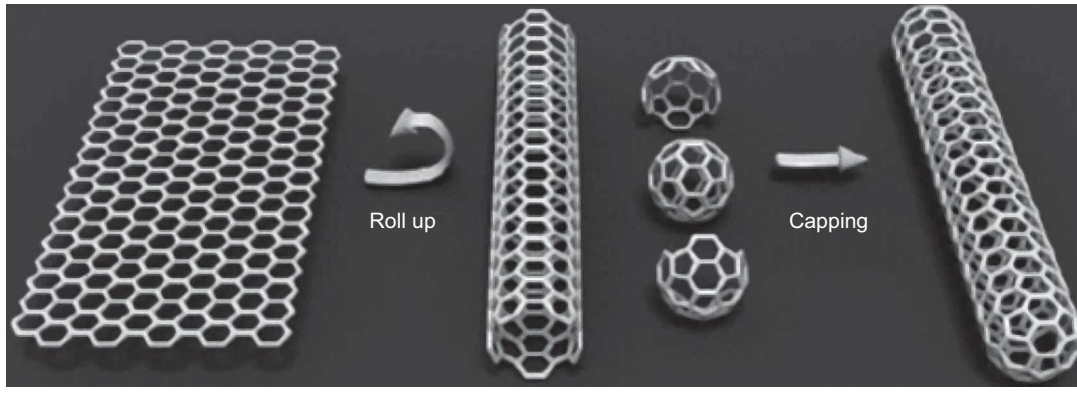


Fig. 4.33 *Single Walled CNT capped with half of a C_{60} fullerene molecule*

Multi Walled Carbon Nanotube

Two or more CNTs are concentrically rolled up to make an MWCNT. The structure of an MWCNT is shown in the Fig 4.34.

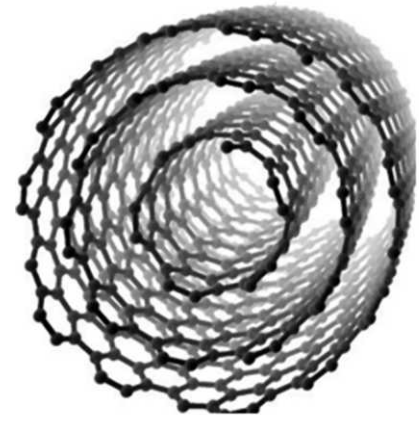


Fig. 4.34 *Multi Walled Carbon Nanotube (MWCNT)*

4.9.3 Carbon Nanotube Field-Effect Transistor

The first transistor utilizing carbon nanotubes was fabricated in the year 1998. A Carbon Nanotube Field Effect Transistor (CNTFET) is a nanoscale device that has excellent mechanical and electrical properties. CNT is used to develop FET for the small-scale transistor. Its construction and operation are similar to those of a MOSFET, hence, it is called Carbon Nanotube Field Effect Transistor (CNTFET). CNTFET is a novel molecular electronic device. MOSFET involves two main electrodes, designated as *source* and *drain*, connected by a semiconducting channel (Gate). In conventional devices, this channel is made of silicon. In CNTFET, the channel is formed by a single semiconducting CNT. This design was kept at research level and experimented on until about 2002. IBM at this point comes out with a new optimized design, providing many characteristic improvements over the first generation CNTFET.

The structure of CNTFET is shown in Fig. 4.35. Since CNTFET is similar to the MOSFET, it requires three terminals: *source*, *drain*, and *gate*. The gate (G) is used to control the current across the source (S) and drain (D) terminals as shown in Fig. 4.35. When the gate is ON, the current is able to flow across the source and drain through a channel. The main difference between CNTFET and MOSFET is that CNTFET uses the carbon nanotubes as the channel, whereas MOSFET channel is made from lightly doped silicon. Both the technologies use complimentary devices, the *P*-type and *N*-type. The *P*-type transistor conducts holes, whereas the *N*-type conducts electrons. CNTFET uses *N*-doped CNT as the contact. Potassium doped source and drain regions are used to fabricate CNTFET.

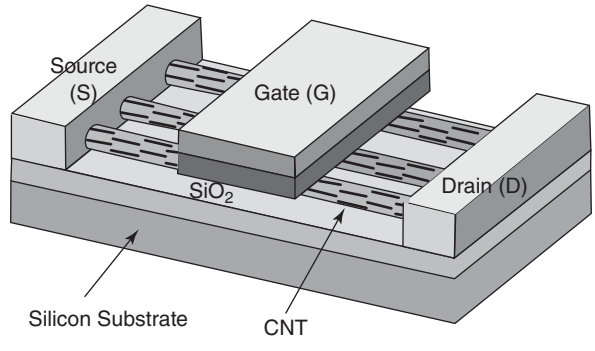


Fig. 4.35 *Structure of CNTFET*

DC Characteristics of CNTFET

The *V-I* characteristics of P-type and N-type CNTFET are shown in Fig. 4.36(a) and (b) respectively, which are similar to that of MOSFET. CNTFET shows excellent characteristics on *V-I* analysis to decide transistor and hence, MOS circuit is implemented by using CNTFET. CNTFET overcomes the constraints of the MOSFET.

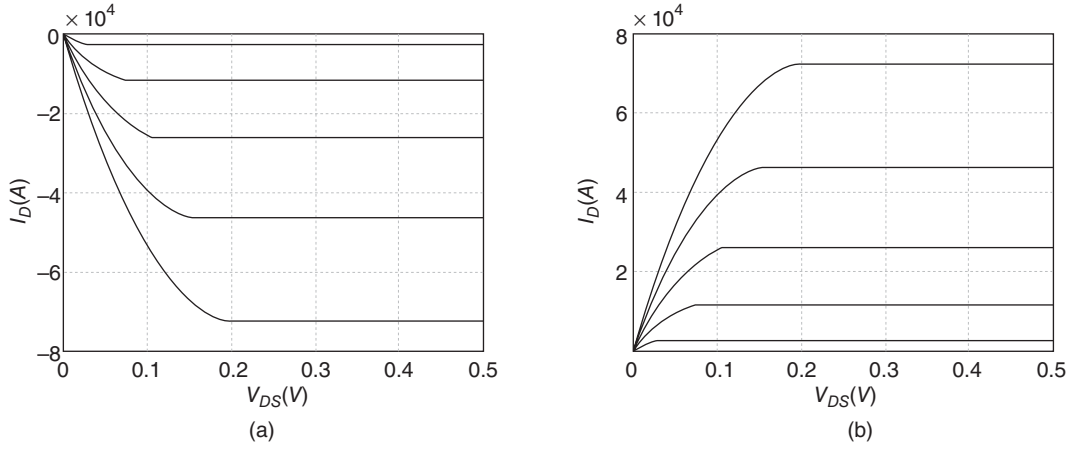


Fig. 4.36 *DC Characteristics of CNTFET (a) P-type and (b) N-type*

Advantages and Applications of CNTFET

The advantage of ON-current performance of the CNTFET is due to the high gate capacitance or the improved channel transport. As the diameter increases, the switching speed increases and the device gets saturated faster. The drain current increases with the diameter of CNT and its difference decreases as the values of diameter go higher. The improved channel velocity for the CNTFET is due to its increased mobility and band structure.

The usage of CNTFET helps to reduce power consumption, has higher gain, is stable, and is easily implanted in logic circuits. CNTFET can reach a higher frequency domain (THz regime) than conventional semiconductor technologies.

4.9.4 Comparison between CNTFET and Si-MOSFET

The comparison between the performance of the CNTFET and Si-MOSFET is as follows:

1. The threshold voltage of CNTFET is inversely proportional to the diameter of CNT. With decreasing diameter, the power reduces and delay increases but it is moderately fast. Hence, CNTFET based circuits are energy efficient.
2. At the voltage of over 1 V, CNTFET is capable of delivering at least three times higher drive currents than the Si-MOSFET.
3. The switching occurs by modulating the channel resistivity in Si-MOSFET, whereas it occurs by modulation of contact resistance in CNTFET.
4. The average carrier velocity in CNTFET is almost double that in Si-MOSFET.
5. CNTFET has four times higher transconductance when compared to Si-MOSFET.

4.9.5 Types of CNTFET

CNTFET has been fabricated with Ohmic or Schottky contact. There are two main types of CNTFET based on the current injection methods:

1. Metal Oxide Semiconductor type Carbon Nanotube Field Effect Transistor (MOS-CNTFET)
2. Schottky Barrier Carbon Nanotube Field Effect Transistor (SB-CNTFET)

Metal Oxide Semiconductor type Carbon Nanotube Field Effect Transistor

Construction The construction of the CNTFET is shown in Fig. 4.37. The device is made up of silicon which forms the substrate of the device. A layer of silicon dioxide (SiO_2) insulator is grown over the silicon substrate. An *ideal* MOSFET like CNTFET is formed by one or more semiconducting CNTs perfectly aligned and well-positioned whose section under the gate is intrinsic and the source or drain extension regions are *N*-type/*P*-type doped depending upon the type of CNTFET.

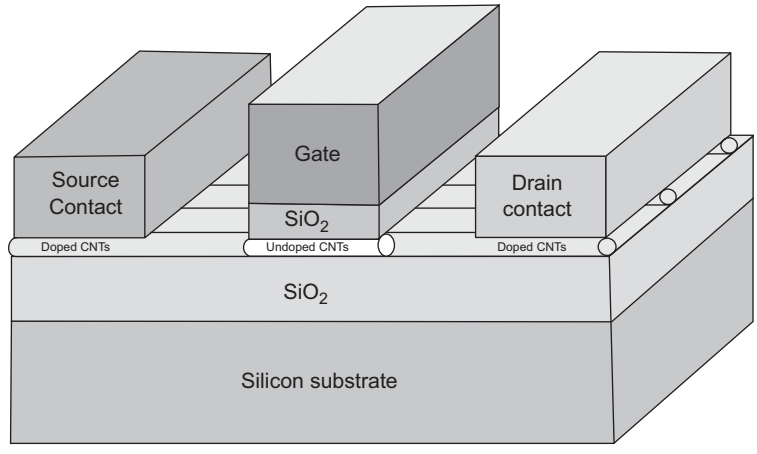


Fig. 4.37 *Construction of MOS CNTFET*

Operation As the source and drain regions are *N*-type doped and the channel is intrinsic, there is a potential barrier for the charge carriers from source to the drain region. This is the OFF state. Carriers (Electrons or holes) do not see any barrier in ON state. Gate electrostatics controls the top of the barrier similar to the conventional MOSFET as shown in Fig. 4.38. When positive voltage is applied to the gate, the barrier decreases and when the applied gate voltage is greater than the threshold voltage, a channel is formed between the source and drain.

Schottky Barrier Carbon Nanotube Field Effect Transistor

In SB-CNTFET, tunneling of electrons and holes from the potential barriers at the source and drain junctions constitutes the current. The barrier width is modulated by the application of gate voltage.

Construction The construction of an SB-CNTFET is shown in Fig. 4.39. The device is made up of intrinsic or lightly doped P-type silicon which forms the substrate of the device. The CNTFET is formed by one or more semiconducting CNTs perfectly aligned and well-positioned, whose section under the gate is intrinsic and the source and drain extension regions are metal contacts instead of doped CNTs.

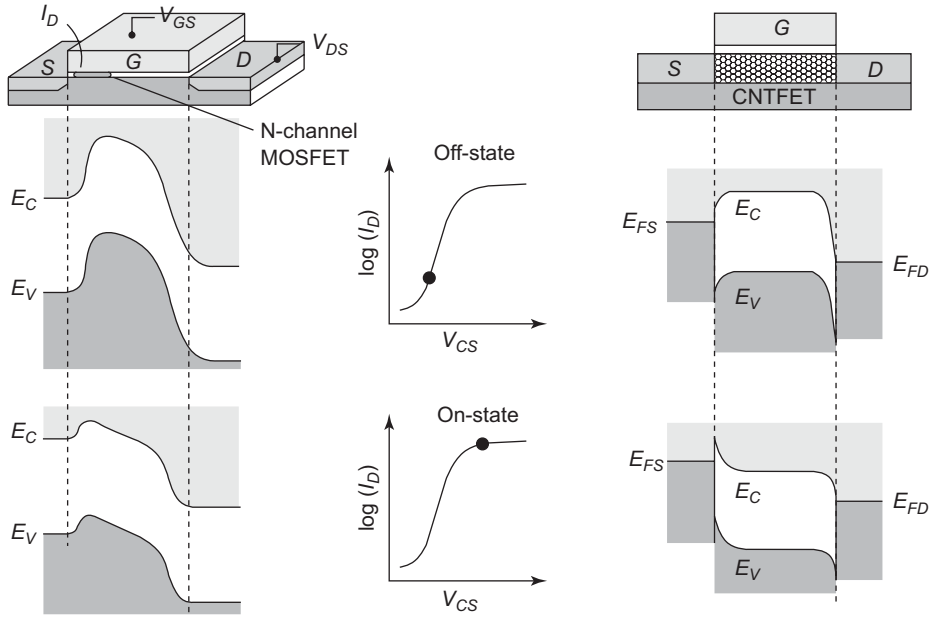


Fig. 4.38 *Operation of a CNTFET compared to Si-MOSFET using band diagrams*

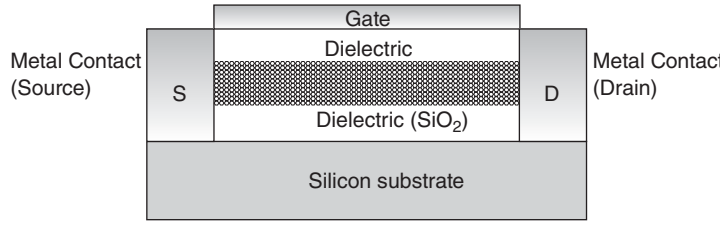


Fig. 4.39 *Construction of SB-CNTFET*

Operation Electrons see a tunneling barrier in ON state. Gate electrostatics will control the tunneling barrier. Gate electrostatics controls tunnelling barrier at the source channel interface as shown in Fig. 4.40. Since the current carrying mechanism is tunnelling, the ON current is limited.

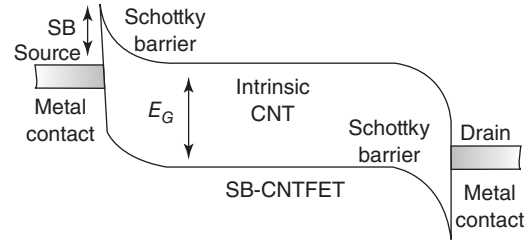


Fig. 4.40 *Band diagram of SB-CNTFET*

4.9.6 Comparison of V-I Characteristics of MOS-CNTFET and SB-CNTFET

The *V-I* characteristics of the MOS-CNTFET and SB-CNTFET are shown in Fig. 4.41 for the comparison of their performances. MOS type CNTFET has low OFF current compared to SB type CNTFET.

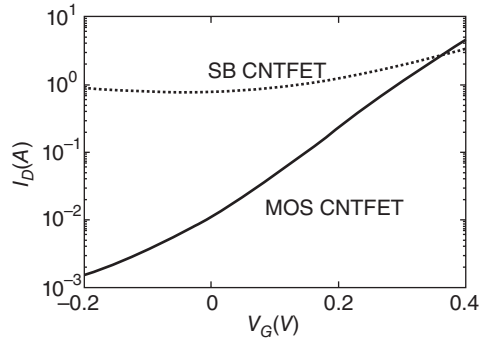


Fig. 4.41 *Comparison of MOS-CNTFET with SB-CNTFET*

4.10 ZENER DIODE [AU May 2017, Nov 2016, Nov 2014, May 2014, 16 marks]

When the reverse voltage reaches breakdown voltage in normal *PN* junction diode, the current through the junction and the power dissipated at the junction will be high. Such an operation is destructive and the diode gets damaged. Whereas diodes can be designed with adequate power dissipation capabilities to operate in the breakdown region. One such diode is known as Zener diode. Zener diode is heavily doped than the ordinary diode.

From the *V–I* characteristics of the Zener diode, shown in Fig. 4.42, it is found that the operation of Zener diode is same as that of ordinary *PN* diode under forward-biased condition. Whereas under reverse-biased condition, breakdown of the junction occurs. The breakdown voltage depends upon the amount of doping. If the diode is heavily doped, depletion layer will be thin and, consequently, breakdown occurs at lower reverse voltage and further, the breakdown voltage is sharp. Whereas a lightly doped diode has a higher breakdown voltage. Thus breakdown voltage can be selected with the amount of doping.

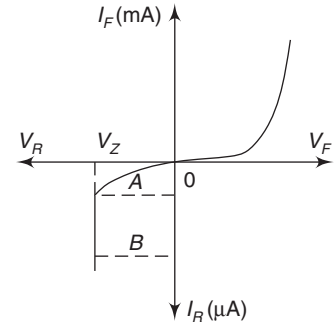


Fig. 4.42 *V–I characteristics of a Zener diode*

The sharp increasing current under breakdown conditions are due to the following two mechanisms.

(1) Avalanche breakdown
(2) Zener breakdown

4.10.1 Avalanche Breakdown

As the applied reverse bias increases, the field across the junction increases correspondingly. Thermally generated carriers while traversing the junction acquire a large amount of kinetic energy from this field. As a result the velocity of these carriers increases. These electrons disrupt covalent bond by colliding with immobile ions and create new electron-hole pairs. These new carriers again acquire sufficient

energy from the field and collide with other immobile ions thereby generating further electron–hole pairs. This process is cumulative in nature and results in generation of avalanche of charge carriers within a short time. This mechanism of carrier generation is known as Avalanche multiplication. This process results in flow of large amount of current at the same value of reverse bias.

4.10.2 Zener Breakdown

When the *P* and *N* regions are heavily doped, direct rupture of covalent bonds takes place because of the strong electric fields, at the junction of *PN* diode. The new electron–hole pairs so created increase the reverse current in a reverse biased *PN* diode. The increase in current takes place at a constant value of reverse bias typically below 6 V for heavily doped diodes. As a result of heavy doping of *P* and *N* regions, the depletion region width becomes very small and for an applied voltage of 6 V or less, the field across the depletion region becomes very high, of the order of 10^7 V/m, making conditions suitable for Zener breakdown. For lightly doped diodes, Zener breakdown voltage becomes high and breakdown is then predominantly by Avalanche multiplication. Though Zener breakdown occurs for lower breakdown voltage and Avalanche breakdown occurs for higher breakdown voltage, such diodes are normally called Zener diodes.

4.10.3 Zener Resistances and Zener Diode Ratings

Zener Resistances

Let us consider two resistances of the Zener diode (i) DC or static resistance and (ii) AC or dynamic resistance.

(a) Zener Static or DC Resistance R_Z It is the ratio of total Zener diode voltage to total diode current measured at the given operating point, i.e., $R_Z = \dfrac{V_{ZQ}}{I_{ZQ}}$. For more precise calculation, a Zener diode can be replated by an ideal battery in series with a small Zener resistance R_Z.

(b) Zener Dynamic or AC Resistance $r_{Z\,max}$ It is defined as voltage difference divided by current difference at the given operating point, i.e., $r_Z = \dfrac{\Delta V_Z}{\Delta I_Z}$. If the dynamic resistance is less, then the Zener diode will be better as a voltage regulator.

Zener Diode Ratings

The Zener diode will perform satisfactorily only if it is operated within certain limiting values. They are the following:

(a) Minimum Zener Current $I_{Z\,min}$ It is the minimum reverse current where the breakdown becomes stable. If a Zener diode has to remain in the breakdown region the current through it has to be more than $I_{Z\,min}$.

(b) Maximum Zener Current $I_{Z\,max}$ It is related to the power rating $P_{Z\,max}$ as

$$I_{Z\,max} = \frac{P_{Z\,max}}{V_Z}$$

where V_Z is the Zener voltage. This parameter gives the maximum current a Zener diode can handle without exceeding its power rating.

(c) Maximum Power of a Zener Diode P_Z The power dissipation of a Zener diode equals the product of its Zener voltage and current, i.e., $P_Z = V_Z \cdot I_Z$. As long as P_Z is less than the maximum power rating $P_{Z\,\max}$, the Zener diode can operate in the breakdown region without being destroyed.

4.10.4 Effect of Temperature on Zener Diode

The Zener voltage V_Z changes with the temperature. The percentage change in the Zener voltage V_Z for every °C change in temperature is called temperature coefficient (TC) of a Zener diode. It is denoted as TC and expressed as % / °C.

Mathematically, it can be defined as

$$TC = \frac{\Delta V_Z}{V_Z (T_1 - T_0)} \times [100]\,\% / °C$$

where T_1 is the final temperature of junction while T_0 is generally 25°C at which normal Zener voltage V_Z is specified. ΔV_Z is the resulting change in the Zener voltage due to the temperature variation.

A positive value of TC indicates that there is an increase in V_Z due to increases in temperature or decrease in V_Z due to decrease in temperature. The negative value of TC indicates that there is an increase in V_Z due to decrease in temperature or decrease in V_Z due to increase in temperature.

From the above equation, we have

$$\Delta V_Z = \frac{V_Z\, TC\, (T_1 - T_0)}{100} = \frac{V_Z\, TC\, \Delta T}{100}$$

where ΔT is the change in temperature.

Sometimes TC for a Zener diode is expressed in mV/°C and in such a case, the corresponding change in V_Z can be obtained as

$$\Delta V_Z = TC\,(T_1 - T_0) = TC \times \Delta T$$

For a Zener diode with V_Z less than 6 V, the temperature coefficient is negative and hence V_Z decreases as temperature increases. While for a Zener diode with V_Z greater than 6 V, the temperature coefficient is positive and hence V_Z increases as temperature increases.

4.10.5 Applications

From the Zener characteristics shown in Fig. 4.42, under the reverse bias condition, the voltage across the diode remains almost constant although the current through the diode increases as shown in region AB. Thus, the voltage across the Zener diode serves as a reference voltage. Hence, the diode can be used as a voltage regulator.

In Fig. 4.43, it is required to provide constant voltage across load resistance R_L, whereas the input voltage may be varying over a range. As shown, Zener diode is reverse biased and as long as the input voltage does not fall below V_Z (Zener breakdown voltage), the voltage across the diode will be constant and hence the load voltage will also be constant.

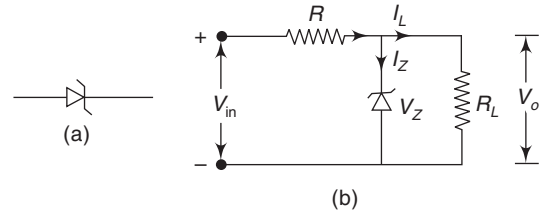


Fig. 4.43 *Zener diode: (a) Circuit symbol, (b) as a voltage regulator*

4.10.6 Voltage Regulators

In an unregulated power supply, the output voltage changes whenever the input voltage or load changes. An ideal regulated power supply is an electronic circuit designed to provide a predetermined DC voltage V_o which is independent of the load current and variations in the input voltage. A voltage regulator is an electronic circuit that provides a stable DC voltage independent of the load current, temperature and AC line voltage variations.

Factors determining the stability The output DC voltage V_o depends on the input unregulated DC voltage V_{in}, load current I_L and temperature T. Hence, the change in output voltage of power supply can be expressed as follows:

$$\Delta V_o = \frac{\partial V_o}{\partial V_{in}} \Delta V_{in} + \frac{\partial V_o}{\partial I_L} \Delta I_L + \frac{\partial V_o}{\partial T} \Delta T$$

or $$\Delta V_o = S_V \Delta V_{in} + R_o \Delta I_L + S_T \Delta T$$

where the three coefficients are defined as

Input regulation factor, $$S_V = \frac{\Delta V_o}{\Delta V_{in}} \mid \Delta I_L = 0; \Delta T = 0$$

Output resistance, $$R_o = \frac{\Delta V_o}{\Delta I_L} \mid \Delta V_{in} = 0; \Delta T = 0$$

Temperature coefficient, $$S_T = \frac{\Delta V_o}{\Delta T} \mid \Delta V_{in} = 0; \Delta I_L = 0$$

Smaller the value of the three coefficients, better the regulation of the power supply.

Line regulation Line regulation is defined as the change in output voltage for a change in line supply voltage, keeping the load current and temperature constant. Line regulation is given by

$$\text{Line regulation} = \frac{\text{change in output voltage}}{\text{change in input voltage}} = \frac{\Delta V_o}{\Delta V_{in}}$$

Load regulation Load regulation is defined as a change in regulated output voltage as the load current changes from no load to full load. It is expressed as a percentage of no load voltage or full load voltage.

$$\% \text{ Load regulation} = \frac{V_{\text{no load}} - V_{\text{full load}}}{V_{\text{no load}}} \times 100$$

or $$\% \text{ Load regulation} = \frac{V_{\text{no load}} - V_{\text{full load}}}{V_{\text{full load}}} \times 100$$

where $V_{\text{no load}}$ the output voltage at zero load current and $V_{\text{full load}}$ the output voltage at rated load current. This is usually denoted in percentage. The plot of the output voltage V_o versus the load current I_L for a full-wave rectifier is given in Fig. 4.44. The drop in the characteristics is a measure of the internal resistance of the power supply.

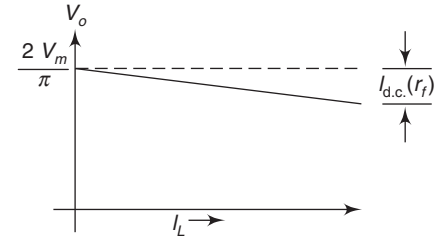


Fig. 4.44 *Load regulation characteristics*

Zener diode shunt regulator A Zener diode, under reverse bias breakdown condition, can be used to regulate the voltage across a load, irrespective of the supply voltage or load current variations. A simple Zener voltage regulator circuit is shown in Fig. 4.45. The Zener diode is selected with V_z equal

to the voltage desired across the load. The Zener diode has a characteristic that under reverse bias condition, the voltage across it practically remains constant, even if the current through it changes by a large extent. Under normal conditions, the input current $I_i = I_L + I_Z$ flows through resistor R. The input voltage V_i can be written as $V_i = I_iR + V_z = (I_L + I_z)\, R + V_z$.

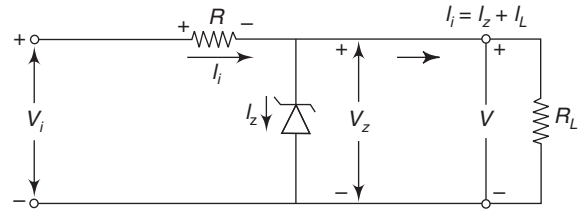


Fig. 4.45 *Zener voltage regulator*

When the input voltage V_i increases (say due to supply voltage variations), as the voltage across Zener diode remains constant, the drop across resistor R will increase with a corresponding increase in $I_L + I_Z$. As V_Z is a constant, the voltage across the load will also remain constant and hence, I_L will be a constant. Therefore, an increase in $I_L + I_Z$ will result in an increase in I_Z which will not alter the voltage across the load.

It must be ensured that the reverse voltage applied to the Zener diode never exceeds PIV of the diode and at the same time, the applied input voltage must be greater than the breakdown voltage of the Zener diode for its operation. The Zener diodes can be used as 'stand-alone' regulator circuits and also as reference voltage sources.

EXAMPLE 4.3

Design a Zener shunt voltage regulator with the following specifications: V_o = 10 V; V_{in} = 20–30 V; I_L = (30–50) mA; I_z = (20–40) mA.

Solution Refer to Fig. 4.45.

Selection of Zener diode

$$V_z = V_o = 10 \text{ V}$$

$$I_{z\,(\max)} = 40 \text{ mA}$$

$$P_z = V_z \times I_{z(\max)} = 10 \times 40 \times 10^{-3} = 0.4 \text{ W}$$

Hence a 0.5 WZ 10 Zener can be selected

Value of load resistance, R_L

$$R_{L\,(\min)} = \frac{V_o}{I_{L(\max)}} = \frac{10}{50 \times 10^{-3}} = 200\ \Omega$$

$$R_{L\,(\max)} = \frac{V_o}{I_{L(\min)}} = \frac{10}{30 \times 10^{-3}} = 333\ \Omega$$

Value of input resistance, R

$$R_{\max} = \frac{V_{\text{in(max)}} - V_o}{I_{L(\min)} + I_{z(\max)}}$$

$$= \frac{30 - 10}{(30 + 40) \times 10^{-3}} = 286\ \Omega$$

$$R_{\min} = \frac{V_{\text{in}(\min)} - V_o}{I_{L(\max)} + I_{z(\min)}}$$

$$= \frac{20 - 10}{(50 + 20) \times 10^{-3}} = 143\ \Omega$$

Therefore, $$R = \frac{R_{\max} + R_{\min}}{2} = 215\ \Omega$$

EXAMPLE 4.4

In a Zener regulator, the DC input is 10V ± 20%. The output requirements are 5 V, 20 mA. Assume $I_{z(\min)}$ and $I_{z(\max)}$ as 5 mA and 80 mA respectively. Design the Zener regulator.

Solution The minimum Zener current is $I_{z(\min)} = 5$ mA when the input voltage is minimum. Here the input voltage varies between 10 V ± 20%, i.e., 8 V and 12 V.

Therefore, the input voltage, $V_{i(\min)} = 8$ V

Given load current $I_L = 20$ mA and the voltage across the load, $V_0 = 5$ V.

Therefore,

$$R_L = \frac{V_0}{I_L} = \frac{5\ \text{V}}{20 \times 10^{-3}} = 250\ \Omega$$

Hence, the series resistance

$$R = \frac{V_{i(\max)} - V_0}{(I_{z(\min)} + I_L)}$$

$$= \frac{(8 - 5)}{(5 + 20) \times 10^{-3}} = 120\ \Omega$$

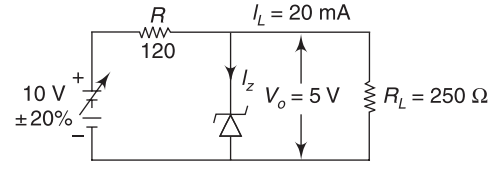


Fig. 4.46

The various values are given in the Zener regulator shown in Fig. 4.46.

EXAMPLE 4.5

If DC unregulated input is 20 V, V_0 = 10 V and load current is 0–20 mA, design the regulator. Assume for the Zener, $I_{z(\min)}$ = 10 mA and $I_{z(\max)}$ = 100 mA.

Solution Given input voltage, $V_i = 20$ V

Output voltage $V_0 = 10$ V

Load current varies from 0 to 20 mA

$$I_{z(\min)} = 10\ \text{mA},\ I_{z(\max)} = 100\ \text{mA}$$

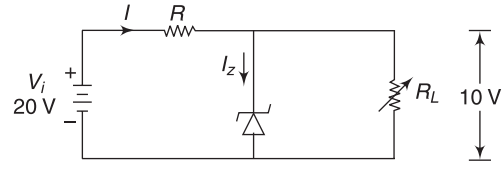


Fig. 4.47

Here, $V_z = V_0 = 10$ V (constant)

Applying KVL to a closed-loop circuit, $V_i = IR + V_z$

Hence, $20 = IR + 10$

or $IR = 10$

Therefore, $R = \frac{10}{I}\ \Omega$, where I is the loop current in amperes.

(i) For $I_z = I_{z\,(\min)} = 10$ mA and $I_L = 0$

The total current $I = I_L + I_z = 10$ mA

Therefore, $R_{\max} = \frac{V_0}{I_{z(\min)}} = \frac{10}{10 \times 10^{-3}} = 1000\ \Omega$

(ii) For $I_z = I_{z(\max)} = 100$ mA and $I_L = 20$ mA

$$I = I_L + I_{z(\max)} = 20 \times 10^{-3} + 100 \times 10^{-3} = 120 \text{ mA}$$

Therefore $R_{\min} = \frac{10}{120 \times 10^{-3}} = 83.33\ \Omega$

(iii) The range of R varies from 83.33 Ω to 1000 Ω.

EXAMPLE 4.6

Design a Zener voltage regulator to meet the following specifications: Output voltage = 5 V, Load current = 10 mA, Zener wattage = 400 mW and Input voltage = 10 V ± 2 V.

Solution Given $V_0 = 5$ V, $I_L = 10$ mA

Here, load resistance is $R_L = \frac{V_0}{I_L} = \frac{5}{10 \times 10^{-3}} = 500\ \Omega$

Maximum Zener current, $I_{z(\max)} = \frac{400 \text{ mW}}{5 \text{ V}} = 80$ mA

The minimum input voltage required will be when $I_z = 0$. Under this condition,

$$I = I_L = 10 \text{ mA}$$

Minimum input voltage, $V_{i(\min)} = V_o + IR$

Hence, $V_{i(\min)} = 10 - 2 = 8$ V

or $8 = 5 + (10 \times 10^{-3})\ R$

Therefore, $R_{\max} = \frac{3}{10 \times 10^{-3}} = 300\ \Omega$

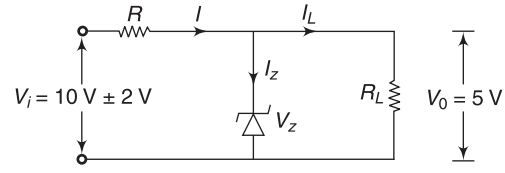


Fig. 4.48

Now, maximum input voltage, $V_{i(\max)} = V_o + [I_{z(\max)} + I_L]R = 5 + [(80 + 10)10^{-3}]R$

or $12 = 5 + (90 \times 10^{-3})R$

$$R_{\min} = \frac{7}{90 \times 10^{-3}} = 77.77\ \Omega$$

The value of R is chosen between 77.77 Ω and 300 Ω.

EXAMPLE 4.7

A 24 V, 600 mW Zener diode is used for providing a 24 V stabilized supply to a variable load. If the input voltage is 32 V, calculate (i) the value of series resistance required, and (ii) diode current when the load is 1200 Ω.

Solution Given $V_0 = 24$ V, $V_i = 32$ V, $P_Z = 600$ mW and $R_L = 1200\ \Omega$.

The load current, $$I_L = \frac{V_0}{R_L} = \frac{24}{1200} = 20 \text{ mA}$$

Max. Zener current, $$I_{z(\max)} = \frac{P_z}{V_o} = \frac{600 \times 10^{-3}}{24} = 25 \text{ mA}$$

$$R_{\max} = \frac{V_i - V_0}{I_{L(\min)} + I_{z(\max)}} = \frac{32 - 24}{(20 + 25) \times 10^{-3}} = \frac{8}{45 \times 10^{-3}} = \frac{1600}{9} = 177.78\ \Omega$$

EXAMPLE 4.8

A Zener voltage regulator circuit is to maintain constant voltage at 60 V, over a current range from 5 to 50 mA. The input supply voltage is 200 V. Determine the value of resistance *R* to be connected in the circuit, for voltage regulation from load current I_L = 0 mA to I_L max, the maximum possible value of I_L. What is the value of I_L max?

Solution Given $V_z = V_o = 60$ V and $V_{\text{in}} = 200$ V

(a) *To find the value of resistance* (*R*):

$$R = \frac{V_{\text{in}} - V_o}{I_{z(\max)} + I_{L(\min)}}$$

$$= \frac{200 - 60}{50 \times 10^{-3}} = \frac{140}{50 \times 10^{-3}} = 2.8 \text{ k}\Omega$$

(b) *To find the value of* $I_{L(\max)}$:

If $I_{z(\max)} = 50$ mA, $I_{L(\min)} = 0$ mA

If $I_{z(\min)} = 5$ mA, $I_{L(\max)} = 45$ mA

Therefore, $I_{L(\max)} = 45$ mA

EXAMPLE 4.9

For the Zener voltage regulation shown, determine the range of R_L and I_L that gives the stabilizer voltage of 10 V.

Solution From the circuit, $I = I_Z + I_L$

But from *R*, $$I = \frac{V_{\text{in}} - V_o}{R} = \frac{40 - 10}{1 \times 10^3} = 30 \text{ mA}$$

When I_L is minimum, I_Z is maximum and vice versa.

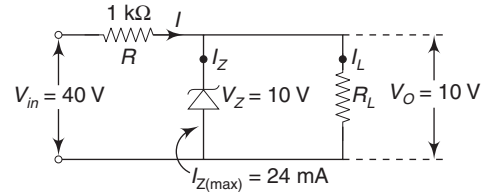


Fig. 4.49

$$I = I_{Z(\max)} + I_{L(\min)}$$

$$30 \text{ mA} = 24 \text{ mA} + I_{L(\min)}$$

Therefore, $$I_{L(\min)} = 6 \text{ mA}$$

But $$I_{L(\min)} = \frac{V_o}{R_{L(\max)}}$$

$$R_{L(\max)} = \frac{V_o}{I_{L(\min)}} = \frac{10}{6 \times 10^{-3}} = 1.667 \text{ k}\Omega$$

and $$I = I_{Z(\min)} + I_{L(\max)}$$

$$30 \text{ mA} = 5 \text{ mA} + I_{L(\max)}$$

$$I_{L(\max)} = 25 \text{ mA}$$

But $$I_{L(\max)} = \frac{V_o}{R_{L(\min)}}$$

$$R_{L(\min)} = \frac{V_o}{I_{L(\max)}} = \frac{10}{25 \times 10^{-3}} = 400 \ \Omega$$

Hence, the range of I_L is 6 mA to 25 mA and that of R_L is 400 Ω to 1.667 kΩ.

EXAMPLE 4.10

Determine the range of input voltage that maintains the output voltage of 10 V, for the regulator circuit shown.

Solution As V_o = 10 V constant and R_L = 10 kΩ constant, we have

$$I_L = \frac{V_o}{R_L} = \frac{10}{10 \times 10^3} = 1 \text{ mA}$$

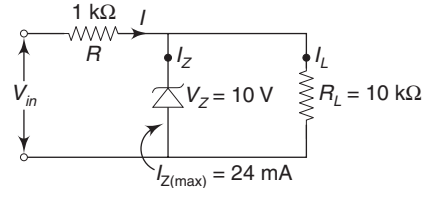


Fig. 4.50

When $V_{\text{in}} = V_{\text{in}(\max)}$, $I_Z = I_{Z(\max)}$

Now $I = I_Z + I_L$

Therefore, $I_{\max} = I_{Z(\max)} + I_L = 24 \text{ mA} + 1 \text{ mA} = 2 \text{ mA}.$

$$\frac{V_{\text{in}(\max)} - V_Z}{R} = 25 \text{ mA}$$

Therefore, $V_{\text{in}(\max)} - 10 = 1 \times 10^3 (25 \times 10^{-3})$

$$V_{\text{in}(\max)} = 35 \text{ V}.$$

When $V_{in} = V_{in(min)}, \quad I_Z = I_{Z(min)} = 5$ mA

Therefore, $I_{min} = I_{Z(min)} + I_L = 5$ mA + 1 mA = 6 mA

$$\frac{V_{in(min)} - V_z}{R} = 6 \times 10^{-3}$$

$$V_{in(min)} - 10 = 1 \times 10^3 (6 \times 10^{-3})$$

$$V_{in(min)} = 16 \text{ V}$$

Thus range of input voltage is 16 V to 35 V for which the output voltage will be of 10 V.

4.11 VARACTOR DIODE

[AU May 2016, Dec 2015, May 2015, Nov 2014, May 2014 and May 2013, 8 marks]

The varactor, also called a varicap, tuning or voltage variable capacitor diode, is a junction diode with a small impurity dose at its junction, which has the useful property that its junction or transition capacitance is easily varied electronically.

When any diode is reverse biased, a depletion region is formed, as seen in Fig. 4.51. The larger the reverse bias applied across the diode, the width of the depletion layer "W" becomes wider. Conversely, by decreasing the reverse bias voltage, the depletion region width "W" becomes narrower. This depletion region is devoid of majority carriers and acts like an insulator preventing conduction between the N and P regions of the diode, just like a dielectric, which separates the two plates of a capacitor. The varactor diode with its symbol is shown in Fig. 4.52(a).

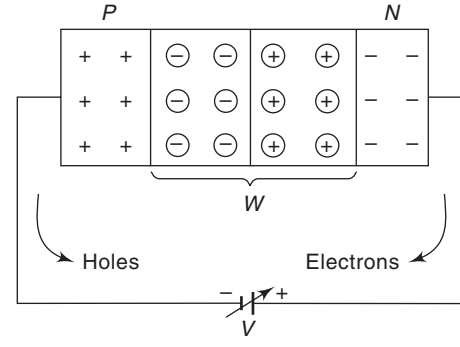


Fig. 4.51 *Depletion region in a reverse biased PN junction*

As the capacitance is inversely proportional to the distance between the plates ($C_T \propto 1/W$), the transition capacitance C_T varies inversely with the reverse voltage as shown in Fig. 4.52. Consequently, an increase in reverse bias voltage will result in an increase in the depletion region width and a subsequent decrease in transition capacitance C_T. At zero volt, the varactor depletion region W is small and the capacitance is large at approximately 600 pF. When the reverse bias voltage across the varactor is 15 V, the capacitance is 30 pF.

The varactor diodes are used in FM radio and TV receivers, AFC circuits, self adjusting bridge circuits and adjustable bandpass filters. With improvement in the type of materials used and construction, varactor diodes find application in tuning of LC resonant circuit in microwave frequency multipliers and in very low noise microwave parametric amplifiers.

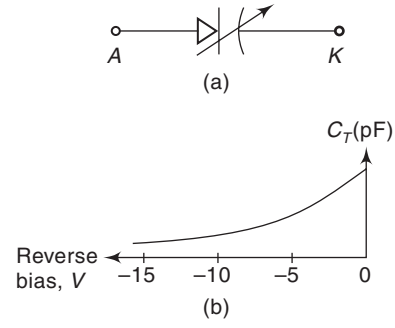


Fig. 4.52 *(a) Circuit symbol of varactor diode (b) Characteristics of varactor diode*

4.12 TUNNEL DIODE [AU Nov 2016, May 2016, May 2015, Nov 2013, May 2013, Nov 2010 and June 2010, 16 marks]

The Tunnel or Esaki diode is a thin-junction diode which exhibits negative resistance under low forward bias conditions.

An ordinary *PN* junction diode has an impurity concentration of about 1 part in 10^8. With this amount of doping, the width of the depletion layer is of the order of 5 microns. This potential barrier restrains the flow of carriers from the majority carrier side to the minority carrier side. If the concentration of impurity atoms is greatly increased to the level of 1 part in 10^3, the device characteristics are completely changed. The width of the junction barrier varies inversely as the square root of the impurity concentration and therefore, is reduced from 5 microns to less than 100 Å (10^{-8} m). This thickness is only about 1/50th of the wavelength of visible light. For such thin potential energy barriers, the electrons will penetrate through the junction rather than surmounting them. This quantum mechanical behavior is referred to as tunneling and hence, these high-impurity-density *PN* junction devices are called tunnel diodes.

The *V–I* characteristic for a typical germanium tunnel diode is shown in Fig. 4.53. It is seen that at first, forward current rises sharply as applied voltage is increased, where it would have risen slowly for an ordinary *PN* junction diode (which is shown as dashed line for comparison). Also, reverse current is much larger for comparable back bias than in other diodes due to the thinness of the junction. The

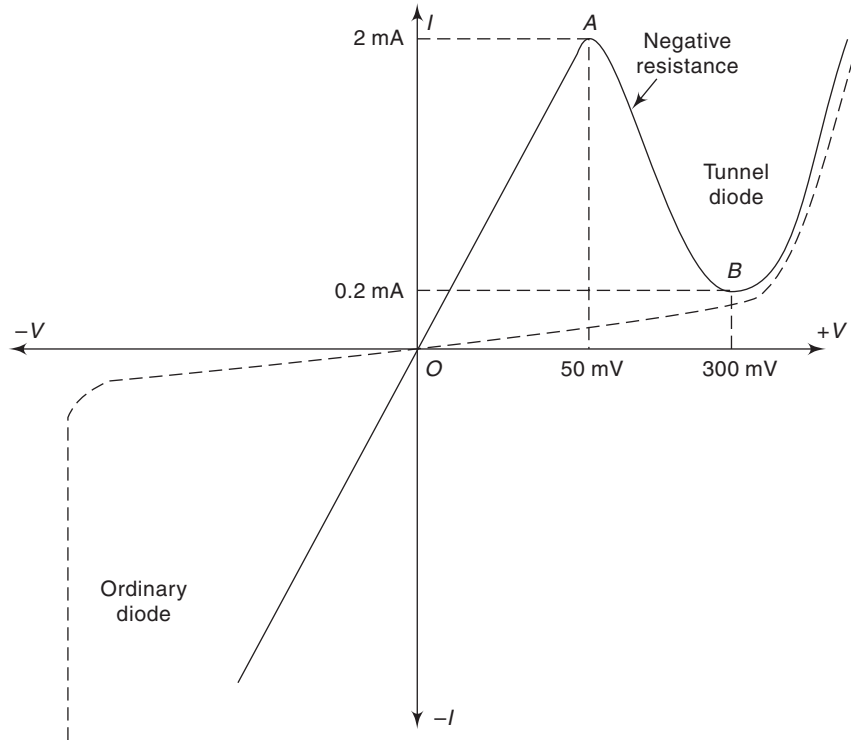


Fig. 4.53 *V–I characteristic of tunnel diode*

interesting portion of the characteristic starts at the point A on the curve, i.e., the peak voltage. As the forward bias is increased beyond this point, the forward current drops and continues to drop until point B is reached. This is the valley voltage. At B, the current starts to increase once again and does so very rapidly as bias is increased further. Beyond this point, characteristic resembles that of an ordinary diode. Apart from the peak voltage and valley voltage, the other two parameters normally used to specify the diode behaviour are the peak current and the peak-to-valley current ratio, which are 2 mA and 10 respectively, as shown.

The V–I characteristic of the tunnel diode illustrates that it exhibits dynamic resistance between A and B. Figure 4.54 shows energy level diagrams of the tunnel diode for three interesting bias levels. The shaded areas show the energy states occupied by electrons in the valence band, whereas the cross hatched regions represent energy states in the conduction band occupied by the electrons. The levels to which the energy states are occupied by electrons on either side of the junctions are shown by dotted lines. When the bias is zero, these lines are at the same height. Unless energy is imparted to the electrons from some external source, the energy possessed by the electrons on the N-side of the junction is insufficient to permit to climb over the junction barrier to reach the P-side. However, quantum mechanics show that there is a finite probability for the electrons to tunnel through the junction to reach the other side, provided there are allowed empty energy states in the P-side of the junction at the same energy level. Hence, the forward current is zero.

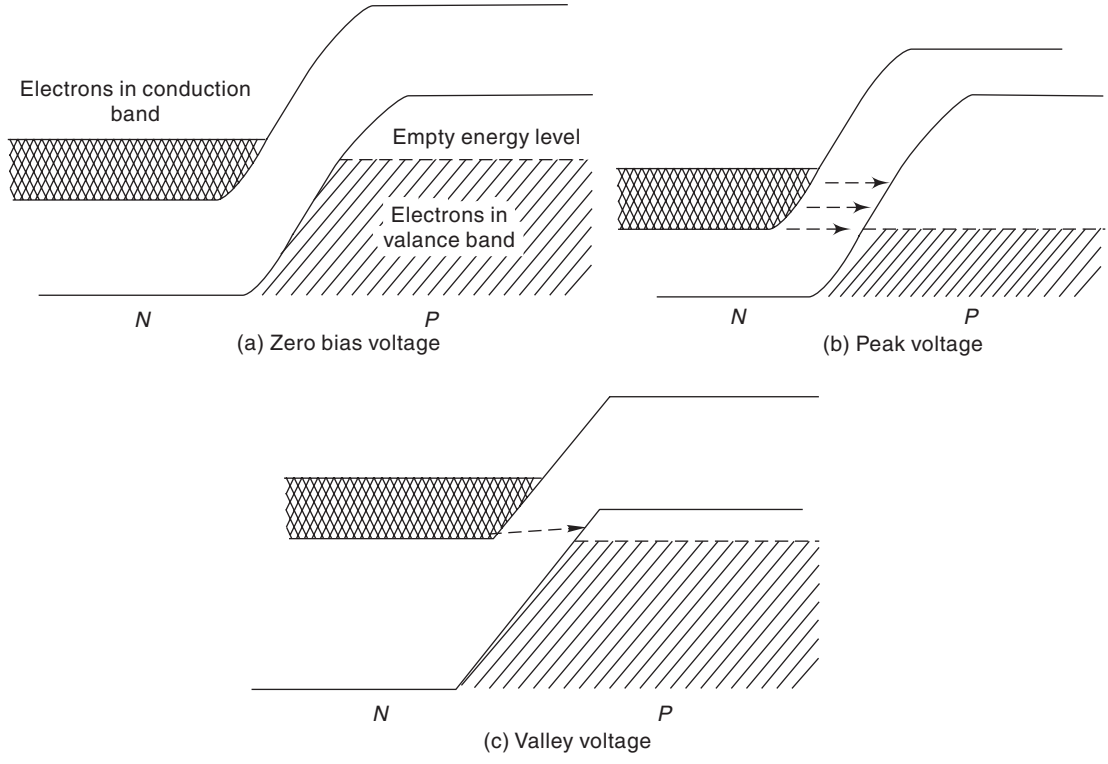


Fig. 4.54 *Energy level diagrams of tunnel diode*

When a small forward bias is applied to the junction, the energy level of the *P*-side is lower as compared with the *N*-side. As shown in Fig. 4.54(b), electrons in the conduction band of the *N*-side see empty energy level on the *P*-side. Hence, tunneling from *N*-side to *P*-side takes place. Tunneling in other directions is not possible because the valence band electrons on the *P*-side are now opposite to the forbidden energy gap on the *N*-side. The energy band diagram shown in Fig. 4.54(b), is for the peak of the diode characteristic.

When the forward bias is raised beyond this point, tunneling will decrease as shown in Fig. 4.54(c). The energy of the *P*-side is now depressed further, with the result that fewer conduction band electrons on the *N*-side are opposite to the unoccupied *P*-side energy levels. As the bias is raised, forward current drops. This corresponds to the negative resistance region of the diode characteristic. As forward bias is raised still further, tunneling stops altogether and it behaves as a normal *PN* junction diode.

4.12.1 Equivalent Circuit

The equivalent circuit of the tunnel diode when biased in the negative resistance region is as shown in Fig. 4.55(a). In the circuit, R_s is the series resistance and L_s is the series inductance which may be ignored except at highest frequencies. The resulting diode equivalent circuit is thus reduced to parallel combination of the junction capacitance C_j and the negative resistance $-R_n$. Typical values of the circuit components are $R_s = 6\ \Omega$, $L_s = 0.1$ nH, $C_j = 0.6$ pF and $R_n = 75\ \Omega$. The symbol of tunnel diode is shown in Fig. 4.55(b).

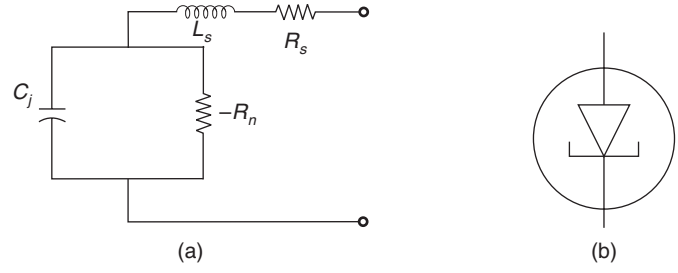


Fig. 4.55 *(a) Equivalent circuit of tunnel diode, (b) Symbol of tunnel diode*

Applications

1. Tunnel diode is used as an ultra-high speed switch with switching speed of the order of ns or ps
2. As logic memory storage device
3. As microwave oscillator
4. In relaxation oscillator circuit
5. As an amplifier

Advantages

1. Low noise
2. Ease of operation
3. High speed
4. Low power

Disadvantages

1. Voltage range over which it can be operated is less than 1 V.
2. Being a two terminal device, there is no isolation between the input and output circuit.

4.13 LASER DIODE

[AU Nov 2016, May 2016, Dec 2015, May 2014 and Nov 2011, 8 marks]

Similar to LED, Lasers are used to convert the electrical signal to light signal. In direct band gap materials where high recombination velocities exist, optical gain can be achieved by creating population inversion of carriers through high-level current injection and by forming a resonant cavity. This cavity is usually produced by the high Fresnel reflectivity obtained from cleaving the material along faces perpendicular to the junction plane.

The structure and characteristics of a typical laser diode is shown in Fig. 4.56. In this diode, opposite ends of the junction are polished to get mirror like surfaces. When free electrons recombine with holes, the emitted photons reflect back and forth between the mirror surfaces. The region between the mirrored ends acts like a cavity that filters the light and purifies its colour. As the photons bounce back and forth, they induce an Avalanche effect that causes all newly created photons to be emitted with the same phase. One of the mirror surfaces is semitransparent. From this surface a fine thread like beam of photons emerge out. All the photons of laser light have same frequency and phase and hence coherent.

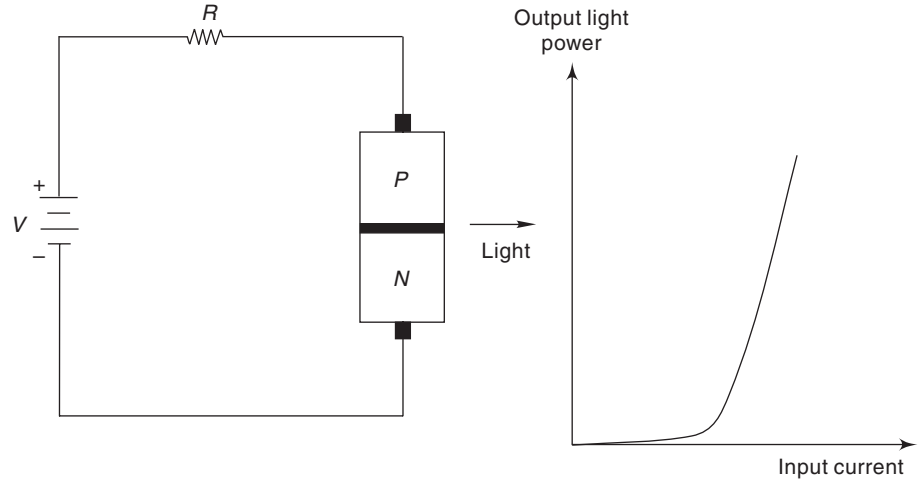


Fig. 4.56 *Structure and characteristics of laser diode*

It has a well defined current threshold as seen from the power output vs. drive current characteristic. Below this threshold the device exhibits low levels of spontaneous emission. At the limiting current density, stimulated emission occurs and the emitted radiation increases linearly with drive current.

4.14 LIGHT DEPENDENT RESISTOR (LDR)

[AU Nov 2016, Nov 2013, Nov 2012, May 2012, Nov 2011 and June 2011, 8 marks]

The bulk type photoresistor, *photoconductive cell* (*PC*) or photodetector is a two terminal device which is used as a Light Dependent Resistor (LDR). It is made of a thin layer of semiconductor material such as cadmium sulphide (CdS), lead sulphide (PbS), or cadmium selenide (CdSe) whose spectral responses are shown in Fig. 4.57. The photoconducting device with the widest applications is the CdS cell, because it has high dissipation capability, with excellent sensitivity in the visible spectrum and low resistance when stimulated by light. The main drawback of CdS cell is its slower speed of response. PbS has the fastest speed of response.

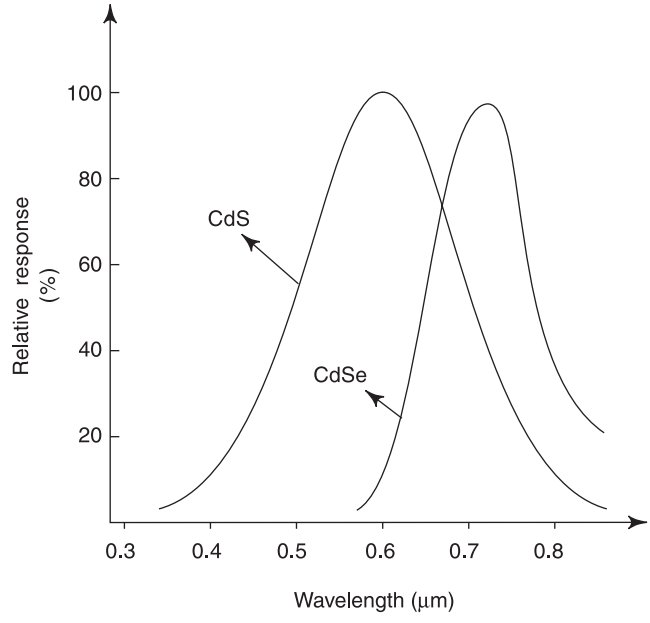


Fig. 4.57 *Spectral responses of CdS and CdSe*

The illumination characteristics of photoconductive detectors are shown in Fig. 4.58(a). It exhibits the peculiar property that its resistance decreases in the presence of light and increases in the absence of light. The cell simply acts as a conductor whose resistance changes when illuminated. In absolute darkness, the resistance is as high as 2 MΩ and in strong light, the resistance is less than 10 Ω.

A simple circuit for a photoconductive detector is shown in Fig. 4.58(b). The semiconductor layer is enclosed in a sealed housing. A glass window in the housing permits light to fall on the active material of the cell. Here, the resistance of the photoconductive detector, in series with R, limits the amount of current I in the circuit. The ammeter A is used to measure the current I. When no light falls on the cell, its resistance is very high and the current I is low. Hence, the voltage drop V_o across R is relatively low. When the cell is illuminated, its resistance becomes very low. Hence, current I increases and voltage V_o increases. Thus, this simple circuit arrangement with slight modification can be used in control circuits to control the current.

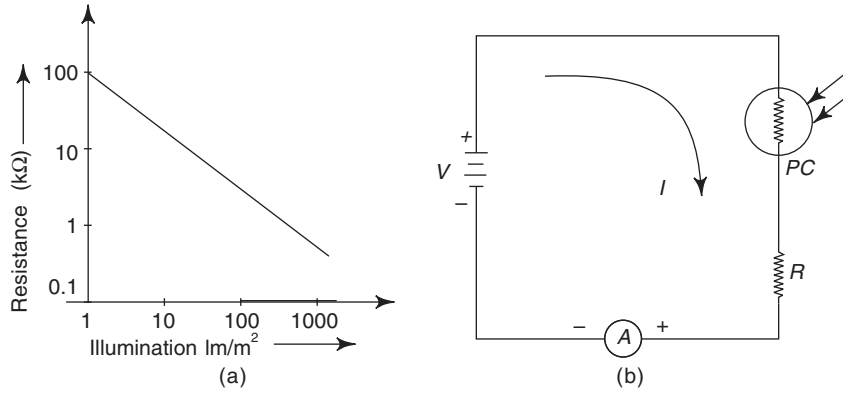


Fig. 4.58 *(a) Illumination characteristics of the photoconductive detector, (b) Photoconductive detector connected in a simple circuit*

Applications The detector is used either as an ON/OFF device to detect the presence or absence of a light source which is used for automatic street lighting or some intermediate resistance value can be used as a trigger level to control relays and motors. Further, it is used to measure a fixed amount of illumination and to record a modulating light intensity.

It is used in counting systems where the objects on a conveyor belt interrupt a light beam to produce a series of pulses which a counter.

It is used in twilight switching circuits. When the day light has faded to a given level, the corresponding resistance of the detector causes another circuit to switch ON the required lights.

It is widely used in cameras to control shutter opening during the flash. Twin photoconductive cells mounted in the same package have been used in optical bridge circuits for position control mechanisms and dual-channel remote volume control circuits.

Two Mark Questions and Answers

1. **Mention a few applications of varactor diode.** **(April/May 2017)**

   Varactor diode finds application in FM radio and TV receivers, AFC circuits, self adjusting bridge circuits, adjustable bandpass filters, tuning of LC resonant circuit in microwave frequency multipliers and low-noise microwave parametric amplifiers.

2. **Mention some advantages and disadvantages of Tunnel Diode.** **(May/June 2016)**

   Advantages:

   (i) Low noise
   (ii) Ease of operation
   (iii) High speed and
   (iv) Low power

Disadvantages:

(i) Voltage range over which it can be operated is less than 1 V.

(ii) Being a two terminal device, there is no isolation between the input and output circuit.

3. Compare Schottky diode and conventional diode. **(Nov/Dec 2016)**

(i) The current in a conventional PN junction diode is controlled by the diffusion of minority carriers whereas the current in the Schottky diode results from the flow of majority carriers over the potential barrier at the metal-semiconductor junction.

(ii) The reverse saturation current for a Schottky diode is larger than that of a PN junction diode.

(iii) The Schottky diode has a smaller turn-on voltage and shorter switching time than the PN junction diode.

4. What is a metal-semiconductor contact? **(Nov/Dec 2016)**

A metal-semiconductor contact is a junction contact of a metal with a piece of semiconductor. The active junction is the interface between the metal, which acts as an anode, and the semiconductor. The other interface between the semiconductor and the metal, which acts as a cathode, is an Ohmic contact and there is no potential drop at this junction.

5. Draw the energy band diagram of metal and semiconductor before and after conduction is made. **(May/June 2014)**

Refer to Fig. 4.2.

6. What is a MESFET?

MESFET is a modified form of MOSFET in which GaAs epitaxial layer was grown on semi-insulating GaAs substrate. Its operation is similar to *PN* junction JFET. The use of GaAs rather than silicon in MESFET provides two more significant advantages: first, the electron mobility at room temperature is more than 5 times larger, while the peak electron velocity is about twice that of silicon. Second, it is possible to fabricate semi-insulating (SI) GaAs substrates, which eliminates the problem of absorbing microwave power in the substrate due to free carrier absorption and provides better operation at high temperature.

7. Define avalanche breakdown? **(May/June 2010)**

Under reverse bias condition, when the reverse voltage approaches a particular value called breakdown voltage, the reverse current suddenly increases because the velocity of thermally generated minority carriers increases and they cross the depletion region and acquire sufficient kinetic energy from the applied potential to produce new charge carriers by removing valence electrons from their bonds. These new carriers will in turn collide with other atoms and will increases the number of electron-holes available for conduction. Because of the cumulative increase in carrier density after each collision, this process is called *avalanche breakdown*.

8. What is Zener breakdown? **(Nov/Dec 2011)**

When the P and N regions are heavily doped, direct rupture of covalent bonds takes place because of the strong electric field at the junction of PN diode. The new electron-hole pair, so created,

increases the reverse current in a reverse biased PN diode. As a result of heavy doping of P and N regions, the depletion width becomes very small and the field across the depletion region becomes very high and it is due to ruptures of the covalent bond. This breakdown is known as *Zener breakdown*.

9. What is meant by Zener effect? (May/June 2012)

Zener effect is a type of electrical breakdown in a reverse biased *PN* diode in which the electric field enables movement of electrons from the valence to the conduction band of a semiconductor, leading to a large number of free minority carriers, which suddenly increase the reverse current. Zener breakdown is employed in a Zener diode.

10. Write any two applications of Zener diode. (May/June 2014)

Zener diode is used in

(a) voltage regulator,
(b) voltage clipper circuits, and
(c) controlling the output amplitude.

11. What are the differences between a tunnel diode and an ordinary *PN* junction diode? (Nov/Dec 2014)

Unlike an ordinary *PN* junction diode, the tunnel diode is heavily doped which results in a very thin depletion region in terms of few Angstrom. The high electric field between the *P* and *N* regions causes tunneling of charge carriers resulting in a current flow with a very small bias voltage. Further, the tunnel diode exhibits negative resistance characteristics wherein the conduction current decreases with increase in the applied voltage.

12. What is tunneling phenomenon? (May/June 2012, Nov/Dec 2012 and Nov/Dec 2011)

According to the classical laws of physics, a charged particle should possess energy at least equal to the energy barrier in order to cross an energy barrier. Hence, the particle will cross the energy barrier if its energy is greater than the barrier and cannot cross the barrier if its energy is less than the energy barrier. But quantum mechanically, there exists non-zero probability that the particle with energy less than the energy barrier will cross the barrier as if it tunnels across the barrier. This is called *tunneling effect.*

13. List out the applications of tunnel diode. (May/June 2014 and June 2010)

Tunnel diode is used:

(i) as an ultra-high speed switch with switching speed of the order of *ns* or *ps*,
(ii) as a logic memory storage device,
(iii) as a microwave oscillators,
(iv) in relaxation oscillator circuit,
(v) as an amplifier,
(vi) microwave and RF power monitors, and
(vii) high-frequency triggers.

14. Expand: LASER, LDR. **(April/May 2015)**

LASER: Light Amplification by Stimulated Emission of Radiation

LDR: Light Dependent Resistor

15. Mention the analog and digital applications of LDR. **(Nov/Dec 2014)**

LDRs have low cost and simple structure. They are often used as light sensors. They are used when there is a need to detect the absence or presence of light as in a camera light meter. They are also used in street lamps, alarm clock, burglar alarm circuits, light intensity meters, for counting the packages moving on a conveyor belt, etc.

16. Draw the symbol for DUAL GATE MOSFET. **(Nov/Dec 2014)**

Refer to Fig. 4.21.

Review Questions

1. Draw the structure of a metal–semiconductor junction and explain the energy band structure before and after contact.
2. Derive an expression for built-in potential in a metal and *N*-type semiconductor junction in terms of barrier height and carrier concentration.
3. Derive an expression for built-in potential in a metal and *P*-type semiconductor junction in terms of barrier height and carrier concentration.
4. What are the advantages of GaAs devices over Si devices?
5. Explain the energy band structure of GaAs.
6. What are the salient features of GaAs technology?
7. Write short notes on MESFET.
8. Explain in detail the fabrication of GaAs depletion mode MESFET.
9. Mention the advantages and disadvantages of MESFET.
10. List the applications of MESFET.
11. Why is P-channel GaAs MESFET not preferred for high speed applications?
12. Explain the construction and working characteristics of N-channel depletion MESFET with neat diagram.
13. Distinguish between MOSFET and MESFET.
14. Discuss the different type of GaAs devices.
15. What is a Schottky diode? Explain the flow of carriers across its junction during forward and reverse biased conditions with energy band diagrams.
16. Briefly explain how the construction of a Schottky diode favours its use in the high frequency region.
17. How does the dual-gate MOSFET provide reduced input capacitance?
18. Explain the construction and operation of dual-gate MOSFET.
19. Give the applications of dual-gate MOSFET.
20. Describe the physical construction, operation and applications of FinFET.
21. Describe the construction of PIN-FET device with its cross-sectional diagram.
22. Write the applications of PIN-FET device.
23. Mention the two types of CNTFET.
24. Differentiate between CNTFET and Si-MOSFET.

25. Give the advantages and applications of CNTFET.
26. Explain the construction and working of MOS-CNTFET and SB-CNTFET.
27. Explain Avalanche breakdown and Zener breakdown.
28. Draw the *V–I* characteristic of Zener diode and explain its operation.
29. Show that the Zener diode can be used as a voltage regulator.
30. Define line regulation and load regulation in a voltage regulator.
31. How does a Zener diode maintain constant output voltage?
32. Design a Zener shunt regulation with the following specifications: V_o = 15 V, V_{in} = 20–25 V, I_L = 25–50 mA, I_Z = 20–45 mA.
33. A Zener diode shunt regulator circuit is to be designed to maintain a constant load current of 400 mA and voltage of 40 V. The input voltage is 90 ± 5 V. The Zener diode voltage is 40 V and its dynamic resistance is 2.5 Ω. Find the following quantities for the regulator: (a) the series dropping resistance, (b) Zener power dissipation, and (c) load resistance. Assume the Zener current to be 10% of load current. [*Ans.* 112.5 Ω, 3.6 Ω, 100 Ω]
34. A Zener diode voltage regulator shown in Fig. 4.44 has the following specifications:
    V_z = 15 V, $I_{z(min)}$ = 2 μA, P_z = 120 W, R_z = 40 Ω, R_L = 5 kΩ and V_{in} = 18–24 V.
    Determine the minimum and maximum value of series dropping resistance *R*. [*Ans.* 158 Ω, 1 kΩ]
35. Design a Zener regulator for the following specifications: output voltage, V_o = 5 V, Load current, I_L = 20 mA, Input voltage, V_i = 12 V ± 3V, Zener wattage, P_z = 500 mW. [*Ans.* R_L = 250 Ω, *R* = 83.33 Ω to 200 Ω]
36. Explain the principle behind the varactor diode and list out its applications.
37. What is tunneling?
38. From the energy band diagram explain the *V–I* characteristic of a tunnel diode.
39. Draw the equivalent circuit of a tunnel diode and explain it.
40. List out the applications of tunnel diode and mention its advantages and disadvantages.
41. Explain the principle behind the laser diode with a neat sketch.
42. Describe the principle involved in the bulk type photoconductive cell (LDR).
43. Describe the applications of LDR.
44. How does the dual-gate MOSFET provide reduced input capacitance?
45. Explain the construction and operation of dual-gate MOSFET.
46. Give the applications of dual-gate MOSFET.
47. Describe the physical construction, operation and applications of FinFET.

POWER DEVICES AND DISPLAY DEVICES

5.1 INTRODUCTION

In general, a thyristor is a semiconductor device having three or more junctions. Such a device acts as a switch without any bias and can be fabricated to have voltage ratings of several hundred volts and current ratings from a few amperes to almost thousand amperes. A family of thyristors consisting of *PNPN* diode (Shockley diode), SCR, LASCR, TRIAC and DIAC is dealt within this chapter. Also, UJT sawtooth generator, Power BJT, Power MOSFET, DMOS, VMOS, LED, LCD, Photodiode and Phototransistor, Optocoupler, Solar Cell and CCD are discussed in this chapter.

5.2 UJT (UNIJUNCTION TRANSISTOR) RELAXATION OSCILLATOR

[AU May 2016, Nov 2014, May 2014, Nov 2013, May 2013, Nov 2012, May 2012 and June 2010, 16 marks]

UJT is a three terminal semiconductor switching device. As it has only one *PN* junction and three leads, it is commonly called as Unijunction transistor.

The basic structure of UJT is shown in Fig. 5.1(a). It consists of a lightly doped *N*-type Silicon bar with a heavily doped *P*-type material alloyed to its one side closer to B_2 for producing single *PN* junction. The circuit symbol of UJT is shown in Fig. 5.1(b). Here, the emitter leg is drawn at an angle to the vertical and the arrow indicates the direction of the conventional current.

Characteristics of UJT Referring to Fig. 5.1(c), the interbase resistance between B_2 and B_1 of the silicon bar is $R_{BB} = R_{B1} + R_{B2}$. With emitter terminal open, if voltage V_{BB} is applied between the two bases, a voltage gradient is established along the *N*-type bar. The voltage drop across R_{B1} is given by $V_1 = \eta V_{BB}$, where the *intrinsic stand-off ratio* $\eta = R_{B1}/(R_{B1} + R_{B2})$. The typical value of η ranges from 0.56 to 0.75. This voltage V_1 reverse biases the *PN* junction and emitter current is cut-off. But a small leakage current flows from B_2 to emitter due to minority carriers. If a positive voltage V_E is applied to the emitter, the *PN* junction will remain reverse biased so long as V_E is less than V_1. If V_E exceeds V_1 by the cut-in voltage V_γ, the diode becomes forward biased. Under this condition, holes are injected into *N*-type bar. These holes are repelled by the terminal B_2 and are attracted by the terminal B_1. Accumulation of holes in E to B_1 region reduces the resistance in this section and hence emitter current I_E is increased and is limited by V_E. The device is now in the 'ON' state.

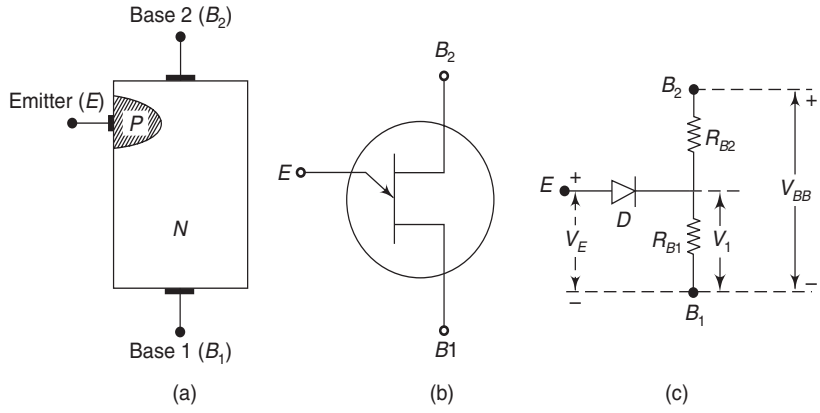


Fig. 5.1 *UJT: (a) Basic structure (b) Circuit symbol (c) Equivalent circuit*

If a negative voltage is applied to the emitter, *PN* junction remains reverse biased and the emitter current is cut-off. The device is now in the 'OFF' state.

Figure 5.2 shows a family of input characteristics of UJT. Here, upto the peak point *P*, the diode is reverse biased and hence, the region to the left of the peak point is called *cut-off region*. The UJT has a stable firing voltage V_P which depends linearly on V_{BB} and a small firing current I_P ($\approx 25\ \mu A$). At *P*, the peak voltage $V_P = \eta V_{BB} + V_\gamma$, the diode starts conducting and holes are injected into *N*-layer. Hence, resistance decreases thereby decreasing V_E for the increase in I_E. So, there is a *negative resistance region* from peak point *P* to valley point *V*. After the valley point, the device is driven into saturation and behaves like a conventional forward biased *PN* junction diode. The region to the right of the valley point is called saturation region. In the valley point, the resistance changes from negative to positive. The resistance remains positive in the saturation region. For very large I_E, the characteristic asymptotically approaches the curve for $I_{B2} = 0$.

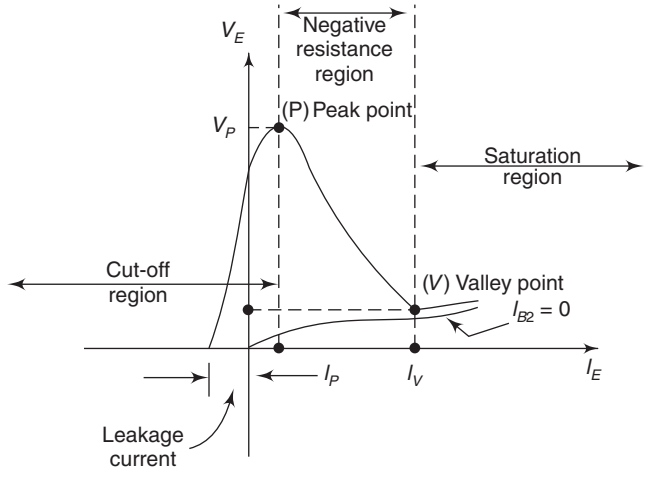


Fig. 5.2 *Input characteristics of UJT*

A unique characteristic of UJT is, when it is triggered, the emitter current increases regeneratively until it is limited by emitter power supply. Due to this negative resistance property, a UJT can be employed in a variety of applications, viz. sawtooth wave generator, pulse generator, switching, timing and phase control circuits.

UJT relaxation oscillator The relaxation oscillator using UJT which is meant for generating sawtooth waveform is shown in Fig. 5.3. It consists of a UJT and a capacitor C_E which is charged through R_E as the supply voltage V_{BB} is switched ON.

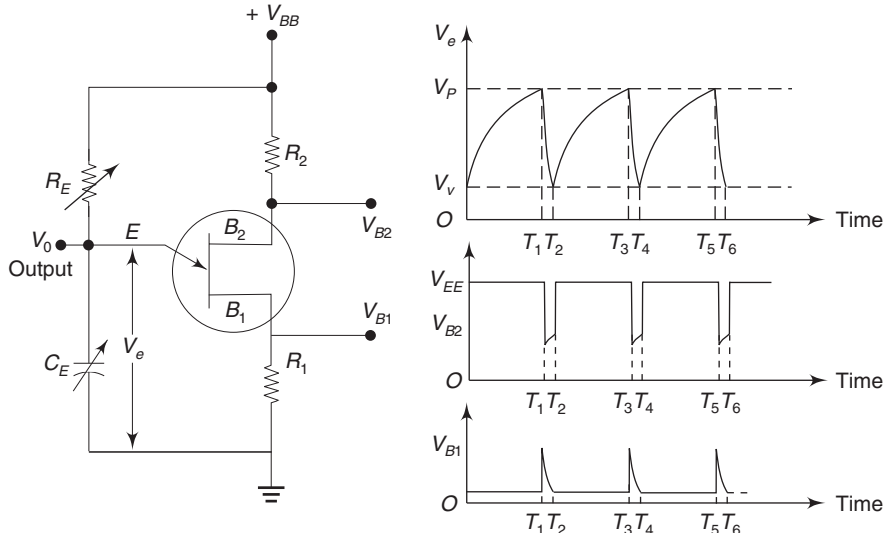


Fig. 5.3 *UJT Relaxation oscillator*

The voltage across the capacitor increases exponentially and when the capacitor voltage reaches the peak point voltage V_P, the UJT starts conducting and the capacitor voltage is discharged rapidly through EB_1 and R_1. After the peak point voltage of UJT is reached, it provides negative resistance to the discharge path which is useful in the working of the relaxation oscillator. As the capacitor voltage reaches zero, the device then cuts off and capacitor C_E starts to charge again. This cycle is repeated continuously generating a *sawtooth waveform* across C_E.

The inclusion of external resistors R_2 and R_1 in series with B_2 and B_1 provides spike waveforms. When the UJT fires, the sudden surge of current through B_1 causes drop across R_1, which provides positive going spikes. Also, at the time of firing, fall of V_{EBI} causes I_2 to increase rapidly which generates negative going spikes across R_2.

By changing the values of capacitance C_E or resistance R_E, frequency of the output waveform can be changed as desired, since these values control the time constant $R_E C_E$ of the capacitor charging circuit.

Frequency of oscillation The time period and hence the frequency of the sawtooth wave can be calculated as follows. Assuming that the capacitor is initially uncharged, the voltage V_C across the capacitor prior to breakdown is given by

$$V_C = V_{BB}(1 - e^{-t/R_E C_E})$$

where $R_E C_E$ = charging time constant of resistor-capacitor circuit, and t = time from the commencement of the waveform.

The discharge of the capacitor occurs when V_C is equal to the peak-point voltage V_P, i.e.,

$$V_P = \eta V_{BB} = V_{BB}(1 - e^{-t/R_E C_E})$$

$$\eta = 1 - e^{-t/R_E C_E}$$

$$e^{-t/R_E C_E} = (1 - \eta)$$

Therefore,

$$t = R_E C_E \log_e \frac{1}{(1-\eta)}$$

$$= 2.303\, R_E C_E \log_{10} \frac{1}{(1-\eta)}$$

If the discharge time of the capacitor is neglected, then $t = T$, the period of the wave.
Therefore, frequency of oscillation of sawtooth wave,

$$f_o = \frac{1}{T} = \frac{1}{2.303\, R_E C_E \log_{10} \frac{1}{(1-\eta)}}$$

EXAMPLE 5.1

Design a UJT relaxation oscillator to generate a sawtooth waveform at a frequency of 500 Hz. Assume the supply voltage V_{BB} = 20 V, V_P = 2.9 V, V_V = 1.118 V, I_P = 1.6 mA and I_V = 3.5 mA.

Solution We know that

$$f_o = \frac{1}{2.303\, R_E C_E \log_{10} \frac{1}{(1-\eta)}}$$

We know that $\eta_{\min} = 0.56$
For determining R_E, we have

$$R_E < \frac{V_{BB} - V_P}{I_P}, \text{ i.e., } R_E < \frac{20 - 2.9}{1.6 \times 10^{-3}} = 10.7\ \text{k}\Omega$$

$$R_E > \frac{V_{BB} - V_V}{I_V}, \text{ i.e., } R_E > \frac{20 - 1.118}{3.5 \times 10^{-3}} = 5.36\ \text{k}\Omega$$

Therefore, R_E is selected as 10 kΩ.

$$\frac{1}{500} = 2.303 \times 10 \times 10^3\, C_E \log_{10}\left(\frac{1}{(1-0.56)}\right)$$

Therefore,

$$C_E = \frac{1}{500 \times 2.303 \times 10^4 \times 0.36} = 0.24\ \mu\text{F}$$

So, C_E is selected as 0.22 μF.

Let the required pulse voltage at B_1 = 5 V

Let the peak pulse current, I_E = 250 mA.

Therefore,

$$R_1 = \frac{V_{R1}}{I_E} = \frac{5}{250 \times 10^{-3}} = 20\ \Omega$$

So, R_1 is selected to be 22 Ω.

We select the voltage characteristics for $V_{B1B2} = 4$ V.

Therefore, $V_{R2} = 20 - (4 + 5) = 11$ V

$$R_2 = \frac{11}{250} \times 10^3 = 44\ \Omega$$

So, R_2 is selected as 100 Ω.

EXAMPLE 5.2

A UJT has a firing potential of 20 V. It is connected across the capacitor of a series *RC* circuit with *R* = 100 kΩ and *C* = 1000 pF supplied by a source of 40 V DC. Calculate the time period of the sawtooth waveform generated.

Solution Given, $R_E = 100$ kΩ and $C_E = 1000$ pF

$$V_{BB} = 40 \text{ V}, V_P = 20 \text{ V}$$

$$V_P = V_{BB}\left(1 - e^{\frac{-t}{R_E C_E}}\right)$$

$$20 = 40\left(1 - e^{\frac{-t}{R_E C_E}}\right)$$

$$e^{-t/R_E C_E} = \frac{1}{2}$$

$$\frac{-t}{R_E C_E} = \ln \frac{1}{2}$$

$$\frac{t}{R_E C_E} = \ln (2)$$

$$t = \ln (2) \times R_E C_E$$

Therefore, $t = 0.693 \times 100 \times 10^3 \times 1000 \times 10^{-12} = 0.693 \times 10^{-4} = 69.3\ \mu s$

5.3 PNPN DIODE (SHOCKLEY DIODE)

As shown in Fig. 5.4, it is a four layer *PNPN* silicon device with two terminals. When an external voltage is applied to the device in such a way that anode is positive with respect to cathode, junctions J_1 and J_3 are forward biased and J_2 is reverse biased. Then the applied voltage appears across the reverse biased junction J_2. Now the current flowing through the device is only reverse saturation current.

However, as this applied voltage is increased, the current increases slowly until the so called firing or breakover voltage (V_{BO}) is reached. Once firing takes place, the current increases abruptly and the voltage drop across the device decreases sharply. At this point, the diode switches over from 'OFF' to 'ON'state. Once the device is fired into conduction, a minimum amount of current known as *holding current*, I_H, is required to flow to keep the device in ON state. To turn the device OFF

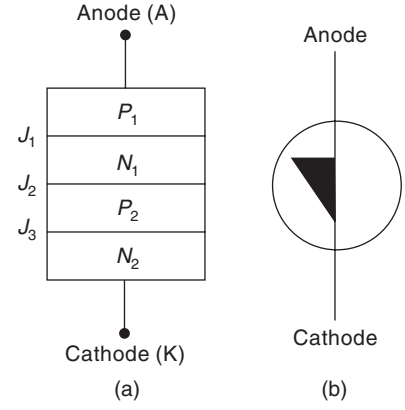


Fig. 5.4 *PNPN diode: (a) Basic structure (b) Circuit symbol*

from ON state, the current has to be reduced below I_H by reducing the applied voltage close to zero, i.e., below *holding voltage*, V_H. Thus, the diode acts as a switch during forward bias condition. The characteristic curve of a PNPN diode is shown in Fig. 5.5.

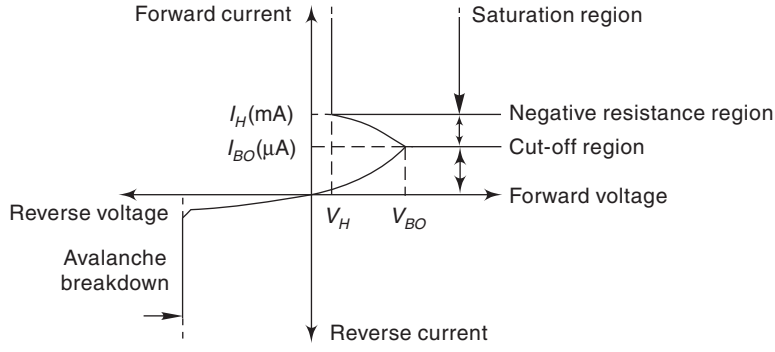


Fig. 5.5 *Characteristic curve of PNPN diode*

5.4 SCR (SILICON CONTROLLED RECTIFIER)

[AU May 2017, Nov 2016, Dec 2015, May 2014, Nov 2013, Nov 2011 and June 2011, 8/16 marks]

The basic structure and circuit symbol of SCR is shown in Fig. 5.6. It is a four layer, three terminal device in which the end *P*-layer acts as anode, the end *N*-layer acts as cathode and *P*-layer nearer to cathode acts as gate. As leakage current in silicon is very small compared to germanium, SCRs are made of silicon and not germanium.

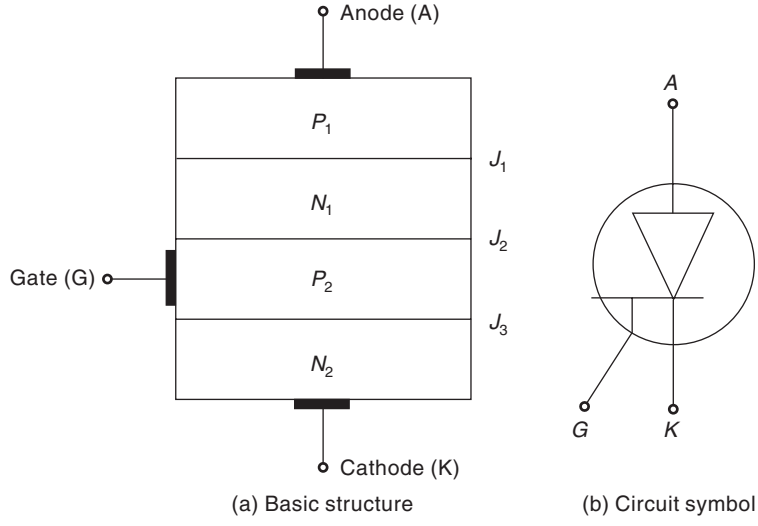


Fig. 5.6 *Basic structure and circuit symbol of SCR*

Characteristics of SCR The characteristics of SCR are shown in Fig. 5.7. SCR acts as a switch when it is forward biased. When the gate is kept open, i.e., gate current $I_G = 0$, operation of SCR is similar to *PNPN* diode. When $I_G < 0$, the amount of reverse bias applied to J_2 is increased. So the breakover

voltage V_{BO} is increased. When $I_G > 0$, the amount of reverse bias applied to J_2 is decreased, thereby decreasing the breakover voltage. With very large positive gate current, breakover may occur at a very low voltage such that the characteristics of SCR is similar to that of ordinary PN diode. As the voltage at which SCR is switched 'ON' can be controlled by varying the gate current I_G, it is commonly called as controlled switch. Once SCR is turned ON, the gate loses control, i.e., the gate cannot be used to switch the device OFF. One way to turn the device OFF is by lowering the anode current below the holding current I_H by reducing the supply voltage below holding voltage V_H, keeping the gate open.

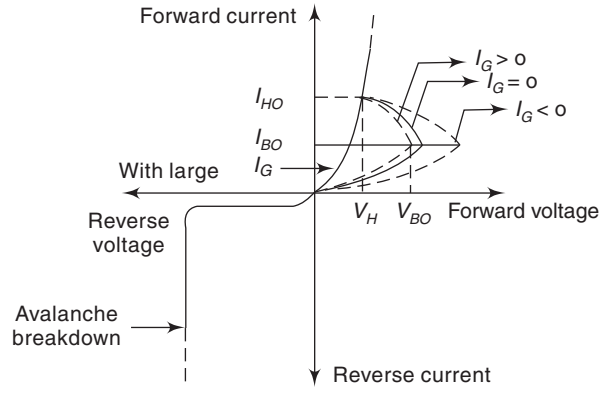


Fig. 5.7 *Characteristics of SCR*

SCR is used in relay control, motor control, phase control, heater control, battery chargers, inverters, regulated power supplies and as static switches.

Two transistor version of SCR The operation of SCR can be explained in a very simple way by considering it in terms of two transistors, called as the two transistor version of SCR. As shown in Fig. 5.8, an SCR can be split into two parts and displaced mechanically from one another, but connected electrically. Thus the device may be considered to be constituted by two transistors T_1 (PNP) and T_2 (NPN) connected back to back.

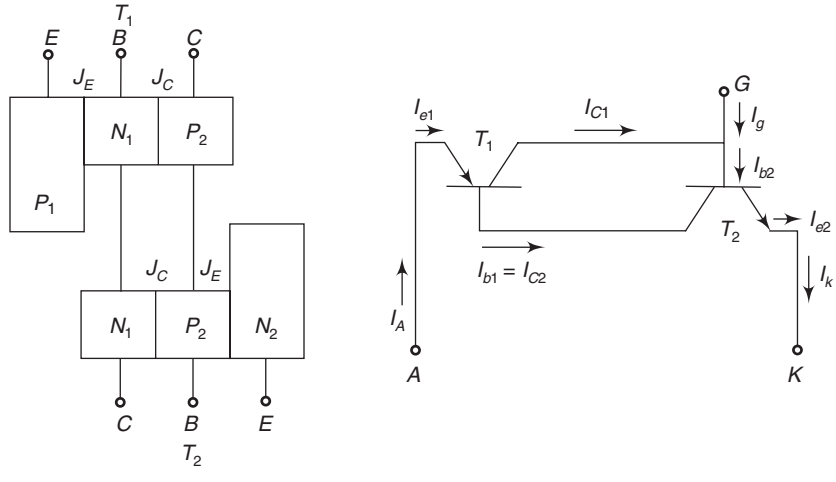


Fig. 5.8 *Two transistor version of SCR*

Assuming the leakage current of T_1 to be negligibly small, we obtain

$$I_{b1} = I_A - I_{c1} = I_A - \alpha_1 I_A = (1 - \alpha_1)\, I_A \tag{5.1}$$

Also, from Fig. 5.8, it is clear that

$$I_{b1} = I_{C2} \tag{5.2}$$

and

$$I_{C2} = \alpha_2 I_K \tag{5.3}$$

Substituting the values given in Eqs (5.2) and (5.3) in Eq. (5.1), we get

$$(1 - \alpha_1)\, I_A = \alpha_2 I_K \tag{5.4}$$

We know that

$$I_K = I_A + I_g \tag{5.5}$$

Substituting Eq. (5.5) in Eq. (5.4), we obtain

$$(1 - \alpha_1)\, I_A = \alpha_2 (I_A + I_g)$$

i.e.,

$$(1 - \alpha_1 - \alpha_2) I_A = \alpha_2 I_g$$

i.e.,

$$I_A = \left[\frac{\alpha_2 I_g}{1 - (\alpha_1 + \alpha_2)}\right] \tag{5.6}$$

Equation (5.6) indicates that if $(\alpha_1 + \alpha_2) = 1$, then $I_A = \infty$, i.e., the anode current I_A suddenly reaches a very high value approaching infinity. Therefore, the device suddenly triggers into ON state from the original OFF state. This characteristic of the device is known as its *regenerative action.*

The value of $(\alpha_1 + \alpha_2)$ can be made almost equal to unity by giving a proper value of positive current I_g for a short duration. This signal I_g applied at the gate which is the base of T_2 will cause a flow of collector current I_{C2} by transferring T_2 to its ON state. As $I_{C2} = I_{b1}$, the transistor T_1 will also be switched ON. Now, the action is regenerative since each of the transistors would supply base current to the other. At this point even if the gate signal is removed, the device keeps on conducting, till the current level is maintained to a minimum value of holding current.

5.5 THYRISTOR RATINGS

Latching current (I_L) Latching current is the minimum current required to latch or trigger the device from its OFF-state to its ON-state.

Holding current (I_H) Holding current is the minimum value of current to hold the device in ON-state. For turning the device OFF, the anode current should be lowered below I_H by increasing the external circuit resistance.

Gate current (I_g) Gate current is the current applied to the gate of the device for control purposes. The minimum gate current is the minimum value of current required at the gate for triggering the device. The maximum gate current is the maximum value of current applied to the device without damaging the gate. More the gate current, earlier is the triggering of the device and vice-versa.

Voltage safety factor (V_f) Voltage safety factor V_f is a ratio which is related to the PIV, the RMS value of the normal operating voltage as,

$$V_f = \frac{\text{Peak Inverse Voltage (PIV)}}{\sqrt{2} \times \text{rms value of the operating voltage}}$$

The value of V_f normally lies between 2 and 2.7. For a safe operation, the normal working voltage of the device is much below its PIV.

5.6 RECTIFIER CIRCUITS USING SCR

SCRs are much superior in performance than ordinary diode rectifiers. They find their main applications as rectifiers. Some of the rectifier circuits have been explained in the following sections.

5.6.1 SCR Half Wave Rectifier

Though the SCR is basically a switch, it can be used in linear applications like rectification. Figure 5.9(a) shows the circuit of an SCR half wave rectifier.

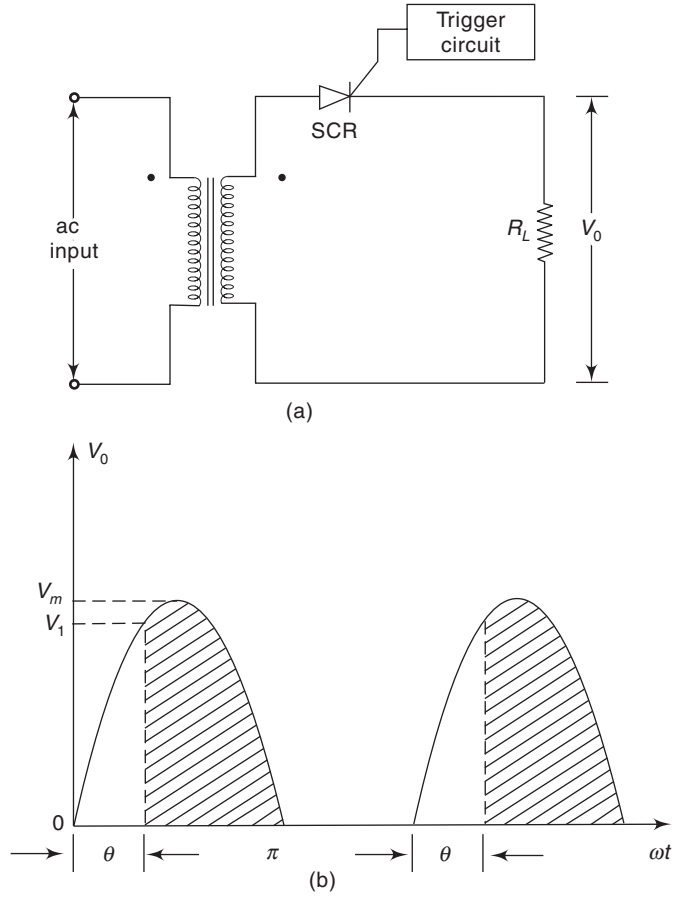


Fig. 5.9 *SCR half wave rectifier*

During the negative halfcycle, the SCR does not conduct irrespective of the gate current, as the anode is negative with respect to cathode and also PIV is less than the reverse breakdown voltage.

During the positive half cycle of AC voltage appearing across secondary, the SCR will conduct provided proper gate current is made to flow. The greater the gate current, the lesser the supply voltage at which the SCR is triggered ON. Referring to Fig. 5.9(b), the gate current is adjusted to such a value that SCR is turned ON at a positive voltage V_1 of AC secondary voltage which is less than the peak voltage V_m. Beyond this, the SCR will be conducting till the applied voltage becomes zero. The angle at which the SCR starts conducting during the positive half cycle is called firing angle θ. Therefore, the conduction angle is $(180° - \theta)$.

The SCR will block not only the negative part of the applied sinusoidal voltage, but will also block the part of the positive waveform up to a point SCR is triggered ON. If the angle θ is zero, this will be an ordinary half wave rectification. Therefore by proper adjustment of gate current, the SCR can be made to conduct full or part of a positive half cycle, thereby controlling the power fed to the load.

Analysis Let $V = V_m \sin \omega t$ be alternating voltage that appear across the secondary of the transformer. In SCR halfwave rectifier, θ is the firing angle and the rectifier conducts from θ to 180° (π radians) during the positive half cycle.

Therefore, average or DC output, $V_{av} = \frac{1}{2\pi}\int_{\theta}^{\pi} V_m \sin \omega t \, d\omega t$

$$= \frac{1}{2\pi}[-V_m \cos \omega t]_{\theta}^{\pi}$$

$$= \frac{V_m}{2\pi}(1 + \cos \theta)$$

For $\theta = 0°$, $V_{av} = \frac{V_m}{\pi}$. Here the full positive half cycle will appear across the load. This is the value of average voltage for ordinary halfwave rectifier.

When $\theta = 90°$, $V_{av} = \frac{V_m}{2\pi}$. This shows that greater the firing angle θ, the smaller is the average voltage and vice-versa.

Similarly, $$V_{rms} = \sqrt{\frac{1}{2\pi}\int_{\theta}^{\pi}(V_m \sin \omega t)^2 \, d\omega t}$$

$$= \sqrt{\frac{V_m^2}{4\pi}\int_{\theta}^{\pi}(1 - \cos 2\omega t) \, d\omega t}$$

$$= \sqrt{\frac{V_m^2}{4\pi}\left[\omega t - \frac{\sin 2\omega t}{2}\right]_{\theta}^{\pi}}$$

$$= \frac{V_m}{2}\left[\frac{1}{\pi}\left(\pi - \theta + \frac{\sin 2\theta}{2}\right)\right]^{\frac{1}{2}}$$

If $\theta = 0$, then $$V_{rms} = \frac{V_m}{2}$$

EXAMPLE 5.3

In an SCR half wave rectifier, the forward breakdown voltage of SCR is 110 V for a gate current of 1 mA. If a 50 Hz sinusoidal voltage of 220 V peak is applied, find firing angle, conduction angle, average voltage, average current, power output and the time during which SCR remains OFF. Assume load resistance is 100 Ω and the holding current to be zero.

Solution We know that $V_1 = V_m \sin\theta$

$110 = 220 \sin\theta$, i.e., $\sin\theta = 0.5$

Therefore, firing angle $\theta = \sin^{-1}(0.5) = 30°$

Conduction angle $= 180° - \theta = 180° - 30° = 150°$

Average voltage $V_{av} = \dfrac{V_m}{2\pi}(1 + \cos\theta)$

$= \dfrac{220}{2\pi}(1 + \cos 30°) = 65.37$ V

Average current, $I_{av} = V_{av}/R_L = 65.37/100 = 0.6537$ A

Power output, $P_o = V_{av} I_{av} = (65.37)(0.6537) = 42.7326$ W

As $V_1 = V_m \sin\theta = V_m \sin\omega t$

$\omega t = \theta = 30° = \dfrac{\pi}{6}$

$2\pi \times 50t = \dfrac{\pi}{6}$

Therefore, the time during which the SCR remains OFF is

$t = 1/(2 \times 6 \times 50) = 1/600 = 1.667$ ms

5.6.2 SCR Full Wave Rectifier

The SCR full wave rectifier is shown in Fig. 5.10(a). It is exactly similar to an ordinary full wave rectifier except that the two diodes have been replaced by two SCRs. The angle of conduction can be changed by adjusting the gate currents.

During the positive half cycle of the input signal, anode of SCR1 becomes positive and at the same time the anode of SCR2 becomes negative. When the input voltage reaches V_1 as shown in Fig. 5.10(b), SCR1 starts conducting and therefore only the shaded portion of positive half cycle will pass through the load.

During the negative half cycle of the input, the anode of SCR1 becomes negative and the anode of SCR2 becomes positive. Hence, SCR1 does not conduct and SCR2 conducts when the input voltage becomes V_1.

The main advantage of this circuit over ordinary full wave rectifier circuit is that any voltage can be made available at the output by simply changing the firing angle of the SCRs.

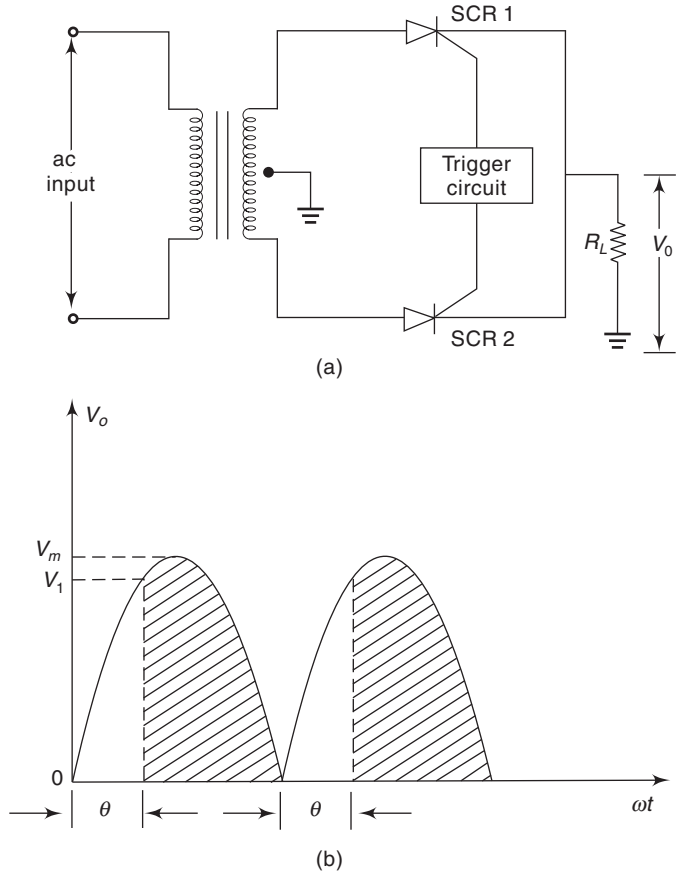


Fig. 5.10 *SCR full wave rectifier*

Analysis Referring to Fig. 5.10(b), let $V = V_m \sin \omega t$ be the alternating voltage that appears between center tap and either end of secondary and θ be the firing angle.

$$V_{\text{av}} = \frac{1}{\pi}\int_{\theta}^{\pi} V_m \sin \omega t \, d\omega t$$

$$= \frac{V_m}{\pi}[-\cos \omega t]_{\theta}^{\pi}$$

$$= \frac{V_m}{\pi}[1 + \cos \theta]$$

This is double that of a half wave rectifier, as negative half cycle is also rectified.

EXAMPLE 5.4

A full wave controlled rectifier employs 2 SCRs and 2 diodes in bridge configuration to rectify 230 V, 50 Hz AC mains and give an output of 150 V to a resistive load of 10 Ω. Find the firing angle, the time during which the SCR remains OFF and the load current.

Solution For an SCR full wave rectifier,

$$V_{DC} = \frac{V_m}{\pi}(1 + \cos\theta)$$

$$150 = \frac{230 \times \sqrt{2}}{\pi}(1 + \cos\theta)$$

Therefore, $\theta = 63.33°$

For 50 Hz, $T = 20$ ms for 360°

Therefore, $t = \frac{20}{360°} \times 63.33° = 3.52$ ms

Load current, $I_{av} = \frac{V_{av}}{R_L} = \frac{150}{10} = 15$ A

EXAMPLE 5.5

When an SCR full wave rectifier is connected across a sinusoidal voltage of 400 sin 314*t*, the RMS value of the current flowing through the device is 20 A. Find the power rating of the SCR.

Solution As the supply voltage is 400 sin 314*t*, $V_m = 400$ V

Peak inverse voltage (PIV) = $\sqrt{3}\ V_m = \sqrt{3} \times 400 = 692.8$ V

rms value of current = 20 A

Average value of current, I_{av} = rms value/form factor = 20/1.11 = 18 A

Power rating of the SCR = PIV × I_{av} = 692.8 × 18 = 12.47 kW

5.6.3 SCR Bridge Rectifier

The SCR bridge rectifier is shown in Fig. 5.11(a). During the positive half cycle of the input AC voltage, SCR1 and diode D_1 conduct whereas SCR2 and diode D_2 do not conduct. During the negative half cycle, SCR2 and diode D_2 conduct. As shown in Fig. 5.11(b), the conduction angle and hence the output voltage can be changed by adjusting the gate currents of SCR1 and SCR2. Here, the DC output voltage, $V_{av} = \frac{V_m}{\pi}[1 + \cos\theta]$, which is equal to that of SCR full wave rectifier.

If the current is lowered below I_H by increasing the external circuit resistance, the SCR will switch OFF.

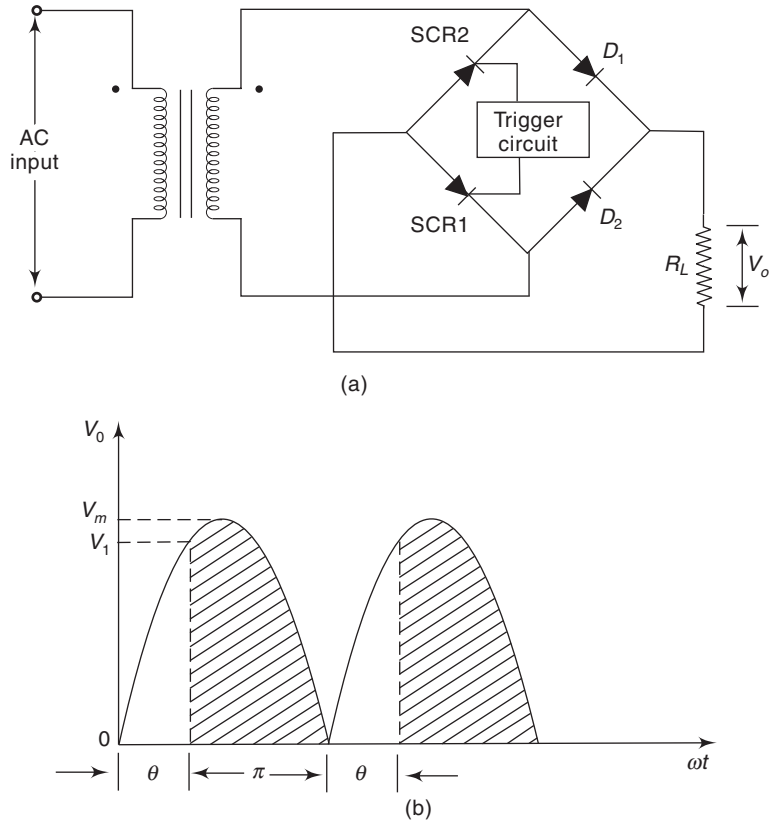


Fig. 5.11 *SCR bridge rectifier*

5.7 LASCR (LIGHT ACTIVATED SCR)

The LASCR shown in Fig. 5.12 is triggered by irradiating with light. The arrows represent incoming light that passes through a window and falls on the depletion layer closer to the middle junction J_2 of SCR. The incident light generates electron-hole pairs in the device thus increasing the number of charge carriers. This leads to the instantaneous flow of current within the device and the device turns ON. For light triggering to occur, the device must have high value of rate of change of voltage with time, dV/dt.

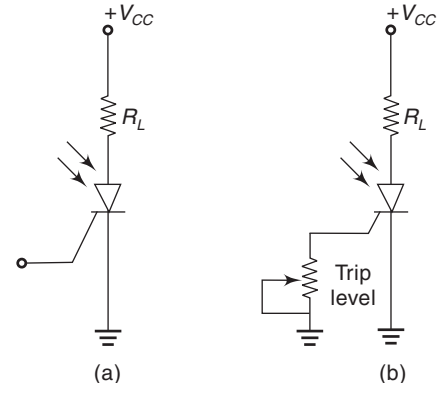


Fig. 5.12 *Light activated SCR*

5.8 TRIAC (TRIODE AC SWITCH)

[AU May 2016, Nov 2011 and Nov 2010, 4 marks]

TRIAC is a three terminal semiconductor switching device which can control alternating current in a load. Its three terminals are MT_1, MT_2 and the gate (G). The basic structure and circuit symbol of a TRIAC are shown in Fig. 5.13. TRIAC is equivalent to two SCRs connected in parallel but in the reverse direction as shown in Fig. 5.14. So, a TRIAC will act as a switch for both directions. The characteristics of a TRIAC are shown in Fig. 5.15.

Like an SCR, a TRIAC also starts conducting only when the breakover voltage is reached. Earlier to that the leakage current which is very small in magnitude flows through the device and therefore remains in the OFF state. The device, when starts conducting, allows very large amount of current to flow through it. The high inrush of current must be limited using external resistance, or it may otherwise damage the device.

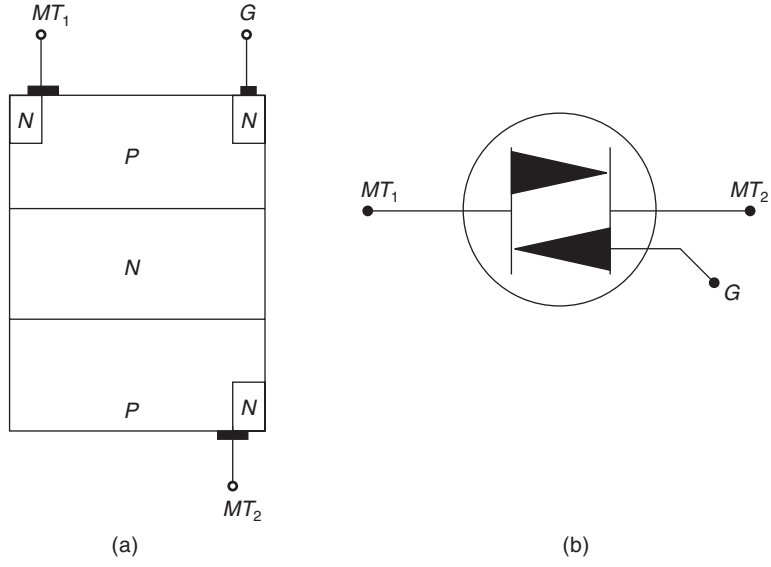


Fig. 5.13 *TRIAC: (a) Basic structure, (b) Circuit symbol*

During the positive half cycle, MT_1 is positive with respect to MT_2, whereas MT_2 is positive with respect to MT_1 during negative half cycle. A TRIAC is a bidirectional device and can be triggered either by a positive or by a negative gate signal. By applying proper signal at the gate, the breakover voltage, i.e., firing angle of the device can be changed; thus phase control process can be achieved.

TRIAC is used for illumination control, temperature control, liquid level control, motor speed control and as static switch to turn AC power ON and OFF. Nowadays, the DIAC–TRIAC pairs are increasingly being replaced by a single component unit known as quadrac. Its main limitation in comparison to SCR is its low power handling capacity.

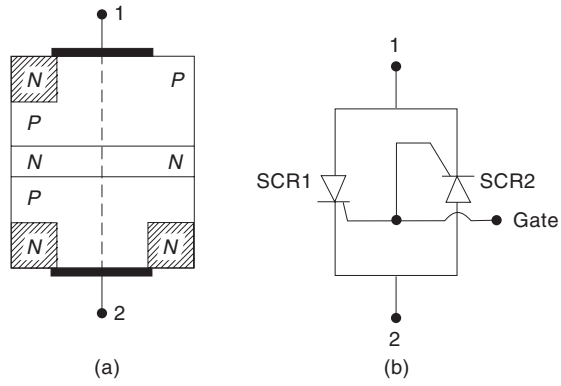


Fig. 5.14 *Two SCR version of TRIAC: (a) Basic structure, (b) Equivalent circuit*

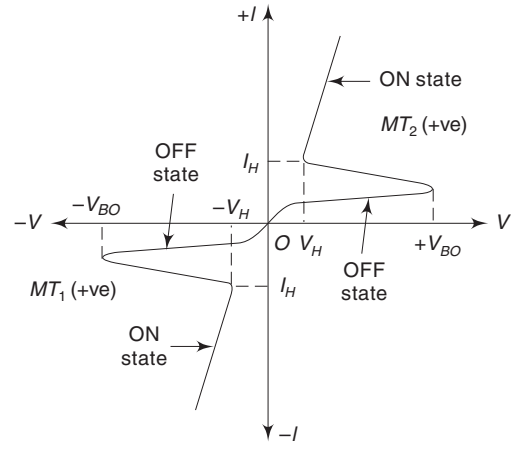


Fig. 5.15 *Characteristics of TRIAC*

5.9 DIAC (DIODE AC SWITCH)

[AU May 2016, Nov 2012, May 2012 and Nov 2010, 8 marks]

The construction and symbol of DIAC are shown in Fig. 5.16. DIAC is a three layer, two terminal semiconductor device. MT_1 and MT_2 are the two main terminals which are interchangeable. It acts as a bidirectional Avalanche diode. It does not have any control terminal. It has two junctions J_1 and J_2. Though the DIAC resembles a bipolar transistor, the central layer is free from any connection with the terminals.

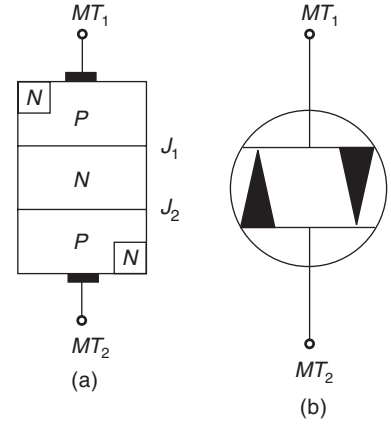


Fig. 5.16 *DIAC: (a) Basic structure, (b) Circuit symbol*

From the characteristic of a DIAC shown in Fig. 5.17, it acts as a switch in both directions. As the doping level at the two ends of the device is the same, the DIAC has identical characteristics for both positive and negative half of an AC cycle. During the positive half cycle, MT_1 is positive with respect to MT_2 whereas MT_2 is positive with respect to MT_1 in the negative half cycle. At voltage less than the breakover voltage, a very small amount of current called the leakage current flows through the device and the device remains in OFF state. When the voltage level reaches the breakover voltage, the device starts conducting and it exhibits negative resistance characteristics, i.e., the current flowing in the device starts increasing and the voltage across it starts decreasing.

The DIAC is not a control device. It is used as triggering device in TRIAC phase control circuits used for light dimming, motor speed control and heater control.

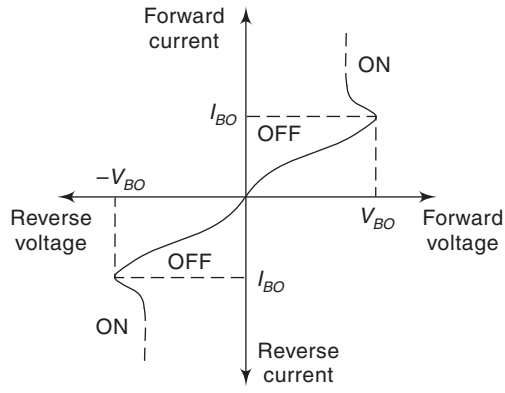


Fig. 5.17 *Characteristic of DIAC*

5.10 POWER AMPLIFIERS

Power dissipation in the output stage of a transistor is a major concern. While analyzing and designing the circuit that delivers a specified power to a load, the output signal has to be linear.

Consider a multistage amplifier that has to deliver a large amount of power to a passive load. This power may be in the form of a large current delivered to a relatively small load resistance (e.g., loudspeaker) or may be in the form of a large voltage delivered to a relatively large load resistance (e.g., switching power supply). The output stage of a power amplifier must be designed to meet the power requirements. Power amplifiers use transistors like power BJTs and power MOSFETs.

One of the major concerns in the design of output stage is to deliver the required signal power to the load efficiently. Moreover, the power dissipated in the transistors of the output stage should be as small as possible. The output transistors must be capable of delivering the required current to the load and sustaining the required output voltage.

Power Transistors

Physical transistors have certain limitations in terms of maximum current, voltage and power. These limitations are not considered for normal transistors, because it was assumed that the transistors which are capable of handling the current, voltage and power dissipation within a transistor does not cause any damage to the circuit.

But in the design of power amplifiers, it is necessary to consider the limitations of a transistor. The limitations are maximum rated current (amp), maximum rated voltage (volts) and maximum rated power (watts). As power amplifiers use BJTs and MOSFETs, the limitation effect is considered on these two transistors. The maximum power limitation is related to maximum allowed temperature of a transistor (BJTs or MOSFETs) which in turn is a function of the rate at which heat is removed.

5.11 POWER BJT

[AU May 2015, 8 marks]

Construction

The structure of a vertically oriented *NPN* power transistor with the doping levels and thickness of the layers is shown in Fig. 5.18. This type of configuration has a large cross sectional area to handle large currents and minimize the thermal resistance of the transistor. Here, the collector terminal is at the bottom.

The primary collector N^- region, called drift region, has a low-doped impurity concentration 10^{20} m^{-3} in such a way that the voltage can be applied across base-collector terminals without initiating breakdown. This region has a thickness of about 50–200 μm. The thickness of the drift region determines the breakdown voltage of the transistor.

Another N^+ region has a higher doping concentration (typically 10^{25} m^{-3}) which reduces the collector resistance and makes contact with the external terminal. This region has a thickness of around 250 μm. The doping in the emitter layer is large (typically 10^{25} m^{-3}), whereas the base doping is comparatively less (10^{22} m^{-3}).

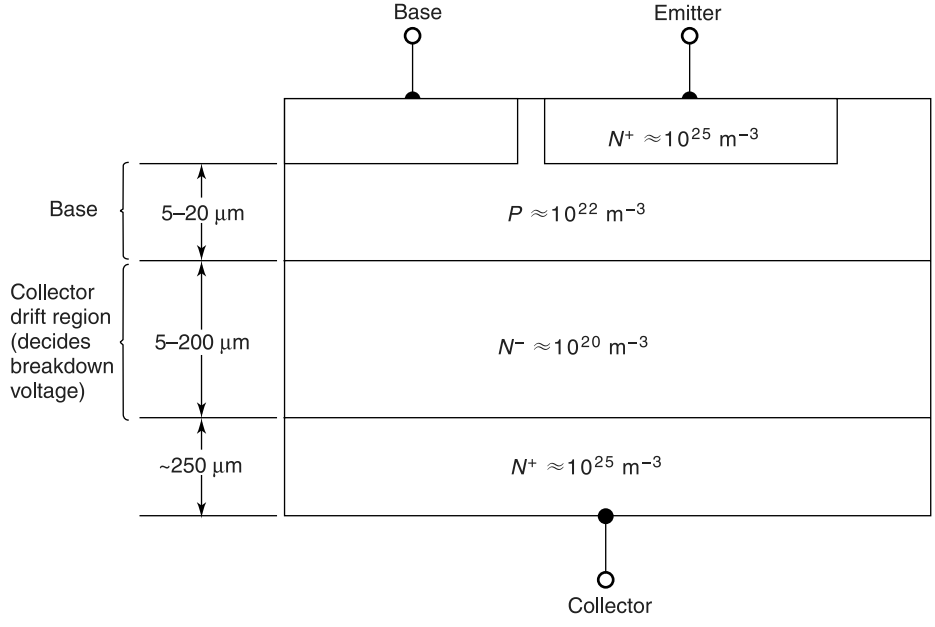


Fig. 5.18 *Cross section of a vertical NPN power BJT*

A sufficient base width is required to prevent punch-through breakdown. Since the small base thickness of about 5–20 μm reduces the breakdown voltage, the amplification capabilities and the breakdown voltages are to be compromised in power transistors. A large base-collector voltage implies a relatively large space-charge width induced in the collector and base regions.

V-I Characteristics

Power transistors are generally large area devices. The properties of power transistors vary from small-signal devices in terms of differences in geometry and doping concentrations. Table 5.1 compares the various parameters of small-signal BJT to those of two power BJTs. The current gain of power transistors is in the range of 5 to 20 which is smaller to that of small-signal BJTs. But the current gain is a strong function of collector current and temperature. The current gain versus collector current characteristics of 2N3055 power BJT at various temperatures is shown in Fig. 5.19. The current gain drops off for high current levels. The parasitic resistances in the base and collector regions may become significant by affecting the transistor terminal characteristics.

Table 5.1

Parameter	*Small-Signal BJT (2N2222A)*	*Power BJT (2N3055)*	*Power BJT (2N6078)*
$V_{CE(\max)}$ in V	40	60	250
$I_{C(\max)}$ in A	0.8	15	7
$P_{D(\max)}$ in W (at T =25°C)	1.2	115	45
β	35–100	5–20	12–70
f_T in MHz	300	0.8	1

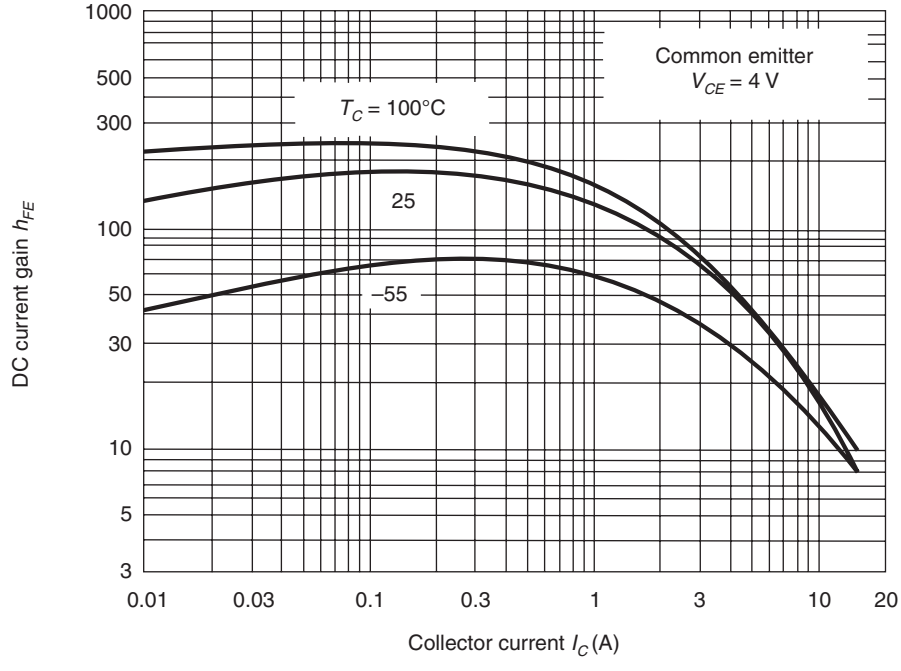


Fig. 5.19 *DC characteristics of β vs I_c*

The maximum rated collector current $I_{c,\text{rated}}$ is related to the maximum current that the wires connecting the semiconductor to the external terminals can handle or the collector current at which the current gain falls below a minimum specified value or the current that leads to the maximum power dissipation when the transistor is in saturation.

In BJTs, the maximum voltage limitation is associated with avalanche breakdown in the reverse-biased base-collector junction. In common-emitter configuration, the breakdown voltage mechanism involves the transistor gain and breakdown phenomenon on the *PN* junction. Typical I_C versus V_{CE} characteristics curve of BJT is shown in Fig. 5.20. When the base terminal is open circuited, i.e., $I_B = 0$, the breakdown voltage is V_{CE0} and from Fig. 5.20, its value is 130 V.

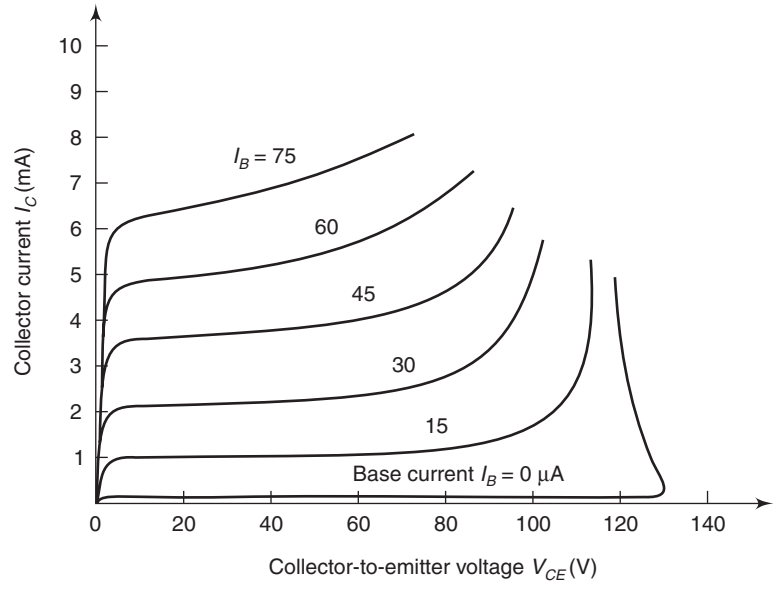


Fig. 5.20 *I_C versus V_{CE} characteristics*

When the transistor is biased in the active region, the collector current begins to increase significantly before breakdown voltage V_{CE0} is reached. Once the breakdown has occurred, all the curves tend to merge to the same collector-emitter voltage. The curves merging voltage is denoted as $V_{CE(\text{sus})}$ and it is the minimum voltage necessary to sustain the transistor in breakdown and from Fig. 5.20, the approximate value of $V_{CE(\text{sus})}$ is 115 V.

When operating BJT at high voltage and at high current, another breakdown effect will occur and it is called *second breakdown*. Slight non-uniformities in the current density produce an increased heating of local region. The resistance level of the semiconductor material is decreased because of increased heating in the local region.

The effect results in positive feedback, and the current continues to increase, which further increases the temperature, until the semiconductor material may actually be melt.

The instantaneous power dissipation in a BJT is given by

$$P_Q = V_{CE}I_C + V_{BE}I_B \tag{5.7}$$

where the base current I_B is generally much smaller than the collector current I_C. Therefore, the instantaneous power dissipation is approximated to

$$P_Q \cong V_{CE}I_C \tag{5.8}$$

By integrating Eq. (5.8) over one cycle, the average power is given by

$$\bar{P}_Q = \frac{1}{T}\int_O^T V_{CE}I_C dt \tag{5.9}$$

The average power dissipated in BJT must be kept below a specified maximum value, to ensure that the temperature of the device remains below the maximum value. If the collector current and collector-emitter voltage are DC quantities, then maximum rated power P_T for BJT can be written as

$$P_T = V_{CE}I_C \tag{5.10}$$

The average power limitation P_T is a hyperbola as per the above equation.

The maximum current, voltage and power limitations are illustrated on the I_C versus V_{CE} characteristics as shown in Fig. 5.21(a). Safe Operating Area (SoA) is the region where a transistor can be operated safely and is bounded by $I_{C\max}$, $V_{CE(\text{sus})}$, P_T and transistor's second breakdown characteristics. Figure 5.21(b) shows the safe operating area in linear scale and in logarithmic scale. The $I_C - V_{CE}$ operating point may move momentarily outside the safe operating area without damaging the transistor. But this depends on how far and how long the Q-point moves outside the area.

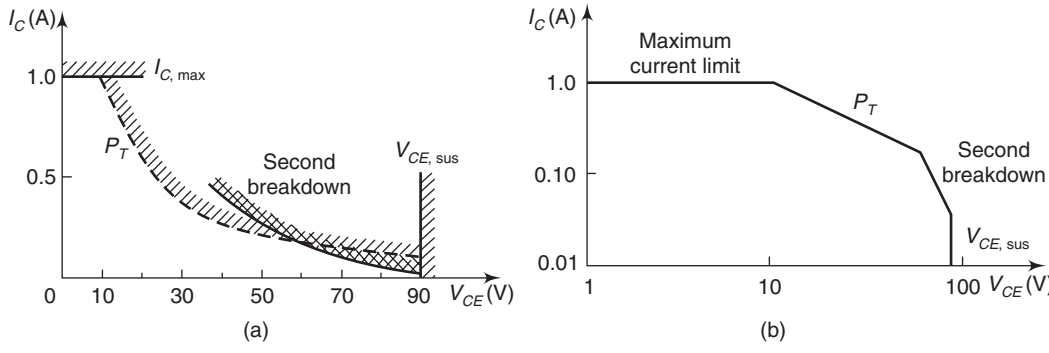


Fig. 5.21 *(a) I_C versus V_{CE} characteristics (b) SoA*

Power transistors which are designed to handle large current require large emitter areas to maintain reasonable current densities. These transistors are usually designed with narrow emitter widths to minimize the parasitic base resistance and fabricated as an interdigitated structure as shown in Fig. 5.22. In each emitter leg, small resistors are incorporated which helps to maintain equal current in each *B-E* junction.

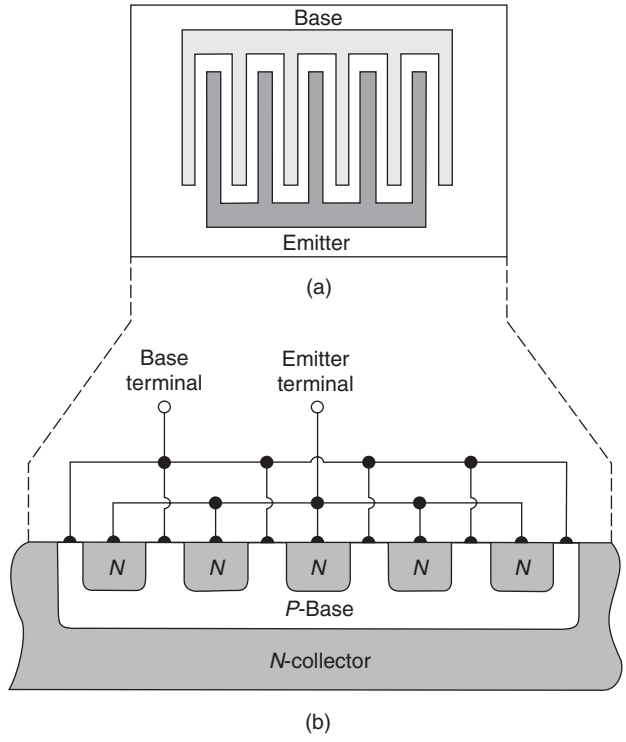


Fig. 5.22 *Interdigitated structure for power transistor*

5.12 POWER MOSFET

[AU May 2015, 8 marks]

Two *N*-channel power MOSFETs with parameters are listed in Table 5.2. The drain currents are in ampere range and breakdown voltages are in the hundreds of volt range. Like BJT, a MOSFET has also to be operated in safe region.

Table 5.2

Parameter	*2N6757*	*2N6792*
V_{DS}(max) in V	150	400
$I_D$max (at $T = 25^{\circ}C$)	8	2
P_D in W	75	20

Power MOSFETS are different from bipolar power transistor in terms of operating principles and performance. The performance characteristics of power MOSFETs include faster switching time, no second breakdown, stable gain and response time over a wide temperature range. Transconductance

versus drain current curves for different temperatures is shown in Fig. 5.23(a). The variation with temperature of the MOSFET transconductance is less.

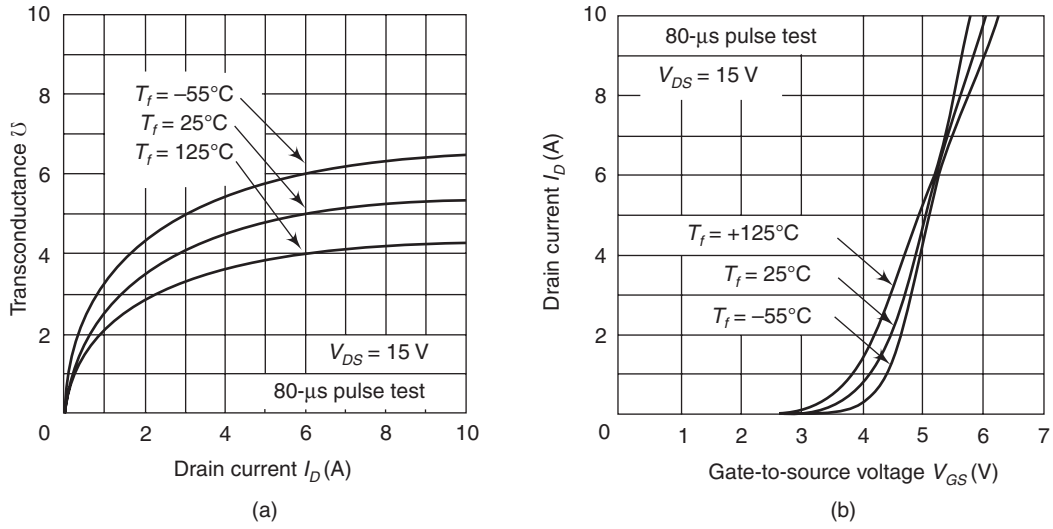


Fig. 5.23 *(a) Transconductance vs drain current characteristics*
(b) I_D versus V_{GS} characteristics for different temperatures

The MOSFET has high input impedance and it is a voltage-controlled device like JFET. Unlike the driver circuit of BJT, a MOSFET driver circuit is simple. The gate of a 10 A power MOSFET may be driven by the output of a standard logic circuit. The MOSFET is a majority carrier device and any increase in temperature of the device affects the mobility of the majority carriers, which results in increase of resistivity of the semiconductor. From this, it is clear that MOSFETs are more immune to the thermal runaway effects and second breakdown phenomena experienced in BJTs. Figure 5.23(b) shows typical I_D versus V_{GS} characteristics at several temperatures. From the characteristics curve, for the given gate-to-source voltage at high current level, the current actually decreases with increase in temperature.

Compared to BJT, typical MOSFET power-handling level is much less. This disadvantage of a MOSFET is compensated by changing the construction mode from planar structure to a vertical one. So, power MOSFETs have different structures comparing to lateral MOSFETs by having a double diffused or vertical diffused process called DMOS or VMOS respectively.

5.13 DMOS

[AU Nov 2016, Dec 2015, 8 marks]

DMOS is an FET structure created specifically for high power applications and it is a planar transistor whose name is derived from the double-diffusion process used to construct it. DMOS is also used in switching applications with high-voltage and high-frequency behaviour, like inkjet printhead power supplies and automobile control electronics.

The cross-sectional view of a DMOS structure is shown in Fig. 5.24. The *P*-substrate region and the N^+ source contact are diffused through a common window defined by the edge of the gate. The *P*-substrate

region is diffused deeper than the N^+ source. The surface channel length is defined as the lateral diffusion distance between the *P*-substrate and the N^+ source.

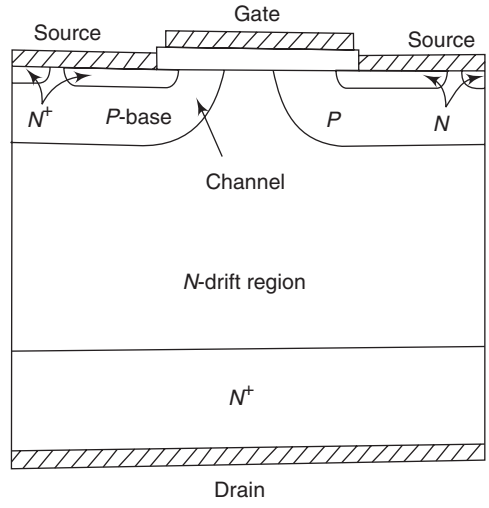


Fig. 5.24 *Cross section of the DMOS structure*

Electrons outside the source terminal flow laterally through the inversion layer under the gate to the *N*-drift region. Then, electrons flow vertically through the *N*-drift region to the drain terminal. The conventional current direction is from the drain to the source. The most important characteristics of DMOS device is their breakdown voltage and on-resistance. DMOS is similar to a BJT, due to the high-voltage and high-frequency characteristics. A lightly doped drift region between the drain contact and the channel region helps to ensure a very high breakdown voltage. Moreover, the *N*-drift region must be moderately doped so that the drain breakdown voltage is sufficiently large. The thickness of the *N*-drift region should be as small as possible to minimize drain resistance.

5.14 VMOS

[AU Dec 2015, 8 marks]

Elements of the planar MOSFET are present in the vertical metal-oxide-silicon FET (VMOS) as shown in Fig. 5.25. The terminals of the device are connected to the metallic surface. The SiO_2 layer is placed between the gate and the *P*-type region and between the drain and source for the growth of the induced *N*-channel (enhancement-mode operation). The term *vertical* in VMOS is because of the fact that the channel is formed in the vertical direction rather than horizontal direction in the planar device. Moreover, the channel has the appearance of a "V" cut in the semiconductor base, which is also an important characteristic of the device. The construction of VMOS is simple by leaving out some of the transition levels of doping.

The *N*-channel is induced or enhanced in the narrow *P*-type region of the device if a positive voltage is applied to the drain, negative voltage to the source and 0 V or same positive voltage level to the gate. The vertical height of the *P*-region defines the length of the channel and this length on a horizontal plane is limited to 1 μm to 2 μm. Diffusion layers such as *P*-region of the device can be controlled to small fraction of μm.

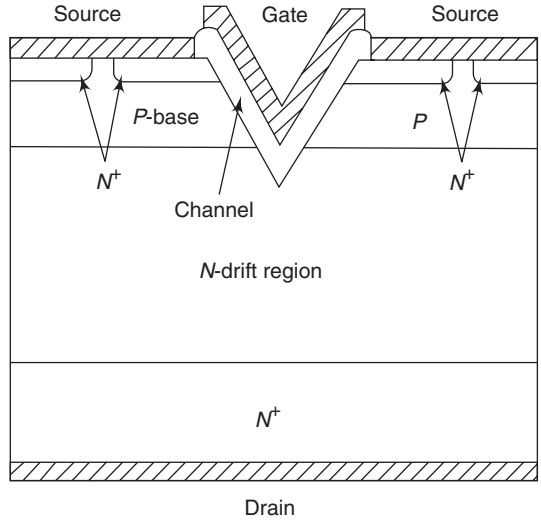


Fig. 5.25 *VMOS structure*

Resistance levels and the power dissipation levels of the device at operating current will be reduced because of reduced channel length. Also, the contact area between the channel and the $n+$ region is much increased by the vertical mode construction, which contributes to a further reduction in the resistance level and increased area between doping layers for current flow. Two conduction paths exist between drain and source. This conduction path further contributes to a higher current rating. The final result is that VMOS is a device with drain currents that can reach the ampere levels with power levels exceeding low. The VMOS FETs have reduced channel resistance levels, higher current rating and higher power rating when compared to planar MOSFET. Due to its high current-handling capability, VMOS transistors are useful in power amplifier applications.

Another important characteristic of vertical construction in VMOS FET is to have positive temperature coefficient which overcomes the possibility of thermal runaway. The resistance level of a device increases with increase in temperature of the device because of surrounding medium and current. This causes a reduction in drain current of the device, whereas for conventional devices, the drain current will increase. Negative temperature coefficient results in decreased levels of resistance with increase in temperature, which increases the current level, which in turn results in temperature stability and thermal runaway.

The reduced charge storage level in VMOS results in faster switching time when compared to conventional planar construction. The switching time of VMOS device is less than one-half that encountered in conventional BJT transistors.

5.15 LIGHT EMITTERS

[AU Dec 2015, May 2014, Nov 2012, 8 marks]

5.15.1 Light Emitting Diode (LED)

The Light Emitting Diode (LED) is a *PN* junction device which emits light when forward biased, by a phenomenon called electroluminescence. In all semiconductor *PN* junctions, some of the energy will

be radiated as heat and some in the form of photons. In silicon and germanium, greater percentage of energy is given out in the form of heat and the emitted light is insignificant. In other materials such as gallium phosphide (GaP) or gallium arsenide phosphide (GaAsP), the number of photons of light energy emitted is sufficient to create a visible light source. Here, the charge carrier recombination takes place when electrons from the *N*-side cross the junction and recombine with the holes on the *P*-side.

LED under forward bias and its symbol are shown in Figs 5.26(a) and (b), respectively. When an LED is forward biased, the electrons and holes move towards the junction and recombination takes place. As a result of recombination, the electrons lying in the conduction bands of *N*-region fall into the holes lying in the valence band of a *P*-region. The difference of energy between the conduction band and the valence band is radiated in the form of light energy. Each recombination causes radiation of light energy. Light is generated by recombination of electrons and holes whereby their excess energy is transferred to an emitted photon. The brightness of the emitted light is directly proportional to the forward bias current.

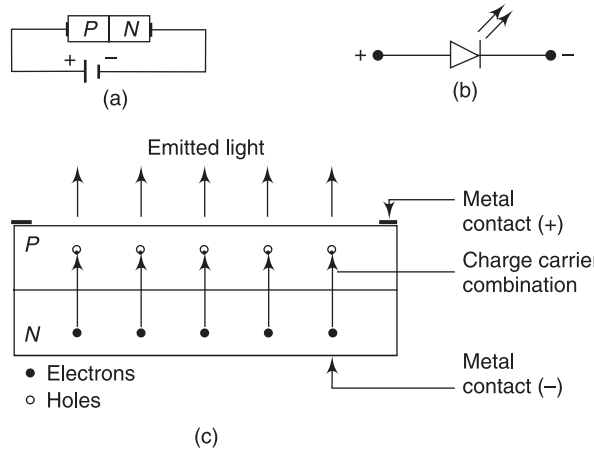


Fig. 5.26 *LED: (a) LED under forward bias (b) Symbol (c) Recombinations and emission of light*

Figure 5.26(c) shows the basic structure of an LED showing recombination of carriers and emission of light. Here, an *N*-type layer is grown on a substrate and a *P*-type is deposited on it by diffusion. Since carrier recombination takes place in the *P*-layer, it is kept uppermost. The metal anode connections are made at the outer edges of the *P*-layer so as to allow more central surface area for the light to escape. LEDs are manufactured with domed lenses in order to reduce the reabsorption problem. A metal (gold) film is applied to the bottom of the substrate for reflecting as much light as possible to the surface of the device and also to provide cathode connection. LEDs are always encased to protect their delicate wires.

The efficiency of generation of light increases with the increases in injected current and with a decrease in temperature. The light is concentrated near the junction as the carriers are available within a diffusion length of the junction.

LEDs radiate different colours such as red, green, yellow, orange, blue and white. Some of the LEDs emit infrared (invisible) light also. The wavelength of emitted light depends on the energy gap

of the material. Hence, the colour of the emitted light depends on the type of material used is given as follows.

Gallium arsenide (GaAs) – infrared radiation (invisible)

Gallium phosphide (GaP) – red or green

Gallium arsenide phosphide (GaAsP) – red or yellow

In order to protect LEDs, resistance of 1 kΩ or 1.5 kΩ must be connected in series with the LED. LEDs emit no light when reverse biased. LEDs operate at voltage levels from 1.5 to 3.3 V, with the current of some tens of milliamperes. The power requirement is typically from 10 to 150 mW with a life time of 1,00,000 + hours. LEDs can be switched ON and OFF at a very fast speed of 1 ns.

They are used in burglar alarm systems, picture phones, multimeters, calculators, digital meters, microprocessors, digital computers, electronic telephone exchange, intercoms, electronic panels, digital watches, solid state video displays and optical communication systems. Also, there are two-lead LED lamps which contain two LEDs, so that a reversal in biasing will change the colour from green to red, or vice-versa.

When the emitted light is coherent, i.e., essentially monocromatic, then such a diode is referred to as an Injection Laser Diode (ILD). The LED and ILD are the two main types used as optical sources. ILD has a shorter rise time than LED, which makes the ILD more suitable for wide-bandwidth and high-data-rate applications. In addition, more optical power can be coupled into a fiber with an ILD, which is important for long distance transmission. A disadvantage of the ILD is the strong temperature dependence of the output characteristic curve.

5.15.2 Infrared Emitters

The infrared emitting diodes are *PN* junction gallium arsenide devices which emit a beam of light when forward biased. When the junction is energised, electrons from the *N*-region will recombine with the excess holes of the *P*-material in a specially formed recombination region sandwiched between the *P*- and *N*-type materials. This recombination, which tends to restore the equilibrium carrier densities, can result in the emission of photons from the junction. The radiant energy from the device is infrared with a typical peak at 0.9 μm, which ideally matches the response of silicon photodiode and phototransistors.

These infrared emitting diodes are used in shaft encoders, data-transmission systems, intrusion alarms, card and paper tape readers, and high density mounting applications. The shaft encoder can produce 150 μW of radiant energy at 1.2 V and 50 mA.

5.16 LIQUID CRYSTAL DISPLAY (LCD)

[AU May 2017, Nov 2016, May 2015, May 2012, 8 marks]

Liquid Crystal Displays (LCDs) are used for display of numeric and alphanumeric character in dot matrix and segmental displays. The two liquid crystal materials which are commonly used in display technology are nematic and cholesteric whose schematic arrangement of molecules is shown in Fig. 5.27(a). The most popular liquid crystal structure is the Nematic Liquid Crystal (NLC). In this type, all the molecules align themselves approximately parallel to a unique axis (director), while

retaining the complete translational freedom. The liquid is normally transparent, but if subjected to a strong electric field, disruption of the well ordered crystal structure takes place causing the liquid to polarise and turn opaque. The removal of the applied electric field allows the crystal structure to regain its original form and the material becomes transparent.

Based on the construction, LCDs are classified into two types. They are (i) Dynamic scattering type, and (ii) Field effect type.

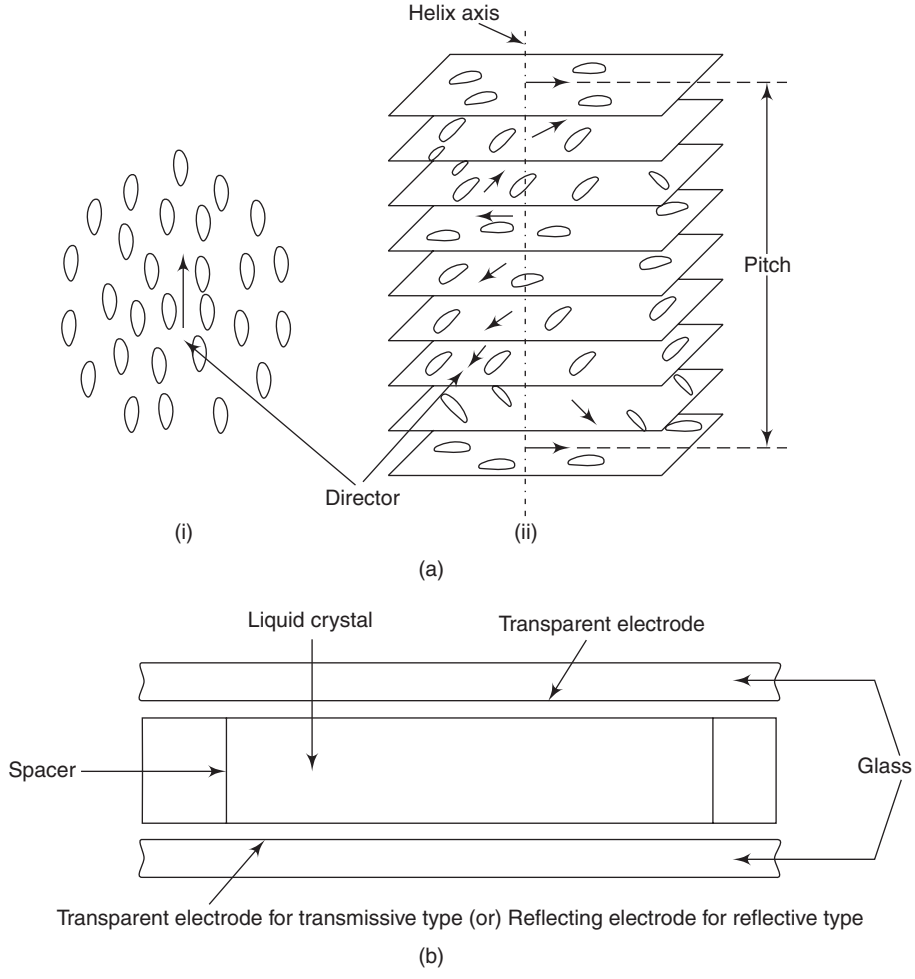


Fig. 5.27 *(a) Schematic arrangement of molecules in liquid crystal (i) Nematic and (ii) Cholesteric (b) Construction of a dynamic scattering LCD*

Dynamic scattering type The construction of a dynamic scattering liquid crystal cell is shown in Fig. 5.27(b). The display consists of two glass plates, each coated with tin oxide (SnO_2) on the inside with transparent electrodes separated by a liquid crystal layer, 5 to 50 μm thick. The oxide coating on the front sheet is etched to produce a single or multi-segment pattern of characters, with each

segment properly insulated from each other. A weak electric field applied to a liquid crystal tends to align molecules in the direction of the field. As soon as the voltage exceeds a certain threshold value, the domain structure collapses and the appearance is changed. As the voltage grows further, the flow becomes turbulent and the substance turns optically inhomogenous. In this disordered state, the liquid crystal scatters light. Thus, when the liquid is not activated, it is transparent. When the liquid is activated, the molecular turbulence causes light to be scattered in all directions and the cell appears to be bright. This phenomenon is called dynamic scattering.

Field effect type The construction of a field effect LCD display is similar to that of the dynamic scattering type, with the exception that two thin polarising optical filters are placed at the inside of each glass sheet. The LCD material is of twisted nematic type which twists the light (change in direction of polarisation) passing through the cell when the latter is not energised. This allows light to pass through the optical filters and the cell appears bright. When the cell is energised, no twisting of light takes place and the cell appears dull.

Liquid crystal cells are of two types: (i) Transmittive type, and (ii) Reflective type. In the transmittive type cell, both glass sheets are transparent so that light from a rear source is scattered in the forward direction when the cell is activated.

The reflective type cell has a reflecting surface on one side of the glass sheet. The incident light on the front surface of the cell is dynamically scattered by an activated cell. Both types of cells appear quite bright when activated even under ambient light conditions.

Liquid crystals consume small amount of energy. In a seven segment display the current drawn is about 25 μA for dynamic scattering cells and 300 μA for field effect cells. LCDs require AC voltage supply. A typical voltage supply to dynamic scattering LCDs is 30 V peak-to-peak with 50 Hz. LCDs are normally used for seven-segmental displays.

Advantages of LCD

(i) The voltages required are small.
(ii) They have a low power consumption. A seven segment display requires about 140 W (20 W per segment), whereas LEDs require about 40 mW per numeral.
(iii) They are economical.

Disadvantages of LCD

(i) LCDs are very slow devices. The turn ON and turn OFF times are quite large. The turn ON time is typically of the order of a few ms, while the turn OFF time is 10 ms.
(ii) When used on DC, their life span is quite small. Therefore, they are used with AC supplies having a frequency less than 50 Hz.
(iii) They occupy a large area.

Table 5.3 *Comparison between LED and LCD* **[AU Nov 2010, 6 marks]**

LED	*LCD*
Consumes more power–requires 10–250 mW power per digit	Essentially acts as a capacitor and consumes very less power–requires 10–200 μW power per digit
Because of high power requirement, it requires external interface circuitry when driven from ICs	Can be driven directly from IC chips
Good brightness level	Moderate brightness level
Operable within the temperature range –40 to 85 °C	Temperature range limited to –20 to 60 °C
Life time is around 100,000 hours	Life time is limited to 50,000 hours due to chemical degradation
Emits light in red, orange, yellow, green, blue and white	Invisible in darkness – requires external illumination
Operating voltage range is 1.5 to 5 V DC Response time is 50 to 500 ns	Operating voltage range is 3 to 20 V AC Has a slow decay time – response time is 50 to 200 ms
Viewing angle 150°	Viewing angle 100°

5.17 ALPHANUMERIC DISPLAYS

Display devices provide a visual display of numbers, letters, and various signs in response to electrical input, and serve as constituents of an electronic display system. Display devices can be classified as passive displays and active displays.

(a) Passive displays: Light controllers—they are modulators of light in which the light pattern gets modified on application of electric field, e.g. LCD.

(b) Active displays: Light emitters—they are generators of light, e.g. LED.

The optical devices described so far were capable of operating in an OFF/ON mode. LEDs are used as low consumption indicator lamps. Also, both LEDs and LCDs are potentially more useful as elements in alphanumeric display panels. There are two possible arrangements of optical displays, viz. seven segment and dot matrix, the choice being based on the display size, definition and allowed circuit complexity.

One way of producing an alphanumeric display is to make a *seven-segment* monolithic device, as shown in Fig. 5.28(a), which can display all numerals and nine letters. Each segment contains LED/LCD which can be turned ON or OFF to form the desired digit. Each segment of the array has to be switched in response to a logic signal. For example, Fig. 5.28(a) shows the response to a logic signal corresponding to 2, in which segments *a*, *b*, *d*, *e* and *g* have been switched ON and *c* and *f* remain OFF. Similarly, when all segments are ON, the digit formed is 8. If only the center segment, *g*, is OFF, the digit will be zero. Common anode and Common cathode seven segment LED displays are shown in Fig. 5.28(b). Common anode type LED displays require an active LOW configuration, whereas an active HIGH circuitry is necessary for the common cathode type LED display.

The seven segment displays are used in digital clocks, calculators, microwave ovens, digital multimeters, microprocessor trainer kits, stereo tuners, etc.

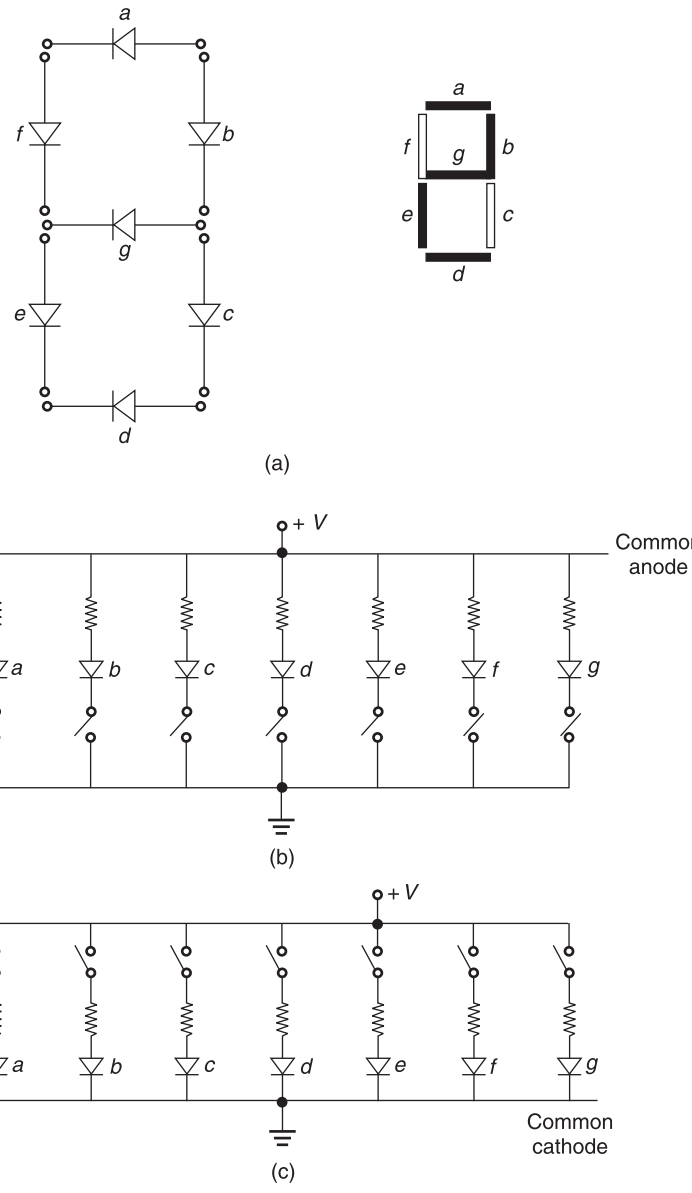


Fig. 5.28 *(a) Seven segment monolithic device, (b) Common anode and (c) Common cathode configurations*

Another method of producing an alphanumeric display is to make a *dot matrix* of LEDs/LCDs in a monolithic structure. Commonly used dot matrices for this display are 5 × 7, 5 × 8 and 7 × 9, which can display 64 different characters including the alphabets, numerals and various symbols, by driving the appropriate horizontal and vertical inputs. A 5 × 7 dot matrix assembly using LEDs and the corresponding wiring pattern is shown in Fig. 5.29.

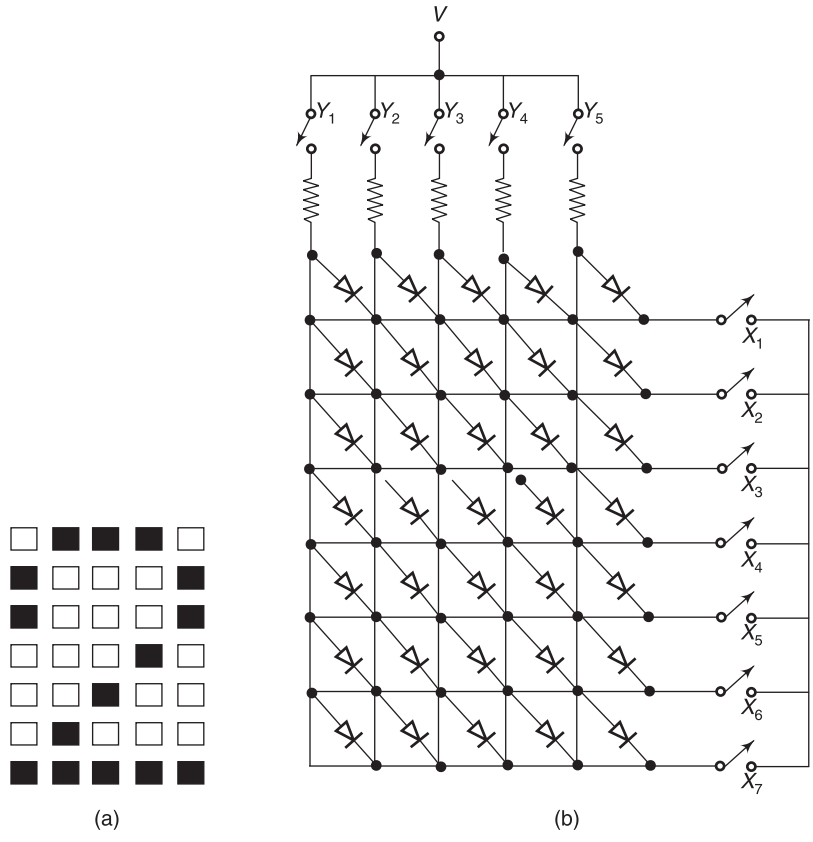


Fig. 5.29 *(a) 5 × 7 dot matrix and (b) Wiring pattern for 5 × 7 dot matrix*

LED displays are available in many different sizes and shapes. The light emitting region is available in lengths from 0.25 to 2.5 cm.

5.18 PHOTODIODE AND PHOTOTRANSISTOR

[AU Nov 2016, Nov 2014, May 2014, May 2013, Nov 2011, Nov 2010, 8 marks]

Photodiode Silicon photodiode is a light sensitive device, also called *photodetector*, which converts light signals into electrical signals. The construction and symbol of a photodiode are shown in Fig. 5.30. The diode is made of a semiconductor *PN* junction kept in a sealed plastic or glass casing. The cover is so designed that the light rays are allowed to fall on one surface across the junction. The remaining sides of the casing are painted to restrict the penetration of light rays. A lens permits light to fall on the junction. When light falls on the reverse biased *PN* photodiode junction, hole-electron pairs are created. The movement of these hole-electron pairs in a properly connected circuit results in current flow. The magnitude of the photocurrent depends on the number of charge carriers generated and hence, on the illumination of the diode element. This current is also affected by the frequency of

the light falling on the junction of the photodiode. The magnitude of the current under large reverse bias is given by

$$I = I_S + I_o\,(1 - e^{V/\eta V_T})$$

where I_o = reverse saturation current
I_S = short-circuit current which is proportional to the light intensity
V = voltage across the diode
V_T = volt equivalent of temperature
η = parameter, 1 for Ge and 2 for Si.

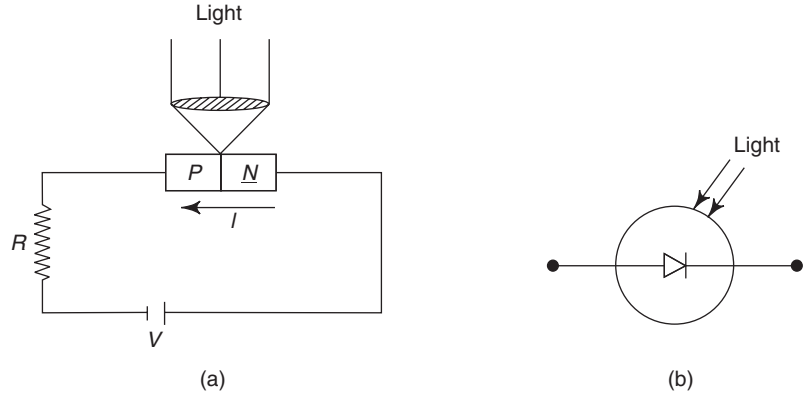


Fig. 5.30 *Photodiode: (a) Construction (b) Symbol*

The characteristics of a photodiode are shown in Fig. 5.31. The reverse current increases in direct proportion to the level of illumination. Even when no light is applied, there is a minimum reverse leakage current called *dark current*, flowing through the device. Germanium has a higher dark current than silicon, but it also has a higher level of reverse current.

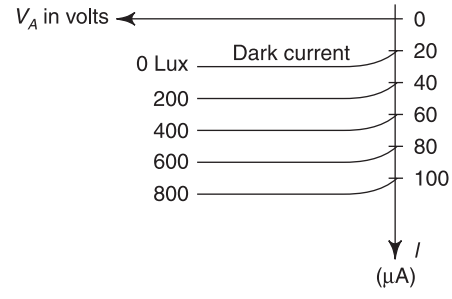


Fig. 5.31 *Classification of photodiode*

Photodiodes are used as light detectors, demodulators and encoders. They are also used in optical communication system, high speed counting and switching circuits. Further, they are used in computer card punching and tapes, light operated switches, sound track films and electronic control circuits.

Phototransistor Phototransistor or Photodiode is a much more sensitive semiconductor photodevice than the *PN* photodiode. The current produced by a photodiode is very low which cannot be directly used in control applications. Therefore, this current should be amplified before applying to control circuits. The phototransistor is a light detector which combines a photodiode and a transistor amplifier. When the phototransistor is illuminated, it permits a larger flow of current.

Figure 5.32 shows the circuit of an *NPN* phototransistor. It is usually connected in a CE configuration with the base open. A lens focuses the light on the base-collector junction. Although the phototransistor

has three sections, only two leads, the emitter and collector leads, are generally used. In this device, base current is supplied by the current created by the light falling on the base-collector photodiode junction.

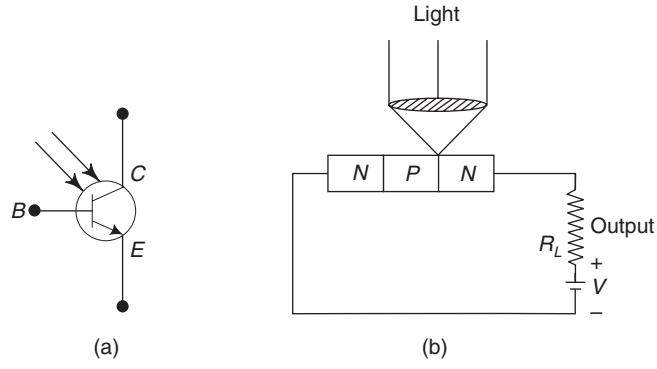


Fig. 5.32 *NPN phototransistor (a) Symbol, (b) Biasing arrangement*

When there is no radiant excitation, the minority carriers are generated thermally, and the electrons crossing from the base to the collector and the holes crossing from the collector to the base constitute the reverse saturation collector current I_{CO}. With $I_B = 0$, the collector current is given by

$$I_C = (\beta + 1)\, I_{CO}$$

When the light is turned ON, additional minority carriers are photogenerated and the total collector current is

$$I_C = (\beta + 1)(I_{CO} + I_L)$$

where I_L is the reverse saturation current due to the light.

Current in a phototransistor is dependent mainly on the intensity of light entering the lens and is less affected by the voltage applied to the external circuit. Figure 5.33 shows a graph of collector current I_C as a function of collector-emitter voltage V_{CE} and as a function of illumination H.

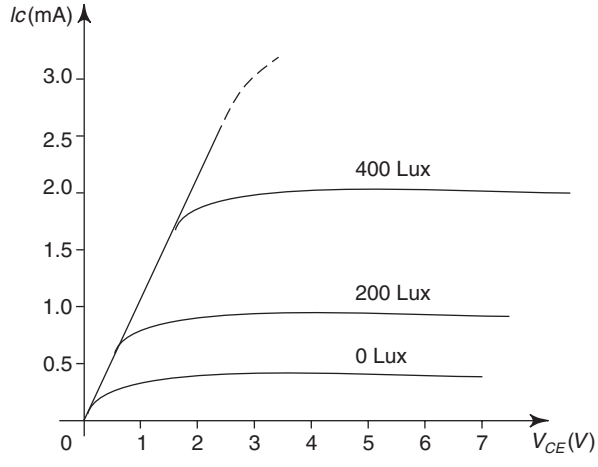


Fig. 5.33 *Characteristics of phototransistor*

The phototransistors find extensive applications in high-speed reading of computer punched cards and tapes, light detection systems, light operated switches, reading of film sound track, production line counting of objects which interrupt a light beam, etc.

5.19 OPTOCOUPLER [AU May 2017, Nov 2014, 5 marks]

An optocoupler is a solid-state component in which the light emitter, the light path and the light detector are all enclosed within the component and cannot be changed externally. As the optocoupler provides electrical isolation between circuits, it is also called *optoisolator*. An optoisolator allows signal transfer without coupling wires, capacitors or transformers. It can couple digital (ON/OFF) or analog (continuous) signals.

The schematic representation for an optocoupler appears in Fig. 5.34. The optoisolator, also referred to as an optoelectronic coupler, generally consists of an infrared LED and a photodetector such as PIN photodiode for fast switching, phototransistor Darlington pair, or photo-SCR combined in a single package. Optoisolators transduce input voltage to proportional light intensity by using LEDs. The light is transduced back to output voltage using light sensitive devices. GaAs LEDs are used to provide spectral matching with the silicon sensors.

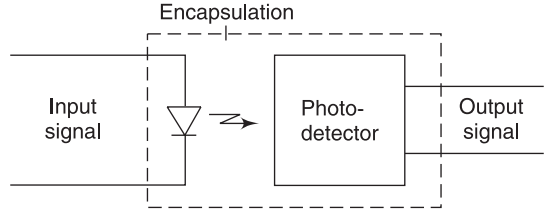


Fig. 5.34 *Schematic representation of an optocoupler*

The wavelength response of each device is made to be as identical as possible to permit the highest measure of coupling possible. There is a transparent insulating cap between each set of elements embedded in the structure (not visible) to permit the passage of light. They are designed with very small response times in such a way that they can be used to transmit data in the MHz range.

The rigid structure of this package permits one-way transfer of the electrical signal from the LED to the photodetector, without any electrical connection between the input and output circuitry. The extent of isolation between input and output depends on the kind of material in the light path and on the distance between the light emitter and the light detector. A significant advantage of the optoisolator is its high isolation resistance, of the order of $10^{11}\ \Omega$ with isolation voltages upto 2500 V between the input and output signals, and this feature allows it to be used as an interface between high voltage and low voltage systems. Application for this device includes the interfacing of different types of logic circuits and their use in level-and-position-sensing circuits.

In the optoisolator, the power dissipation of LED and phototransistor are almost equal and I_{CEO} is measured in nano-amperes. The relative output current is almost constant when the case temperature varies from 25 to 75 °C. The V_{CE} voltage affects the resulting collector current only very slightly. The

switching time of an optoisolator decreases with increased current, while for many devices it is exactly the reverse. It is only 2 μs for a collector current of 6 mA and a load resistance of 100 Ω.

The schematic diagrams for a photodiode, photo-Darlington pair and photo SCR optoisolator are illustrated in Fig. 5.35.

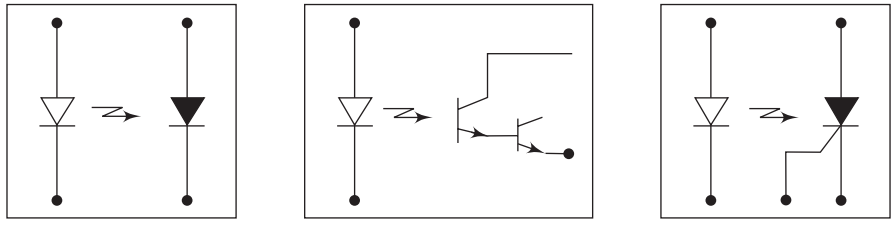


Fig. 5.35 *Optoisolators: (a) Photodiode (b) Photo-darlington (c) Photo-SCR*

5.20 PHOTOVOLTAIC CELL/SOLAR CELL

Photovoltaic effect If the *PN* junction is open circuited, the light energy is used to create a potential difference which is proportional to the frequency and intensity of the incident light. This phenomenon is called *photovoltaic effect*.

5.20.1 Photovoltaic Cell

Photovoltaic cell, a light-sensitive semiconductor device, produces a voltage when illuminated which may be used directly to supply small amounts of electric power. In the photovoltaic device without any applied voltage, the junction generates a voltage depending upon the illumination and the load. The voltage generated is due to the accumulation of carriers produced by photon excitation.

The photovoltaic potential is the voltage at which zero resultant current is obtained under open circuited conditions. The photovoltaic emf is 0.5 V for either silicon or selenium cell and 0.1 V for germanium cell the short circuit cell current is of the order of 1 mA.

The magnitude of the current under large reverse bias is given by

$$I = I_S + I_o\,(1 - e^{V/\eta V_T})$$

The photovoltaic voltage $V_{\max}$ which corresponds to an open circuited diode can be obtained by substituting $I = 0$ in the above equation. Hence,

$$V_{\max} = \eta V_T \ln\left(1 + \frac{I_S}{I_o}\right)$$

As $I_S >> I_o$, $V_{\max}$ increases logarithmically with short circuit current, I_S, and hence with illumination.

The voltage increases as the intensity of light falling on the semiconductor junction of this cell increases. A photovoltaic cell consists of a piece of semiconductor material such as silicon, germanium or selenium which is bonded to a metal plate, as shown in Fig. 5.36(a). The circuit symbol for photovoltaic cell is shown in Fig. 5.36(b).

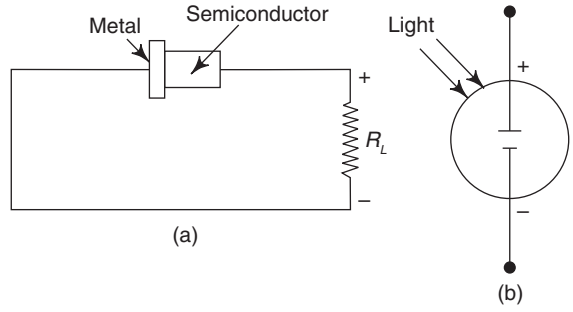


Fig. 5.36 *Photovoltaic cell (a) Construction and (b) Circuit symbol*

The spectral response of silicon, germanium and selenium are shown in Fig. 5.37, indicating that photoconductor is a frequency-selective device. As the spectral response of silicon and germanium extends well into infrared region, its efficiency is quite high. Selenium cell has two advantages over silicon, viz. (i) its spectral response is almost similar to that of the human eye, and (ii) it has the ability to withstand damaging radiation environments, lasting up to 10,000 times longer than silicon.

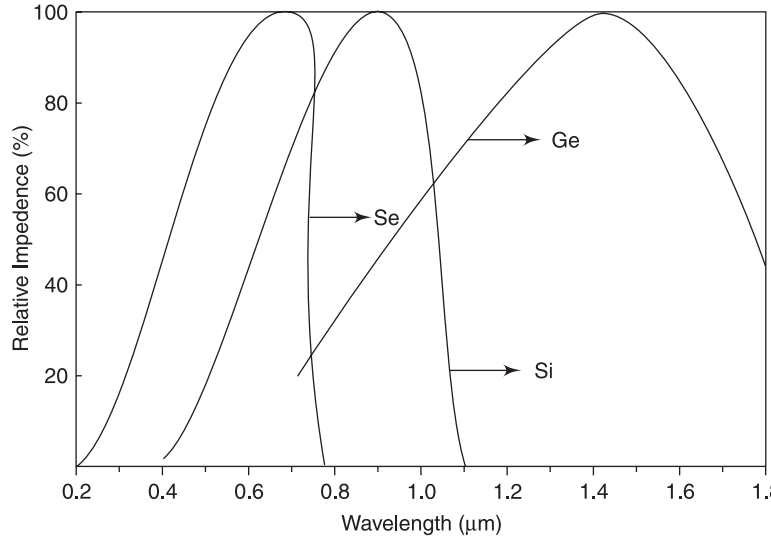


Fig. 5.37 *Spectral response of Si, Ge and Se*

The characteristic curves of output voltage versus light intensity and output current versus light intensity are shown in Figs 5.38(a) and (b), respectively.

Photovoltaic cells are used in low-power devices such as light meters. Nowadays, with an improvement in the efficiency of these cells, more power is produced, as in solar cells which are photovoltaic devices. When operated in the short-circuit mode, the current is proportional to the illumination and photovoltaic cell is used to construct a direct-reading foot-candle meter.

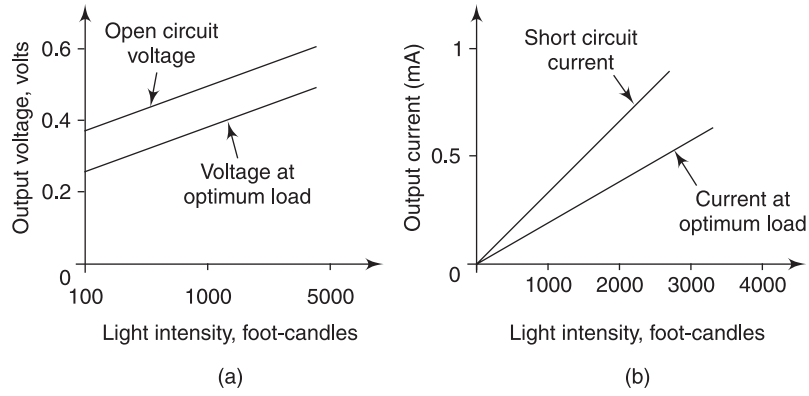


Fig. 5.38 *Characteristic of photovoltaic cell (a) Output voltage vs. light intensity (b) Output current vs. light intensity*

5.20.2 Solar Cell

[AU May 2014, 8 marks]

When sunlight is incident on a photovoltaic cell, it is converted into electric energy. Such an energy converter is called Solar cell or Solar battery and is used in satellites to provide the electrical power. This cell consists of a single semiconductor crystal which has been doped with both *P*- and *N*-type impurities, thereby forming a *PN* junction. The basic construction of a *PN* junction solar cell is shown in Fig. 5.39. Sunlight incident on the glass plate *G* passes through it and reaches the junction. An incident light photon at the junction may collide with a valence electron and impart sufficient energy to make a transition to the conduction band. As a result, an electron-hole pair is formed. The newly formed electrons are minority carriers in the *P*-region. They move freely across the junction. Similarly, holes formed in the *N*-region cross the junction in the opposite direction. The flow of these electrons and holes across the junction is in a direction opposite to the conventional forward current in a *PN* junction. Further, it leads to the accumulation of a majority carriers on both sides of the junction. This

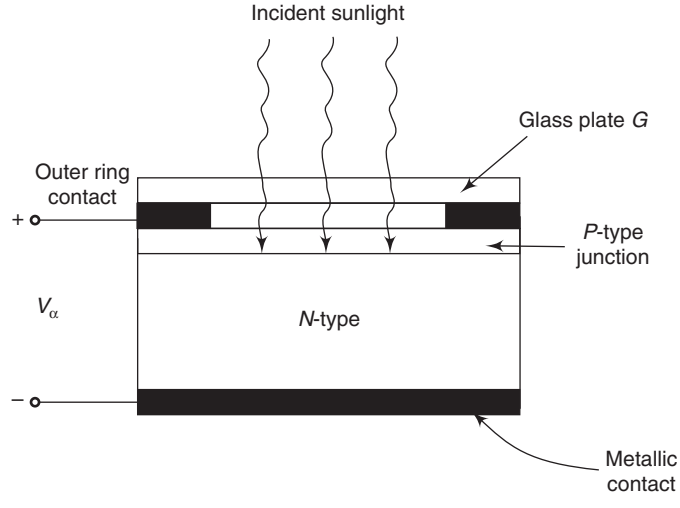


Fig. 5.39 *Basic construction of a PN junction solar cell*

gives rise to a photovoltaic voltage across the junction in the open circuit condition. This voltage is a logarithmic function of illumination.

In bright sunlight, about 0.6 V is developed by a single solar cell. The amount of power the cell can deliver depends on the extent of its active surface. An average cell will produce about 30 mW per square inch of surface, operating in a load of 4 Ω. To increase the power output, large banks of cells are used in series and parallel combinations. The efficiency of the solar cell is measured by the ratio of electric energy output to the light energy input expressed as a percentage. At present, an efficiency in the range of 10 to 40% is obtained. Silicon and selenium are the materials used widely in solar cells because of their excellent temperature characteristics.

5.21 CHARGE COUPLED DEVICE (CCD)

[AU Dec 2015, May 2015, Nov 2014 and May 2014, 8 marks]

A Charge Transfer Device (CTD) is a semiconductor structure in which discrete charge packets are removed. It finds wide applications in shift registers, imaging systems, dynamic memories and high speed filtering. There are two main methods of constructing CTD, namely, (i) Charge-Coupled Device (CDD), and (ii) Bucket Brigade Device (BBD).

5.21.1 Charge-Coupled Device (CCD)

A Charge-Coupled Device (CCD) is a shift register formed by a string of closely spaced MOS capacitors. A CCD can store and transfer analog signals, either electrons or holes, which may be introduced electrically or optically. A cross-sectional view of a three-phase charge-coupled device (CCD) is illustrated in Fig. 5.40. The structure consists of a series of metal gate electrodes, separated from a *P*- (or an *N*-) type semiconducting silicon substrate (for an *N*-channel device) by a thin silicon dioxide layer. On top of the silicon dioxide is an array of metallised electrodes which are connected to signal voltages V_1, V_2 and V_3.

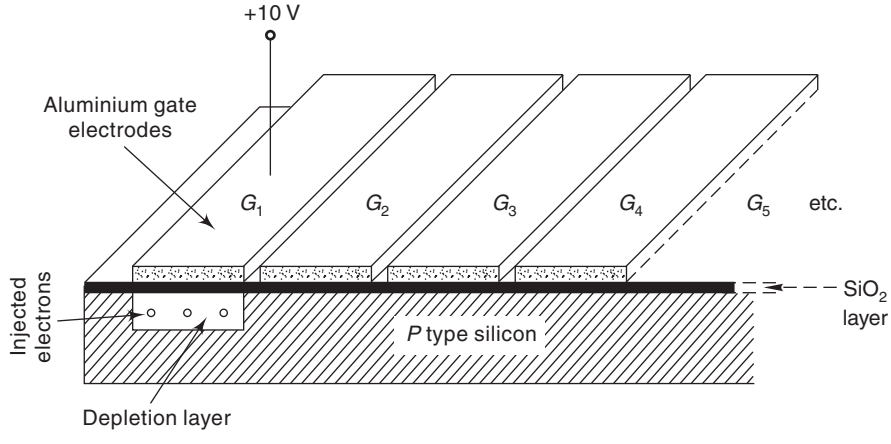


Fig. 5.40 *Basic structure of CCD*

A three-phase clocked voltage pulse system supplied to the gates ensures that the charge is transferred serially between gates and its direction is controlled, as given below. The first phase connects a positive voltage V_1, say +10 V, to G_1, where V_1 is greater than either V_2 or V_3, a depletion layer is formed, in typically less than 1 μs. This produces the potential well into which information, in the form of minority electrons, is stored. During the second phase, the adjacent gate G_2 is biased to a greater positive voltage V_2, say +15 V, to produce a deeper well under it as shown in Fig. 5.41(a). The stored charge then transfers into the deeper potential well by diffusion down the potential gradient, which incidentally can be a relatively slow process. In order to ensure the charge transfer, the potential wells must physically overlap. As depletion layers are typically only a few micrometers deep, the spacing between neighbouring gates must be as small as possible, to ensure sufficient overlap. The charge is then completely stored in the well under G_2 and hence, the voltage on G_1 is reduced to a low value, say +5 V and that on G_2 to a sustaining level of say +10 V as shown in Fig. 5.41(b).

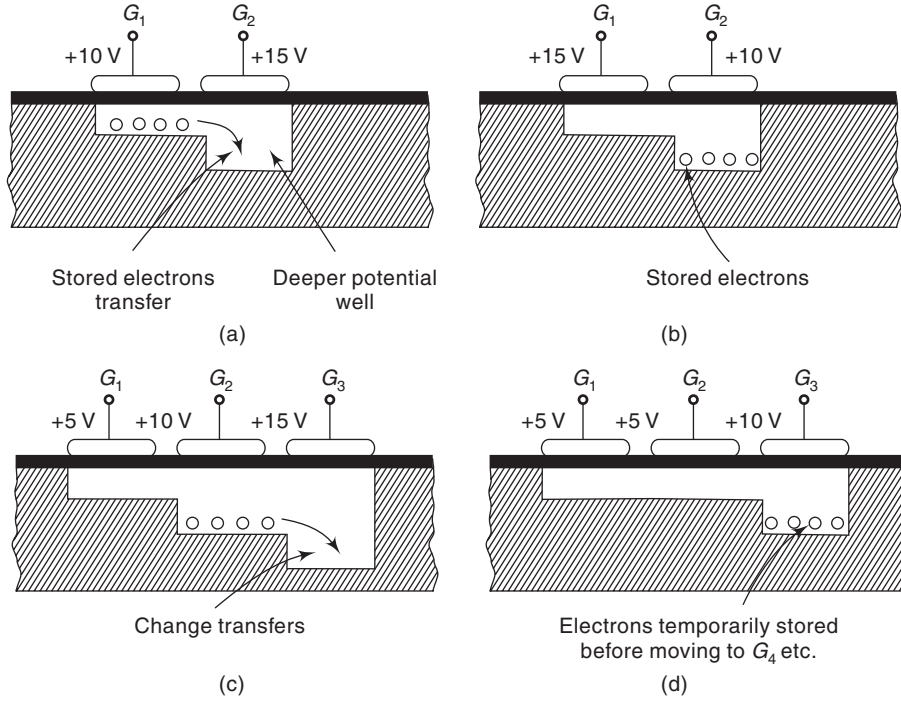


Fig. 5.41 *The mechanism of charge transfer in a three phase CCD*

A third-phase transfers a +15 V voltage pulse to the next gate G_3 and the charge is transferred from G_2 to the well under it, as shown in Fig. 5.41(c). The voltages on G_2 and G_3 can be relaxed as before as shown in Fig. 5.41(d), to complete one cycle of the clock frequency. The charge has been transferred from under G_1 to under G_3 in one cycle of the clocked three-phase pulse which causes a series of voltages in the sequence of +15, +10, +5, +15, etc., to be applied to each gate electrode. In this manner, the charge in the substrate is transferred under one electrode to the next, and so on. Here, each storage cell of three adjacent of three electrodes accommodates the *bit* of information. As soon as charge is moved out of one set of three electrodes, say from G_3 to G_4, then the input gate is again put in a state to receive

a further *bit* of information. The CCD structure thus behaves as a dynamic shift register, and charge has to be transferred to less than 1 ms. This process is only possible provided the charge is not stored in any well for long time for an inversion layer to form, in which case the charge would disappear and the corresponding information will be lost.

Applications The CCD can be used as memories by storing charge corresponding to full and empty wells (1 and 0). However, as the charge storage is limited to the storage time of the associated capacitor, the charge has to be periodically refreshed. The commercial CCDs can transfer up to 20 MHz.

In solid-state imaging, the light may be shone in the CCD, either from the top or the bottom, through a tinned substrate. There may be a separate photosensing element which is connected to the CCD. When the incident light falls on the CCD, it generates charge in the depletion region, which is proportional to the intensity of light incident at that point. This is the signal charge which can be shifted out by suitable clocking pulses for further processing. A matrix version of a CCD operated in this mode has considerable potential in the imaging field and a solid-state colour TV camera.

The CCD finds application as a dynamic shift register in computers and in solid-state imaging, such as video cameras. CCD is used in photosensor arrays and such signal processing components as variable delay lines and signal correlators. In delay lines, the time delay T_d, can be electronically varied over a wide range by varying the clock frequency f_c. For an N-bit CCD, $T_d = N/f_c$. The CCDs offer longer delays than Surface Acoustic Wave (SAW) devices. The delay line can be used as building block for analog signal processing applications like transversal filters, recursive filters, pulse compressors and pulse expanders. The packing density (the number of devices occupying the substrate) is in the order of 100 times as great as for the other semiconductor devices. The CCD consumes very little power and is capable of operating at high frequencies.

Two Mark Questions and Answers

1. **Draw the basic structure of DIAC and its symbol.** **(April/May 2017)**
   Refer to Fig. 5.16

2. **What is the advantage of TRIAC over SCR?** **(Dec 2015/Jan 2016)**
   TRIAC can act as a switch in both directions whereas SCR act as a switch in only one direction.

3. **Give the symbol, structure and equivalent circuit of DIAC.** **(Nov/Dec 2016)**
   Refer to Fig. 5.16

4. **Compare SCR with TRIAC.** **(Nov/Dec 2016)**
   TRIAC is equivalent to two SCRs connected in parallel but in the reverse direction. So, a TRIAC will act as a switch in both directions whereas SCR act as a switch in only one direction. The main limitation of TRIAC in comparison to SCR is its low power handling capacity.

5. **What is the major difference between a bipolar and unipolar device?** **(Nov/Dec 2014)**
   In Bipolar devices, the current flowing is controlled by both minority and majority charge carriers whereas in unipolar devices only the majority charge carriers are responsible for the current flow.

6. Sketch the V-I characteristics of an UJT. **(April/May 2015)**

Refer to Fig. 5.2.

7. Draw the two transistor model of an SCR with its characteristic curve. **(Nov/Dec 2010, May/June 2013 and Nov/Dec 2014)**

Refer to Fig. 5.8.

8. Draw the basic structure of TRIAC and its symbol. **(May/June 2014)**

Refer to Fig. 5.14.

9. Draw the VI characteristics of TRIAC. **(June 2011)**

Refer to Fig. 5.15.

10. Mention any two applications of DIAC. **(May/June 2014)**

(i) DIAC can be used in an oscillator circuits.

(ii) DIAC is used as a low power triggering device.

11. Compare LED and LCD. **(May/June 2013)**

Refer Table 5.3.

12. What is (a) LED (b) LCD? **(June 2011, 2013)**

(a) Light Emitting Diode (LED) is a *PN* junction device which emits light when forward biased, by a phenomenon called electroluminescence. It is often used as a light source. The colour of the emitted light depends on the type of material used as given below:

Gallium Arsenide (GaAs) – infrared radiation (invisible)

Gallium Phosphide (GaP) – red or green

Gallium Arsenide Phosphide (GaAsP) – red or yellow.

(b) Liquid Crystal Displays (LCDs) are used for display of numeric and alphanumeric character in dot matrix and seven segment displays.

13. Write the principle of operation of photodiode. **(Nov/Dec 2013)**

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

14. Write down the significance of Optocoupler. **(May/June 2014)**

Optocouplers are most often used to separate two circuit elements that are operating on extremely different voltages, which prevent damage to the part working at a lower voltage. They also work to keep the two elements from being damaged by reverse voltage or power surges. Because of these characteristics, optocouplers are best utilized in associated with on/off switches and the transfer of digital data. They are commonly found between a transmitter and a receiver in an electric circuit.

15. What is photovoltaic effect? (May/June 2014)

If the *PN* junction is open circuited, the light energy is used to create a potential difference which is proportional to the frequency and intensity of the incident light. This phenomenon is called photovoltaic effect.

16. What is meant by photovoltaic cell? (June 2010 and Nov./Dec. 2011)

A solar cell, or photovoltaic cell, is an electrical device that converts the energy of light directly into electricity by the photovoltaic effect, which is a physical and chemical phenomenon.

17. Under what principle does a photovoltaic cell work? (Nov./Dec. 2010)

The photovoltaic effect is the creation of voltage or electric current in a material upon exposure to light and is a physical and chemical phenomenon.

18. Name two applications of photoconductive cells. (May/June 2012 and Nov./Dec. 2012)

Photoconductive cells are used in different types of circuits and applications.

Analog Applications

- Camera Exposure Control
- Auto Slide Focus - dual cell
- Photocopy Machines - density of toner
- Colorimetric Test Equipment
- Densitometer
- Electronic Scales - dual cell
- Automatic Gain Control - modulated light source
- Automated Rear View Mirror

Digital Applications

- Automatic Headlight Dimmer
- Night Light Control
- Oil Burner Flame Out
- Street Light Control
- Absence / Presence (beam breaker)
- Position Sensor

19. "A solar cell is a *PN* junction device with no voltage directly applied across the junction". If it is so, how does a solar cell deliver power to a load? (April/May 2015)

Solar cell consists of a single semiconductor crystal which has been doped with both *P*- and *N*-type impurities, thereby forming a *PN* junction. The incident sunlight passes through glass plate and reaches the *PN* junction. An incident light photon at the junction may collide with the valence electron and impart sufficient energy to make a transition to the conduction band. As a result, a electron-hole pair is formed. The newly formed electrons are minority carriers in the *P*-region. They move freely across the junction. Similarly, holes formed in the *N*-region cross the junction in the opposite direction. The flow of these electrons and holes across the junction is in a direction opposite to the conventional forward current in a *PN* junction. Further, it leads to the accumulation of a majority carrier on both sides of the junction. This gives rise to a photovoltaic voltage

across the junction in the open circuit condition. In bright sunlight, about 0.6 V is developed by a single solar cell. The amount of power, the cell can deliver depends on the extent of its active surface. An average cell will produce about 30 mW per square inch of surface, operating in a load of 4 Ω.

Review Questions

1. Draw the equivalent circuit of UJT and explain its operation with the help of emitter characteristics.
2. Explain the *V–I* characteristics of a UJT.
3. Define intrinsic stand-off ratio of a UJT.
4. Explain the terms (a) peak point voltage V_P and (b) valley point voltage V_V of a UJT.
5. Mention some of the applications of UJT.
6. Explain the difference between UJT and a conventional bipolar transistor.
7. Explain with the help of a circuit diagram the working of a UJT relaxation oscillator.
8. A silicon UJT has an interbase resistance R_{BB} = 10 kΩ and R_{B1} = 6 kΩ with I_E = 0. If V_{BB} = 20 V and $V_E < V_P$, find UJT current, η and V_P. [*Ans.* 2 mA, 0.6, 12.7 V]
9. A UJT has R_{BB} = 10 Ω and R_{B_2} = 3.5 kΩ. Find its intrinsic stand-off ratio. [*Ans.* 0.65]
10. Design a UJT relaxation oscillator to generate a sawtooth waveform at a frequency of 600 Hz. Assume the supply voltae V_{BB} = 18 V; V_P = 2.9 V, V_V = 1.118 V.
11. The base one of a UJT has resistance of 4.7 kΩ and the value of intrinsic stand-off ratio of the device is 0.58. If an inter-base voltage of 10 V is applied across the two bases, calculate the value of I_B.

    [*Ans.* 1.23 mA]
12. What is a thyristor? Mention some of them.
13. Describe the operation of Shockley diode.
14. Describe the working principle of an SCR with *V–I* characteristics.
15. Draw the two transistor model of an SCR and explain its breakdown operation.
16. Explain why an SCR is operated only in the forward biased condition.
17. Explain how triggering of an SCR can be controlled by the gate signal supplied.
18. Explain the terms (a) firing angle, and (b) conduction angle of an SCR.
19. Once the SCR is triggered, the gate looses its control. Explain.
20. Explain the two transistor analogy of an SCR.
21. A half wave rectifier circuit employing an SCR is adjusted to have a gate current of 1 mA and its forward breakdown voltage is 150 V. If a sinusoidal voltage of 400 V peak is applied, determine (i) firing angle, (ii) average output voltage, (iii) average current for a load resistance of 200 Ω, and (iv) power output.

    [*Ans.* (i) 22° (ii) 122.6 V (iii) 0.613 A (iv) 75.15 W]
22. A sinusoidal voltage V_1 = 200 sin 314*t* is applied to an SCR whose forward breakdown voltage is 150 V. Determine the time during which SCR remains OFF. [*Ans.* 2.7 ms]
23. An SCR full wave rectifier is connected to 230 V, 50 Hz mains to supply AC voltage to a resistive load of 10 Ω for firing angle of 90°. Find the DC output voltage and load current. [*Ans.* 103.53 V, 10.353 A]
24. The brightness of a 100 W, 110 V lamp is to be varied by controlling firing angle of SCR full wave circuit. The RMS value of AC voltage appearing across each SCR is 110 V. Find the rms voltage and current in the lamp and firing angle of 60°. [*Ans.* 98.9 V, 0.82 A]
25. What is a TRIAC? Sketch its characteristics and describe its operation.

26. DIAC is a bidirectional device. Explain.
27. Draw the *V–I* characteristics of a DIAC and explain its working principle.
28. What is the advantage of TRIAC over SCR?
29. Define Safe Operating Area (SOA) in power transistors.
30. Describe briefly about power BJT and power MOSFET.
31. Explain in detail about DMOS structure.
32. Write briefly about VMOS structure and its characteristics.
33. Describe with the help of a relevant diagram, the construction of an LED and explain its working.
34. List the applications of an LED.
35. Compare the working principle of LED with solar cell.
36. Describe the principle of operation of an LCD.
37. What are the advantages and disadvantages of LCD?
38. What are the relative advantages of LCD over LED?
39. In what respect is an LED different from an ordinary *PN* junction diode?
40. Draw and explain a seven-segment LED display.
41. Draw and explain a dotmatrix (35 elements) LED display.
42. Explain the principle and working of photodiode.
43. Write the equation for the volt-ampere characteristics of a photodiode. Define each term in the equation.
44. Explain the volt-ampere characteristics of a semiconductor photodiode.
45. List the applications of a photodiode.
46. Describe with neat diagrams the operation of a phototransistor and state its applications.
47. With output characteristics, explain how phototransistor responds to the incident light.
48. Explain the difference between a photodiode and photovoltaic cell.
49. Give the constructional features of a solar cell and explain its principle of operation. Mention two of its applications.
50. Distinguish between photovoltaic cell and solar cell.
51. List the applications of optocoupler.
52. Explain the construction and operation of a three phase charge coupled device with necessary diagrams.
53. What are the applications of CCD?