2e

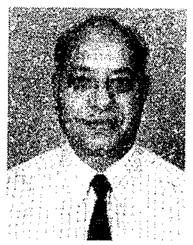
Basic Electronics and Linear Circuits

N N Bhargava
D C Kulshreshtha
S C Gupta

NITTTR Chandigarh

About the Authors

N N Bhargava was Assistant Professor in electronics at TTTI, Chandigarh and was the one who initiated and coordinated the writing of this book. After completing his MSc (Tech.) in electronics engineering from the Birla Institute of Technology and Science, Pilani, he taught electronics to undergraduate students at the University of Roorkee and at the Delhi College of Engineering.



D C Kulshreshtha is working as Professor in the department of Electronics and Communication Engineering, Jaypee University of Information Technology, Waknaghat, Solan (HP) since July 2003. He served for more than 25 years in Delhi College of Engineering, Delhi. From time to time, he has been a visiting faculty member at Netaji Subhash Institute of Technology, New Delhi; Institute of Technology and Management, Gurgaon; Guru Tegh Bahadur College of Engineering, New Delhi; Maharaja Agrasen College of Engineering, Delhi; Galgotia College of Engineering and Technology, Greater Noida and IEC College of Engineering and Technology, Greater Noida.

Dr Kulshreshtha obtained his BTech (Hons) in Electronics and Electrical Communication Engineering from IIT Kharagpur in 1967, and a master's degree in Advanced Electronics from Delhi College of Engineering, Delhi in 1980. For three years, he was at IIT Delhi, pursuing research in Digital Signal Processing. He secured second rank in the All India Merit in Engineering Services Examination conducted by the UPSC. He was also in-charge of Frequency Section, responsible for the management of electromagnetic spectrum on Station, Srinagar (J&K), and was responsible for installation of equipments and antenna system for the monitoring station.

He has conducted or attended short-term refresher courses in Microelectronics at IIT Kharagpur; Measurement Techniques at University of Roorkee; Modern Process Control Techniques at DCE, Delhi; Digital Signal Processing at REC, Warangal; and Microwave Techniques and Systems at REC, Surathkal.

He has authored/co-authored many textbooks namely, *Electronic Devices, Applications and Integrated Circuits* (1980), *Basic Electronics and Linear Circuits* (1984), *Engineering Network Analysis* (1989), *Elements of Electronics and Instrumentation* (1992), *Electronic Devices and Circuits* (2005) and *Electronics Engineering* (for UPTU) (2006).

Subhash Chandra Gupta is Dean (Academics) and Assistant Professor in the Department of Computer Science at the Hindustan Institute of Technology and Management, Keetham, Agra. After acquiring his diploma degree Electrical Engineering, Prof. Gupta started teaching diploma students in 1969. He received his Diploma in Technical Teaching from TTTI, Chandigarh in 1974 and thereafter, served various polytechnics in UP as Assistant Lecturer and Lecturer. After completing AMIE (I) in EC branch in 1981, he served as HOD, Electronics Engineering in various Government Polytechnics in UP for several years. He also worked as Assistant Director, Technical Education, UP for five years. After completing his ME, Computer Science and Engineering in 1998, he was appointed as HOD, Computer Engineering in Dr. Ambedkar Institute of Technology for Handicapped (AITH), UP, Kanpur. Prof. Gupta retired as Principal in a Government Polytechnic. His areas of expertise include subjects like Electrical, Electronics and Computer science disciplines.

BASIC ELECTRONICS AND LINEAR CIRCUITS

Second Edition

(Late) N N Bhargava
National Institute of Technical Teachers' Training and Research
Chandigarh

D C Kulshreshtha
Jaypee University
Solan

S C Gupta
Hindustan Institute of Technology and Management
Mathura

McGraw Hill Education (India) Private Limited
NEW DELHI

McGraw Hill Education Offices

New Delhi 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

CONTENTS

PREFACE

About the Book

The Curriculum Development Center (CDC), National Institute of Technical Teachers' Training and Research (NITTTR), Chandigarh, initiated the task of developing basic instructional material for the polytechnics in 1980. A team of authors worked very hard to prepare the material in a very simple language to clarify each and every concept. Finally, the first edition of this book was published in 1984 by McGraw Hill Education (India). This book proved to be extremely useful and appropriate support material for students as well as teachers. Hence, it was widely accepted as a basic textbook.

Early development in the field of electronics used vacuum tube devices. But in 1947, the wonder device, called *transistor*, was invented. Soon people started using this tiny device in place of bulky vacuum tubes. However, these semiconductor devices had some limitations; they could not work at high power and high frequencies. Therefore, the use of vacuum tube devices continued in many applications. At the time when the text of first edition was developed, both vacuum tubes and semiconductor devices were in use. The knowledge of vacuum tubes was as important as that of semiconductor devices.

However, with the rapid development in the field of electronics in last two decades, semiconductor devices are now available for almost all applications. The final blow to phase out the vacuum tubes came only a few years back, when voluminous CRT TVs were replaced by sleek, large-screen, wall-mountable LCD, LED or plasma TVs.

This book is the result of above-mentioned changes. We hope that the current edition enjoys as much popularity as the first one.

Salient Features of the Book

- Basic principles developed without recourse to advance mathematics
- Emphasis on semiconductor devices and applications
- Simplistic presentation of the fundamentals of electronics
- Each unit beginning with *Objectives*, thus, enabling students to know what is expected out of them after going through the unit
- Use of illustrations to make complex concepts easily understandable
- Lucid language to provide a smooth reading
- Solved examples included to help students in applying the principles in practice

1 UNIT
INTRODUCTION TO ELECTRONICS

"When I was a teenager in the late 30s and early 40s, electronics wasn't a word. You were interested in radio if you were interested in electronics."

Ken Olsen (1926–2011)

American Engineer and Co-founder of Digital Equipment Corporation in 1957

After completing this unit, students will be able to:

- define the scope of electronics
- state some of the applications of electronics in day-to-day life
- state the latest trends in the field of electronics
- draw the symbols, and state the main applications of some of the important *active devices* such as transistors, FETs, SCR, UJT, etc.
- recognise resistors, capacitors and inductors of various types from their physical appearance
- read the values of resistors and capacitors from the code marks usually found on such components
- draw the symbol of different passive components as per ISI specifications
- explain the important specifications of resistors, capacitors and inductors
- write SI units of various physical quantities used in electronics

1.1 WHAT IS ELECTRONICS?

The word 'electronics' is derived from *electron* mechan*ics* which means the study of the behaviour of an electron under different conditions of externally applied fields.

The Institution of Radio Engineers (IRE) has given a standard definition of electronics in the *Proceedings of IRE*, Vol. 38, (1950) as "that field of science and engineering, which deals with electron devices and their utilisation." Here, an electron device is "a device in which conduction takes place by the movement of electrons—through a vacuum, a gas or a semiconductor".

Compared to the more established branches such as civil, mechanical, electrical, etc., electronics is a newcomer in the field of engineering. Until recently, it was considered an integral part of electrical engineering, but due to the tremendous advancement during the last few decades, it has now gained its rightful place.

We shall study, in the chapters that follow, how electronic devices function, and how they could be used to advantage in our daily life.

1.2 APPLICATIONS OF ELECTRONICS

Life today offers many conveniences which involve the use of electronic devices. As can be seen from Table 1.1, electronics plays a major role in almost every sphere of our life.

1.2.1 Communications and Entertainment

The progress of a country depends upon the availability of economical and rapid means of communication. During the earlier part of last century, the main application of electronics was in the field of telegraphy and telephony. This utilises a pair of wires. However, it is now possible with the help of radio waves to transmit any message from one place to another, thousands of kilometres away, without any wires. With such *wireless* communication (radio broadcasting), people in any part of the world can know what is happening in other parts. With the help of a teleprinter, it is possible to type the message on a typewriter kept in another city. Photographs of events occurring somewhere, can be transmitted on facsimile (radiophoto). They can then be printed in the newspapers all over the world.

Radio and TV broadcasting provide a means of both communication as well as entertainment. With the help of satellites it has become possible to establish instant communication between places very far apart. Electronic gadgets like CD players, stereo systems, video games, public-address systems, etc., are widely used for entertainment.

1.2.2 Defence Applications

One of the most important developments during World War II was the RADAR (which is the short form for 'RAdio Detection And Ranging'). By using radar, it is possible not only to detect, but also to find the exact location of the enemy aircraft. The anti-aircraft guns can then be accurately directed to shoot down the aircraft. In fact, the radar and the anti-aircraft guns can be linked by an automatic control system to make a complete unit.

Table 1.1 *Applications of electronics*

ELECTRONICS USED IN

- ***Communications and Entertainment***
  - *Line Communication* → Telegraphy, Telephony, Telex, Teleprinters
  - *Wireless Communications* → Radio broadcasting, TV broadcasting, Facsimile (radiophoto), Satellite communication
  - *Audio Systems* → PA systems, Stereo amplifiers, CD players, Tape recorders, i-pods
- ***Defence*** → Rader, Guided missiles, Coded communication
- ***Industry*** → Automatic control systems, Heating and welding systems, Computers (digital and analogue)
- ***Medical Sciences*** → X-rays, Electrons microscope, Electrocardiogram (ECG), Electrotherapy
- ***Instrumentation*** → Precision measuring instruments, e.g., DVM, CRO, Frequency counters, pH-meter, Strain gauge, etc.

Guided missiles are completely controlled by electronic circuits. In a war, success or defeat for a nation depends on the reliability of its communication system. In modern warfare, communication is almost entirely electronic.

1.2.3 Industrial Applications

Use of automatic control systems in industries is increasing day by day. Electronic circuits are used in industrial applications like control of thickness, quality, weight and moisture content of a material. Electronic amplifier circuits are used to amplify signals and thus control the operations of automatic door-openers, lighting systems, power systems and safety devices, etc. Electronic circuits are used to produce stroboscopic lights of any desired frequency. When this is directed on a fast rotating object, it can be made to appear stationary or to be in slow motion by adjusting the frequency of light. This principle makes it possible to study the movement of various parts of a machine under normal running conditions.

For quick arithmetical calculations, desk calculators are commonly used in banks, departmental stores, etc. calculators are sometimes used in classrooms while solving problems. Electronic computers, also called 'electronic brains', are used for automatic record keeping and solving of complicated problems.

Electronically controlled systems, using suitable timers, are used for heating and welding in the industry. Even the power stations, which generate thousands of megawatts of electricity are controlled by tiny electronic devices and circuits.

1.2.4 Medical Sciences

Doctors and scientists are constantly finding new uses for electronic systems in the diagnosis and treatment of various diseases. Some of the instruments which have been in use are:

1. *X-rays*, for taking pictures of internal bone structures and also for treatment of some diseases.
2. *Electrocardiographs* (*ECG*), to find the condition of the heart of a patient.
3. *Short-wave diathermy units*, for healing sprains and fractures.
4. *Oscillographs* for studying muscle action.

The use of electronics in medical science has expanded so enormously as to start a new branch of study, called 'bioelectronics'. Electronics is proving useful in saving mankind from a lot of suffering and pain.

1.2.5 Instrumentation

Instrumentation plays a very important role in any industry and research organisation, for precise measurement of various quantities. It is only due to electronic instruments that an all-round development in every walk of life has been possible. DVM, cathode-ray oscilloscopes, frequency counters, signal generators, pH-meters, strain-gauges, etc., are some of the electronic instruments without which no research laboratory is complete.

1.3 MODERN TRENDS IN ELECTRONICS

The real beginning in electronics was made in 1906, when Lee De Forest invented the *vacuum triode*. Without this device, the amplifier (which is the heart of all intricate and complex electronic gadgets) would not have been possible. Until the end of World War II, vacuum tubes (valves) dominated the field of electronics.

In 1948, the invention of the transistor by three Nobel laureates—John Bardeen, Walter Brattain and William Shockley at the Bell Laboratory, completely revolutionised the electronics industry. Transistors opened the floodgate to further developments in electronics. Within almost 10 years of its discovery, the process of miniaturisation of electronic equipments had gained momentum. The first integrated circuits (ICs) appeared in the market during the early sixties. Man's desire to conquer space accelerated this growth even further. The electronic age had truly begun. During the eighties, this tremendous growth rate not only continued but also accelerated with each passing year. The use of valves nearly became obsolete during the sixties. Due to the rapid developments in integrated circuit technology—starting from the small scale integration (SSI), then medium scale integration (MSI), large scale integration (LSI) and now with the most recent, very large scale integration (VLSI) technique—even the use of individual transistors is becoming unnecessary. The vast changes that have taken place during the last 30 years can best be understood by noting the reduction in size and price of modern digital computers. A small, modern minicomputer is more than 100 times smaller in size and 1/100th of the price of a computer designed 30 years ago to do similar jobs.

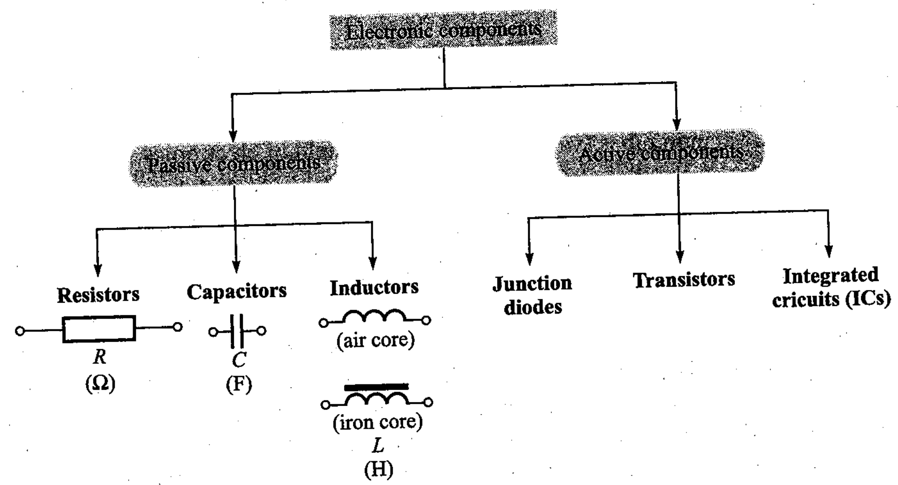
From an ordinary wristwatch to the control room of 400 000 tonne supertanker carrying cargo across the sea; from the telephone repeaters buried deep under the ocean to the spaceships far out in space; from a modern household to the gigantic steel mills and powerhouses, electronics has penetrated everywhere.

Electronics deals in the *micro* and *milli* range of voltage, current and power, but it is capable of controlling *kilo* and *mega* volts, amperes and watts. Therefore, it is not surprising to find the fundamentals of electronics as a core subject in all branches of engineering nowadays.

1.4 ELECTRONIC COMPONENTS

An electronic circuit may appear quite complicated and may be capable of performing fantastic functions. But, all electronic circuits, however complicated, contain a few basic components. Generally speaking, there are only five components—three passive and two active (see Table 1.2). An integrated circuit (for example, a microprocessor) may contain thousands of transistors, a few thousand resistors, etc., on a very small chip. The total number of components used in an electronic circuit may run into thousands—yet each component will be one of the above *five* types.

Table 1.2 *Types of electronic components*



1.4.1 Passive Components

Resistors, capacitors and inductors are called *passive components*. These components by themselves are not capable of amplifying or processing an electrical signal. However, these components are as important, in an electronic circuit, as active (such as transistors) components are. Without the aid of these components a transistor cannot be made to amplify signals.

Resistors The flow of charge (or current) through any material, encounters an opposing force similar in many respects to mechanical friction. This 'opposing force' is called the *resistance* of the material. It is measured in *ohms*, for which the symbol is Ω (the greek capital letter omega). The circuit symbol for resistance (R) is shown in Table 1.2.

In some parts of an electronic circuit, resistance is deliberately introduced. The device or component to do this is called a *resistor*. Resistors are made in many forms. But all belong to either of two groups—*fixed* or *variable*.

1. Fixed resistors The most common of the low wattage, fixed-type resistors is the *moulded-carbon composition resistor*. The basic construction is shown in Fig. 1.1. The resistive material is of carbon-clay composition. The leads are made of tinned

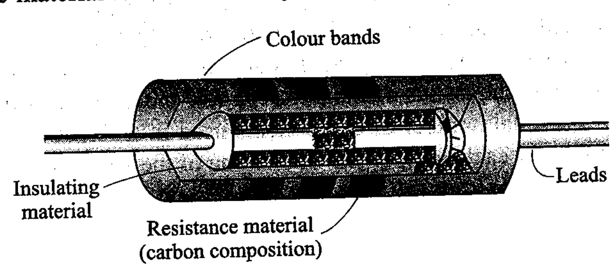


Fig. 1.1 *The basic construction of a fixed, moulded-carbon composition resistor*

copper. Resistors of this type are readily available in values ranging from a few ohms to about 22 MΩ, having a tolerance range of 5 to 20 %. They are quite inexpensive. A resistor may cost only a rupee.

The relative sizes of all fixed (and also variable) resistors change with the wattage (power) rating. The size increases for increased wattage rating in order to withstand higher currents and dissipation losses. The relative sizes of moulded-carbon composition resistors for different wattage ratings are shown in Fig. 1.2.

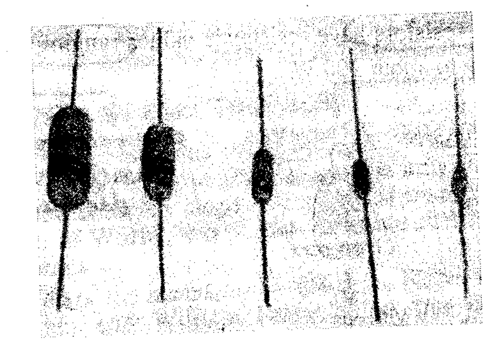


Fig. 1.2 *Moulded-carbon composition resistors of different wattage ratings*

Another variety of carbon composition resistors is the metallised type. Its basic structure is shown in Fig. 1.3. It is made by depositing a homogeneous film of pure carbon (or some metal) over a glass, ceramic or other insulating core. The carbon film can be deposited by pyrolysis of some hydrocarbon gas (e.g., benzyne) on the ceramic core. Only approximate values of resistance can be obtained by this method. Desired values are obtained by either trimming the layer thickness or by cutting helical grooves of suitable pitch along its length. During this process, the value of resistance is monitored constantly. The cutting of grooves is stopped as soon as the desired value of resistance is obtained. Contact caps are fitted on both ends. The lead wires, made of tinned copper, are then welded to these end caps. This type of film-resistor is sometimes called *precision type*, since it can be obtained with an accuracy of ±1 %.

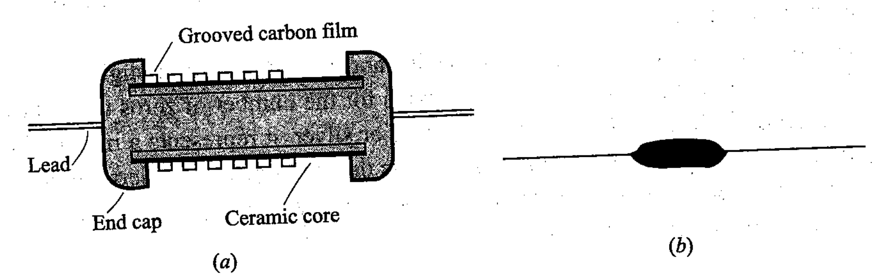


Fig. 1.3 *Carbon-film resistor: (a) Construction; (b) A carbon-film resistor*

A *wire-wound resistor* uses a length of resistance wire, such as nichrome. This wire is wound onto a round, hollow porcelain core. The ends of the winding are attached to metal pieces inserted in the core. Tinned copper wire leads are attached to these metal pieces. This assembly is coated with an enamel containing powdered glass. It is then heated to develop a coating known as *vitreous enamel*. This coating is very smooth and gives mechanical protection to the winding. It also helps in conducting heat away from the unit quickly. In other wire-wound resistors, a ceramic material is used for the inner core and the outer coating (see Fig. 1.4). Commonly available wire-wound resistors have resistance values ranging from 1 Ω to 100 kΩ, and wattage ratings up to about 200 W.

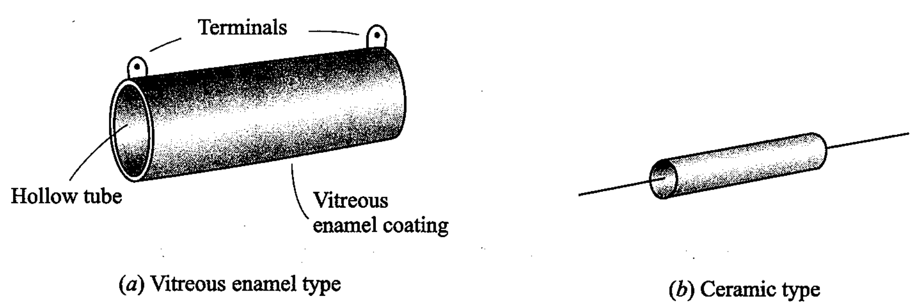


Fig. 1.4 *Wire-wound fixed resistors*

Colour coding and standard resistor values Some resistors are large enough in size to have their resistance (in Ω) printed on the body. However, there are some resistors that are too small in size to have numbers printed on them. Therefore, a system of *colour coding* is used to indicate their values. For the fixed, moulded composition resistor, four colour bands are printed on one end of the outer casing as shown in Fig. 1.5*a*.

The numerical value associated with each colour is indicated in Table 1.3. The colour bands are always read left to right from the end that has the bands closest to it, as shown in Fig. 1.5*a*.

The first and second bands represent the first and second significant digits, respectively, of the resistance value. The third band is for the number of zeros that follow the second digit. In case the third band is gold or silver, it represents a multiplying factor of 0.1 or 0.01. The fourth band represents the manufacturer's tolerance. It is a measure of the precision with which the resistor was made. If the fourth band is not present, the tolerance is assumed to be ±20 %.

Table 1.3 *Colour coding*

Colour	*Digit*	*Multiplier*	*Tolerance*
Black	0	$10^0 = 1$	
Brown	1	$10^1 = 10$	
Red	2	10^2	
Orange	3	10^3	
Yellow	4	10^4	
Green	5	10^5	
Blue	6	10^6	
Violet	7	10^7	
Gray	8	10^8	
White	9	10^9	
Gold	–	$0.1 = 10^{-1}$	±5 %
Silver	–	$0.01 = 10^{-2}$	±10 %
No colour	–	–	±20 %

Mnemonics: As an aid to memory in remembering the sequence of colour codes given above, the student can remember the following sentence (all the capital letters stand for colours):

(*a*) Bill Brown Realised Only Yesterday Good Boys Value Good Work.
(*b*) Bye Bye Rosie Off You Go Bristol Via Great Western.

Example 1.1 A resistor has a colour band sequence: yellow, violet, orange and gold. Find the range in which its value must lie so as to satisfy the manufacturer's tolerance.

Solution: With the help of the colour coding table (Table 1.3), we find

1st band	*2nd band*	*3rd band*	*4th band*
Yellow	Violet	Orange	Gold
4	7	10^3	±5 % = 47 kΩ ± 5 %

$$\text{Now, 5 \% of 47 k}\Omega = \frac{47 \times 10^3 \times 5}{100}\ \Omega = 2.35\ \text{k}\Omega$$

Therefore, the resistance should be within the range 47 kΩ ± 2.35 kΩ, or between **44.65 kΩ** and **49.35 kΩ**.

Example 1.2 A resistor has a colour band sequence: gray, blue, gold and gold. What is the range in which its value must lie so as to satisfy the manufacturer's tolerance?

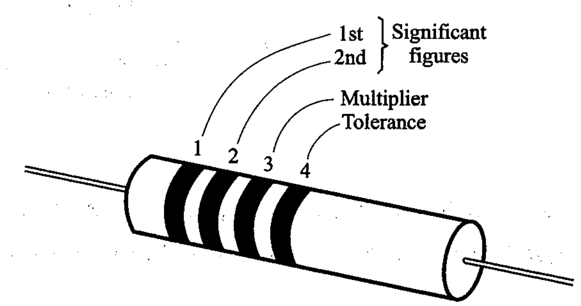
Solution: The specification of the resistor can be found by using the colour coding table as follows:

1st band	*2nd band*	*3rd band*	*4th band*
Gray	Blue	Gold	Gold
8	6	10^{-1}	$\pm 5\ \% = 86 \times 0.1\ \Omega \pm 5\ \%$
			$= 8.6\ \Omega \pm 5\ \%$

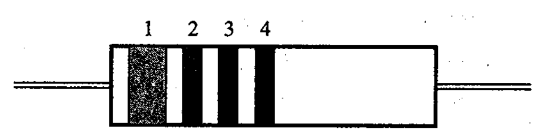
$$5\ \%\ \text{of}\ 8.6\ \Omega = \frac{8.6 \times 5}{100} = 0.43\ \Omega$$

The resistance should lie somewhere between the values (8.6 – 0.43) Ω and (8.6 + 0.43) Ω or **8.17 Ω** and **9.03 Ω**.

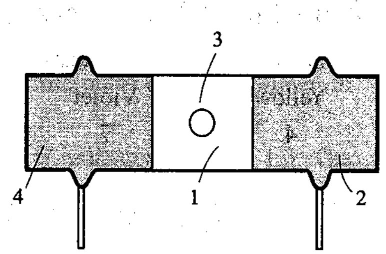
The colour coding for wire-wound resistors, and composition resistors with radial leads is shown in Figs. 1.5*b* and *c*, respectively. Note that the first band in Fig. 1.5*b* is of double the width compared to the rest. The system of colour coding used for the moulded resistors with radial leads is called *body-end-dot system*. The numerical values associated with each colour is the same for all the three methods of colour coding.



(*a*) Moulded composition resistor



(*b*) Wire-wound resistor



(*c*) Moulded composition resistor with radial leads

Fig. 1.5 *Colour coding*

In practical electronic circuits, the values of the resistors required may lie within a very wide range (say, from a few ohms to about 20 MΩ). In most of the circuits, it is not necessary to use resistors of exact values. Even if a resistor in a circuit has a value which differs from the desired (designed) value by as much as 20 %, the circuit still works quite satisfactorily. Therefore, it is not necessary to manufacture resistors of all the possible values. A list of readily available *standard values of resistors* appears in Table 1.4.

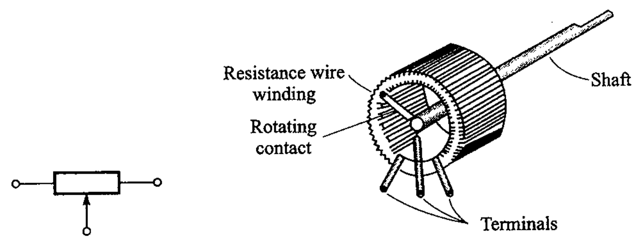
Table 1.4 *Standard values of commercially available resistors (having 10 % tolerance)*

Ohms (Ω)			*Kilohms* (kΩ)			*Megohms* (MΩ)	
1.0	10	100	1.0	10	100	1.0	10
1.2	12	120	1.2	12	120	1.2	12
1.5	15	150	1.5	15	150	1.5	15
1.8	18	180	1.8	18	180	1.8	18
2.2	22	220	2.2	22	220	2.2	22
2.7	27	270	2.7	27	270	2.7	
3.3	33	330	3.3	33	330	3.3	
3.9	39	390	3.9	39	390	3.9	
4.7	47	470	4.7	47	470	4.7	
5.6	56	560	5.6	56	560	5.6	
6.8	68	680	6.8	68	680	6.8	
8.2	82	820	8.2	82	820	8.2	

If resistors of very precise values are required for some specific application, special requests are to be made to the manufacturer.

2. Variable resistors In electronic circuits, sometimes it becomes necessary to adjust the values of currents and voltages. For example, it is often desired to change the volume (or loudness) of sound, the brightness of a television picture, etc. Such adjustments can be done by using variable resistors.

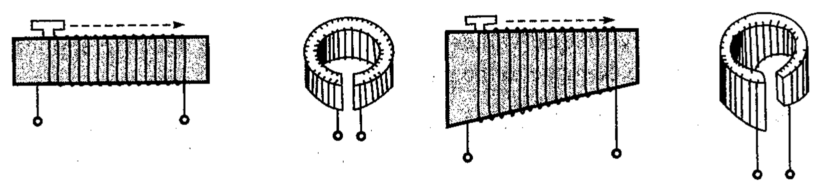
Although the variable resistors are usually called *rheostats* in other applications, the smaller variable resistors commonly used in electronic circuits are called *potentiometers* (usually abbreviated to 'pots'). The symbol for potentiometer is shown in Fig. 1.6*a*. The arrow in the symbol is a contact movable on a continuous resistive element. The moving contact will determine whether the resistance in the circuit is minimum (0 Ω) or maximum value, *R*. The construction of all potentiometers is basically the same. Some have a wire-wound resistance as their primary element, while others have a carbon-film element. The basic construction of a wire-wound potentiometer is shown in Fig. 1.6*b*. The resistance wire is wound over a dough-shaped core of bakelite or ceramic. There is a rotating shaft at the centre of the core. The shaft moves an arm and a contact point from end to end of the resistance element. There are three terminals coming out of a potentiometer. The outer two are the end points of the resistance element and the middle leads to the rotating contact.



(a) Symbol (b) Basic construction of a wire-wound potentiometer

Fig. 1.6 *Potentiometer*

A potentiometer can be either *linear* or *non-linear*. Figure 1.7 shows the construction of both a linear and non-linear (tapered type) potentiometer. In the linear type, the former (the part over which the wire is wound) is of uniform height and that is why the resistance varies linearly with the rotation of the contact. In a non-linear potentiometer, the height of the former is not uniform. To make a potentiometer of this type, a tapered strip is taken and the resistance wire is wound over it, ensuring a uniform pitch. The strip is then bent into a round shape. The tapered strip gives a non-linear variation of resistance with the rotation of the moving contact. The strip can be tapered suitably so as to obtain a desired variation in resistance per unit rotation of moving contact. The 'pots' used as volume control in sound equipment are generally of the non-linear type (logarithmic variation).



(a) Linear type (b) Non-linear type

Fig. 1.7 *Wire-wound potentiometer*

Capacitors Capacitors of different kinds are found in nearly every electronic circuit. A capacitor is basically meant to store electrons (or electrical energy), and release them whenever desired. The circuit symbol of a capacitor is shown in Table 1.2. *Capacitance* is a measure of a capacitor's ability to store charge. It is measured in farads (F). However, the unit farad being too large, practical capacitors are specified in microfarads (μF), or picofarads (pF).

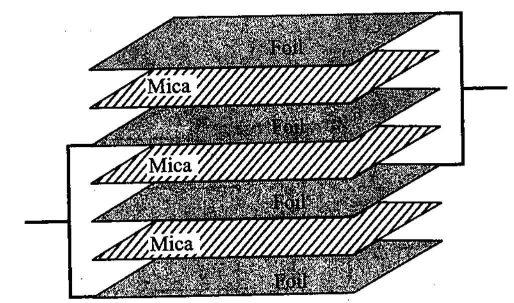
A capacitor offers low impedance to ac, but very high impedance to dc. So, capacitors are used when we want to couple alternating voltage from one circuit to another, while at the same time blocking the dc voltage from reaching the next circuit. In this role, the capacitor is called a *coupling capacitor* or a *blocking capacitor*. It is also used as a *bypass capacitor*, where it bypasses the ac through it without letting the ac to go through the circuit across which it is connected. Also, a capacitor forms a tuned circuit in series or parallel with an inductor.

A capacitor consists of two conducting plates, separated by an insulating material known as *dielectric*. Since the two plates of a capacitor can be of many different

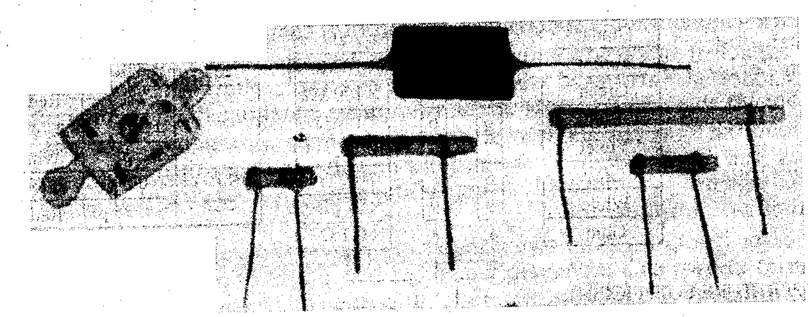
conducting materials and the dielectric may be of many different insulating materials, there are many types of capacitors.

Capacitors, like resistors, can either be fixed or variable. Some of the most commonly used fixed capacitors are mica, ceramic, paper and electrolytic. Variable capacitors are mostly air-gang capacitors.

1. Mica capacitors Mica capacitors are constructed from plates of aluminium foil separated by sheets of mica as shown in Fig. 1.8. The plates are connected to two electrodes. The mica capacitors have excellent characteristics under stress of temperature variations and high voltage applications. Available capacitances range from 5 to 10 000 pF. Mica capacitors are usually rated at 500 V. Its leakage current is very small ($R_{leakage}$ is about 1000 MΩ).



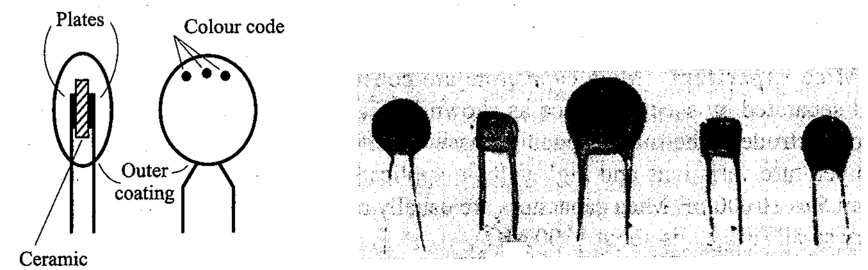
(a) Construction



(b) Mica capacitors

Fig. 1.8 *Mica capacitors*

2. Ceramic capacitors Ceramic capacitors are made in many shapes and sizes. However, the basic construction is the same for each. A ceramic disc is coated on two sides with a metal, such as copper or silver. These coatings act as the two plates (see Fig. 1.9). During the manufacture of the capacitor, tinned wire leads are also attached to each plate. Then the entire unit is coated with plastic and marked with its capacitance value—either using numerals or a colour code. The colour coding is similar to that used for resistances. Figure 1.10 explains the colour code used for resistors and capacitors. Besides the value of the resistor (or capacitor), it also indicates the tolerance and temperature coefficients. Ceramic capacitors are very versatile. Their working voltage ranges from 3 V (for use in transistors) up to 6000 V. The capacitance ranges from 3 pF to about 2 μF. Ceramic capacitors have a very low leakage currents ($R_{leakage}$ is about 1000 MΩ) and can be used in both dc and ac circuits.



(*a*) Construction (*b*) Ceramic capacitors

Fig. 1.9 *Ceramic capacitor*

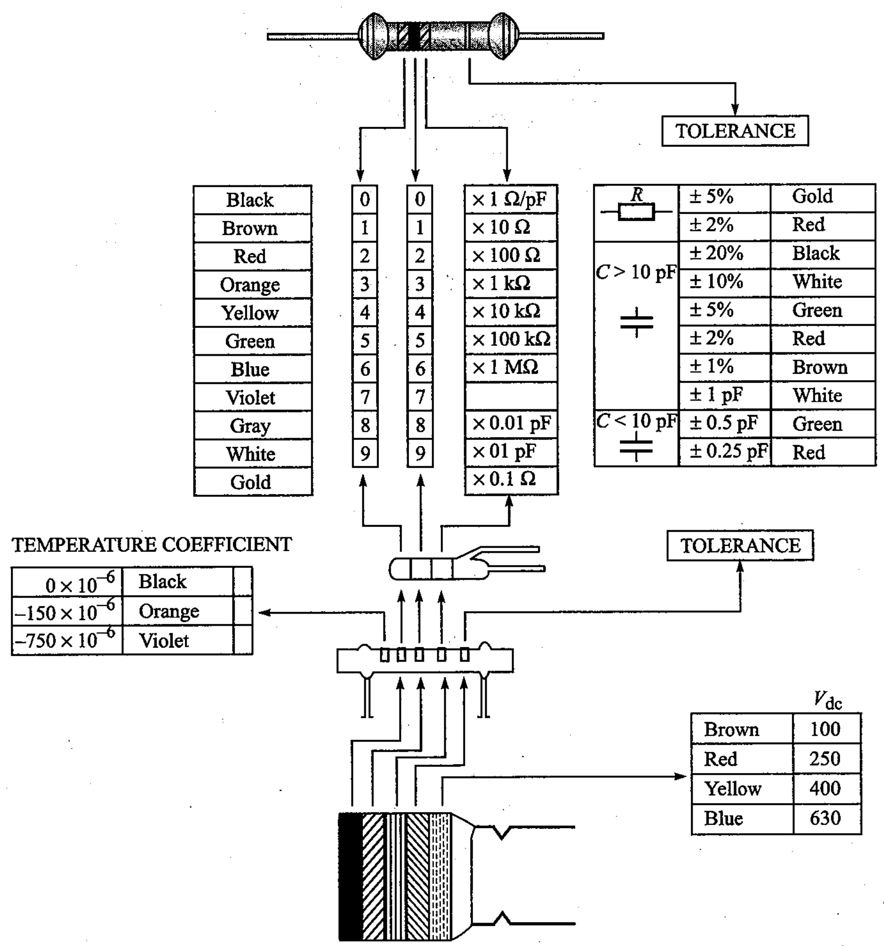
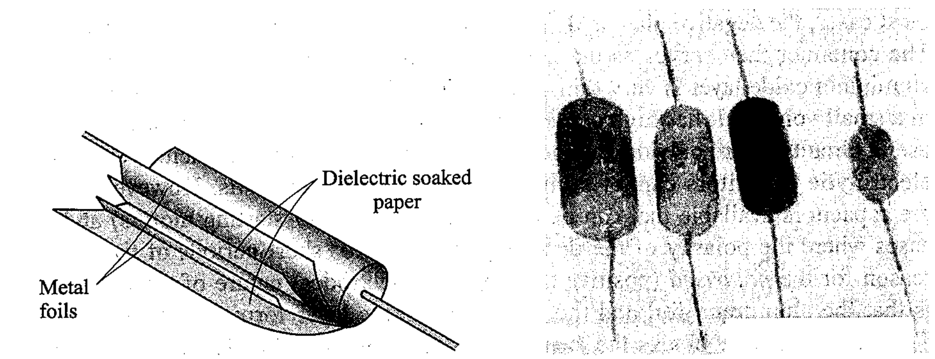


Fig. 1.10 *Explanation of colour code used for resistors and capacitors*

3. Paper capacitors The basic construction of a paper capacitor is shown in Fig. 1.11. Since paper can be rolled between two metals foils, it is possible to concentrate a large plate area in a small volume. The capacitor consists of two metal foils separated by strips of paper. This paper is impregnated with a dielectric material such as wax, plastic or oil.



(*a*) Construction (*b*) Some tubular paper capacitors

Fig. 1.11 *Tubular paper capacitors*

Paper capacitors have capacitances ranging from 0.0005 μF to several μF, and are rated from about 100 V to several thousand volts. They can be used for both dc and ac circuits. Its leakage resistance is of the order of 100 MΩ.

4. Electrolytic capacitors Electrolytic capacitors are extremely varied in their characteristics. The capacitance value may range from 1 μF to several thousand microfarads. The voltage ratings may range from 1 V to 500 V, or more. These capacitors are commonly used in situations where a large capacitance is required. Various types of electrolytic capacitors are shown in Fig. 1.12.

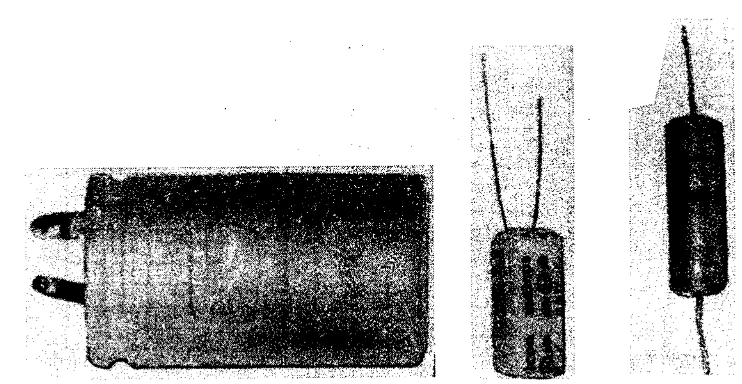


Fig. 1.12 *Electrolytic capacitors*

The electrolytic capacitor consists of an aluminium-foil electrode which has an aluminium-oxide film covering on one side. The aluminium plate serves as the positive plate and the oxide as the dielectric. The oxide is in contact with a paper or gauze saturated with an *electrolyte*. The electrolyte forms the second plate (negative) of the capacitor. Another layer of aluminium without the oxide coating is also provided for making electrical contact between one of the terminals and the electrolyte. In most cases, the negative plate is directly connected to the container of the capacitor. The container then serves as the negative terminal for external connections. The aluminium oxide layer is very thin. Therefore, the capacitor has a large capacitance in a small volume. It has high *capacitance-to-size ratio*. It is primarily designed for use in circuits where only dc voltages will be applied across the capacitor. Ordinary electrolytic capacitors cannot be used with alternating currents. However, there are capacitors available that can be used in ac circuits (for starting motors) and in cases where the polarity of the dc voltage reverses for short periods of time. The reason for the *polarised* (positive and negative electrodes) nature of the capacitor is that the aluminium foil and the aluminium oxide layer form a *semiconductor*. This semiconductor blocks the current coming through the oxide film toward the electrode, but it readily passes current in the opposite direction. The capacitor should be properly connected so that the applied voltage encounters the high resistance.

A new type of electrolytic capacitor is the *tantalum capacitor*. It has an excellent capacitance-to-size ratio.

5. Variable capacitors In some circuits, such as a tuning circuit, it is desirable to be able to change the value of capacitance readily. This is done by means of a variable capacitor. The most common variable capacitor is the *air-gang capacitor* shown in Fig. 1.13. The dielectric for this capacitor is air. By rotating the shaft at one end, we can change the common area between the movable and fixed set of plates. The greater the common area, the larger the capacitance.

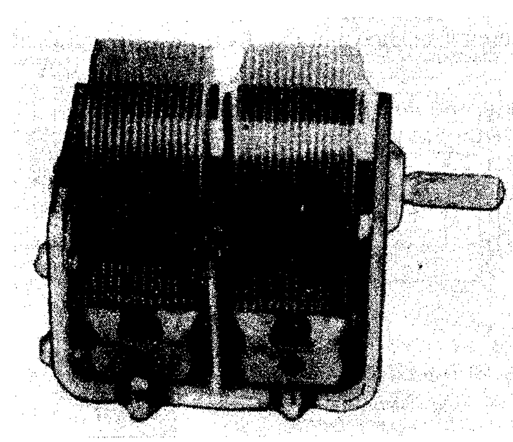


Fig. 1.13 *Air-gang capacitor (variable)*

In some applications, the need for variation in the capacitance is not frequent. One setting is sufficient for all normal operations. In such situations we use a variable capacitor called a *trimmer* (sometimes called *padder*). Both mica and ceramic are used as the dielectric for trimmer capacitors. Figure 1.14 shows the basic construction of a mica trimmer.

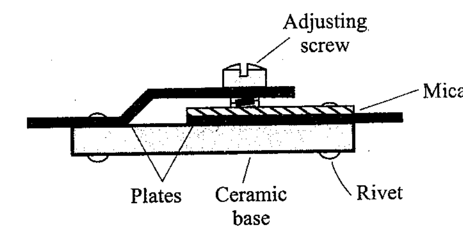
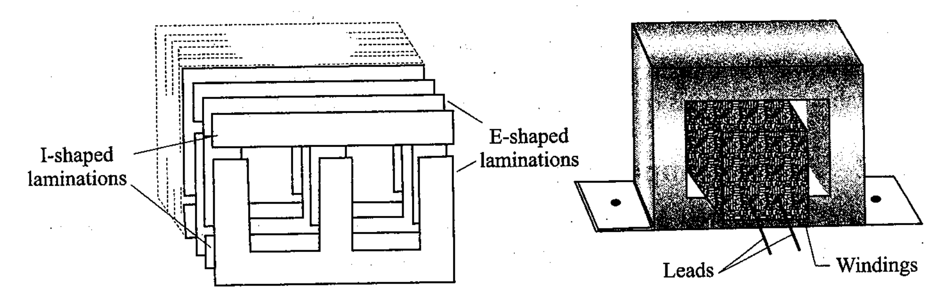


Fig. 1.14 *Construction of variable capacitor*

Inductors When current flows through a wire that has been coiled, it generates a magnetic field. This magnetic field reacts so as to oppose any change in the current. This reaction of the magnetic field, trying to keep the current flowing at a steady rate, is known as *inductance*; and the force it develops is called an *induced emf*. The electronic component producing inductance is called an *inductor*. The symbols of an air-core and an iron-core inductor are shown in Table 1.2. The inductance is measured in henrys (H).

All inductors, like resistors and capacitors, can be listed under two general categories: fixed and variable. Different types of inductors are available for different applications.

1. Filter chokes These are the inductors used in smoothing the pulsating current produced by rectifying ac into dc. A typical filter choke has many turns of wire wound on an iron core. To avoid power losses, the core is made of laminated sheets of E- and I-shapes (Fig. 1.15). Many power supplies use filter chokes of 5 to 20 H, capable of carrying current up to 0.3 A.



(*a*) Exploded view of laminated core (*b*) Assembly choke

Fig. 1.15 *Typical filter choke*

2. Audio-frequency chokes (AFCs) They are used to provide high impedance to audio frequencies (say, 60 Hz to 5 kHz). Compared to filter chokes, they are smaller in size and have lower inductance. Chokes having still smaller inductances are used to block the radio frequencies. Such chokes are called *radio-frequency chokes* (RFCs). Variable inductors are used in tuning circuits for radio frequencies. The *permeability-tuned variable coil* has a ferromagnetic shaft. This shaft can be moved within the coil to vary the inductance as shown in Fig. 1.16.

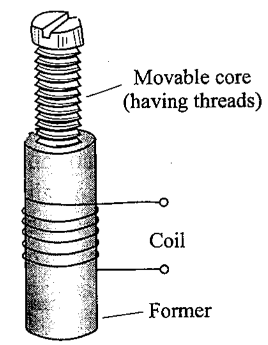


Fig. 1.16 *Permeability-tuned variable coil*

3. Transformers A transformer is quite similar in appearance to an inductor. It consists basically of two inductors having the same core (Fig. 1.17). One of these inductors, or windings, is called *primary*. The other is called *secondary*.

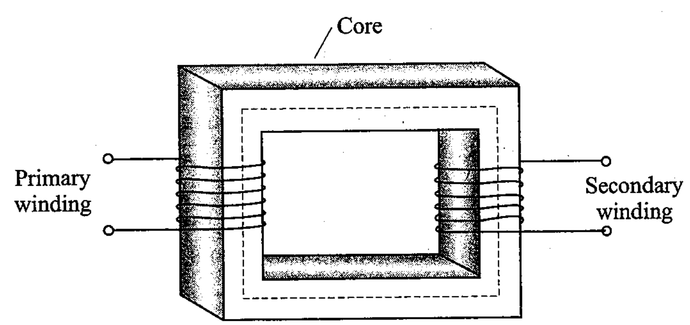


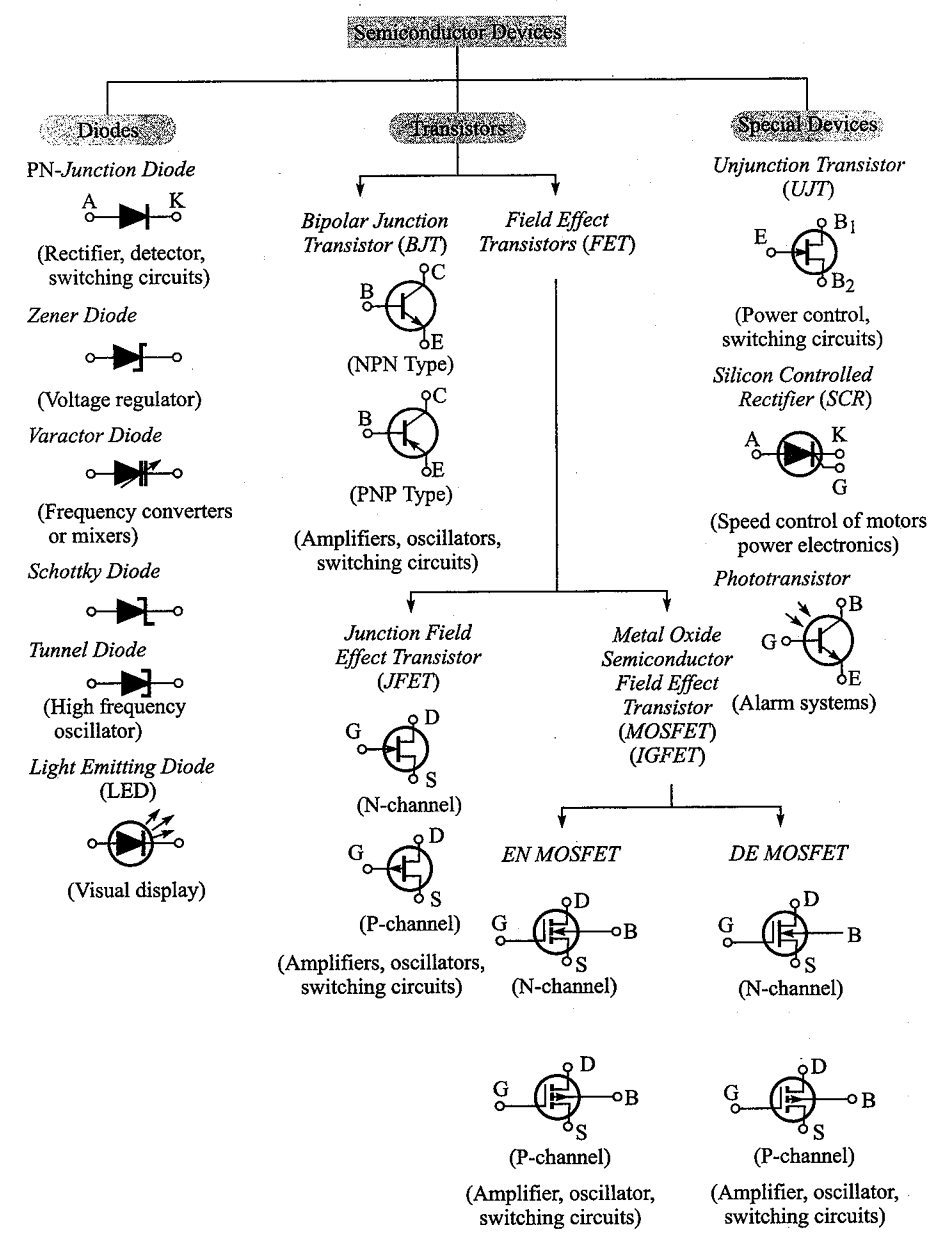
Fig. 1.17 *Basic structure of a transformer*

When an alternating current is applied at the primary, an induced voltage appears in the secondary. In a *step-up transformer*, the number of turns in the secondary is more than that in the primary. The secondary voltage is more than the primary. If the number of turns in the secondary is less than that in the primary, the voltage will be stepped down. The transformer is then called a *step-down transformer*. A transformer of suitable *turns-ratio* is often used in electronic circuits for impedance matching.

1.4.2 Active Components

Table 1.5 gives a brief information about commonly used active components and Fig. 1.18 shows photographs of some active components.

Table 1.5 *Active components*



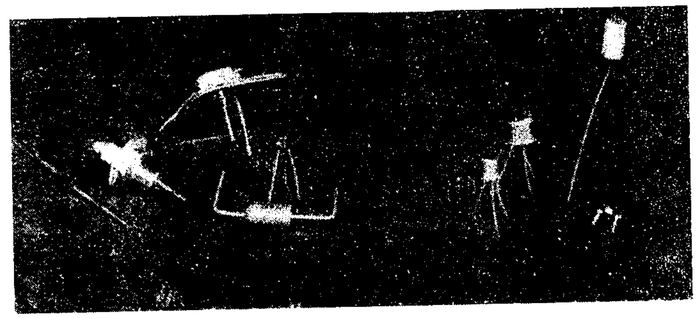


Fig. 1.18 *Some active components*

1.5 SI UNITS

SI is the abbreviation for "Systemé International d'Unités" in French, and is the modern form of the metric system introduced at the Eleventh International Conference of Weights and Measures, 1960. This system of units possesses features that make it logically superior to any other system and also more convenient, as it is *coherent, rational* and *comprehensive*.

A system of units is said to be coherent if the product or quotient of any two unit quantities in the system is the unit of the resultant quantity without the introduction of any numerical factor. For example, unit velocity will result when unit length is divided by unit time.

In 1956, India, by an Act of Parliament No. 89, switched over to the metric system of weights and measures. The definitions of various units given in the Act conform to the definitions of the SI units.

The SI units are based on seven base units with a unit symbol assigned to each of them as given in Table 1.6. The definitions of these base units are as follows:

Table 1.6 *Base units*

Physical quantity	*Name of SI unit*	*Symbol*	*Dimensional notation*
Length	metre	m	[L]
Mass	kilogram	kg	[M]
Time	second	s	[T]
Electric current	ampere	A	[I]
Thermodynamic temperature	kelvin*	K	[θ]
Amount of substance	mole	mol	[mol]
Luminous intensity	candela	cd	[ϕ]

* It should be written as kelvin only and not degree kelvin or °K.

1. **Length** The metre is the length equal to 1 650 760.73 wavelengths, in vacuum, of the radiation corresponding to the transition between the levels $2p^{10}$ and $5d^5$ of the krypton-86 atom.
2. **Mass** The kilogram is equal to the mass of the international prototype kilogram stored at Sevrés, France.
3. **Time** The second is the duration of 9 192 631 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the caesium-133 atom.
4. **Electric current** The ampere is that current which, if maintained in two straight parallel conductors of infinite length and of negligible circular cross section and placed one metre apart in vacuum, would produce between these conductors a force equal to 2×10^{-7} newton per metre of length (N/m).
5. **Thermodynamic temperature** The kelvin is the 1/273.16 fraction of the thermodynamic temperature of the triple point of water**.
6. **Amount of substance** The mole is the amount of substance in a system which contains as many elementary entities as there are atoms in 0.012 kg of carbon-12.
7. **Luminous intensity** The candela is the luminous intensity, in the perpendicular direction, of a surface of 1/600 000 m^2 of a blackbody at the temperature of freezing platinum (2046 K), under a pressure of 101 325 newtons per square metre.

Table 1.7 *Some derived units*

Physical quantity	*Name of SI unit*	*Symbol*
Frequency	hertz	Hz = cycles/s = 1/s
Force	newton	N = kg m/s^2
Work, energy, quantity of heat	joule	J = N m
Power	watt	W = J/s
Electric charge	coulomb	C = A s
Electric potential	volt	V = W/A
Electric capacitance	farad	F = A s/V
Electric resistance	ohm	Ω = V/A
Electric conductance	siemens*	S = A/V
Magnetic flux	weber	Wb = V s
Magnetic flux density	tesla	T = Wb/m^2
Inductance	henry	H = V s/A
Customary temperature	degree celsius	°C
Pressure	pascal	Pa = N/m

*The unit *siemens* is same as *mho* (℧) which was used earlier.

** The temperature at which ice, water and water vapours coexist.

The two dimensionless quantities, plane angle and solid angle, are treated as independent quantities with SI units radian (rad) and steradian (sr), respectively. These are known as *supplementary units*. The *radian* is the plane angle between two such radii of a circle which cuts off, on the circumference, an arc equal to the length of the radius. Thus,

$$\theta(\text{in radians}) = \frac{\text{arc}}{\text{radius}}$$

The *steradian* is the solid angle which, with its vertex at the centre of a sphere, cuts off an area of the surface of the sphere equal to that of a square having sides equal to the radius of the sphere. Thus, if S is the area cut off on the surface of a sphere of radius r, the solid angle at the centre of the sphere is

$$\Omega(\text{in steradians}) = \frac{S}{r^2}$$

All other units are known as *compound* or *derived SI units*, some of which may have special names as given in Table 1.7. The SI units cover all fields of physics and engineering.

1.5.1 Decimal Multiple and Submultiple Factors

Since all the coherent units are not of a convenient size for all applications, provision had to be made for multiples and submultiples of the coherent units. A complete list of such factors is given in Table 1.8. The guidelines for the application of these prefixes are as follows:

Table 1.8 *SI prefixes*

Factor	*Prefix*	*Symbol*	*Factor*	*Prefix*	*Symbol*
10^1	deca	da	10^{-1}	deci	d
10^2	hecto	h	10^{-2}	centi	c
10^3	kilo	k	10^{-3}	milli	m
10^6	mega	M	10^{-6}	micro	µ
10^9	giga*	G	10^{-9}	nano	n
10^{12}	tera	T	10^{-12}	pico	p
10^{15}	peta	P	10^{-15}	femto	f
10^{18}	exa	E	10^{-18}	atto	a

*Pronounced as jeega.

Note

(*i*) The prefixes for factors greater than unity have Greek origin; those for factors less than unity have Latin origin (except femto and atto, recently added, which have Danish origin).

(*ii*) Almost all abbreviations of prefixes for magnitudes < 1, are English lower-case letters. An exception is *micro* (Greek letter μ).

(*iii*) Abbreviations of prefixes for magnitudes > 1 are English upper-case letters. Exceptions are kilo, hecto and deca.

(*iv*) The prefixes hecto, deca, deci and centi should not be used unless there is a strongly felt need.

1. Multiples of the fundamental unit should be chosen in powers of $\pm 3n$ where n is an integer. Centimetre, owing to its established usage and its convenient size, cannot be given up lightly.
2. Double or compound prefixes should be avoided, e.g., instead of micromicrofarad (µµF) or millinanofarad (mnF), use picofarad (pF).
3. To simplify calculations, attach the prefix to the numerator and not to the denominator. For example, use MN/m^2 instead of N/mm^2; even though mathematically, both forms are equivalent.
4. The rules for *binding-in indices* are not those of ordinary algebra, e.g., cm^2 means $(\text{cm})^2 = (0.01)^2\ \text{m}^2 = 0.0001\ \text{m}^2$, and not $\text{c} \times (\text{m})^2 = 0.01\ \text{m}^2$.

1.5.2 Other Accepted Units

It has been recognised at the international level, that some departures from strict purity and coherence are acceptable for practical reasons. For instance, pure SI would acknowledge only decimal multiples and submultiples of the second for time measurement; whereas minute, hour, day, month and year are in everyday use internationally and will clearly continue to be used. Similarly, the division of the circle into 360 degrees is an internationally recognised practice. Some symbols, other than SI, that are commonly used to express physical quantities are given in Table 1.9.

Table 1.9 *Symbols other than SI that are commonly used*

Name	*Abbreviation*	*Name*	*Abbreviation*
angström	Å	inch	in
British thermal unit	Btu	kilowatt-hour	kW h
calorie	cal	mile	mi
day	d	minute (of arc)	′
degree	°	minute (of time)	min
dyne	dyn	pound	lb
electron volt	eV	revolution	rev
foot	ft	second (of arc)	″
gauss	G	standard atmosphere	atm
horsepower	hp	atomic mass unit	amu
hour	h	year	y

1.5.3 Guidelines for Using SI Units

Following are the rules and conventions regarding the use of SI units:

1. Full names of units, even when they are named after a person, are not written with a capital (or upper case) initial letter, e.g., kelvin, newton, joule, watt, volt, ampere, etc.
2. The symbols for a unit, named after a person, has a capital initial letter, e.g., W for watt (after James Watt) and J for joule (after James Prescott Joule).
3. Symbols for other units are not written with capital letter, e.g., m for metre.
4. Units may be written out in full or using the agreed symbols, but no other abbreviation may be used. They are printed in full or abbreviated, in roman (upright) type, e.g., amp. is not a valid abbreviation for ampere.
5. Symbols for units do not take a plural form with added 's'; the symbol merely names the unit in which the preceding magnitude is measured, e.g., 50 kg, and not 50 kgs.
6. No full stops or hyphens or other punctuation marks should be used within or at the end of the symbols for units. However, when a unit symbol prefix is identical to a unit symbol, a raised dot may be used between the two symbols to avoid confusion. For example, while writing, say, metre second it should be abbriviated as m·s to avoid confusion with ms, the symbol for millisecond.
7. There is a mixture of capital and lower-case letters in the symbols for the prefixes as shown in Table 1.8, but the full names of the prefixes commence with lower-case letters only, e.g., 5 MW (5 megawatt), 2 ns (2 nanosecond).
8. A space is left between a numeral and the symbol except in case of the permitted non-SI units for angular measurements, e.g., 57° 16′ 44″.
9. A space is left between the symbols for compound units, e.g., N m for newtons × metres and kW h for killowatt hour. This reduces the risk of confusion when an index notation instead of the solidus (/) is used. In the former notation, a velocity in metres per second is written as m s^{-1} instead of m/s, but ms^{-1} may mean 'per millisecond'. This type of confusion will not occur if we follow the rule that the denominators of compound units are always expressed in the base units and not in their multiples or submultiples. Thus, a heat flow rate will not be given as J/ms but only as kJ/s = kW.
10. When a compound unit is formed by dividing one unit by another, this may be indicated in one of the two forms as m/s or m s^{-1}. In no case, should more than one solidus sign (/) on the same line be included in such a combination unless a parenthesis be inserted to avoid all ambiguity. In complicated cases, negative powers or parenthesis should be used.
11. Algebraic symbols representing "quantities" are written in *italics*, while symbols for "units" are written as upright characters, e.g.,

a current	$I = 3$ A
an energy	$E = 2.75$ J
a terminal voltage	$V = 1.5$ V

12. When expressing a quantity by a numerical value and a certain unit, it has been found suitable in most applications to use units resulting in numerical values between 1.0 and 1000. To facilitate the reading of numerals, the digits may be separated into groups of three—counting from the decimal sign towards the left and the right. The groups should be separated by a small space, but not by a comma or a point. In numerals of four digits, the space is usually not necessary. (It is recognised, however, that to drop the comma from commercial accounting will involve difficulties, particularly with the adding machines in use at present). A few examples are given below:

Incorrect	*Correct*
(*a*) 40,000 or 40000	(*a*) 40 000
(*b*) 81234.765	(*b*) 81 234.765
(*c*) 764213.87629	(*c*) 764 213.876 29
(*d*) 6 543.21	(*d*) 6543.21

Note

(*i*) The recommended decimal sign is a full stop (.). The sign of multiplication of numbers is a cross (×).

(*ii*) If the magnitude of a number is less than unity, the decimal sign should be preceded by a zero.

• Review Questions •

1. What is meant by electronics?
2. How has electronics affected our daily life?
3. Write at least two important applications of electronics in the field of (*a*) communications and entertainment, (*b*) industry and (*c*) medical sciences.
4. State what is meant by radar? Mention some of its important applications.
5. What are the modern trends in electronics?
6. Before understanding electronic circuits, one must first have an understanding of the components that make up those circuits. Justify the statement (in about 7-8 lines).
7. Write the unit of resistance? If a resistor is rated at 1000 Ω and 10 W, what is the maximum current it can carry?
8. Explain constructional features of a wire-wound resistor. What is the range of wattage for wire-wound resistors?
9. Explain in brief, what is (*a*) a capacitor, and (*b*) a dielectric.
10. Name three primary uses of capacitors?
11. Explain briefly the basic construction of a ceramic capacitor. What is the range of capacitance values available in ceramic capacitors?
12. Why are paper capacitors not used in filters of rectifier power supplies?

13. What forms the dielectric of an electrolytic capacitor? Why is the electrolytic capacitor polarised?
14. While tuning your radio receiver to a desired station, which component inside the set are you varying?
15. When you adjust the volume control knob of your radio receiver, which component is varied inside the set?
16. What is a trimmer capacitor? Describe the basic construction of a mica trimmer capacitor.
17. What is an inductor? What is the unit of inductance?
18. Give some important applications of inductors.
19. For what purpose can a transformer be used in an electronic circuit?
20. Name a few active components (devices) used in electronic circuits.
21. Write down the seven base units in SI units.

• Objective-Type Questions •

Below are some incomplete statements. Four alternatives are provided for each. Choose the alternative that completes the statement correctly.

1. Electronics is that branch of engineering which deals with the application of
   (*a*) high-current machines
   (*b*) production of electronic components
   (*c*) electronic devices
   (*d*) fission of uranium nuclei
2. One of the examples of an active device is
   (*a*) an electric bulb
   (*b*) a transformer
   (*c*) a loudspeaker
   (*d*) a silicon controlled rectifier (SCR)
3. Which one of the following is used as a passive component in electronic circuits?
   (*a*) Resistor (*b*) Transistor
   (*c*) Zener diode (*d*) Tunnel diode
4. The term IC, as used in electronics, denotes
   (*a*) internal combustion (*b*) integrated circuits
   (*c*) industrial control (*d*) Indian culture
5. A 100-μF capacitor is required in fabricating an electronic circuit. Such a large value of capacitance is possible if the capacitor is
   (*a*) a mica capacitor (*b*) a ceramic capacitor
   (*c*) an air-gang capacitor (*d*) an electrolytic capacitor

6. A resistor has a colour band sequence: brown, black, green, and gold. Its value is
   (*a*) 1 kΩ ± 10 % (*b*) 10 kΩ ± 5 %
   (*c*) 1000 kΩ ± 5 % (*d*) 1 MΩ ± 10 %
7. We need a resistor of value 47 kΩ with ±5 % tolerance. The sequence of the colour band on this resistor should be
   (*a*) yellow, violet, yellow, and gold
   (*b*) yellow, violet, orange, and gold
   (*c*) yellow, violet, orange, and silver
   (*d*) yellow, violet, brown, and silver
8. By rotating the volume control in a radio receiver, you can change the volume (level) of sound. When you rotate this control, a resistance is varied inside the receiver. Similarly, you can tune in any desired station by rotating the tuning control. When we rotate the tuning control, we vary
   (*a*) a resistance
   (*b*) a capacitance
   (*c*) an inductance
   (*d*) only the position of the indicating needle
9. With the help of radar, we can
   (*a*) listen to more melodious music
   (*b*) perform mathematical calculations very fast
   (*c*) cure the damaged tissues in the human body
   (*d*) detect the presence of an aircraft as well as locate its position
10. With the help of a computer, we can
    (*a*) perform mathematical calculations very fast
    (*b*) transmit messages to a distant place
    (*c*) amplify very weak signals
    (*d*) see the details of a photograph by magnifying it more than million times
11. Ratings on a capacitor are given as 25 μF, 12 V. Also, a plus sign is written near one of its terminals. From this information, we can definitely say that the capacitor is
    (*a*) a mica capacitor (*b*) a ceramic capacitor
    (*c*) an electrolytic capacitor (*d*) any of these
12. The colour bands on a fixed carbon resistor are: brown, red, and black (given sequentially). Its value is
    (*a*) 12 Ω (*b*) 120 Ω
    (*c*) 21 Ω (*d*) 210 Ω

13. An ac voltage can be converted into a unidirectional voltage by using
(*a*) a power amplifier circuit (*b*) an oscillator circuit
(*c*) a multivibrator circuit (*d*) a rectifier circuit

Answers

1. (*c*)	2. (*d*)	3. (*a*)	4. (*b*)	5. (*d*)	6. (*c*)
7. (*b*)	8. (*b*)	9. (*d*)	10. (*a*)	11. (*c*)	12. (*a*)
13. (*d*)					

2 UNIT CURRENT AND VOLTAGE SOURCES

"In the present state of our knowledge, it would be useless to attempt to speculate on the remote cause of the electrical energy... its relation to chemical affinity is, however, sufficiently evident. May it not be identical with it, and an essential property of matter?"

Humphry Davy (1778–1829)
British Chemist and Inventor

After completing this unit, students will be able to:

- name a few sources of electrical energy
- draw the symbols of ideal voltage source, practical voltage source ideal current source and practical current source
- name examples of some voltage sources
- name examples of some current sources
- draw *V-I* characteristics for the voltage and current sources (both ideal as well as practical)
- explain the difference between ideal and practical voltage source
- explain the difference between ideal and practical current source
- convert a given practical voltage source into an equivalent practical current source and vice versa
- solve simple problems involving voltage and current sources

2.1 SOURCES OF ELECTRICAL POWER

The basic purpose of a source is to supply power to a load. A source is, therefore, connected to the load as shown in Fig. 2.1. The source may supply either dc (direct current) or ac (alternating current). The terminology dc as employed here stands for any quantity that is steady, unchanging and unidirectional in nature. Similarly, the termi-

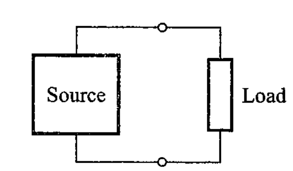


Fig. 2.1 *Transfer of energy from source to load*

nology ac stands to specify any quantity which is alternating in nature, i.e., its magnitude is changing in both the positive and negative directions with time. Unless stated otherwise, the term ac represents *sinusoidal* variations. Some dc sources are battery, dc generator and rectification-type dc supply. Similarly, examples of ac sources are alternators and oscillators or signal generators.

2.1.1 Batteries

The battery is the most common dc voltage source. The term *battery* is derived from the expression "battery of cells". A battery consists of a series or parallel combination of two or more similar cells. A *cell* is the fundamental source of electrical energy. Cells can be divided into *primary* and *secondary* types. The secondary cell is rechargeable, whereas the primary is not. The battery used in a car is of secondary type, since it can be recharged. But the cells used in a torch are of primary type, as they cannot be recharged. Figure 2.2 shows a battery and some typical cells.

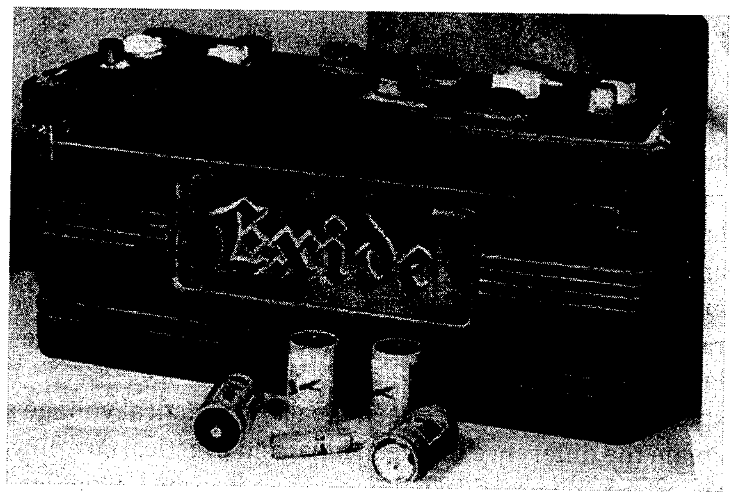


Fig. 2.2 *A battery and some cells*

Cells* and batteries produce electrical power at the expense of chemical energy, and all have the same basic construction. Each has two electrodes (one positive and the other negative) which are immersed in an electrolyte. Electrolytes are chemical compounds. When dissolved in a solution, they decompose into positive and negative ions. These ions carry the charge inside the cell from one electrode to the other.

* An exception is the *solar cell*, which converts light energy into electrical energy. Solar cells are in the developmental stage; and very soon, inexpensive solar will be available.

2.1.2 Generators

The dc generator is quite different from the battery. It has a rotating shaft. When this shaft is rotated at the specified speed by some external agency (such as a steam turbine or water turbine), a voltage of rated value appears across its terminals (see Fig. 2.3). Generally speaking, a generator is capable of giving higher voltage and power than a battery.

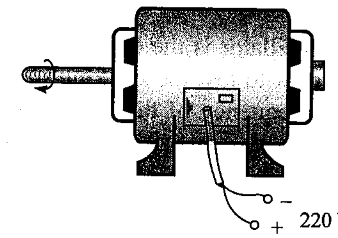


Fig. 2.3 *DC generator*

2.1.3 Rectification-Type Supply

The dc supply most frequently used in an electronics laboratory is of this type. It contains a rectifier which converts time-varying voltage, i.e., ac (such as that available from the domestic power-mains) into a voltage of fixed value. This process will be discussed in detail in Unit 4. A dc laboratory supply of this type is shown in Fig. 2.4. The adaptor used with a laptop is also of this type.

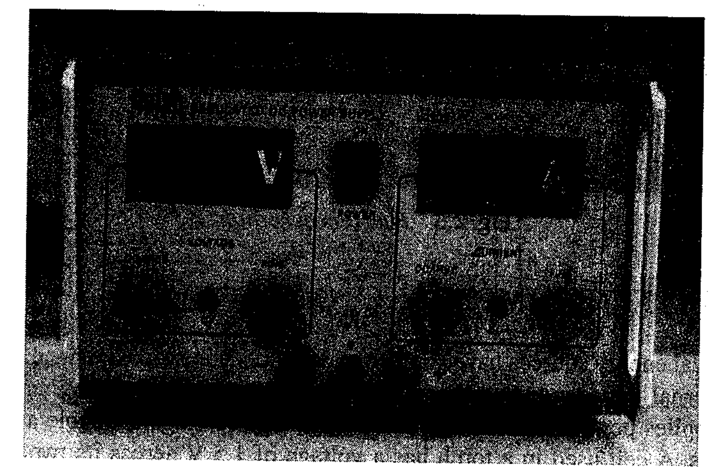


Fig. 2.4 *Regulated DC power supply*

2.1.4 Alternators

The alternator is quite similar, both in construction and mode of operation, to the dc generator. When its shaft is rotated, an alternating (sinusoidal) voltage is generated across its terminals. These type of alternators are used in most electric power stations. Electronics has hardly anything to do with such alternators, except that in most of the cases it is these alternators, in the power generating station, which give power for the operation of the electronic equipment.

2.1.5 Oscillators or Signal Generators

An oscillator is the equipment which supplies ac voltages. This voltage is used as a signal to test the working of different electronic circuits (such as an amplifier). The frequency of the ac signal supplied by this instrument can be varied. Some signal generators are capable of giving other type of waveforms, such as triangular, square, etc., in addition to the sinusoidal wave. Figure 2.5 shows a laboratory signal (function) generator.

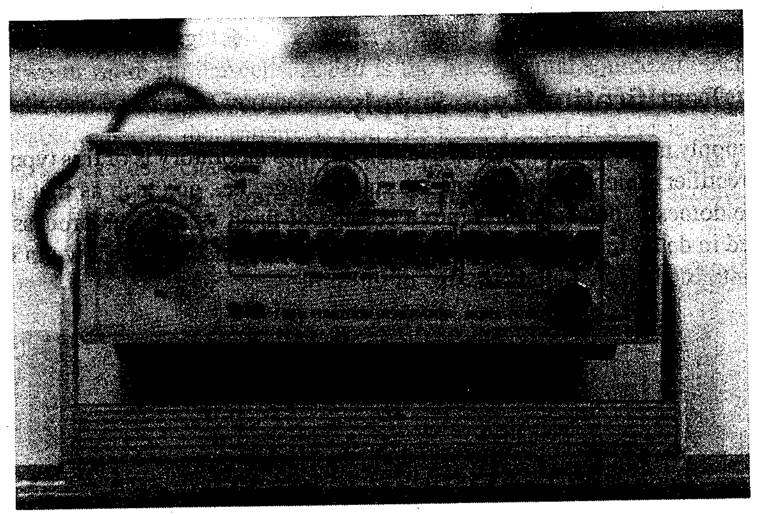


Fig. 2.5 *Function generator*

2.2 INTERNAL IMPEDANCE OF A SOURCE

All electrical energy sources have some internal impedance (or resistance*). It is due to this internal impedance that the source does not behave ideally. When a voltage source supplies power to a load, its terminal voltage (voltage available at its terminals) drops. A cell used in a torch has a voltage of 1.5 V across its two electrodes when nothing is connected to it. However, when connected to a bulb, its voltage becomes less than 1.5 V. Such a reduction in the terminal voltage of the cell may be explained as follows.

Figure 2.6*a* shows a cell of 1.5 V connected to a bulb. When we say "cell of 1.5 V", we mean a cell whose open-circuit voltage is 1.5 V. In the equivalent circuit of Fig. 2.6*b*, the bulb is replaced by a load resistor R_L (of, say, 0.9 Ω), and the cell is replaced by a constant voltage source of 1.5 V in series with the internal resistance R_S (of, say, 0.1 Ω). The total resistance in the circuit is now 0.1 + 0.9 = 1.0 Ω. Since the net voltage that sends current into the circuit is 1.5 V, the current in the circuit is

$$I = \frac{V}{R} = \frac{1.5}{1.0} = 1.5 \text{ A}$$

* In case of dc circuits, the impedance simply reduces to resistance.

The terminal voltage (the voltage across the terminals AB) of the cell is same as the voltage across the load resistor R_L. Therefore,

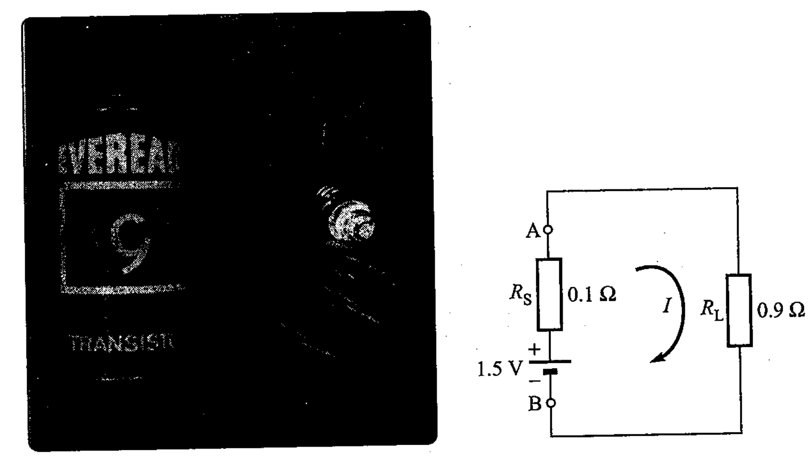
$$V_{AB} = I \times R_L = 1.5 \times 0.9 = 1.35 \text{ V}$$

The voltage that drops because of the internal resistance is

$$= 1.5 - 1.35 = 0.15 \text{ V}$$

Note that, if the internal resistance of the cell were smaller (compared to the load resistance), the voltage drop would also have been smaller than 0.15 V. The internal resistance (or impedance in case of ac source) of a source may be due to one or more of the following reasons:

1. The resistance of the electrolyte between the electrodes, in case of a cell.
2. The resistance of the armature winding in case of an alternator or a dc generator.
3. The output impedance of the active device like a transistor in case of an oscillator (or signal generator) and rectification-type dc supply.



(*a*) (*b*)

Fig. 2.6 *A cell connected to a bulb*

2.3 CONCEPT OF VOLTAGE SOURCE

Consider an ac source. Let V_S be its open-circuit voltage (i.e., the voltage which exists across its terminals when nothing is connected to it), and Z_S be its internal impedance. Let it be connected to a load impedance Z_L whose value can be varied, as shown in Fig. 2.7.

Now, suppose Z_L is infinite. It means that the terminals AB of the source are open-circuited. Under this condition, no current can flow. The terminal voltage V_T is obviously the same as the emf V_S, since there is no voltage drop across Z_S. Let us now connect a finite load impedance Z_L, and then go on reducing its value. As we do this, the current in the circuit goes on increasing. The voltage drop across Z_S also goes on increasing. As a result, the terminal voltage V_T goes on decreasing.

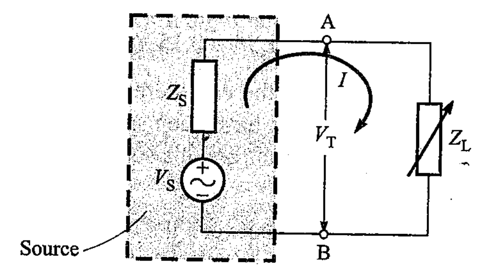


Fig. 2.7 *A variable load connected to an ac source*

For a given value of Z_L, the current in the circuit is given as

$$I = \frac{V_S}{Z_S + Z_L}$$

Therefore, the terminal voltage of the source, which is the same as the voltage across the load, is

$$V_T = I \times Z_L = \frac{V_S}{Z_S + Z_L} \times Z_L = \frac{V_S}{1 + Z_S/Z_L} \tag{2.1}$$

From the above equation, we find that if the ratio Z_S/Z_L is small compared to unity, the terminal voltage V_T remains almost the same as the voltage V_S. Under this condition, the source behaves as a good voltage source. Even if the load impedance changes, the terminal voltage of the source remains practically constant (provided ratio Z_S/Z_L is quite small). Such a source can then be said to be a "*good (but not ideal) voltage source*".

2.3.1 Ideal Voltage Source

It would have been *ideal*, if the terminal voltage of a source remains fixed whatever be the load connected to it. In other words, a voltage source should ideally provide a fixed terminal voltage even though the current drain (or load resistance) may vary. In Eq. (2.1), to make the terminal voltage V_T fixed for any value of Z_L, the only way is to make the internal impedance Z_S zero. Thus, we infer that *an ideal voltage source must have zero internal impedance.* The symbolic representation of dc and ac ideal voltage source are given in Fig. 2.8. Figure 2.9 gives the characteristics of an ideal voltage source. The terminal voltage V_T is seen to be constant at V_S for all values of load current*.

* Load current varies as the load impedance is changed. When we reduce the value of load impedance, the current increases.

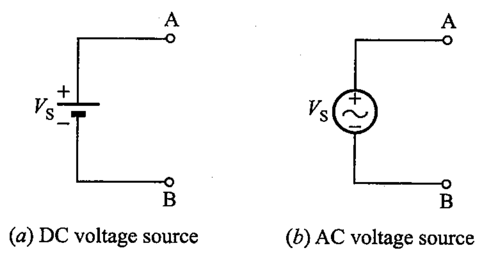


Fig. 2.8 *Symbolic representation of an ideal voltage source*

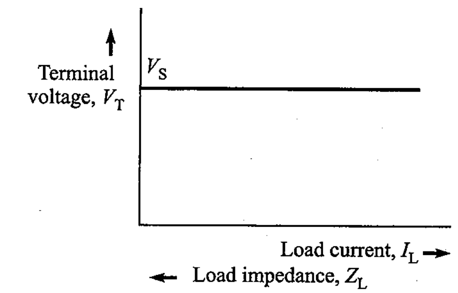


Fig. 2.9 *V-I characteristics of an ideal voltage source*

2.3.2 Practical Voltage Source

An ideal voltage source is not practically possible. There is no source which can maintain its terminal voltage constant when its terminals are short-circuited. It it could do so, it would mean that it can supply an infinite amount of power to a short circuit. This is not possible. Hence, an ideal voltage source does not exist in practice. However, the concept of an ideal voltage source is very helpful in understanding the circuits containing a practical voltage source.

A practical voltage source can be considered to consist of an ideal voltage source in series with an impedance. This impedance is called the *internal impedance* of the source. The symbolic representation of practical voltage sources are shown in Fig. 2.10.

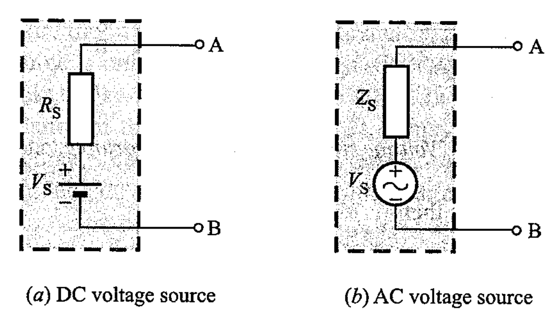


Fig. 2.10 *Practical voltage source*

It is not possible to reach any other terminal except A and B. These are the terminals available for making external connections. In the dc source, since the upper terminal of the ideal voltage source is marked positive, the terminal A will be positive with respect to terminal B. In the ac source in Fig. 2.10*b*, the upper terminal of the ideal voltage source is marked as positive and lower as negative. The marking of positive and negative on an ac source does not mean the same thing as the markings on a dc source. Here (in ac), it means that the upper terminal (terminal A) of the ideal voltage source is positive with respect to the lower terminal at *that particular instant*. In the next half-cycle of ac, the lower terminal will be positive and the upper negative. Thus, the positive and negative markings on an ac source indicate the polarities at a given instant of time. In some books, you will find the reference polarities marked by—instead of positive and negative signs—an arrow pointing towards the positive terminal.

The question naturally arises: What should be the characteristics of a source so that it may be considered a good enough constant voltage source? An ideal voltage source, of course, must have zero internal impedance. In practice, no source can be an ideal one. Therefore, it is necessary to determine how much the value of the internal impedance Z_S should be, so that it can be called a practical voltage source. Let us consider an example. A dc source has an open-circuit voltage of 2 V, and internal resistance of only 1 Ω. It is connected to a load resistance R_L as shown in Fig. 2.11*a*. The load resistance can assume any value ranging from 1 Ω to 10 Ω. Let us now find the variation in the terminal voltage of the source.

When the load resistance R_L is 1 Ω, the total resistance in the circuit is 1 + 1 = 2 Ω. The current in the circuit is

$$V_1 = \frac{V_S}{R_S + R_L} = \frac{2}{1+1} = 1\ \text{A}$$

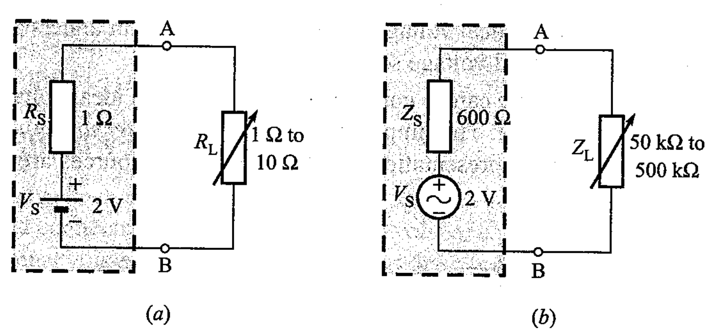


Fig. 2.11 *Voltage sources connected to variable loads*

The terminal voltage is then

$$V_{T1} = I_1 \times R_{L1} = \frac{V_S}{R_S + R_{L1}} \times R_{L1}$$
$$= \frac{2}{1+1} \times 1 = 1.0\ \text{V}$$

When the load resistance becomes 10 Ω, the total resistance in the circuit becomes 10 + 1 = 11 Ω. We can again find the terminal voltage as

$$V_{T2} = \frac{V_S}{V_S + R_{L2}} \times R_{L2}$$
$$= \frac{2}{1+10} \times 10 = \frac{20}{11} = 1.818\ \text{V}$$

Thus, we find that the maximum voltage available across the terminals of the source is 1.818 V. When the load resistance varies between its extreme limits—from 1 Ω to 10 Ω—the terminal voltage varies from 1 V to 1.818 V. This is certainly a large variation. The variation in the terminal voltage is more than 40 % of the maximum voltage.

Let us consider another example. A 600-Ω, 2-V ac source is connected to a variable load, as shown in Fig. 2.11*b*. The load impedance Z_L can vary from 50 kΩ to 500 kΩ—again a variation having the same ratio of 1 : 10, as in the case of the first example. We can find the variation in the terminal voltage of the source. When the load impedance is 50 kΩ, the terminal voltage is

$$V_{T1} = \frac{V_S}{Z_S + Z_{L1}} \times Z_{L1}$$
$$= \frac{2}{600 + 50\,000} \times 50\,000 = 1.976\ \text{V}$$

When the load impedance is 500 kΩ, the terminal voltage is

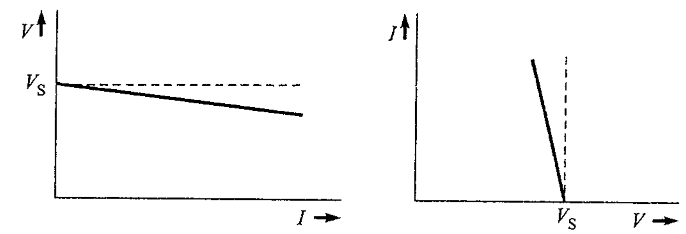
$$V_{T2} = \frac{V_S}{Z_S + Z_{L2}} \times Z_{L2}$$
$$= \frac{2}{600 + 500\,000} \times 500\,000 = 1.997\ \text{V}$$

With respect to the maximum value, the percentage variation in terminal voltage

$$= \frac{1.997 - 1.976}{1.997} \times 100 = 1.05\ \%$$

We can now compare the two examples. In the first case, although the internal resistance of the dc source is only 1 Ω, yet it is not justified to call it a constant voltage source. Its terminal voltage varies by more than 40 %. In the second case, although the internal impedance of the ac source is 600 Ω, it may still be called a practical constant voltage source, since the variation in its terminal voltage is quite small (only 1.05 %). Thus, we conclude that it is not the absolute value of the internal impedance that decides whether a source is a good constant voltage source or not. It is the value of the internal impedance relative to the load impedance that is important. The lesser the ratio Z_S/Z_L (in the first example, this ratio varies from 1 to 0.1, whereas in the second example it varies from 0.012 to 0.0012), the better is the source as a constant voltage source.

No practical voltage source can be an ideal voltage source. Thus, no practical voltage source can have the *V-I* characteristic as shown in Fig. 2.9. When the load current increases, the terminal voltage of a practical voltage source decreases. The characteristic is then modified to that shown in Fig. 2.12*a*. It is sometimes preferred to take voltage on the *x*-axis and current on the *y*-axis. The *V-I* characteristic of a practical voltage source then looks like the one shown in Fig. 2.12*b*.



(*a*) (*b*)

Fig. 2.12 *Two ways of drawing V-I characteristics of a practical voltage source*

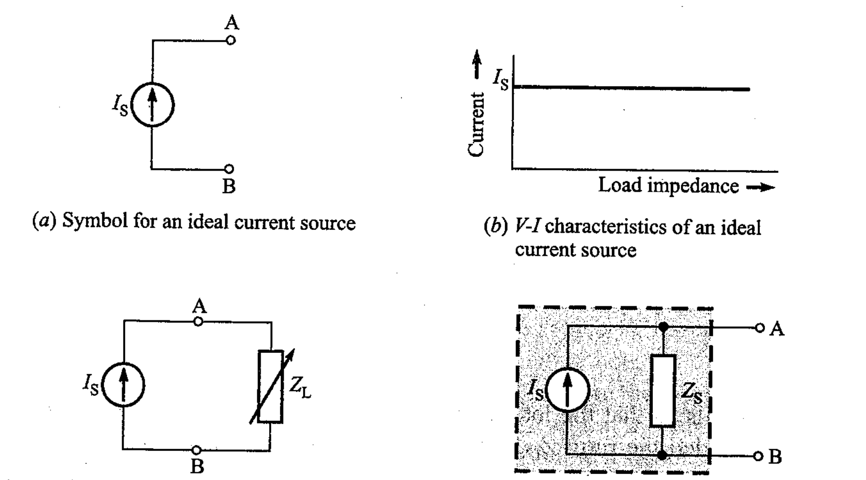
2.4 CONCEPT OF CURRENT SOURCE

Like a constant voltage source, there may be a constant current source—a source that supplies a constant current to a load even if its impedance varies. Ideally, the current supplied by it should remain constant, no matter what the load impedance is. A symbolic representation of such an ideal current source is shown in Fig. 2.13*a*. The arrow inside the circle indicates the direction in which current will flow in the circuit when a load is connected to the source. Figure 2.13*b* shows the *V-I* characteristic of an ideal current source.

Let us connect a variable load impedance Z_L to a constant current source as shown in Fig. 2.13*c*. As stated above, the current supplied by the source should remain constant at I_S for all values of load impedance. It means even if Z_L is made infinity, the current through this should remain I_S. Now, we must see if any practical current source could satisfy this condition. The load impedance $Z_L = \infty$ means no conducting path, external to the source, exists between the terminals A and B. Hence, it is a physical impossibility for current to flow between terminals A and B. If the source could maintain a current I_S through an infinitely large load impedance, there would have been an infinitely large voltage drop across the load. It would then have consumed infinite power from the source. Of course, no practical source could even supply infinite power.

A practical current source supplies current I_S to a short circuit (i.e., when $Z_L = 0$). That is why, the current I_S is called *short-circuit current*. But, when we increase the load impedance, the current falls below I_S. When the load impedance Z_L is made infinite (i.e., the terminals A and B are open-circuited), the load current reduces to zero. It means there should be some path (inside the source itself) through which the current I_S can flow. When some finite load impedance is connected, only a part of this current I_S flows through the load. The remaining current goes through the path

inside the source. The inside path has an impedance Z_S, and is called the *internal impedance*. The symbolic representation of such a practical current source is shown in Fig. 2.13*d*.



(*a*) Symbol for an ideal current source

(*b*) *V-I* characteristics of an ideal current source

(*c*) A variable load connected to an ideal current source

(*d*) Representation of a practical current source

Fig. 2.13 *Current source*

Now, if terminals AB are open-circuited ($Z_L = \infty$) in Fig. 2.13*d*, the terminal voltage does not have to be infinite. It is now an infinite value, $V_T = I_S Z_S$. It means that the source does not have to supply infinite power.

Do not be alarmed if the concept of a current source is strange and somewhat confusing at this point. It will become clearer in later chapters. The introduction of semiconductor devices such as the transistor is responsible, to a large extent, for the increasing interest in current sources.

2.4.1 Practical Current Source

An ideal current source is merely an *idea*. In practice, an ideal current source cannot exist. Obviously, there cannot be a source that can supply constant current even if its terminals are open-circuited. The reason why an actual source does not work as an ideal current source is that its internal impedance is not infinite. A practical current source is represented by the symbol shown in Fig. 2.13*d*. The source impedance Z_S is put in parallel with the ideal current source I_S. Now, if we connect a load across the terminals A and B, the load current will ne different from the current I_S. The current I_S now divides itself between two branches—one made of the source impedance Z_S inside the source itself, and the other made of the load impedance Z_L external to the source.

Let us find the conditions under which a source can work as a good (practical) current source. In Fig. 2.14*a*, load impedance Z_L is connected to a current source.

Let I_S be the short-circuit current of the source, and Z_S be its internal impedance. The current I_S is seen to be divided into two parts—I_1 through Z_S and I_L through Z_L. That is,

$$I_S = I_1 + I_L$$

or

$$I_1 = I_S - I_L$$

Since the impedance Z_S and Z_L are in parallel, the voltage drop across each should be equal, i.e.,

$$I_1 Z_S = I_L Z_L$$

or

$$(I_S - I_L) Z_S = I_L Z_L$$

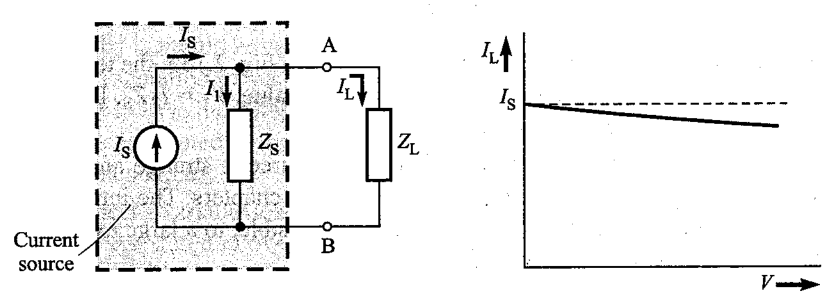
or

$$I_L = \frac{I_S Z_S}{Z_S + Z_L}$$

or

$$I_L = \frac{I_S}{1 + (Z_L/Z_S)} \tag{2.2}$$

This equation tells us that the load current I_L will remain almost the same as the current I_S, provided the ratio Z_L/Z_S is small compared to unity. The source then behaves as a good current source. In other words, the larger the value of internal impedance Z_S (compared to the load impedance Z_L), the smaller is the ratio Z_L/Z_S, and the better it works as a constant current source.



(*a*) Practical current source feeding current to a load impedance

(*b*) *V-I* characteristic of a practical current source

Fig. 2.14

From Eq. (2.2), we see that the current $I_L = I_S$, when $Z_L = 0$. But, as the value of load impedance is increased, the current I_L is reduced. For a given increase in load impedance Z_L, the corresponding reduction in load current I_L is much smaller. Thus, with the increase in load impedance, the terminal voltage ($V = I_L Z_L$) also increases. The *V-I* characteristic of a practical current source is shown in Fig. 2.14*b*.

2.5 EQUIVALENCE BETWEEN VOLTAGE SOURCE AND CURRENT SOURCE

Practically, a voltage source is not different from a current source. In fact, a source can either work as a current source or as a voltage source. It merely depends upon

its working conditions. If the value of the load impedance is very large compared to the internal impedance of the source, it proves advantageous to treat the source as a voltage source. On the other hand, if the value of the load impedance is very small compared to the internal impedance, it is better to represent the source as a current source. From the circuit point of view, it does not matter at all whether the source is treated as a current source or a voltage source. In fact, it is possible to convert a voltage source into a current source and vice versa.

2.5.1 Conversion of Voltage Source into Current Source and Vice Versa

Consider an ac source connected to a load impedance Z_L. The source can either be treated as a voltage source or a current source, as shown in Fig. 2.15. The voltage-source representation consists of an ideal voltage source V_S in series with a source impedance Z_{S1}. And the current-source representation consists of an ideal current source I_S in parallel with source impedance Z_{S2}. These are the two representations of the same source. Both types of representations must appear the same to the externally connected load impedance Z_L. They, must give the same results.

In Fig. 2.15*b*, if the load impedance Z_L is reduced to zero (i.e., the terminals A and B are short-circuited), the current through this *short* is given as

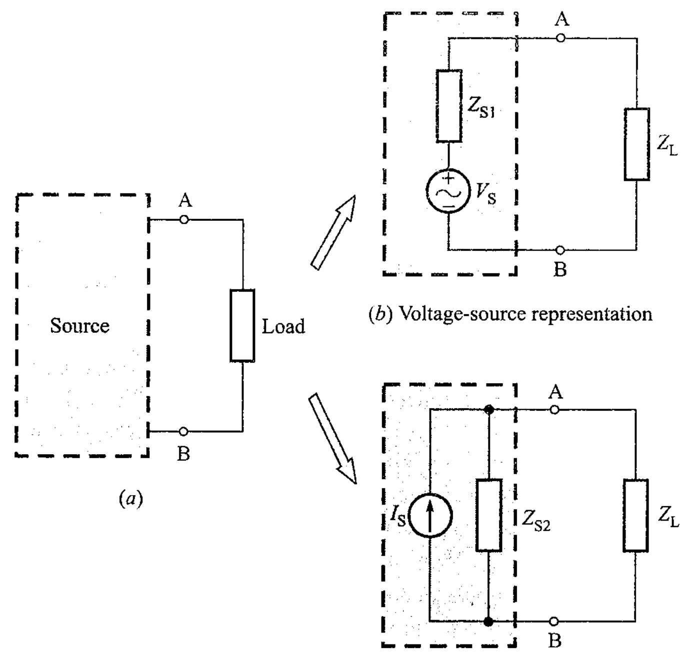
$$I_L(\text{short circuit}) = \frac{V_S}{Z_{S1}} \tag{2.3}$$

We want both the representations (voltage-source and current-source) to give the same results. This means that current source in Fig. 2.15*c* must also give the same current (as given by Eq. (2.3)) when terminals A and B are shorted. But the current obtained by shorting the terminals A and B of Fig. 2.15*c* is simply the source current I_S (the source impedance Z_{S2} connected in parallel with a short circuit is as good as not being present). Therefore, we conclude that the current I_S of the equivalent current source must be the same as that given by Eq. (2.3). Thus,

$$I_L(\text{short circuit}) = I_S = \frac{V_S}{Z_{S1}} \tag{2.4}$$

Again, the two representations of the source must give the same terminal voltage when the load impedance Z_L is disconnected from the source (i.e., when the terminals A and B are open-circuited). In Fig. 2.15*b*, the open-circuit terminal voltage is simply V_S. There is no voltage drop across the internal impedance Z_{S1}. Let us find out the open-circuit voltage in the current-source representation of Fig. 2.15*c*. When the terminals A and B are open-circuited, the whole of the current I_S flows through the impedance Z_{S2}. The terminal voltage is then the voltage drop across this impedance. That is

$$V_T(\text{open circuit}) = I_S Z_{S2} \tag{2.5}$$



(c) Current-source representation

Fig. 2.15 *A source connected to a load*

Therefore, if the two representations of the source are to be equivalent, we must have

$$V_T = V_S$$

Using Eqs. (2.4) and (2.5), we get

$$I_S Z_{S1} = I_S Z_{S2}$$

or $$Z_{S1} = Z_{S2} = Z_S \text{ (say)}$$

Then both Eqs. (2.4) and (2.5) reduce to

$$V_S = I_S Z_S \tag{2.6}$$

It may be noted (see Eq. (2.5)) that in both the representations of the source, the source impedance as faced by the load impedance at the terminals AB, is the same (impedance Z_S). Thus, we have established the equivalence between the voltage-source representation and current-source representation of Fig. 2.15, for short circuits and for open circuits. But, we are not sure that the equivalence is valid for any other value of load impedance. To test this, let us check whether a given impedance Z_L draws the same amount of current when connected either to the voltage-source representation or to the current-source representation.

In Fig. 2.15*b*, the current through the load impedance is

$$I_{L1} = \frac{V_S}{Z_S + Z_L} \tag{2.7}$$

In Fig. 2.15*c*, the current I_S divides into two branches. Since the current divides itself into two branches in inverse proportion of the impedances, the current through the load impedance Z_L is

$$I_{L2} = I_S \times \frac{Z_S}{Z_S + Z_L} = \frac{I_S Z_S}{Z_S + Z_L}$$

By making use of Eq. (2.6), the above equation can be written as

$$I_{L2} = \frac{V_S}{Z_S + Z_L} \tag{2.8}$$

We now see that the two currents I_{L1} and I_{L2} as given by Eqs. (2.7) and (2.8) are exactly the same. Thus, the equivalence between the voltage-source and current-source representations of Fig. 2.15 is completely established. *We may convert a given voltage source into its equivalent current source by using Eq. (2.6). Similarly, any current source may be converted into its equivalent voltage source by using the same equation.*

Example 2.1 Figure 2.16 shows a dc voltage source having an open-circuit voltage of 2 V and an internal impedance of 1 Ω. Obtain its equivalent current-source representation.

Solution: If we short circuit the terminal A and B of the voltage source, the current supplied by the source is

$$I(\text{short circuit}) = \frac{V_S}{R_S} = \frac{2}{1} = 2\text{ A}$$

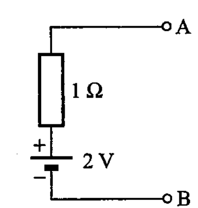


Fig. 2.16 *A voltage source*

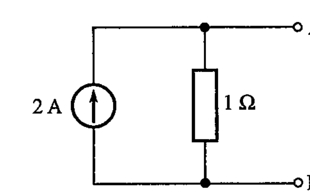


Fig. 2.17 *Equivalent current source*

In the equivalent current-source representation, the current source is of 2 A. The source impedance of 1 Ω is connected in parallel with this current source. The equivalent current source obtained is shown in Fig. 2.17.

Example 2.2 Obtain an equivalent voltage source of the ac current source shown in Fig. 2.18.

Solution: The open-circuit voltage across terminals A and B is given as

$$V(\text{open circuit}) = I_S Z_S$$
$$= 0.2 \times 100 = 20 \text{ V}$$

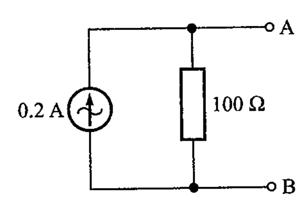


Fig. 2.18 *An ac current source*

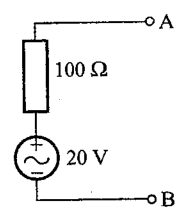


Fig. 2.19 *Equivalent voltage source*

This will be the value of the "ideal voltage source" in the equivalent voltage-source representation. The source impedance Z_S is put in series with the ideal voltage source. Thus, the equivalent voltage-source representation of the given current source is as given in Fig. 2.19.

Example 2.3 In the circuit of Fig. 2.20, an ac current source of 1.5 mA and 2 kΩ is connected to a load consisting of two parallel branches; one of 10 kΩ and other of 40 kΩ. Determine the current I_4 flowing in the 40 kΩ impedance. Now convert the given current source into its equivalent voltage source and then again calculate the current I_4 in the 40 kΩ impedance. Check whether you get the same results in the two cases.

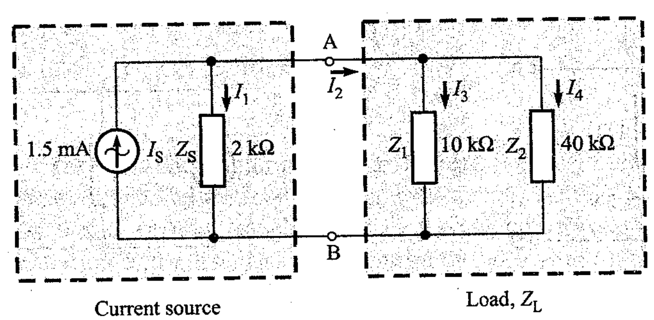


Fig. 2.20 *A current source connected to a load*

Solution: Let us first determine the net load impedance that is connected across the source terminals A and B. This would be the parallel combination of the two impedances. Thus,

$$Z_L = \frac{Z_1 Z_2}{Z_1 + Z_2} = \frac{10 \times 40}{10 + 40} \text{ k}\Omega = 8 \text{ k}\Omega$$

Now the circuit of Fig. 2.20 can be redrawn as given in Fig. 2.21. A net impedance of 8 kΩ is shown to be connected across the source terminals A and B.

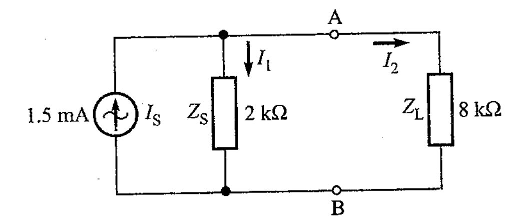


Fig. 2.21

It is clear that the current I_S divides itself into two branches—one consisting of Z_S and other consisting of Z_L. Therefore, using current divider concept the current I_2 is given by

$$I_2 = I_S \times \frac{Z_S}{Z_S + Z_L} = 1.5 \times 10^{-3} \times \frac{2 \times 10^3}{(2+8) \times 10^3} = 0.3 \text{ mA}$$

Again, look at Fig. 2.20. The current I_2 divides into two parallel branches. The current in the 40 kΩ impedance can be determined as follows:

$$I_4 = I_2 \times \frac{Z_1}{Z_1 + Z_2} = 0.3 \times 10^{-3} \times \frac{10 \times 10^3}{(10+40) \times 10^3} = \mathbf{60\ \mu A}$$

We shall again solve this problem following another approach. Here we convert the given current source into its equivalent voltage source. The open-circuit voltage of the source is given as

$$V_S = I_S Z_S = 1.5 \times 10^{-3} \times 2 \times 10^3 = 3.0 \text{ V}$$

Therefore, the equivalent voltage-source representation will be an ideal voltage source of 3.0 V in series with an impedance of 2 kΩ. We can connect the net load impedance Z_L (of 8 kΩ, as calculated above) to this voltage source, as in Fig. 2.22.

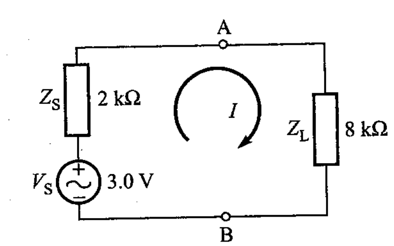


Fig. 2.22 *Single-loop circuit*

The circuit in Fig. 2.22 is a single-loop circuit. The loop current can be calculated by applying Kirchhoff's voltage law.

$$3 = I(2 \times 10^3 + 8 \times 10^3)$$

$$\therefore \quad I = \frac{3}{10 \times 10^3} = 0.3 \text{ mA}$$

(Note that the current I turns out to be the same as current I_2 of Fig. 2.21). This current gets divided into two parallel branches of the load impedance Z_L (see Fig. 2.20). The current through the 40 kΩ impedance is

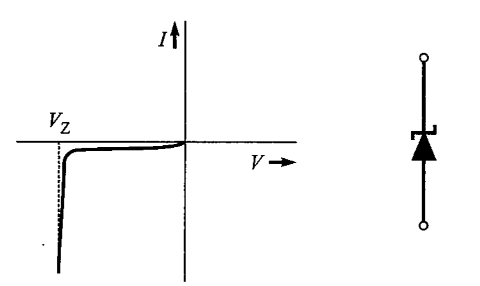
$$I_4 = I \times \frac{Z_1}{Z_1 + Z_2} = 0.3 \times 10^{-3} \times \frac{10 \times 10^3}{(10 + 40) \times 10^3} = \mathbf{60\ \mu A}$$

This is the same result as obtained earlier. Thus, we find that solving an electrical circuit gives the same result whether we treat the source in the circuit as voltage source or current source. However, as we shall see later, the solving of a particular circuit sometimes becomes simpler if we treat the source as one type rather than the other.

2.6 USEFULNESS OF THE CONCEPT OF VOLTAGE AND CURRENT SOURCE IN ELECTRONICS

As we proceed with the study of electronics, we find that the concept of voltage and current sources are of great help. For example, in determining the performance of an electronic circuit (such as an amplifier, which is used for amplifying electrical signals), we convert the original circuit into its equivalent ac circuit. In this equivalent circuit, the active device (such as a BJT or an FET) is replaced by its current-source equivalent or voltage-source equivalent. We can now apply the basics of circuit theory to determine the characteristic behaviour of the electronic circuit.

Figure 2.23*a* shows the *V-I* characteristics of a semiconductor device called the *zener diode*. Its symbol is shown in Fig. 2.23*b*. For the time being, we will ignore the details of this device. But, if you compare its characteristics with that of a dc voltage source (as shown in Fig. 2.12*b*), you will find a marked similarity. The only difference is that the characteristic curve of the zener diode is inverted. It is *shown* inverted to emphasise that the zener diode is operated with reverse bias (the term *reverse bias* is explained in detail in Unit 4). This means that the current through the zener diode flows in a direction opposite to that of the arrow (in its symbol).



(*a*) Characteristics (*b*) Symbol

Fig. 2.23 *Zener diode*

Let us see what happens if we connect a zener diode across a practical voltage source, as in Fig. 2.24. It the load impedance R_L varies, the current I_L through it also varies. If the zener diode were not there, the terminal voltage V_T would also vary, because the voltage drop across the source impedance varies. But now, since a zener diode is connected across terminals A and B, and it has characteristics quite similar to that of *an ideal voltage source*, the situation is different. The terminal voltage V_T remains constant at V_Z whatever be the current flowing through the zener diode. When the load current varies, the zener diode current adjusts itself so that its terminal voltage remains constant. This is an example of simple *voltage regulator circuit.* The resistance R_2 is put in the circuit so as to *limit* the current through the zener diode. It ensures safe operation of the zener diode.

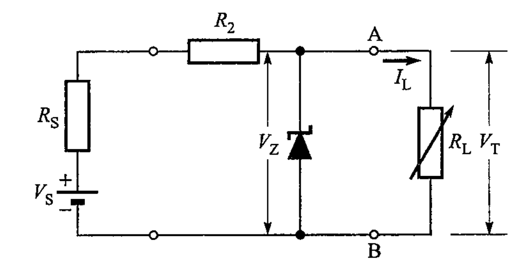
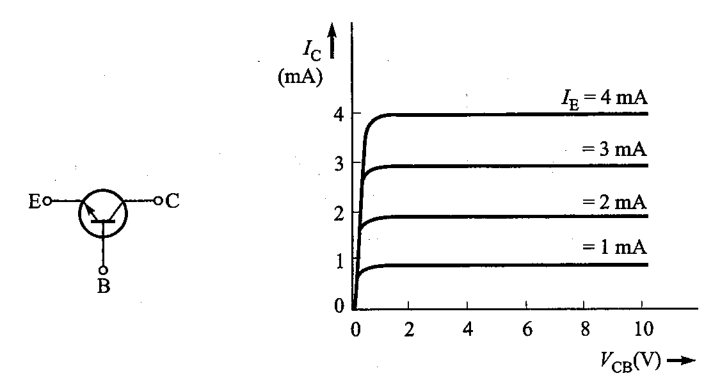


Fig. 2.24 *Zener diode connected across a practical voltage source*

A zener diode when connected in the circuit of Fig. 2.24 works as a voltage source. Strictly speaking, it is not a source, because it cannot supply any power of its own. We need another voltage source for its operation. Once it is connected in an electrical circuit, it has *V-I* characteristics similar to that of a constant voltage source. Loosely speaking, we can say that a zener diode is a constant dc-voltage source.

Another important device is the *transistor*. It is a three-terminal device. These terminals are called *emitter, base* and *collector.* Its symbol is shown in Fig. 2.25*a*. The transistor is extensively used as an amplifying device. When connected in the amplifier circuit, one of its terminals can be made common between input and output. Figure 2.25*b* shows the output characteristics of a transistor connected in a common-base mode.

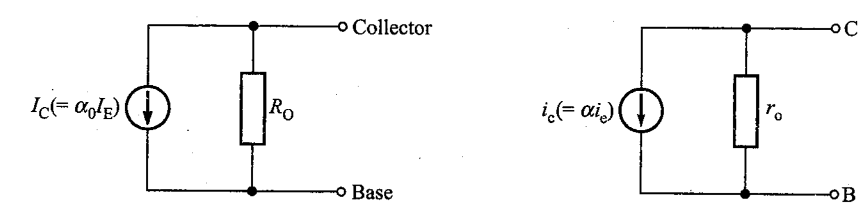


(*a*) Symbol of a transistor (NPN) (*b*) Output characteristics

Fig. 2.25 *Transistor*

The characteristics of the transistor are almost horizontal lines. For a given value of the input current (emitter current I_E), the collector current I_C remains constant when the collector voltage V_{CB} is varied. Such characteristics are very similar to the characteristics of a current source as shown in Fig. 2.14*b*. Thus, we say that a transistor behaves as a current source. Between its output terminals (collector and base), it can be represented by a current source as shown in Fig. 2.26*a*. Here, the current source $I_C (= \alpha I_E)$ is dependent upon the input current I_E. The resistance R_O represents static (dc) output resistance.

The equivalent representation of the transistor as shown in Fig. 2.26*a* is not of much significance to us. It is meant for the dc operation of the transistor. Since, in an amplifier circuit, the voltages and currents are changing all the time (because of the input signal), we are interested in the transistor's ac behaviour. The ac behaviour of the transistor can be represented by the circuit shown in Fig. 2.26*b*. Here, r_o is the ac resistance of the transistor between its collector and base. This resistance has high value (typically, 1 MΩ). This resistance represents the source resistance when we look upon a transistor as a current source. The value of the ac current source i_c depends upon the input ac current i_e.



(*a*) DC representation (*b*) AC representation

Fig. 2.26 *A transistor behaves like a current source between its output terminals*

Example 2.4 Figure 2.27 shows the ac equivalent of an amplifier using a transistor. Calculate the output voltage v_o.

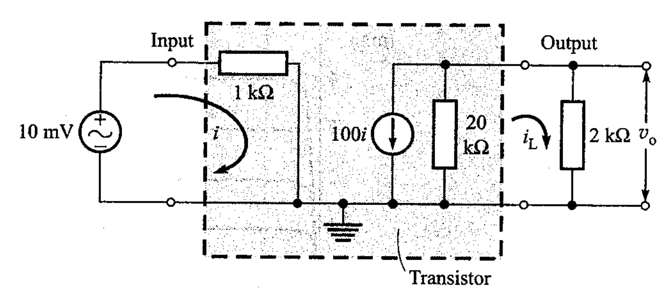


Fig. 2.27 *AC equivalent of a transistor amplifier*

Solution: The value of the current source in ac equivalent of the transistor is $100i$. The current i can be calculated from the input circuit.

$$i = \frac{10\text{ mV}}{1\text{ k}\Omega} = \frac{10 \times 10^{-3}}{1 \times 10^{3}} = 10 \times 10^{-6}\text{ A}$$

Therefore, the current source in the output circuit is

$$100i = 100 \times 10 \times 10^{-6} = 10^{-3}\text{ A}$$

This current divides itself into two branches. The current through the 2-kΩ resistance is

$$i_L = 10^{-3} \times \frac{20 \times 10^{3}}{(20+2) \times 10^{3}} = 0.909 \times 10^{-3}\text{ A}$$

Therefore, the output voltage v_o is given as

$$v_o = i_L \times 2 \times 10^{3} = 0.909 \times 10^{-3} \times 2 \times 10^{3}$$
$$= \mathbf{1.818\ V}$$

Review Questions

1. Name two sources of electrical power. Are they voltage sources or current sources?
2. Draw the symbol of an ideal dc voltage source.
3. Draw the symbolic representation of a practical ac voltage source. Explain the necessity of including an impedance in the representation.
4. Explain the condition under which a practical voltage source is considered to be good voltage source.
5. Name at least one electronic device whose characteristics are very close to that of an ideal voltage source. Explain in one or two lines its characteristics.
6. State an application of an electronic device whose characteristics are similar to that of an ideal voltage source.
7. Draw the symbol of an ideal current source.
8. Draw the symbolic representation of a practical current source. Explain the reason for putting an impedance in this symbolic representation.
9. Name an electronic device which has characteristics similar to that of an ideal current source. Justify your answer in about three lines.
10. A practical source can be represented either as a voltage source or as a current source. How can you convert one representation to the other?

Objective-Type Questions

I. Below are some incomplete statements. Four alternatives are provided for each. Choose the alternative that completes the statement correctly.

1. An ideal voltage source is one which has
   (a) very high internal resistance
   (b) very low internal resistance
   (c) zero internal resistance
   (d) infinite internal resistance
2. An ideal current source is one whose internal resistance is
   (a) very high (b) very low
   (c) zero (d) infinite
3. In a practical voltage source, the source resistance is
   (a) very low compared to load resistance
   (b) very high compared to load resistance
   (c) equal to the load resistance
   (d) zero
4. A device whose characteristics are very close to that of an ideal voltage source is a
   (a) silicon signal diode
   (b) transistor in common-base mode
   (c) field-effect transistor
   (d) zener diode
5. A device whose characteristics are very close to that of an ideal current source is a
   (a) transistor in common-base mode
   (b) crystal diode
   (c) gas diode
   (d) zener diode
6. Figure O.2.1 shows the circuit of a simple constant-voltage supply, using a zener diode. The constant voltage available across the zener diode is 5 V. The current flowing through the 1-kΩ load is

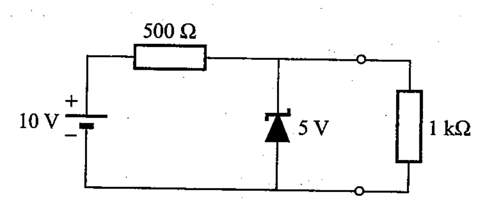


Fig. O.2.1

   (a) 20 mA (b) 15 mA
   (c) 25 mA (d) 5 mA
7. The current flowing through the zener diode in Fig. O.2.1 is
   (a) 20 mA (b) 15 mA
   (c) 25 mA (d) 5 mA

8. In Fig. O.2.1, the current drain from the battery is
   (a) 20 mA (b) 15 mA
   (c) 10 mA (d) 5 mA
9. A constant current source supplies a current of 300 mA to a load of 1 kΩ. When the load is changed to 100 Ω, the load current will be
   (a) 3 A (b) 30 mA
   (c) 300 mA (d) 600 mA
10. An ideal voltage source of 12 V provides a current of 120 mA to a load. If the load impedance is doubled, the new load current becomes
   (a) 60 mA (b) 120 mA
   (c) 240 mA (d) none of the above

II. Below are some statements. Indicate against each, whether it is TRUE(T) or FALSE(F).

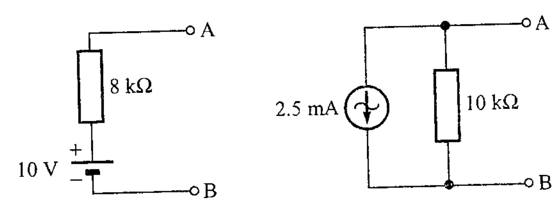
1. A practical current source has low internal resistance. ________
2. An ideal current source has low internal resistance. ________
3. An ideal voltage source has low internal resistance. ________
4. An ideal voltage source has zero internal resistance. ________
5. An ideal current source has infinite internal resistance. ________
6. A practical voltage source has very high internal resistance. ________
7. Solving an electrical circuit will give the same results whether the source is treated as voltage source or as current source. ________
8. A resistance is connected to a practical source. For finding the current through this resistance, the only way is to represent the source as a current source. ________
9. The output side of a transistor connected in common-base mode should be treated as a constant voltage source, since its output impedance is very high. ________
10. A zener diode has characteristics similar to that of an ideal current source. ________

Answers

I.	1. (c)	2. (d)	3. (a)	4. (d)	5. (a)	6. (d)
	7. (d)	8. (c)	9. (c)	10. (a)		
II.	1. F	2. F	3. F	4. T	5. T	6. F
	7. T	8. F	9. F	10. F		

Tutorial Sheet 2.1

1. Figure T. 2.1.1*a* represents a voltage source. Convert it into an equivalent current source. **[Ans.** I_S = 1.25 mA; R_S = 8 kΩ]

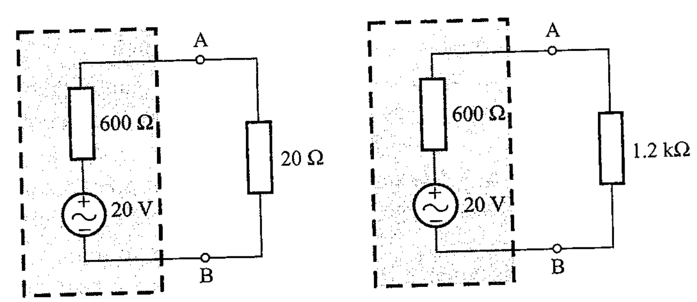


(*a*) (*b*)

Fig. T. 2.1.1

2. Figure T. 2.1.1*b* represents an ac current source. Convert it into an equivalent voltage source. **[Ans.** V_S = 25 V, Z_S = 10 kΩ]
3. Calculate the voltage available between the points A and B in the two situations represented in Fig. T. 2.1.2*a* and *b*. State which situation represents a better voltage source condition.

[Ans. (*a*) 0.645 V; (*b*) 13.3 V; second case represents a better voltage-source condition]

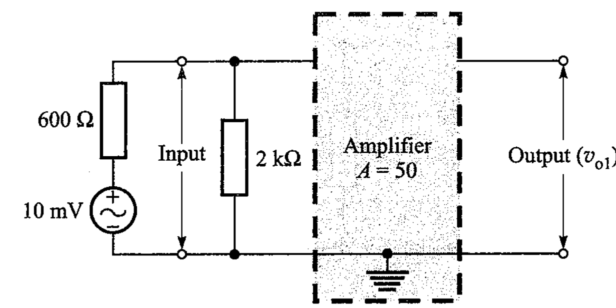


(*a*) (*b*)

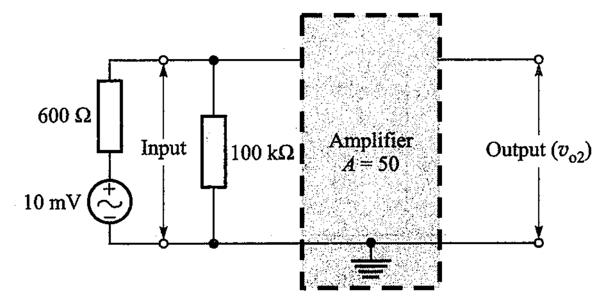
Fig. T. 2.1.2

4. An electronic amplifier is used to amplify electrical signals. The voltage gain of an amplifier is defined as the ratio of its output voltage to its input voltage. Figure T. 2.1.3 represents two amplifiers having different values of input impedances Z_{in}. Calculate the output voltage in the two cases. Assume a voltage gain A = 50 in both the cases.

[Ans. v_{o1} = 384.6 mV, v_{o2} = 497.02 mV]



(*a*)



(*b*)

Fig. T. 2.1.3

Tutorial Sheet 2.2

1. Calculate the current through the 10-kΩ resistor shown in Fig. T. 2.2.1. Assume that the zener diodes have ideal voltage-source characteristics.

[Ans. 1 mA]

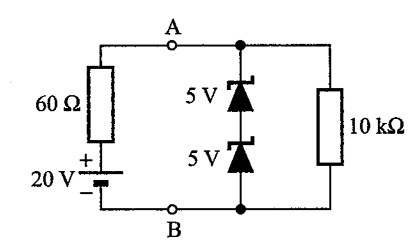


Fig. T. 2.2.1

2. Calculate the voltage v_o in Fig. T. 2.2.2. Identify the source impedance Z_S.

[Ans. 7.5 V, Z_S = 5 kΩ]

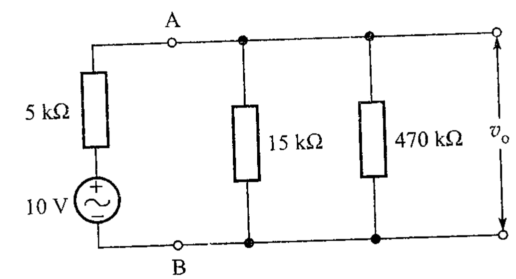


Fig. T. 2.2.2

3. Calculate the voltage v_o in Fig. T. 2.2.3. [**Ans.** 161.29 mV]

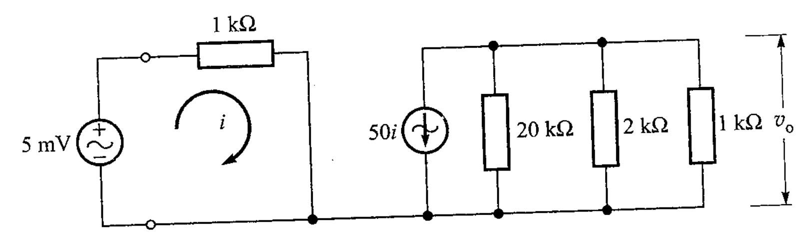


Fig. T. 2.2.3

3 UNIT SEMICONDUCTOR PHYSICS

"Electronic aids, particularly domestic computers, will help the inner migration, the opting out of reality. Reality is no longer going to be the stuff out there..."

J G Ballard (1930–2009)
English novelist, short story writer

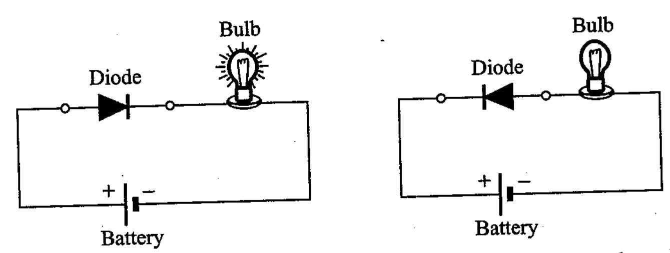
After completing this unit, students will be able to:

- state the names of a few conductors, insulators and semiconductors
- state the differences in conductors, insulators and semiconductors using energy-band diagrams
- state the meaning of covalent bond, thermal generation, lifetime of charge carriers, recombination, forbidden-energy gap, valence band and conduction band, doping, donor impurity, acceptor impurity, majority and minority carriers, drift current and diffusion current
- describe the mechanism of flow of current in an intrinsic semiconductor on the basis of movement of electrons and holes
- state how extrinsic (P- and N-type) semiconductor material is obtained from intrinsic semiconductor material
- describe the mechanism of flow of current in an extrinsic (P- and N-type) semiconductor
- state the effect of temperature on the conductivity of an extrinsic semiconductor

3.1 WHY STUDY SEMICONDUCTOR PHYSICS

All of us are familiar with some of the simple applications of electronics like the radio, television and calculator. If one looks inside any electronic equipment, one will find resistors, capacitors, inductors, transformers, semiconductor diodes, transistors and ICs. We already know something about resistors, capacitors, inductors, and transformers, but small semiconductor devices like diodes and transistors are new to most of us. In modern electronic systems, the whole electronic circuit, containing many diodes, transistors, resistors, etc., is fabricated on a single chip. This is known as an *integrated circuit* (IC).

Let us take a simple example of a semiconductor diode. It is a two-terminal device. It has a very important property of conducting in one direction only. Figure 3.1*a* shows a circuit having a battery, small bulb and diode, all connected in series. When the diode is connected in this manner, the bulb glows. This means the diode is conducting and the current is flowing in the circuit. In Fig. 3.1*b*, the two terminals of the diode are reversed. When the diode is connected in this manner, the lamp does not glow. It means no current is flowing in the circuit, although the battery is present. The diode does not permit the current to flow.



(*a*) Diode conducts (*b*) Diode does not conduct

Fig. 3.1 *Unidirectional conducting property of a diode*

Thus, we find that a diode conducts in one direction only. The *unidirectional* conducting property of a diode finds great applications in electronics. The power available at the mains in our homes is generally ac. Quite often, we require dc power to operate some appliances. The diode makes it possible to convert ac into dc.

The diode is one of the many components used in electronic circuits. Another important component is the transistor. It is used for *amplifying* weak electrical signals. Relatively newer devices, like junction field-effect transistor (JFET), metal-oxide semiconductor field-effect transistor (MOSFET), silicon controlled rectifier (SCR), unijunction transistor (UJT), etc., are finding wide applications in electronics. *All these devices are made of semiconductor materials.* To understand the operation of these devices (and many more that are likely to come in future), it is necessary to study the semiconductor materials in some detail.

3.2 SEMICONDUCTOR MATERIALS

We are familiar with conducting and insulating materials. Conducting materials are good conductors of electricity. Examples of good conductors are copper, silver, aluminium, etc. Insulating materials are bad conductors of electricity. Examples of insulators are porcelain, glass, quartz, rubber, bakelite, etc. The electrical wire used in houses consists of a core made of conducting material like copper or aluminium. The core provides an easy path for the flow of electric current. This core is covered with some insulating material such as rubber, cotton, PVC, plastic, etc. These coverings provide protection against short circuits and also against electrical-shock hazards.

There is another group of materials, such as germanium and silicon. These are neither good conductors nor good insulators. At room temperature these materials have conductivities considerably lower than that of conductors, but much higher than that of insulators. It is for this reason that these materials are classified as *semiconductors*. Table 3.1 gives the resistivities of some commonly used conductors, semiconductors and insulators at room temperature.

Table 3.1 *Resistivities of some conductors, semiconductors and insulators (at room temperature)*

Materials	*Conductivity* (S/m)	*Resistivity* (Ω m)	*Classification*
Silver	6.25×10^{7}	1.6×10^{-8}	Conductors
Copper	5.88×10^{7}	1.7×10^{-8}	
Aluminium	3.85×10^{7}	2.6×13^{-8}	
Germanium (pure)	1.54	6.5×10^{-1}	Semiconductors
Silicon (pure)	5.0×10^{-4}	2.0×10^{3}	
Porcelain	3.33×10^{-10}	3.0×10^{9}	Insulators
Glass	5.88×10^{-12}	1.7×10^{11}	
Hard rubber	1.0×10^{-16}	1.0×10^{16}	

When we increase the temperature of a metal conductor such as copper, its resistivity increases. In this respect, the semiconductors behave in an opposite way. When we raise the temperature of a semiconductor, its resistivity decreases (at higher temperature, a semiconductor conducts better), i.e., the semiconductors have *negative temperature coefficient of resistance*.

The semiconductors have another very important property. The conductivity (or resistivity) of a semiconductor can be changed, to a very large extent, by adding a very small amount of some specific materials (called *impurities*). The conductivity of the semiconductor can also be controlled by controlling the amount of impurity added to it. To understand the important properties of semiconductors, it is necessary to study the structure of atom.

3.3 STRUCTURE OF AN ATOM

We know that the most fundamental unit of matter which is capable of independent existence is the *atom*. According to a simplified picture, an atom consist of a central body, called the *nucleus*, about which a number of smaller particles (called *electrons*) move in approximately elliptical orbits. The nuclei of all elements (except that of hydrogen, which has only one proton in its nucleus) contains two types of particles, called *protons* and *neutrons*. A proton and a neutron have almost same mass. But the proton is a positively charged particle whereas the neutron is electrically neutral. Almost all the mass of the atom is concentrated in its nucleus. The electrons revolving around the nucleus are very light in weight. An electron is about 1850 times lighter than a proton (or neutron). An electron has the same amount of charge as a proton. But the charge on an electron is *negative*. Since matter in its normal state is electrically neutral, the atom (which is the basic building block of matter) should also be neutral. It means, that in an atom (in its normal state), the number of orbiting electrons must be the same as the number of protons in its nucleus.

There is no difference between an electron in an atom of copper and an electron in an atom of aluminium or any other element. Similarly, there is no difference between a proton in one atom and a proton in another atom of a different element. Likewise, the neutrons in the atoms of various elements are identical. Thus, it follows that electrons, protons and neutrons are the fundamental particles of the universe. If it is so, then why do various elements behave differently? This is because of the difference in the *number* and *arrangement* of the electrons, protons, and neutrons of which each atom is composed.

The number of protons in an atom (or of electrons in a neutral atom) is called its *atomic number*. All the electrons of an atom do not move in the same orbit. The electrons are arranged in different *orbits* or *shells*. The maximum number of electrons that can exist in the first shell (the one nearest to the nucleus) is two. The second shell can accommodate not more than eight electrons, the third not more than 18, the fourth not more than 32, and so on. In general, a shell can contain a maximum of $2n^2$ electrons, where n is the number of the shell. But to this rule, there is an exception. The outermost orbit in an atom cannot accommodate more than eight electrons. The electrons present in the outermost orbit are called *valence electrons*.

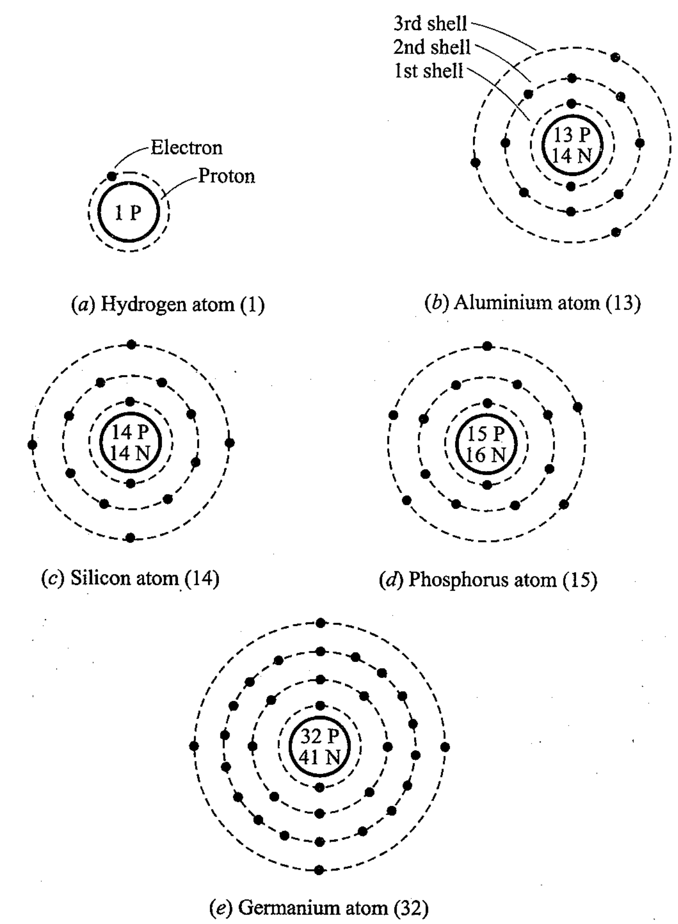
All the elements have been arranged in a *periodic table* according to the electronic arrangements in their atoms. The elements placed in one vertical column (called *group*) have very similar properties. A part of this periodic table is shown in Table 3.2. The atomic number of each element is shown within brackets along with its symbol. We shall study in the next section, the atomic structure of some elements useful for semiconductor devices.

Table 3.2 *A part of the periodic table of elements*

Group No.→	III	IV	V
	B (5)	C (6)	
	Al (13)	Si (14)	P(15)
	Ga (31)	Ge (32)	As(33)
	In (49)		Sb (51)

3.3.1 Atomic Structure of Some Elements

Figure 3.2 shows the representations of the atomic structures of different elements. Figure 3.2*a* is the simplest of all. It represents the hydrogen atom. It contains one electron revolving around one proton which is the nucleus. Note that the nucleus contains no neutrons.



(*a*) Hydrogen atom (1) (*b*) Aluminium atom (13)

(*c*) Silicon atom (14) (*d*) Phosphorus atom (15)

(*e*) Germanium atom (32)

Fig. 3.2 *Diagrammatic representation of a few atoms (figures inside brackets represent atomic numbers)*

Figure 3.2*b* shows the structure of an aluminium atom. The nucleus of the aluminium atom contains 13 protons (13 P) and 14 neutrons (14 N). The positive charge of 13 protons is just balanced by the negative charge of the 13 electrons (13 e) revolving around the nucleus in different shells. The atom as a whole is electrically neutral. Note that there are two electrons in the first shell, eight electrons in the second, and only *three* electrons in the third (the outermost). The importance of the outermost shell having only three electrons is explained in Section 3.6.2.

The electrons in the inner shells do not normally leave the atom. But the electrons which revolve at a great speed in the outermost shell (near the edge of the atoms) do not always remain confined to the same atom. Some of them move in a random manner and may travel from atom to atom. Figure 3.3 shows how the electrons may move from one atom to another in a random manner. Electrons that are able to move in this fashion are known as *free electrons*. It is due to the presence of these free electrons in a material, that it is able to conduct electric current. The electrons in the inner orbits remain bound to the nucleus and are therefore, called *bound electrons*.

Figure 3.2*c* represents a silicon atom. Its nucleus contains 14 protons and 14 neutrons. There are 14 electrons revolving around the nucleus—two in the first shell, eight in the second and *four* in the third. Thus, there are four valence electrons. The importance of this arrangement is explained in Section 3.5.1.

Figure 3.2*d* represents a phosphorus atom. There are 15 protons and 16 neutrons in its nucleus. It has *five* valence electrons. The importance of this arrangement is explained in Section 3.6.1.

Figure 3.2*e* shows a more complex structure. It represents a germanium atom (atomic number 32). Its nucleus contains 32 protons and 41 neutrons. There are two electrons in the first shell, eight in the second, 18 in the third and *four* in the fourth (outermost) shell. Thus, germanium, like silicon, has four valence electrons. Because of this similarity in atomic structure, many properties of the two materials are similar.

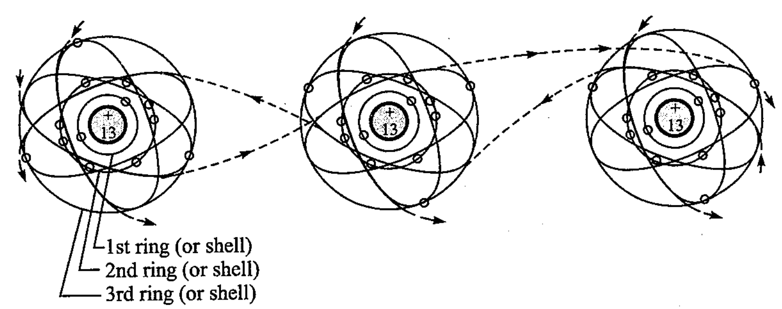


Fig. 3.3 *Random movement of outermost electrons in aluminium atoms*

In Table 3.2, the elements B(5), Al(13), Ga(31) and In(49) are placed in one group (group III) because all these have *three* valence electrons. Similarly, the elements P(15), As(33) and Sb(51) are all in group V, as they have *five* valence electrons. The elements Si(14) and Ge(32) are in group IV, since they have *four* valence electrons.

3.3.2 Electron Energies

Each isolated atom has only a certain number of orbits available. These available orbits represent energy levels for the electrons. Modern physics tells us that only discrete values of electron energies are possible. An electron cannot have *any* value

of energy (usually expressed in eV*), but only certain permissible values. No electron can exist at an energy level other than a permissible one. For a single atom, a diagram can be drawn showing the different energy levels available for its electrons. Figure 3.4 is the energy-level diagram for the hydrogen atom. The permissible energy levels are numbered $n = 1, 2, \ldots$ in increasing order of energy.

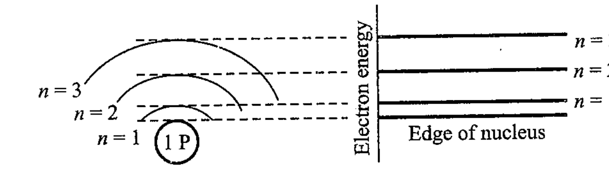


Fig. 3.4 *Permissible energy levels in an isolated hydrogen atom*

In an atom, the greater the distance of an electron from the nucleus, the greater is its total energy (the total energy includes kinetic and potential energies). An electron orbiting very close to the nucleus in the first shell is tightly bound to the nucleus and possesses only a small amount of energy. It would be difficult to knock out this electron. On the other hand, an electron orbiting far from the nucleus would have a greater energy, and hence it could easily be knocked out of its orbit. This is why it is the valence electrons (i.e., the electrons in the outermost orbit, having maximum energy) that take part in chemical reactions and in bonding the atoms together to form solids.

When energy like heat, light or other radiations impinge on an atom, the energy of the electrons increases. As a result, they are lifted to higher energy levels (larger orbits). The atom is then said to be *excited*. This state does not last long. Very soon, the electrons fall back to the original energy level. In this process, the electrons give out energy in the form of heat, light or other radiations.

3.3.3 Energy Bands in Solids

When atoms bond together to form a solid, the simple diagram of Fig. 3.4, for the electron energies, is no longer applicable. In a solid, the orbit of an electron is influenced not only by the charges in its own atom but by nuclei and electrons of every atom in the solid. Since each electron occupies a different position inside the solid, no two electrons can see exactly the same pattern of surrounding charges. As a result, the orbits of the electrons are different.

The simple energy-level diagram in Fig. 3.4 now modifies to that shown in Fig. 3.5. There are millions of electrons, belonging to the first orbits of atoms in the solid. Each of them has different energy. Since there are millions of first-orbit electrons, the

* eV is the abbreviation of *electron volt*; a unit of energy. It is defined as that energy which an electron acquires in moving through a potential rise of 1 V. This unit is commonly used in electronics and particle physics. In terms of *joules*, a more common unit of energy, an electron volt is equivalent to 1.6×10^{-19} J.

closely spaced energy levels differing very slightly in energy, form a cluster or *band*. Similarly, the second-orbit and higher-orbit electrons also form bands. We now have first energy band, second energy band, third energy band, etc.

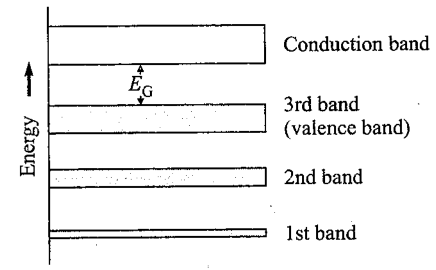


Fig. 3.5 *Energy-band diagram of a solid (silicon)*

Silicon is a material commonly used in making transistors. Since, the atomic number of silicon is 14, and each of its atoms has only four electrons in the third (outermost) orbit, the third band becomes the *valence band*. Figure 3.5 represents the energy-band diagram of silicon. An additional band, called *conduction band*, is also shown above the valence band. All the three lower bands, including the valence band, are shown completely filled. Although the third shell of an isolated atom of silicon is not completely filled (it has only four electrons whereas it could accommodate a maximum of eight electrons), the third energy band (valence band) of solid silicon is completely filled. It is so, because in solid silicon each atom positions itself between four other silicon atoms and each of these neighbours share an electron with the central atom. In this way, each atom now has eight electrons filling the valence band completely (for details see Section 3.5.1). When we say that a band is filled, it means that all the permissible energy levels in the band are occupied by electrons. No electron in a filled band can move, because there is no place to move. Thus, *an electron in a completely filled band cannot contribute to electric current*.

The conduction band represents the next larger group of permissible energy levels. There is an energy gap, E_G, between the valence band and the conduction band. An electron can be lifted from the valence band to the conduction band by adding to silicon some energy. This energy must be more than the energy gap E_G. If we add energy less than E_G, silicon will not accept it because no permissible energy levels exists between the conduction band and valence band to which an electron can be lifted. For this reason, the gap between the valence band and the conduction band is called the *forbidden energy gap*. (For silicon, E_G = 1.12 eV and for germanium it is 0.72 eV).

The orbits in the conduction band are very large. An electron in the conduction band experiences almost negligible nuclear attraction. In fact, an electron in the conduction band does not belong to any particular atom. But, it moves randomly throughout the solid. This is why the electrons in the conduction band are called *free electrons*.

3.4 METALS, INSULATORS AND SEMICONDUCTORS

A material is able to conduct electricity, if it contains movable charges in it. The free electrons (that is, the electrons that exist in the conduction band) move randomly inside a solid and can carry charge from one point to another, when an external field is applied. The free electrons thus work as *charge carriers*.

A metal such as copper or silver contains a large number of free electrons at room temperature. In fact, there is no forbidden energy gap between the valence and conduction bands. The two bands actually overlap as shown in Fig. 3.6*a*. The valence-band energies are the same as the conduction-band energies in the metal. It is very easy for a valence electron to become a conduction (free) electron. Therefore, without supplying any additional energy such as heat or light, a metal already contains a large number of free electrons and that is why it works as a good conductor.

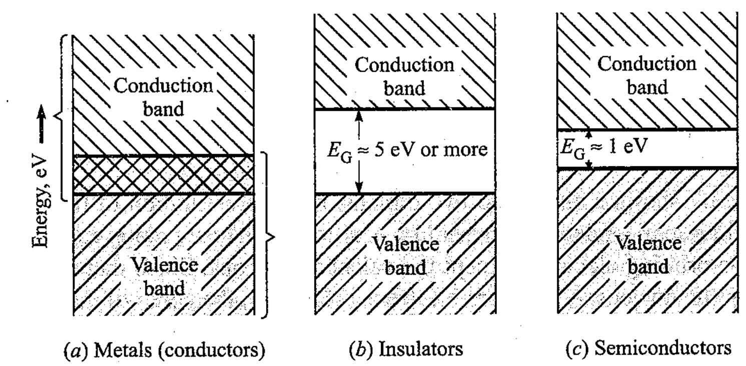


Fig. 3.6 *Energy-band diagram for the three types or materials*

An insulating material has an energy-band diagram as shown in Fig. 3.6*b*. It has a very wide forbidden-energy gap (5 eV or more). Because of this, it is practically impossible for an electron in the valence band to jump the gap, to reach the conduction band. Only at very high temperatures or under very stressed (electrically) conditions, can an electron jump the gap. At room temperature, an insulator does not conduct because there are no conduction electrons in it. However, it may conduct if its temperature is very high or if a high voltage is applied across it. This is termed as the *breakdown* of the insulator.

The energy-band diagram for a semiconductor is shown in Fig. 3.6*c*. In this case, the forbidden energy gap is not wide. It is of the order of 1 eV (for germanium, E_G = 0.72 eV and for silicon E_G = 1.12 eV). The energy provided by heat at room temperature is sufficient to lift electrons from the valence band to the conduction band. Some electrons do jump the gap and go into the conduction band. Therefore, at room temperature, semiconductors are capable of conducting some electric current.

3.5 INTRINSIC SEMICONDUCTORS

Semiconductor devices, such as diodes and transistors, are made from a single crystal of semiconductor material (germanium or silicon). To make a semiconductor device, the very first step is to obtain a sample of semiconductor in its purest form. Such a semiconductor (in pure form) is called an *intrinsic semiconductor*. A semiconductor is not truly intrinsic unless its impurity content is less than one part impurity in 100 million parts of semiconductor. To understand the phenomenon of conduction of current in a semiconductor, it is necessary to study its crystal structure.

3.5.1 Crystal Structure of Semiconductors

When atoms bond together to form molecules of matter, each atom attempts to acquire eight electrons, in its outermost shell. If the outermost shell of an atom has eight electrons, it is said to be filled. It then becomes a stable structure. An intrinsic semiconductor (such as pure Ge or Si), has only four electrons in the outermost orbit of its atoms. To fill the valence shell, each atom requires four more electrons. This is done by *sharing* one electron from each of the four neighbouring atoms. The atoms align themselves in a uniform three-dimensional pattern so that each atom is surrounded by four atoms. Such a pattern is called a *crystal*.

Figure 3.7 shows a simplified two-dimensional representation of the crystalline structure of a semiconductor (germanium or silicon). The *core* represents the nucleus and all the orbiting electrons except the valence electrons. Since there are as many protons in the nucleus as there are electrons orbiting it, the core will have an excess +4 charge since the valence electrons are four in number. (For silicon, the core will contain 14 protons but only 10 electrons). The valence electrons are shown around each core. Each of the four valence electrons take part in forming covalent bonds with the four neighbouring atoms. A covalent bond consists of two electrons, one from each adjacent atom. Both the electrons are shared by the two atoms. At absolute zero, *all* the valence electrons are tightly bound to the parent atoms. No free electrons are available for electrical conduction. *The semiconductor therefore behaves as a perfect insulator at absolute zero.*

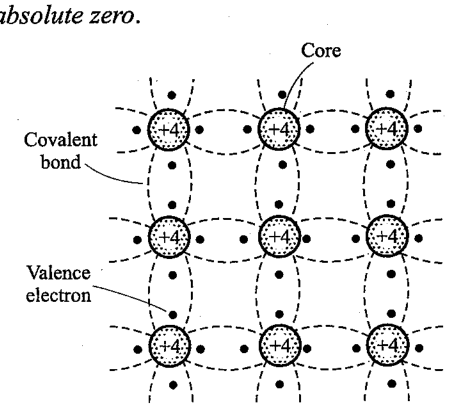
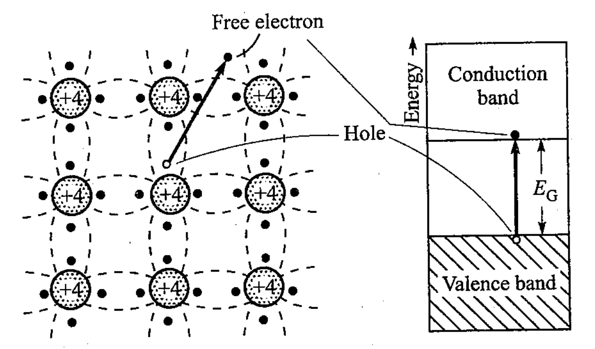


Fig. 3.7 *Simplified representation of the crystalline structure of a semiconductor at absolute zero*

3.5.2 Charge Carriers in Intrinsic Semiconductors

We have seen that an intrinsic semiconductor behaves as an insulator at absolute zero, because all the electrons are bound to the atoms. Let us now see what happens at room temperature. Room temperature (say, 300 K) may be sufficient to make a valence electron of a semiconductor atom to move away from the influence of its nucleus. Thus, a covalent bond is broken. When this happens, the electron becomes free to move in the crystal. This is shown in Fig. 3.8*a*.



(*a*) Crystal structure (*b*) Energy-band diagram

Fig. 3.8 *Generation of electron-hole pair in an intrinsic semiconductor*

When an electron breaks a covalent bond and moves away, a vacancy is created in the broken covalent bond. This vacancy is called a *hole*. Whenever a free electron is generated, a hole is created simultaneously. That is, free electrons and holes are always generated in *pairs*. Therefore, the concentration of free electrons and holes will always be equal in an intrinsic semiconductor. This type of generation of free electron-hole pairs is referred to as *thermal generation*.

Let us examine whether a hole has any charge associated with it. The crystal is electrically neutral. As soon as an electron-hole pair is generated, the electron leaves the covalent bond and moves away from it. Since, an electron is negatively charged, the site of a hole will be left with a net positive charge (equal in magnitude to the charge of an electron). Thus, we say that a positive charge is associated with a hole or *a hole is positively charged*. We shall see in the next section how a hole moves randomly in the same way as does a free electron. The hole too carries charge from one point to another. Although, strictly speaking, a hole is not a particle; for all practical purposes we can view it as a positively charged particle capable of conducting current. This concept of a hole as a positively charged particle merely helps in simplifying the explanation of current flow in semiconductors.

Figure 3.8*a* shows the generation of an electron-hole pair in a crystal. The amount of energy required to break a covalent bond is 0.72 eV in case of germanium and 1.12 eV in case of silicon. Equivalently, we say that the energy needed for lifting an electron form the valence band to the conduction band is 0.72 eV for germanium and

1.12 eV for silicon. When an electron jumps the forbidden gap, it leaves a hole in the valence band as shown in Fig. 3.8*b*.

Note that the value of E_G is more in case of silicon (E_G = 1.12 eV) than in case of germanium (E_G = 0.72 eV). Therefore, less number of electron-hole pairs will be generated in silicon than in germanium at room temperature. *The conductivity of silicon will be less than that of germanium at room temperature.*

3.5.3 Random Movement of Carriers

Both types of charge carriers move randomly or haphazardly in the crystal. The random movement of a free electron is easy to understand. A free electron moves in the crystal because of the thermal energy. Its path deviates whenever it collides with a nucleus (or other free electrons). This gives rise to a zigzag or random motion similar to gas molecules moving in a gas container.

Let us see how the hole movement takes place. A hole is generated whenever an electron breaks a covalent bond and becomes free. Consider that an electron-hole pair is generated at point A in Fig. 3.9*a*. The free electron goes elsewhere in the crystal leaving behind a hole at point A. The broken bond now has only one electron. This unpaired electron has a tendency to acquire an electron (whenever it can) and to complete the pair, forming the covalent bond. Due to thermal energy, an electron from the neighbouring bond may get sufficiently excited to break its own bond. It may then jump into the hole. In Fig. 3.9*b*, the valence electron at B breaks its bond and jumps into the hole at A. When this happens, the original hole at A vanishes and a new hole appears at B. The original hole has apparently moved from A to B, as shown in Fig. 3.9*c*. An instant later, the hole at B attracts and captures the valence electron from the neighbouring bond at C (see Fig. 3.9*d*). Apparently the hole has moved from B to C, as shown in Fig. 3.9*e*.

Figure 3.9*f* shows, by means of solid arrows, how the valence electrons successively jump from B to A and then from C to B. The net effect is as though the hole at A has moved through the crystal from A to C. This movement of holes is shown by dotted arrows in Fig. 3.9*f*. Thus, we find that the movement of a hole in a particular direction actually consists of a series of discontinuous electron movements in the opposite direction. It is for this reason that the holes appear to travel more slowly than the free electrons.

Although the movement of holes actually consists of the movement of electrons, this movement of electrons is different from the movement of free electrons. The free electrons move in the conduction band, but the holes move because of the movement of electrons in the valence band. The movement of the hole from A to C in the crystal can be shown in the energy-band diagram as in Fig. 3.9*g*. An electron jumps from the valence band to the conduction band leaving behind a hole at A. The electron at B moves to hole at A. An instant later, another electron at C moves to point B. The effect is as though the hole has moved from A to C. Actually the holes move because of the jumpy movement of valence electrons from one position to the other. This jumpy movement of valence electrons need not be considered at all, since we are concerned about the net effect (i.e., the movement of holes). Therefore, in future discussions,

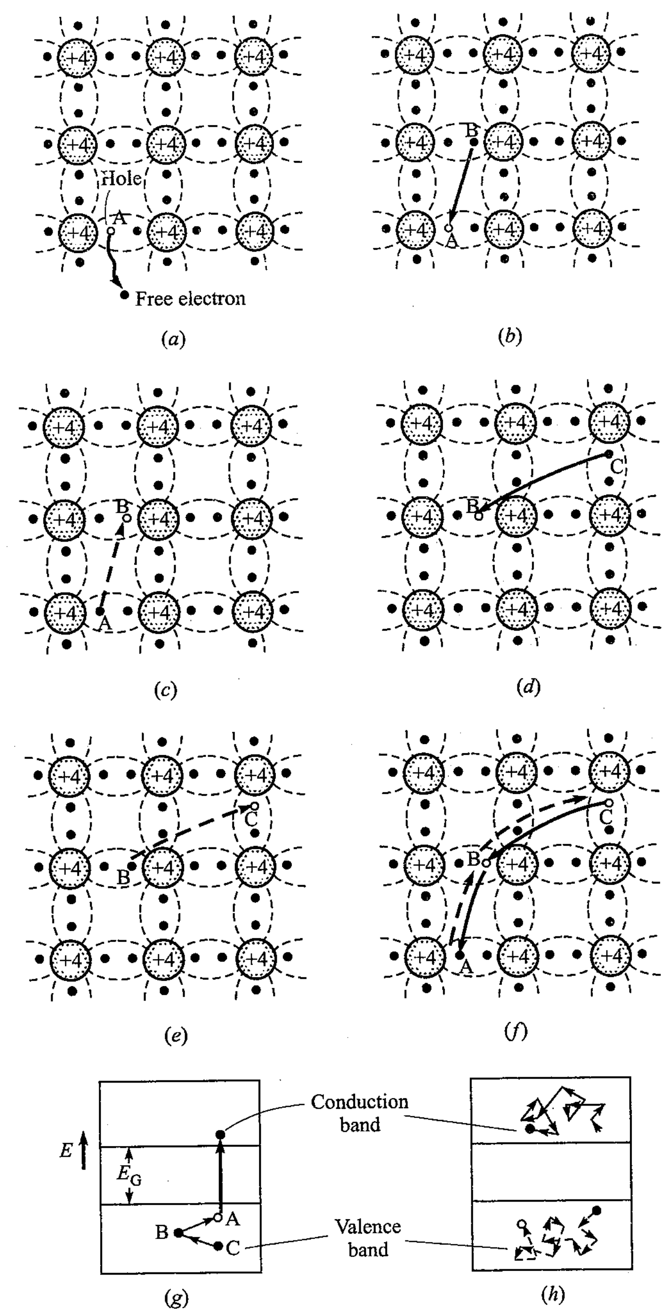


Fig. 3.9 *Movement of a hole through a semiconductor crystal*

whenever we talk of electron movement, it would imply the movement of free electrons and not of the valence electrons. Free electrons move randomly in the conduction band, whereas holes move randomly in the valence band as shown in Fig. 3.9*h*.

3.5.4 Recombination of Electrons and Holes

In an intrinsic semiconductor, electrons and holes are produced continuously on account of thermal agitation. Since the electrons and holes move in the crystal in a random manner, there is a possibility of an electron meeting a hole. When it happens, both the electron and hole disappear because the electron occupies the position of a hole in a broken covalent bond. The covalent bond is again established. At any temperature, the rate of this recombination is equal to the rate of generation of electrons and holes. However, an electron (or a hole) travels some distance before it recombines. The average time an electron (or hole) remains free is called its *lifetime*. At any instant, both types of charge carriers are present in equal numbers at a given temperature.

3.5.5 Conduction in Intrinsic Semiconductors

Let us see what happens when we connect a battery across a semiconductor, as in Fig. 3.10. The electrons experience a force towards the positive terminal of the battery; and holes towards the negative terminal. The random motion of electrons and holes gets modified. Over and above the random motion, there also occurs a net movement called *drift*. Since the random motion (or electrons or holes) does not contribute to any electric current, we need not consider it. The free electrons drift towards the positive terminal of the battery and the holes towards the negative terminal. The electric current flows through the semiconductor in the same direction as in which the holes are moving (the holes have positive charge). Since the electrons are negatively charged, the direction of electric current (conventional) is opposite to the direction of their motion. Although, the two types of charge carriers move in opposite directions, the two currents are in the same direction, i.e., they add together.

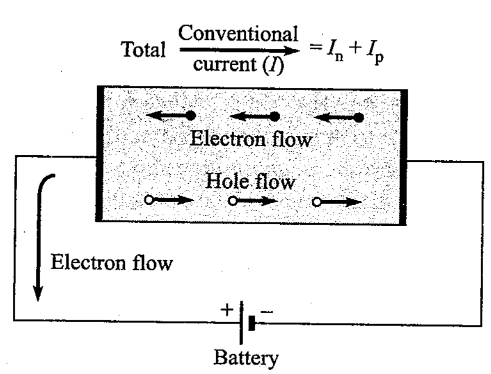


Fig. 3.10 *Conduction of current in an intrinsic semiconductor*

When the flow of carriers is due to an applied voltage (as in Fig. 3.10), the resultant current is called a *drift current*. A second type of current may also exist in a semiconductor. This current is called *diffusion current* and it flows as a result of a gradient of carrier concentration (i.e., the difference of carrier concentration from one region to another). A gradient of carrier concentration arises near the boundary of a PN-junction (as we shall see in next chapter). The diffusion current is also due to the motion of both holes and electrons.

3.5.6 Effect of Temperature on Conductivity of Intrinsic Semiconductors

A semiconductor (germanium or silicon) at absolute zero, behaves as a perfect insulator. At room temperature, because of thermal energy, some electron-hole pairs are generated. For example, in a sample of germanium at room temperature (300 K) the intrinsic carrier concentration (i.e., the concentration of free electrons or of holes) is 2.5×10^{19} per m^3. The semiconductor has a small conductivity. Now, if we raise the temperature further, more electron-hole pairs are generated. The higher the temperature, the higher is the concentration of charge carriers. As more charge carriers are made available, the conductivity of an intrinsic semiconductor increases with temperature. In other words, the resistivity (inverse of conductivity) decreases as the temperature increases. That is, *the semiconductors have negative temperature coefficient of resistance.*

3.6 EXTRINSIC SEMICONDUCTORS

Intrinsic (pure) semiconductors are of little use (it may only be used as a heat or light-sensitive resistance). Practically all the semiconductor devices are made of a semiconductor material to which certain specified types of impurities have been added. The process of deliberately adding impurities to a semiconductor material is called *doping*. Doping is done after the semiconductor material has been refined to a high degree of purity. A doped semiconductor is called an *extrinsic semiconductor*.

3.6.1 N-Type Semiconductors

Let us consider what happens if a small amount of pentavalent impurity, for example, phosphorus is added to a sample of intrinsic silicon. The size of the impurity atoms is roughly the same as that of silicon. An impurity atom replaces a silicon atom in its crystalline structure. If the amount of impurity is very small (say, one part in one million), we can safely assume that each impurity atom is surrounded, all around, by silicon atoms. This is shown in Fig. 3.11, which represents a part of the crystal.

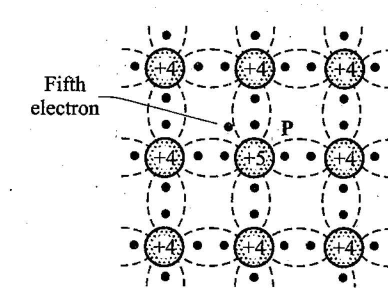


Fig. 3.11 N-*type semiconductor*

Let us now focus our attention on an impurity atom in the crystalline structure. Unlike a silicon atom, the phosphorus atom has five valence electrons. Four of these form covalent bonds with four neighbouring silicon atoms. The fifth electron has no chance of forming a covalent bond. It is this electron that is important to us. Since it is not associated with any covalent bond and is quite far from the nucleus, it is very loosely bound. It requires very little energy to free itself from the attractive force of its nucleus (this energy is only 0.01 eV in the case of germanium and 0.05 eV in the case of silicon). This energy is so small that at room temperature practically all such electrons become free. In other words, at room temperature each impurity atom *donates* one electron to the conduction band. That is the reason why this type of impurity is called *donor* type. These donated electrons are called *excess electrons*, since they are in excess to the electrons which are thermally generated (by breaking covalent bonds).

All the electrons which have been donated by the impurity atoms can take part in the conduction of electric current. Besides, there will also be some electron-hole pairs generated because of the breaking of covalent bonds. The number of thermally generated electron-hole pairs will be very small compared to the number of free electrons due to the impurity atoms. Further, as the number of electrons is very large, the chances of their recombination with holes also increases. Consequently, the net concentration of holes is much less than its intrinsic value. Thus, the number of free electrons becomes far greater than the number of holes. That is why we say that an N-type semiconductor has electrons (negatively charged) as *majority* carriers, and holes as *minority* carriers.

Now, let us see what happens to the core of the impurity atom, when the fifth electron leaves it. The core represents the atom without the valence electrons. Since there are five valence electrons in the impurity atom, a charge of +5 is shown in its core. When the fifth electron leaves the impurity atom, it then has +1 excess charge. It then becomes a positively charged *immobile ion*. It is immobile because it is held tightly in the crystal by the four covalent bonds.

Representation of* N*-type semiconductor In the designation "N-type semiconductor", the letter N stands for negative charges (electrons), because the electrons are the *major* charge carriers. But it does not mean that a sample of N-type semiconductor is negatively charged. It is important to note that whether a semiconductor is intrinsic or doped with impurity, it remains electrically neutral. Free electrons and holes are generated in pairs due to thermal energy. The negative charge of free electrons thus generated is exactly balanced by the positive charge of the holes. In an N-type semiconductor, there are additional free electrons created because of the addition of donor atoms. *The negative charge of these electrons is again balanced by the positive charge of the immobile ions.* (The total number of holes and immobile ions is exactly same as the total number of free electrons created.)

As we shall see in the chapters that follow, the N-type semiconductor (and also P-type semiconductor, which is explained in the next section) is used in the fabrication of diodes and transistors. To understand the mechanism of current flow through

these devices, we should consider all type of charged particles in the semiconductor. In an N-type semiconductor, there are a large number of free electrons, a few holes, and a sufficiently large number of immobile positive ions. In this book, we shall be representing an electron by a black circle, a hole by a white circle and an immobile positive donor ion by an encircled plus sign. Thus, we can represent an N-type semiconductor as shown in Fig. 3.12.

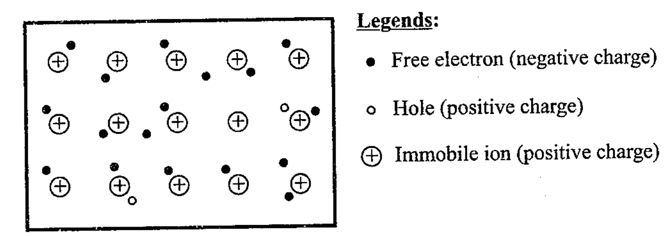
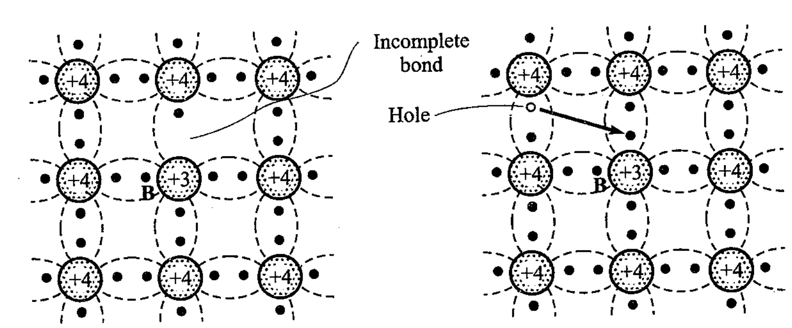


Fig. 3.12 *Representation of an* N*-type semiconductor*

Note that no silicon (or germanium) atoms are shown in this figure. They should be assumed as a continuous structure over the whole background. The fixed ions are regularly distributed in the crystal structure. But the holes and electrons, being free to move, are randomly distributed at that moment.

3.6.2 P-Type Semiconductor

For making an N-type semiconductor, we add a pentavalent impurity to an intrinsic semiconductor. Instead, if we add a trivalent impurity (such as boron, aluminium, gallium and indium) to the intrinsic semiconductor, the result is a P-type semiconductor. As an example, let us consider a sample of intrinsic (pure) silicon to which a very small amount of boron is added. Since the impurity ratio is of the order of one part in one million, each impurity atom is surrounded by silicon atoms. The boron atom in the crystal has only three valence electrons. These electrons form covalent bonds with the three neighbouring silicon atoms (Fig. 3.13). The fourth neighbouring silicon atom is unable to form a covalent bond with the boron atom because the boron atom does not have the fourth electron in its valence orbit.



(*a*) Boron added to silicon (*b*) Creation of a hole

Fig. 3.13 P*-type semiconductor*

There is a deficiency of an electron around the boron atom. The single electron in the incomplete bond has a great tendency to snatch an electron from the neighbouring atom. This tendency is so great that an electron in an adjacent covalent bond, having very small additional energy, can jump to occupy the vacant position. This electron then completes the covalent bond around boron atom. The additional energy required for this is of the order of 0.01 eV. At room temperature, the thermal energy is sufficient of provide this energy so as to fill the incomplete bonds around all the boron atoms.

When an electron from the adjacent covalent bond jumps to fill the vacancy in the incomplete bond around the boron atom, two things happen. First, a vacancy is created in the adjacent bond from where the electron had jumped. This vacancy has a positive charge associated with it, hence it is a *hole* (see Fig. 3.13*b*). Second, due to the filling of the incomplete bond around boron, it now becomes a *negative ion*. It is immobile, since it is held tightly in the crystal structure by covalent bonds. The boron atom becomes negative ion by *accepting* one electron from the crystal. That is why this type of impurity is called *acceptor* type.

Besides the excess holes created due to the addition of acceptor-type impurity, there are some holes (and also equal number of free electrons) generated by breaking covalent bonds. Summarising; a P-type material has holes (positively charged carriers) in *majority*, and free electrons in *minority*. In addition, there are also negative immobile ions.

Representation of P-type semiconductor Following the same convention as explained earlier, we can represent a sample of P-type semiconductor by a diagram (Fig. 3.14). The white circles represent holes, black circles the electrons, and encircled minus signs the immobile negative ions. The majority charge carriers in a P-type semiconductor are holes which are positively charged.

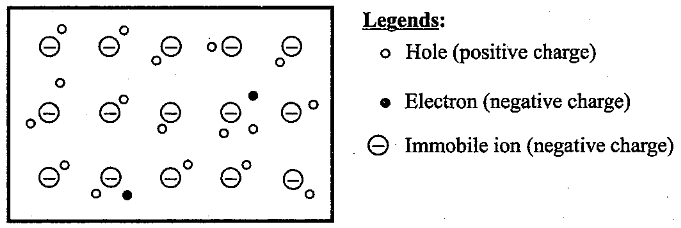


Fig. 3.14 *Representation of a P-type semiconductor*

3.6.3 Effect of Temperature on Extrinsic Semiconductors

We have seen that addition of a small amount of donor or acceptor impurity produces a large number of charge carriers in an extrinsic semiconductor. In fact, this number is so large that the conductivity of an extrinsic semiconductor is many times that of an intrinsic semiconductor at room temperature.

Let us see what happens if we raise the temperature of an N-type semiconductor. Since all the donors have already donated their free electrons (at room temperature), the additional thermal energy only serves to increase the thermally generated carriers. As a result, the concentration of minority carriers increase. Eventually, a temperature is reached when the number of covalent bonds that are broken is very large, so that the number of holes is approximately the same as the number of electrons. The extrinsic semiconductor now behaves essentially like an intrinsic semiconductor (of course, with higher conductivity). This *critical temperature* is 85 °C for germanium and 200 °C for silicon.

This concludes our study of semiconductor physics. In the chapters that follow, we will study some important semiconductor devices. Practically, all of these contain extrinsic semiconductors of both types in one crystal. The simplest combination, called a PN-junction, is the subject for the next chapter.

• Review Questions •

1. Name at least two conductors. Give the order of their conductivities.
2. Name any two insulators and give the order of their conductivities.
3. Name two commonly used semiconductors. Give the order of conductivities of these materials.
4. When atoms share electrons, what type of bonding is it called?
5. Explain why the discrete energy levels of an isolated atom split into a band of energy when atoms combine together to form a crystal.
6. Explain the difference between conductors, insulators and semiconductors using the energy-band diagrams.
7. What will happen to the number of electrons in the conduction band of a semiconductor as the temperature of the material is increased?
8. Sketch the two-dimensional crystal structure of intrinsic silicon at absolute zero of temperature. Also, sketch its energy-band diagram. Sketch the same crystal structure at room temperature. Also sketch its energy-band diagram.
9. Explain the reason why the conductivity of germanium is more than that of silicon at room temperature.
10. Why is a valence electron at the top of the valence band more apt to thermal excitation than the one at the lower level in that band?
11. Explain what a hole is? How do they move in an intrinsic semiconductor?
12. Explain why the temperature coefficient of resistance of a semiconductor is negative?
13. What process is the opposite of thermal generation of electron-hole pairs?
14. Explain, say within five lines, why the concentration of free electrons and holes is equal in an intrinsic semiconductor.
15. Explain what is the need of adding an impurity to an intrinsic semiconductor.
16. Which of the following atoms could be used as N-type impurities and which P-type impurities?

    (*a*) Phosphorous; (*b*) Antimony; (*c*) Boron; (*d*) Arsenic; (*e*) Aluminium; (*f*) Indium

17. Explain why a pentavalent impurity atom is known as donor-type impurity.
18. In an N-type semiconductor, does it take more energy to excite a valence electron thermally or to liberate an electron from the impurity atom? (5-7 lines).
19. What are the majority current carries in an N-type semiconductor? Why should there be any holes in this material?
20. Explain how holes are created in a P-type semiconductor.
21. Of what polarity are the impurity ions in N-type and P-type semiconductors? Justify your answer in brief (within 8 lines).
22. Explain what happens to the concentration of the majority carriers when the temperature of an extrinsic semiconductor is increased.
23. Explain why at high temperatures, an extrinsic semiconductor behaves like an intrinsic semiconductor.
24. In a P-type semiconductor, can the electrons ever outnumber the holes? Explain within 5-6 lines.
25. Explain, within 8 lines, why electrons are the majority carriers in an N-type semiconductor.

Objective-Type Questions

I. Below are some incomplete statements. Four alternatives are provided for each. Choose the alternative that completes the statement correctly.

1. The conductivity of materials found in nature varies between extreme limits of, say, 10^{-18} S/m to 10^{9} S/m. The probable value of conductivity of silicon is
   (*a*) 0.5×10^{-3} S/m (*b*) 1.0×10^{2} S/m
   (*c*) 0.7×10^{5} S/m (*d*) 1.8×10^{-12} S/m
2. A germanium atom contains
   (*a*) four protons (*b*) four valence electrons
   (*c*) only two electron orbits (*d*) five valence electrons
3. When atoms are held together by the sharing of valence electrons,
   (*a*) they from a covalent bond
   (*b*) the valence electrons are free to move away from the atom
   (*c*) each atom becomes free to move
   (*d*) each shared electron leaves a hole
4. An electron in the conduction band
   (*a*) is bound to its parent atom
   (*b*) is located near the top of the crystal
   (*c*) has no charge
   (*d*) has higher energy than an electron in the valence band
5. An intrinsic semiconductor at absolute zero of temperature

   (*a*) behaves like an insulator
   (*b*) has a large number of holes
   (*c*) has a few holes and same number of electrons
   (*d*) behaves like a metallic conductor
6. When a voltage is applied to an intrinsic semiconductor which is at room temperature,
   (*a*) electrons move to the positive terminal and holes move to the negative terminal
   (*b*) holes move to the positive terminal and electrons move to the negative terminal
   (*c*) both holes and electrons move to the positive terminal
   (*d*) both holes and electrons move to the negative terminal
7. When temperature of an intrinsic semiconductor is increased,
   (*a*) resistance of the semiconductor increases
   (*b*) heat energy decreases the atomic radius
   (*c*) holes are created in the conduction band
   (*d*) energy of the atoms is increased
8. The movement of a hole is brought about by
   (*a*) the vacancy being filled by a free electron
   (*b*) the vacancy being filled by a valence electron from a neighbouring atom
   (*c*) the movement of an atomic core
   (*d*) the atomic core changing from a +4 to a +5 charge
9. If a small amount of antimony is added to germanium
   (*a*) the resistance is increased
   (*b*) the germanium will be a P-type semiconductor
   (*c*) the antimony becomes an acceptor impurity
   (*d*) there will be more free electrons than holes in the semiconductor
10. Donor-type impurities
   (*a*) create excess holes
   (*b*) can be added to germanium, but not to silicon
   (*c*) must have only five valence electrons
   (*d*) must have only three valence electrons
11. The conduction band
   (*a*) is always located at the top of the crystal
   (*b*) is also called the forbidden energy gap
   (*c*) is a range of energies corresponding to the energies of the free electrons
   (*d*) is not an allowed energy band
12. The forbidden energy gap in semiconductors
   (*a*) lies just below the valence band

(*b*) lies just above the conduction band
(*c*) lies between the valence band and the conduction band
(*d*) is the same as the valence band

13. In an N-type semiconductor, the concentration of minority carriers mainly depends upon
(*a*) the doping technique
(*b*) the number of donor atoms
(*c*) the temperature of the material
(*d*) the quality of the intrinsic material, Ge or Si

14. If the amount of impurity, either P-type or N-type, added to the intrinsic is controlled to 1 part in one million, the conductivity of the sample
(*a*) increases by a factor of 10^6
(*b*) increases by a factor of 10^3
(*c*) decreases by a factor of 10^{-3}
(*d*) is not affected at all

15. A semiconductor that is electrically neutral
(*a*) has no majority carriers
(*b*) has no free charges
(*c*) has no minority carriers
(*d*) has equal amounts of positive and negative charges

16. When a normal atom loses an electron, the atom
(*a*) becomes a positive ion
(*b*) becomes a negative ion
(*c*) becomes electrically neutral
(*d*) is then free to move about

17. Excess minority carriers are the minority carriers that
(*a*) are thermally generated
(*b*) are impurity generated
(*c*) are in excess of the equilibrium number
(*d*) are in excess of the number of majority carriers

18. Resistivity is a property of a semiconductor that depends on
(*a*) the shape of the semiconductor
(*b*) the atomic nature of the semiconductor
(*c*) the shape and the atomic nature of the semiconductor
(*d*) the length of the semiconductor

II. Read each of the following statements and complete them by filling in the blanks with appropriate words:

1. The electrons in the outermost orbit are called ________ electrons.

2. The larger the orbit, the ________ is the energy of the electron.

3. The forces holding the silicon atoms together in a crystal are called ________ bonds.

4. The merging of a free electron and a hole is called ________.

5. A pure germanium crystal is an ________ semiconductor and a doped crystal is an ________ semiconductor.

6. To get excess electrons in an intrinsic semiconductor, we can add ________ atoms. These atoms have ________ valence electrons.

7. Free electrons are the ________ carriers in N-type semiconductors, and holes are the ________ carriers.

Answers

I.

1. (*a*)	2. (*b*)	3. (*a*)	4. (*d*)	5. (*a*)	6. (*a*)
7. (*d*)	8. (*b*)	9. (*d*)	10. (*c*)	11. (*c*)	12. (*c*)
13. (*c*)	14. (*b*)	15. (*d*)	16. (*a*)	17. (*c*)	18. (*b*)

II.

1. valence	2. greater	3. covalent
4. recombination	5. intrinsic, extrinsic	6. pentavalent, five
7. majority, minority		

4 UNIT
SEMICONDUCTOR DIODE

"In this electronic age we see ourselves being translated more and more into the form of information, moving toward the technological extension of consciousness."

Marshall McLuhan (1911–1980)
Canadian Communications Theorist, Educator, Writer and Social Reformer

After completing this unit, students will be able to:

- explain how barrier potential is set up in a PN-junction diode
- state the approximate value of barrier potential in a germanium and a silicon diode
- explain the meaning of space-charge region (depletion region), zener breakdown, avalanche breakdown, static resistance and dynamic resistance
- explain the conduction property of a forward-biased and reverse-biased diode
- draw the forward and reverse characteristics of germanium and silicon diodes
- explain the difference between germanium diodes and silicon diodes on the basis of forward voltage drop and reverse saturation current
- explain the effect of temperature on the reverse saturation current
- calculate the static and the dynamic resistance of a diode from its *V-I* characteristics
- draw the characteristics of an ideal diode
- explain the need of rectifiers in electronics
- draw the circuit diagram and explain the working of half-wave rectifier, centre-tapped full-wave rectifier and bridge rectifier

- compare the performance of half-wave rectifier and full-wave rectifier (both centre-tapped and bridge type)
- derive in case of half-wave and full-wave rectifiers, the expressions for output dc voltage, average or dc current, rms current, ripple factor, and rectification efficiency
- state the important ratings of a rectifying diode
- explain the need of filters in dc power supply
- explain with the help of suitable waveforms, the working of half-wave and full-wave rectifiers using shunt-capacitor, series inductor and π-filter
- state typical applications of light-emitting diode (LED), varactor diode and zener diode

4.1 PN-JUNCTION

By themselves, P-type and N-type materials taken separately are of very limited use. If we join a piece of P-type material to a piece of N-type material *such that the crystal structure remains continuous at the boundary*, a PN-junction is formed. Such a PN-junction makes a very useful device. It is called a *semiconductor* (or *crystal*) *diode*.

A PN-junction cannot be made by simply pushing the two pieces together; this would not lead to a single crystal structure. Special fabrication techniques are needed to form a PN-junction. For the time being, let us not bother how a PN-junction is formed.

A PN-junction itself is an important device. Furthermore, practically all semiconductor devices contain at least one PN-junction. For this reason, it is very necessary to understand how a PN-junction behaves when connected in an electrical circuit. In this chapter, we shall discuss the properties of a PN-junction. It will help us to understand even those devices which have more than one PN-junction.

4.2 JUNCTION THEORY

The most important characteristic of a PN-junction is its ability to conduct current *in one direction only*. In the other (reverse) direction, it offers very high resistance. How this happens is explained in the following sections.

4.2.1 PN-Junction with no External Voltage

Figure 4.1 shows a PN-junction just immediately after it is formed. Note that it is a single crystal. Its left half is P-type and right half is N-type. The P region has holes and negatively charged impurity ions. The N region has free electrons and positively charged impurity ions. Holes and electrons are the mobile charges, but the ions are immobile. The sample as a whole is electrically neutral and so are the P region and N region considered separately. Therefore in the P region, the charge of moving holes equal the total charges on its free electrons and immobile ions. Similarly, in the N region, the negative charge of its majority carriers is compensated by the charge of its minority carriers and immobile ions.

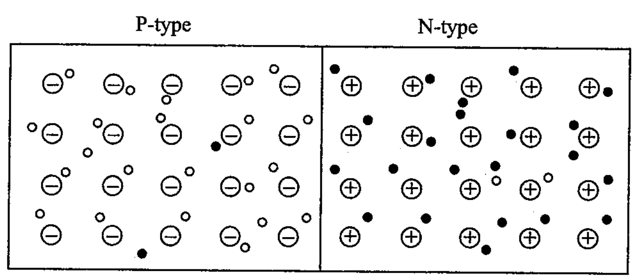


Fig. 4.1 *A PN-Junciton when just formed*

Note that no external voltage has been connected to the PN-junction of Fig. 4.1. As soon as the PN-junction is formed, the following processes are initiated:

1. Holes from the P region diffuse into the N region. They then combine with the free electrons in the N region.
2. Free electrons from the N region diffuse into the P region. These electrons combine with the holes.
3. The diffusion of holes (from P region to N region) and electrons (from N region to P region) takes place because they move haphazardly due to thermal energy and also because there is a difference in their concentrations in the two regions. The P region has more holes and the N region has more free electrons.
4. One would normally expect the holes in the P region and free electrons in the N region to flow towards each other and combine. Thus, all the holes and free electrons would have been eliminated. But in practice this does not occur. The diffusion of holes and free electrons across the junction occurs for a very short time. After a few recombinations of holes and electrons in the immediate neighbourhood of the junction, a restraining force is set up automatically. This force is called a *barrier*. Further diffusion of holes and electrons from one side to the other is stopped by this barrier. How this barrier force is developed is explained in the succeeding paragraphs.
5. Some of the holes in the P region and some of the free electrons in N region diffuse towards each other and recombine. Each recombination eliminates a hole and a free electron. In this process, the negative acceptor ions in the P region and positive donor ions in the N region in the immediate neighbourhood of the junction are left uncompensated. This situation is shown in Fig. 4.2. Additional holes trying to diffuse into the N region are repelled by the uncompensated positive charge of the donor ions. The electrons trying to diffuse into the P region are repelled by the uncompensated negative charges on the acceptor ions. As a result, *total* recombination of holes and electrons cannot occur.

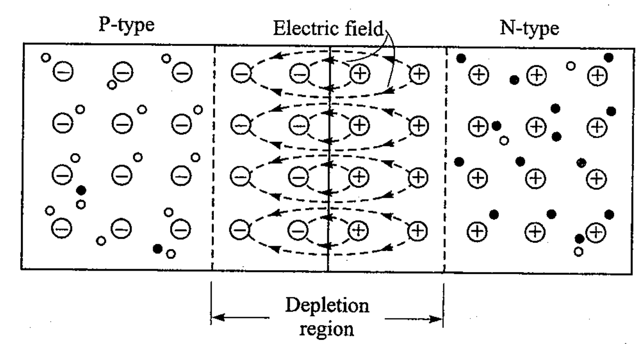


Fig. 4.2 *Space-charge region or depletion region is formed in the vicinity of the junction*

6. The region containing the uncompensated acceptor and donor ions is called *depletion region*. That is, there is a depletion of mobile charges (holes and free electrons) in this region. Since this region has immobile (fixed) ions which are electrically charged it is also referred to as the *space-charge region*. The electric field between the acceptor and the donor ions is called a *barrier*. The physical distance from one side of the barrier to the other is referred to as the *width* of the barrier. The difference of potential from one side of the barrier to the other side is referred to as the *height* of the barrier. With no *external* batteries connected, the barrier height is of the order of tenths of a volt. For a silicon PN-junction, the barrier potential is about 0.7 V, whereas for a germanium PN-junction it is approximately 0.3 V.
7. The barrier discourages the diffusion of majority carrier across the junction. But what happens to the minority carriers? There are a few free electrons in the P region and a few holes in the N region. The barrier helps these minority carriers to drift across the junction. The minority carriers are constantly generated due to thermal energy. Does it mean there would be a current due to the movement of these minority carriers? Certainly not. Electric current cannot flow since no circuit has been connected to the PN-junction. The drift of minority carriers across the junction is counterbalanced by the diffusion of the same number of majority carriers across the junction. These few majority carriers have sufficiently high kinetic energy* to overcome the barrier and cross the junction. In fact, the barrier height adjusts itself so that the flow of minority carriers is exactly balanced by the flow of majority carriers across the junction.

* In a semiconductor, all the charge carriers do not have same kinetic energy. Some have very high energy, whereas some have very low energy. The average energy depends upon the temperature of the sample.

Thus, we conclude that a barrier voltage is developed across the PN-junction even if no external battery is connected.

4.2.2 PN-Junction with Forward Bias

Suppose we connect a battery to the PN-junction diode such that the positive terminal of the battery is connected to the P-side and the negative terminal to the N-side as shown in Fig. 4.3. In this condition, the PN-junction is said to be *forward biased.*

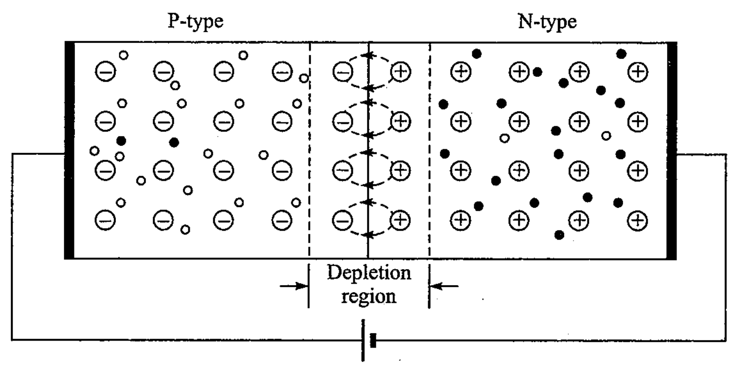


Fig. 4.3 *PN-junction showing forward bias*

When the PN-junction is forward biased, the holes are repelled from the positive terminal of the battery and are compelled to move towards the junction. The electrons are repelled from the negative terminal of the battery and drift towards the junction. Because of their acquired energy, some of the holes and the free electrons penetrate the depletion region. This reduces the potential barrier. The width of the depletion region reduces and so does the barrier height. As a result of this, more majority carriers diffuse across the junction. These carriers recombine and cause movement of charge carriers in the space-charge region.

For each recombination of free electron and hole that occurs, an electron from the negative terminal of the battery enters the N-type material. It then drifts towards the junction. Similarly, in the P-type material near the positive terminal of the battery, an electron breaks a bond in the crystal and enters the positive terminal of the battery. For each electron that breaks its bond, a hole is created. This hole drifts towards the junction. Note that there is a continuous electron current in the external circuit. The current in the P-type material is due to the movement of holes. The current in the N-type material is due to the movement of electrons. The current continues as long as the battery is in the circuit. If the battery voltage is increased, the barrier potential is further reduced. More majority carriers diffuse across the junction. This results in an increased current through the PN-junction.

4.2.3 PN-junction with Reverse Bias

Figure 4.4 shows what happens when a battery with the indicated polarity is connected to a PN-junction. Note that the negative terminal of the battery is connected to the

P-type material and the positive terminal of the battery to the N-type material. The holes in the P region are attracted towards the negative terminal of the battery. The electrons in the N region are attracted to the positive terminal of the battery. Thus the majority carriers are drawn away from the junction. This action widens the depletion region and increases the barrier potential (compare this with the unbiased PN-junction of Fig. 4.1).

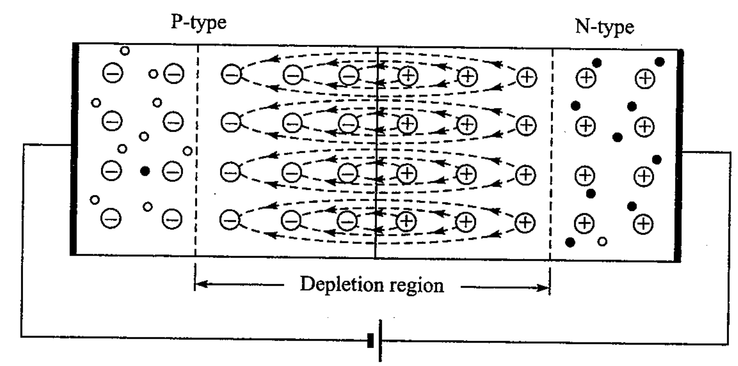


Fig. 4.4 *PN-junction showing reverse bias*

The increased barrier potential makes it more difficult for the majority carriers to diffuse across the junction. However, this barrier potential is helpful to the minority carriers in crossing the junction. In fact, as soon as a minority carrier is generated, it is swept (or drifted) across the junction because of the barrier potential. The rate of generation of minority carriers depends upon temperature. If the temperature is fixed, the rate of generation of minority carriers remains constant. Therefore, the current due to the flow of minority carriers remains the same whether the battery voltage is low or high. For this reason, this current is called *reverse saturation current.* This current is very small as the number of minority carriers is small. It is of the order of nanoamperes in silicon diodes and microamperes in germanium diodes.

There is another point to note. The reverse-biased PN-junction diode has a region of high resistivity (space charge or depletion region) sandwiched inbetween two regions (P and N regions away from the junction) of relatively low resistivity. The P and N regions act as the plates of a capacitor and the space-charge region acts as the dielectric. Thus, the PN-junction in reverse bias has an effective capacitance called *transition* or *depletion capacitance.*

Reverse breakdown We have seen that a PN-junction allows a very small current to flow when it is reverse biased. This current is due to the movement of minority carriers. It is almost independent of the voltage applied. However, if the reverse bias is made too high, the current through the PN-junction increases abruptly (see Fig. 4.5). The voltage at which this phenomenon occurs is called *breakdown voltage.* At this voltage, the crystal structure breaks down. In normal applications, this condition is avoided. The crystal structure will return to normal when the excess reverse bias is removed, provided that overheating has not permanently damaged the crystal.

There are two processes which can cause junction breakdown. One is called *zener breakdown* and the other is called *avalanche breakdown*. When reverse bias is increased, the electric field at the junction also increases. High electric field causes covalent bonds to break. Thus a large number of carriers are generated. This causes a large current to flow. This mechanism of breakdown is called *zener breakdown*.

In case of avalanche breakdown, the increased electric field causes increase in the velocities of minority carriers. These high energy carriers break covalent bonds, thereby generating more carriers. Again, these generated carriers are accelerated by the electric field. They break more covalent bonds during their travel. A chain reaction is thus established, creating a large number of carriers. This gives rise to a high reverse current. This mechanism of breakdown is called *avalanche breakdown*.

4.3 *V-I* CHARACTERISTICS OF A PN-JUNCTION DIODE

We would like to know how a device responds when it is connected to an electrical circuit. This information is obtained by means of a graph, known as its *V-I* characteristics, or simply characteristics. It is a graph between the voltage applied across its terminals and the current that flows through it. For a typical PN-junction diode, the characteristic is shown in Fig. 4.5. It tells us how much diode current flows for a particular value of diode voltage.

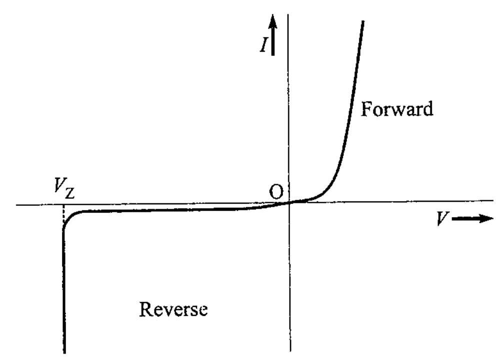
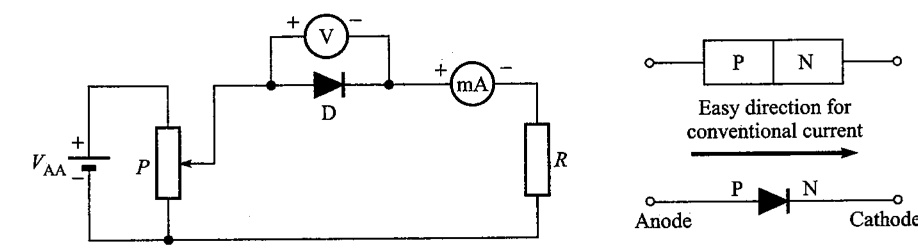


Fig. 4.5 *V-I characteristic of a* PN-*junction diode*

To obtain this graph, we set up a circuit in the laboratory. This circuit is shown in Fig. 4.6*a*. Note that in this circuit, the PN-junction is represented by its schematic *symbol*. The details of the diode symbol appear in Fig. 4.6*b*. The P region of the diode is called the *anode*, and the N region the *cathode*. The symbol looks like an arrow pointing from the P region to the N region. It serves as a reminder to us that the *conventional* current flows easily from the P region to the N region of the diode.

In the circuit (Fig. 4.6*a*), the dc battery V_{AA} is connected to the diode through the potentiometer *P*. Note that the dc battery is pushing the conventional current in the

same direction as the diode arrow. Hence, the diode is *forward biased*. Since current flows easily through a forward biased diode, a resistance *R* is included in the circuit so as to limit the current. If excessive current is permitted to flow through the diode, it may get permanently damaged. The potentiometer helps in varying the voltage applied to the diode. The milliammeter measures the current in the circuit. The voltmeter measures the voltage across the diode.



(*a*) Circuit used to obtain the *V-I* characteristic of a diode for forward bias

(*b*) Symbol of the diode

Fig. 4.6

Figure 4.7 shows the magnified view of a silicon-diode characteristic when the diode is forward biased. Note that the voltage is plotted along the horizontal axis, as voltage is the independent variable. Each value of the diode voltage produces a particular current. The current, being the dependent variable, is plotted along the vertical axis.

From the curve of Fig. 4.7, we find that the diode current is very small for the few tenths of a volt. The diode does not conduct well until the external voltage overcomes the barrier potential. As we approach 0.7 V larger number of free electrons and holes start crossing the junction. Above 0.7 V, even a small increase in the voltage produces a sharp increase in the current. The voltage at which the current starts to increase rapidly is called the *cut-in* or *knee voltage* (V_0) of the diode. For a silicon diode, it is approximately 0.7 V, whereas for a germanium diode, it is about 0.3 V. That is

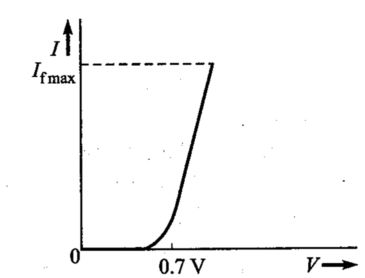
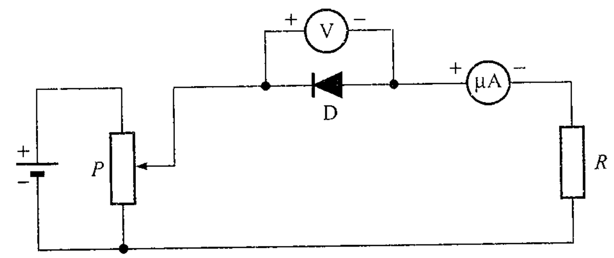


Fig. 4.7 *Forward characteristics of a silicon diode*

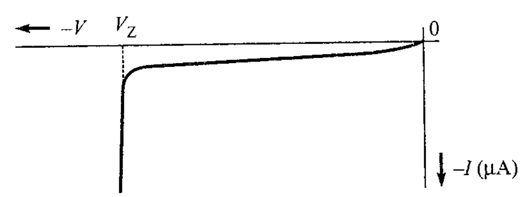
$$V_0 = 0.7\text{ V} \quad \text{for Si}$$
$$V_0 = 0.3\text{ V} \quad \text{for Ge}$$

If too large a current passes through the diode, excessive heat will destroy it. For this reason, the manufacturer's data sheet specifies the maximum current $I_{F\,max}$ that a diode can safely handle. For instance, the silicon junction diode BY126 has a maximum current rating of 1 A.

To obtain the reverse-bias characteristics, we use the same circuit as in Fig. 4.6*a*, except for a few changes. First, we reverse the terminals of the diode. Second, the milliammeter is replaced by a microammeter. The resulting circuit is as shown in Fig. 4.8*a*. The magnified view of the reverse characteristics of the diode is shown in Fig. 4.8*b*.



(*a*) Circuit to plot reverse-bias characteristics of a diode



(*b*) Reverse-bias characteristics

Fig. 4.8

In the reverse bias, the diode current is very small—only few μA for germanium diodes and only a few nA for silicon diodes. It remains small and almost constant for all voltages less than the breakdown voltage V_Z. At breakdown, the current increases rapidly for small increase in voltage. (See Sec. 4.2.3)

4.4 THE IDEAL DIODE

We have seen that a diode has a very important property. It permits only *unidirectional conduction*. It conducts well in the forward direction and poorly in the reverse direction. It would have been ideal if a diode acted as a perfect conductor (with zero voltage across it) when forward biased, and as a perfect insulator (with no current through it) when reverse biased. The *V-I* characteristics of such an *ideal diode* would be as shown in Fig. 4.9*a*. An ideal diode acts like an *automatic switch*. When the current tries to flow in the forward direction, the switch is *closed*. On the other hand, when the current tries to flow the other way (against the direction of the diode arrow) the switch is *open*.

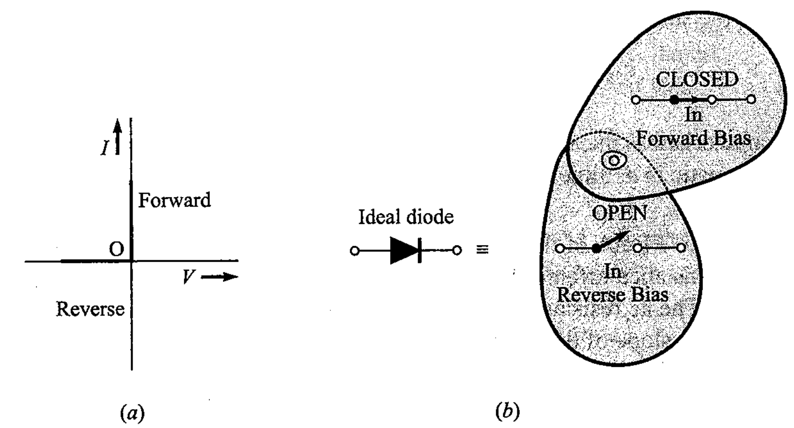


Fig. 4.9 (*a*) *Ideal-diode characteristics;* (*b*) *Switch analogy*

4.5 STATIC AND DYNAMIC RESISTANCE OF A DIODE

No diode can act as an ideal diode. An actual diode does not behave as a perfect conductor when forward biased, and as a perfect insulator when reverse biased. It does not offer zero resistance when forward biased. Also its reverse resistance, though very large, is not infinite.

Figure 4.10 shows the forward characteristics of a typical silicon diode. This diode may be connected in a dc circuit. When forward biased, it offers a definite resistance in the circuit. This resistance is known as the *dc* or *static resistance* (R_F) of the diode. It is simply the ratio of the dc voltage across the diode to the dc current flowing through it. For instance, if the dc voltage across the diode is 0.7 V, the current through it can be found from Fig. 4.10. The operating point of the diode is at point P, and the corresponding current can be read as 14 mA. The static resistance of the diode at this operating point will be given as

$$R_F = \frac{OA}{AP} = \frac{0.7\text{ V}}{14\text{ mA}} = 50\ \Omega$$

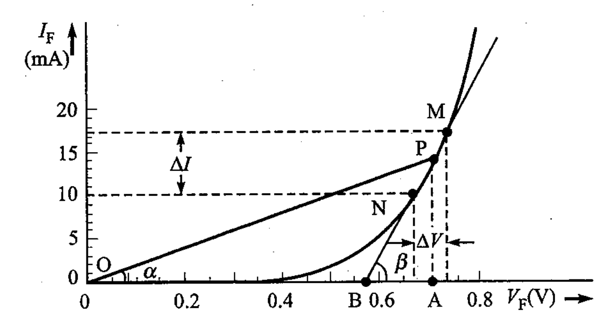


Fig. 4.10 *Calculation of static and dynamic resistance of a diode*

In general, the static resistance is given by the cotangent of the angle α. That is

$$R_F = \frac{OA}{AP} = \cot \alpha \qquad (4.1)$$

If the characteristic is linear, this ratio OA/AP will be a constant quantity. But, in case the characteristic is nonlinear, the dc resistance will vary with the point of measurement.

In addition to 14 mA of dc current, small ac current may be superimposed in the circuit. The resistance offered by the diode to this ac signal is called its *dynamic* or *ac resistance*. The ac resistance of a diode, at a particular dc voltage, is equal to the reciprocal of the slope of the characteristic at the point, i.e.,

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

Note

The Greek letter Δ (delta) means "a change of", wherever it appears in formulae. So, ΔI is a change in current. Generally, it indicates a small-scale (or incremental) change.

We can calculate the ac resistance of a diode as follows. Around the operating point P, take two points M and N very near to it, as shown in Fig. 4.10. These two points will then indicate incremental changes in voltage and current. The dynamic resistance is related to the slope of the line MN and is calculated as follows.

$$r_f = \frac{\Delta V}{\Delta I} = \frac{(0.73 - 0.66)\text{ V}}{(17.5 - 10)\text{ mA}} = \frac{0.07\text{ V}}{7.5\text{ mA}} = 9.46\ \Omega$$

The smaller the incremental changes ΔV and ΔI, the closer is the above result to the exact value of the dynamic resistance. For making these incremental values smaller, the points M and N have to be closer. It then becomes difficult to read the voltage and current values accurately from the graph. We can circumvent this difficulty if we remember that as ΔV becomes smaller and smaller, the slope of the line MN becomes the same as that of the tangent to the curve at point P. In this alternative approach, we first draw a tangent to the curve at point P. This tangent meets the x-axis at point B (see Fig. 4.10). The dynamic resistance of the diode is then given as

$$r_f = \frac{BA}{AP} = \cot \beta \qquad (4.2)$$

From the graph, we can calculate the dynamic resistance as

$$r_f = \frac{BA}{AP} = \frac{(0.7 - 0.57)\text{ V}}{14\text{ mA}} = \frac{0.13\text{ V}}{14\text{ mA}} = 9.3\ \Omega$$

This may be seen to be almost the same as the value obtained earlier.

Now look at the reverse characteristic of the PN-junction diode (Fig. 4.8*b*). We find that even for a large reverse voltage (but below breakdown) the current is very small. The reverse current may be 1 μA at a voltage of 5 V. Then static resistance of the diode is

$$R_R = \frac{5\text{ V}}{1\ \mu\text{A}} = 5\text{ M}\Omega$$

This is sufficiently high. It is much higher than the forward resistance R_R. Since the diode curve in the reverse bias is almost horizontal, its dynamic resistance r_r will be extremely high in this region of operation.

4.6 USE OF DIODES IN RECTIFIERS

Electric energy is available in homes and industries in India, in the form of alternating voltage. The supply has a voltage of 220 V (rms) at a frequency of 50 Hz. In the USA, it is 110 V at 60 Hz. For the operation of most of the devices in electronic equipments, a dc voltage is needed. For instance, a transistor radio requires a dc supply for its operation. Usually, this supply is provided by dry cells. But sometimes we use a battery eliminator in place of dry cells. The battery eliminator converts the ac voltage into dc voltage and thus *eliminates* the need for dry cells. Nowadays, almost all electronic equipments include a circuit that converts ac voltage of mains supply into dc voltage. This part of the equipment is called *power supply*. In general, at the input of the power supply, there is a power transformer. It is followed by a diode circuit called *rectifier*. The output of the rectifier goes to a *smoothing filter*, and then to a *voltage* regulator circuit. A block diagram of such a power supply is shown in Fig. 4.11. The rectifier circuit is the heart of a power supply.

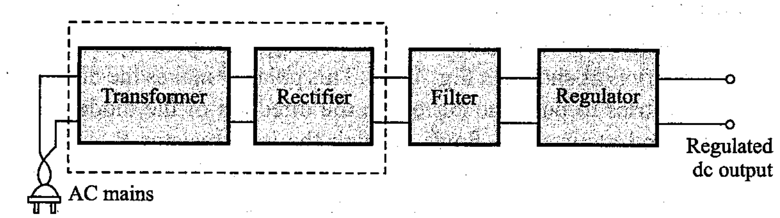


Fig. 4.11 *Block diagram of a power supply*

4.6.1 Half-Wave Rectifier

The unidirectional conducting property of a diode finds great application in rectifiers. These are the circuits which convert an ac voltage into dc voltage.

Figure 4.12 shows the circuit of a half-wave rectifier. Most electronic equipments have a transformer at the input. The transformer serves two purposes. First, it allows us to step the voltage up or down. This way we can get the desired level of dc voltage. For example, the battery eliminator used with a transistor radio gives a dc voltage of about 6 V. We can use a step-down transformer to get such a low ac voltage at the input of the rectifier. On the other hand, the cathode-ray tube used in an oscilloscope needs a very high dc voltage—of the order of a few kV. Here, we may use a step-up transformer. The second advantage of the transformer is the isolation it provides from the power line. It reduces the risk of electrical shock.

In Fig. 4.12, the diode forms a series circuit with the secondary of the transformer and the load resistor R_L. Let us see how this circuit rectifies ac into dc.

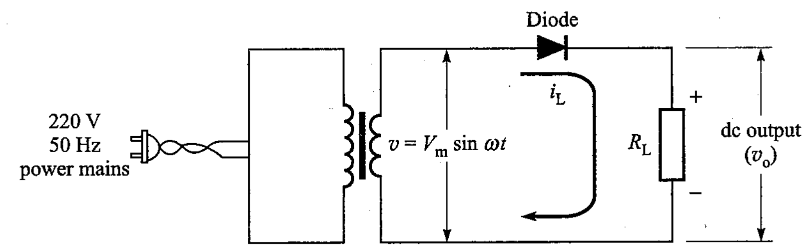


Fig. 4.12 *Half-wave rectifier circuit*

The primary of the transformer is connected to the power mains. An ac voltage is induced across the secondary of the transformer. This voltage may be less than, or equal to, or greater than the primary voltage depending upon the turn ratio of the transformer. We can represent the voltage across the secondary by equation

$$v = V_m \sin \omega t \tag{4.3}$$

Figure 4.13*a* shows how this voltage varies with time. It has alternate positive and negative half-cycles. Voltage V_m is the peak value of this alternating voltage.

During the positive half-cycles of the input voltage, the polarity of the voltage across the secondary is as shown in Fig. 4.14*a*. This polarity makes the diode forward biased, because it tries to push the current in the direction of the diode arrow. The diode conducts, and a current i_L flows through the load resistor R_L. This current makes the terminal A positive with respect to terminal B. Since a forward-biased diode offers a very low resistance, the voltage drop across it is also very small (about 0.3 V for Ge diode and about 0.7 V for Si diode). Therefore, the voltage appearing across the load terminals AB is practically the same as that the voltage v_i at every instant.

During the negative half-cycle of the input voltage, the polarity gets reversed. The voltage v tries to send current against the direction of diode arrow. See Fig. 4.14*b*. The diode is now reverse biased. It is shown shaded in the figure to indicate that it is non-conducting. Practically no current flows through the circuit. Therefore, almost no voltage is developed across the load resistance. All the input voltage appears across the diode itself. This explains how we obtain the output waveshape as shown in Fig. 4.13*b*.

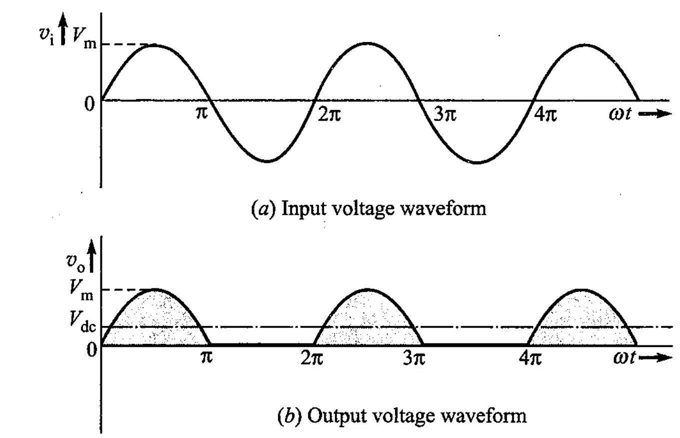


Fig. 4.13 *Half-wave rectifier*

To sum up, when the input voltage is going through its positive half-cycle, the voltage of the output is almost the same as the input voltage. During the negative half-cycle, no voltage is available across the load. The complete waveform of the voltage v_o across the load is shown in Fig. 4.13b. This voltage, though not a perfect dc, is at least *unidirectional.*

Peak inverse voltage Let us again focus our attention on the diode in Fig. 4.14*b*. During the negative half-cycle of the input, the diode is reverse biased. The whole of the input voltage appears across the diode (as there is no voltage across the load resistance). When the input reaches its peak value V_m, in the negative half-cycle, the voltage across the diode is also maximum. This maximum voltage is known as the *peak inverse voltage* (PIV). It represents the maximum voltage the diode must withstand during the negative half-cycle of the input. Thus, for a half-wave rectifier,

$$\text{PIV} = V_m \tag{4.4}$$

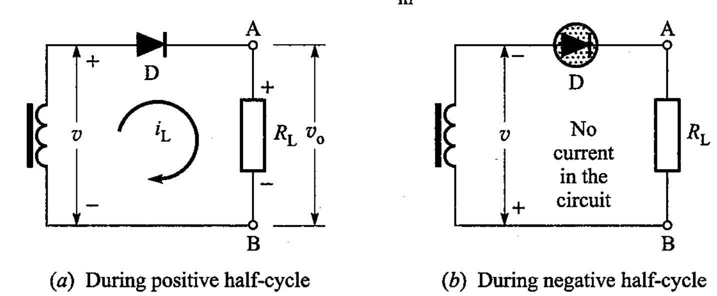


Fig. 4.14 *Half-wave rectifier circuit*

Output dc voltage The average value of a sine wave (such as that in Fig. 4.13*a*) over one complete cycle is zero. If a dc ammeter (moving coil type) is connected in an ac circuit, it will read zero. (The dc meter reads average value of current in a circuit.) Now, if the dc ammeter is connected in the half-wave rectifier circuit

(Fig. 4.12), it will show some reading. This indicates that there is some dc current flowing through the loading R_L. We can find the value of this current in a half-wave rectifier circuit.

In Fig. 4.13*b*, we had plotted the waveform of the voltage across the load resistor R_L. If we divide each ordinate of this curve by the value of resistance R_L, we get the current waveform. This is shown in Fig. 4.15. Note that the two waveforms (for current and for voltage) are similar. Mathematically, we can describe the current waveform as follows:

$$i_L = I_m \sin \omega t; \quad \text{for} \quad 0 \leqslant \omega t \leqslant \pi \tag{4.5}$$

and

$$i_L = 0; \quad \text{for} \quad \pi \leqslant \omega t \leqslant 2\pi \tag{4.6}$$

Here, I_m is the peak value of the current i_L. It is obviously related to the peak value of voltage V_m as

$$I_m = \frac{V_m}{R_L} \tag{4.7}$$

since the diode resistance in the conducting state is assumed to be zero. To find the dc or average value of current, we find the net area under the curve in Fig. 4.15 over one complete cycle, i.e., from 0 to 2π (curve repeats itself after the first cycle), and then divide this area by the base, i.e., 2π. We first integrate and then use Eqs. (4.5) and (4.6) to find the area.

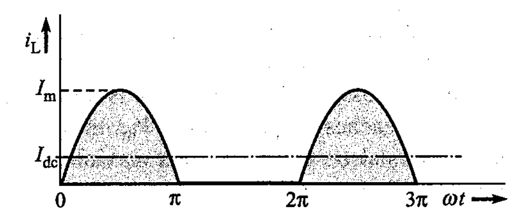


Fig. 4.15 *Waveform of the current flowing through load R_L in a half-wave rectifier*

$$\begin{aligned}
\text{Area} &= \int_0^{2\pi} i_L \, d(\omega t) \\
&= \int_0^{\pi} I_m \sin \omega t \, d(\omega t) + \int_{\pi}^{2\pi} 0 \, d(\omega t) \\
&= I_m [-\cos \omega t]_0^{\pi} + 0 \\
&= I_m [-\cos \pi - (-\cos 0)] \\
&= 2I_m
\end{aligned}$$

Average value of the load current is then

$$I_{avg} = I_{dc} = \frac{\text{area}}{\text{base}} = \frac{2I_m}{2\pi}$$

or

$$I_{dc} = \frac{I_m}{\pi} \tag{4.8}$$

The dc voltage developed across the load R_L is

$$V_{dc} = I_{dc} \times R_L = \frac{I_m}{\pi} \times R_L \tag{4.9}$$

While writing Eq. (4.7), we had assumed that

1. the diode resistance in forward bias is zero and
2. the secondary winding of transformer has zero resistance.

The second assumption is often very near the truth. The winding resistance is almost zero. But, the forward diode resistance r_d is sometimes not so small. If it is comparable to the load resistance R_L, we must take it into consideration. Equation (4.7) for peak current then gets modified to

$$I_m = \frac{V_m}{(R_L + r_d)} \tag{4.10}$$

The dc voltage across the load resistor R_L can now be written with the help of Eq. (4.9) as

$$\begin{aligned}
V_{dc} &= \frac{V_m R_L}{\pi (R_L + r_d)} = \frac{V_m}{\pi (1 + r_d / R_L)} \\
&\cong \frac{V_m}{\pi} \; (\text{if } r_d \ll R_L)
\end{aligned} \tag{4.11}$$

Example 4.1 The turns ratio of a transformer used in a half-wave rectifier (such as in Fig. 4.12) is $n_1 : n_2 = 12 : 1$. The primary is connected to the power mains: 220 V, 50 Hz. Assuming the diode resistance in forward bias to be zero, calculate the dc voltage across the load. What is the PIV of the diode?

Solution: The maximum (peak value) primary voltage is

$$V_p = \sqrt{2} V_{rms} = \sqrt{2} \times 220 = 311 \text{ V}$$

Therefore, the maximum secondary voltage is

$$V_m = \frac{n_2}{n_1} V_p = \frac{1}{12} \times 311 = 25.9 \text{ V}$$

The dc load voltage is

$$V_{dc} = \frac{V_m}{\pi} = \frac{25.9}{\pi} = \mathbf{8.24 \text{ V}}$$

The peak inverse voltage is

$$\text{PIV} = V_m = \mathbf{25.9 \text{ V}}$$

4.6.2 Full-Wave Rectifier

In a half-wave rectifier, discussed above, we utilise only one half-cycle of the input wave. In a full-wave rectifier we utilise both the half-cycles. Alternate half-cycles are inverted to give a unidirectional load current. There are two types of rectifier circuits that are in use. One is called *centre-tap rectifier* and uses two diodes. The other is called *bridge rectifier* and uses four diodes.

Centre-tap rectifier The circuit of a centre-tap rectifier is shown in Fig. 4.16*a*. It uses two diodes D1 and D2. During the positive half-cycles of secondary voltage, the diode D1 is forward biased and D2 is reverse biased. The current flows through the diode D1, load resistor R_L and the upper half of the winding as shown in Fig. 4.16*b*. During negative half-cycles diode D2 becomes forward biased and D1 reverse biased. Now D2 conducts and D1 becomes open. The current flows through diode D2, load resistor R_L and the lower half of the winding, as shown in Fig. 4.16*c*. Note that the load current in both Figs. 4.16*b* and *c* is in the *same direction*. The waveform of the current i_L, and hence of the load voltage v_o is shown in Fig. 4.16*d*.

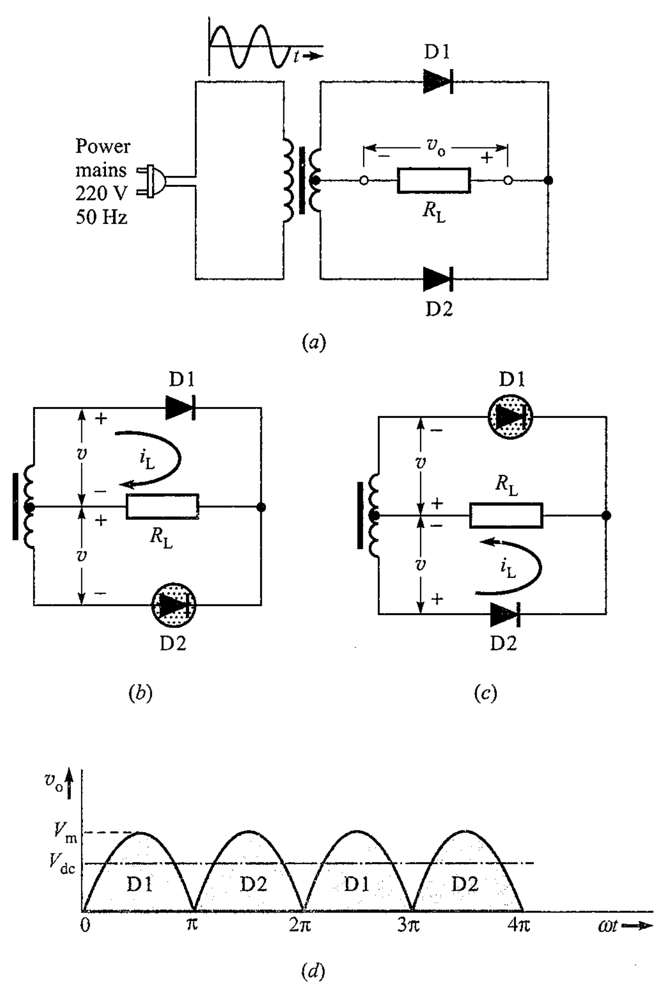


Fig. 4.16 *Centre-tap full-wave rectifier*

Peak inverse voltage Figure 4.17 shows the centre-tap rectifier circuit at the instant the secondary voltage reaches its positive maximum value. The voltage V_m is the maximum (peak) voltage across half of the secondary winding. At this instant, diode D1 is conducting and it offers almost zero resistance. The whole of the voltage V_m across the upper half winding appears across the load resistor R_L. Therefore, the reverse voltage that appears across the non-conducting diode is the summation of the voltage across the lower half winding and the voltage across the load resistor R_L. From the figure, this voltage is $V_m + V_m = 2V_m$. Thus,

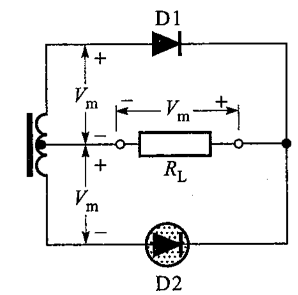


Fig. 4.17 *The PIV across the non-conducting diode D2 in a centre-tap rectifier is* $2V_m$

$$\text{PIV} = 2V_m \tag{4.12}$$

Bridge rectifier A more widely used full-wave rectifier circuit is the bridge rectifier shown in Fig. 4.18. It requires four diodes instead of two, but avoids the need for

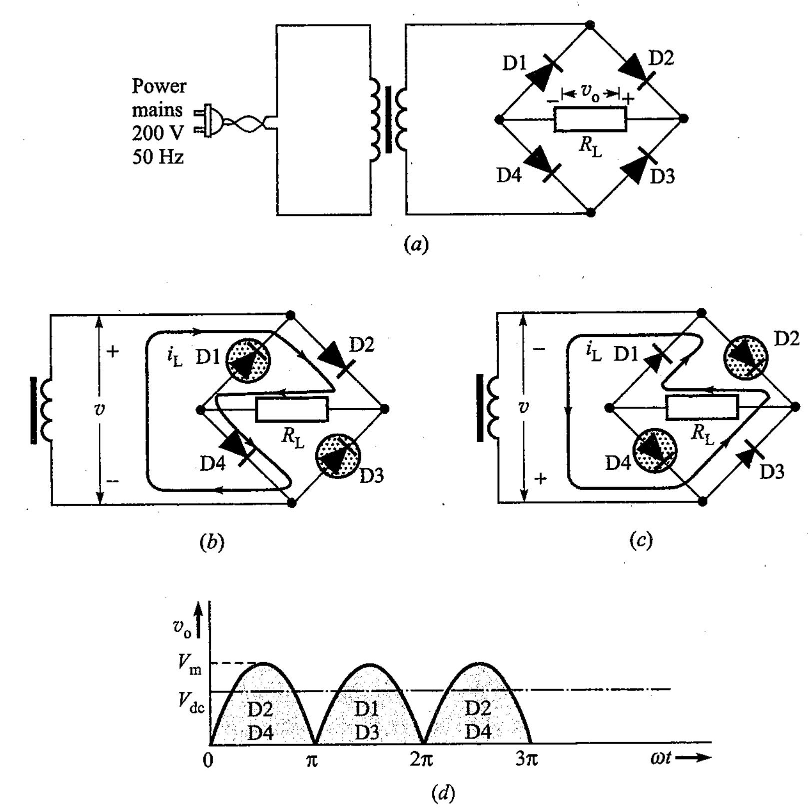


Fig. 4.18 *Bridge rectifier*

a centre-tapped transformer. During the positive half-cycles of the secondary voltage, diodes D2 and D4 are conducting and diodes D1 and D3 are non-conducting. Therefore, current flows through the secondary winding, diode D2, load resistor R_L and D4 as shown in Fig. 4.18*b*. During negative half-cycles of the secondary voltage, diodes D1 and D3 conduct and the diodes D2 and D4 do not conduct. The current therefore flows through the secondary winding, diode D1, load resistor R_L and diode D3 as shown in Fig. 4.18*c*. In both cases, the current passes through the load resistor *in the same direction*. Therefore, a fluctuating, unidirectional voltage is developed across the load. The load voltage waveform is shown in Fig. 4.18*d*.

Peak inverse voltage Let us now find the peak inverse voltage that appears across a non-conducting diode in a bridge rectifier. Figure 4.19 shows the bridge rectifier circuit at the instant the secondary voltage reaches its positive peak value, V_m. The diodes D2 and D4 are conducting, whereas diodes D1 and D3 are reverse biased and are nonconducting. The conducting diodes D2 and D4 have almost zero resistance (and hence zero voltage drops across them). Point B is at the same potential as the point A. Similarly, point D is at the same potential as the point C. The entire voltage V_m across the secondary winding appears across the load resistor R_L. The reverse voltage across the non-conducting diode D1 (or D3) is also V_m. Thus,

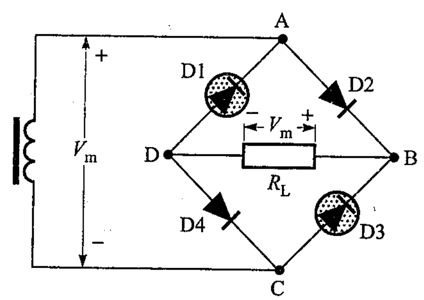


Fig. 4.19 *The PIV across the non-conducting diode D1 or D3 is V_m*

$$\text{PIV} = V_m \tag{4.13}$$

Output dc voltage in various rectifiers The voltage waveform in Fig. 4.18*d* is exactly the same as that in Fig. 4.16*d*. In both the rectifier circuits, the load voltage is the same. However, there is one difference. In the bridge rectifier, V_m is the maximum voltage across the secondary winding. But in the centre-tap rectifier, V_m represents the maximum voltage across half the secondary winding.

Now let us compare the full-wave rectified voltage waveform (of Fig. 4.18*d* or Fig. 4.16*d*) with the half-wave rectified voltage waveform (of Fig. 4.13*b*). In a half-wave rectifier, only positive half-cycle are utilised for the dc output. But a full-wave rectifier utilises both the half-cycles. Therefore, the dc or average voltage available in a full-wave rectifier will be double the dc voltage available in a half-wave rectifier. If the resistance of a forward-biased diode is assumed zero, the dc voltage of a full-wave rectifier (refer to Eq. 4.11) is

$$V_{dc} = \frac{2V_m}{\pi} \tag{4.14}$$

We can mathematically derive Eq. (4.14). The output voltage of a full-wave rectifier (see Fig. 4.18*b*) is described as

$$\begin{aligned} v_o &= V_m \sin \omega t \qquad 0 \leqslant \omega t \leqslant \pi \\ &= -V_m \sin \omega t \qquad \pi \leqslant \omega t \leqslant 2\pi \end{aligned}$$

A minus sign appears in the second equation because during the second half-cycle the wave is still sinusoidal, but inverted. The average or the dc value of voltage is

$$\begin{aligned} V_{dc} &= \frac{1}{2\pi}\int_0^{2\pi} v_o \, d(\omega t) \\ &= \frac{1}{2\pi}\left[\int_0^{\pi} (V_m \sin \omega t)\, d(\omega t) + \int_{\pi}^{2\pi} (-V_m \sin \omega t)\, d(\omega t)\right] \\ &= \frac{1}{2\pi}\left[\left|-V_m \cos \omega t\right|_0^{\pi} + \left|V_m \cos \omega t\right|_{\pi}^{2\pi}\right] \\ &= \frac{V_m}{2\pi}[-\cos\pi + \cos 0 + \cos 2\pi - \cos\pi] = \frac{2V_m}{\pi} \end{aligned}$$

This is same as Eq. (4.14).

Why bridge rectifier circuits are preferred As mentioned earlier, the bridge rectifier is the most widely used full-wave rectifier. It has many advantages over a centre-tap rectifier. It does not require centre-tapped secondary winding. (If stepping up or stepping down of voltage is not needed, we may even do away with the transformer.) The peak inverse voltage of each diode is equal to the peak secondary voltage V_m, whereas the PIV of the nonconducting diode in a centre-tap rectifier is $2V_m$. This fact is of vital importance when higher dc voltages are required.

Suppose we need a certain dc output voltage (say, $2V_m/\pi$) from a full-wave rectifier. If it is a bridge rectifier, the transformer secondary voltage need have a peak value of only V_m. But if it is a centre-tap rectifier, the secondary must have $2V_m$ as its peak voltage. This is twice the value needed for a bridge rectifier. It means that for a centre-tap rectifier, the transformer secondary must have double the number of turns. Such a transformer is costlier. Furthermore, each of the two diodes in a centre-tap rectifier must have a PIV rating of $2V_m$. But the diode in a bridge rectifier is required to have PIV rating of only V_m. Hence, the diodes meant for use in a centre-tap rectifier are costlier than those meant for a bridge rectifier.

The main disadvantage of a bridge rectifier is that it requires four diodes, two of which conduct on alternate half-cycles. This creates a problem when low dc voltages are required. The secondary voltage is low and the two diode voltage drops (1.4 V, in case of Si diodes) become significant. These diode voltage drops may be compensated by selecting a transformer with slightly higher secondary voltage. But then the voltage regulation becomes poor. For this reason, in low-voltage applications we prefer the centre-tap rectifier which has only one diode drop (= 0.7 V). By using germanium diodes instead of silicon, the diode drop may further be reduced to 0.3 V.

Example 4.2 The turns ratio of the transformer used in a bridge rectifier is $n_1 : n_2 = 12 : 1$. The primary is connected to 220 V, 50 Hz power mains. Assuming that the diode voltage drops to be zero, find the dc voltage across the load. What is the PIV of each diode? If the same dc voltage is obtained by using a centre-tap rectifier, what is the PIV?

Solution: The maximum primary voltage is

$$V_p = \sqrt{2}\, V_{rms} = \sqrt{2} \times 220 = 311 \text{ V}$$

Therefore, the maximum secondary voltage is

$$V_m = \frac{n_2}{n_1} V_p = \frac{1}{12} \times 311 = 25.9 \text{ V}$$

The dc voltage accross the load is

$$V_{dc} = \frac{2V_m}{\pi} = \frac{2 \times 25.9}{\pi} = \mathbf{16.48\ V}$$

The PIV (for bridge rectifier) is

$$\text{PIV} = V_m = \mathbf{25.9\ V}$$

For the centre-tap rectifier, the PIV is

$$\text{PIV} = 2V_m = 2 \times 25.9 = \mathbf{51.8\ V}$$

4.7 HOW EFFECTIVELY A RECTIFIER CONVERTS AC INTO DC

If we connect a load resistor R_L directly accross an ac power mains, the current flowing through it will be purely ac (sinusoidal having zero average value). This current is shown in Fig. 4.20*a*. In some applications, we require a dc current to flow through the load. The dc current* is unidirectional and, ideally, has no fluctuations with time. The ideal dc current is shown in Fig. 4.20*b*. To see how effectively a rectifier converts ac into dc, we compare its output current waveshape with the ideal dc current.

If the load takes current from a half-wave rectifier, the current waveform will be as in Fig. 4.20*c*. It is unidirectional, *but fluctuates greatly with time*. The waveform of the load current, when the load is connected to a full-wave rectifier, is shown in Fig. 4.20*d*. This too is unidirectional and *fluctuates with time*. A unidirectional, fluctuating waveform may be considered as consisting of a number of components. It has an average or dc value over which are superimposed a number of ac (sinusoidal)

* The terms ac and dc were originally used as the abbreviation of *alternating current* and *direct current*, respectively. Therefore, it may seem odd from the language point of view to use terms "an ac current, a dc current, a dc voltage, etc." However, the adjectives ac and dc have now been adopted for referring to any quantity whose variation with time is of "alternating" and "direct" type, respectively.

components of different frequencies. These undesired ac components are called *ripples*. The lowest ripple frequency in case of a half-wave rectifier is the same as the power-mains frequency. But, for full-wave rectifier it is not so. As can be seen from Figs. 4.20*d* and *a*, the period of the output wave of a full-wave rectifier is half the period of the input wave. The variation in current (or voltage) repeats itself after each angle π of the input wave. Therefore, the lowest frequency of the ripple in the output of a full-wave rectifier is twice the input frequency. That is, the ripple frequency,

$$f_r = f_i = 50 \text{ Hz} \quad \text{(half-wave rectifier)} \tag{4.15}$$

and

$$f_r = 2f_i = 100 \text{ Hz} \quad \text{(full-wave rectifier)} \tag{4.16}$$

How effectively a rectifier converts ac power into dc power is described quantitatively by terms such as *ripple factor, rectification efficiency*, etc.

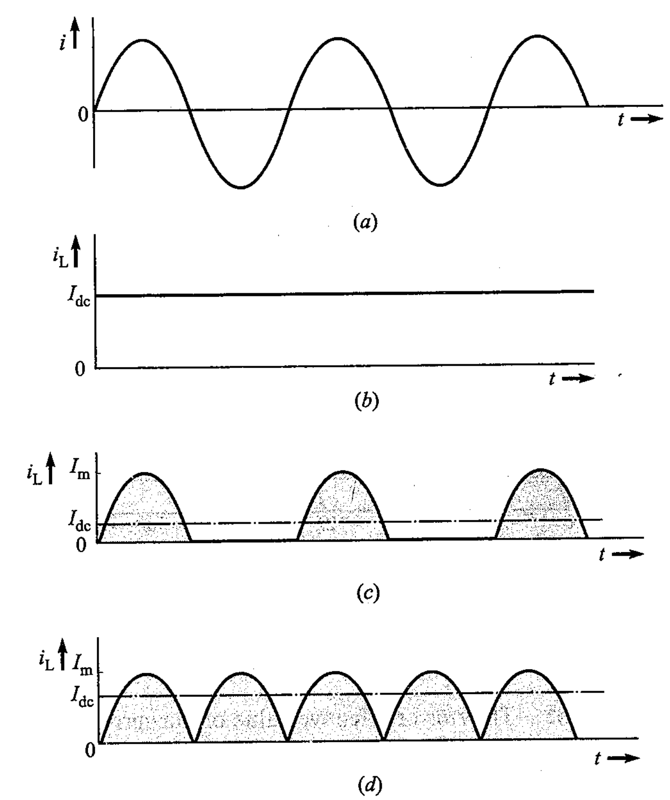


Fig. 4.20 *Comparison of half-wave and full-wave rectifiers with an ideal ac-to-dc converter*

The ***ripple factor*** is a measure of purity of the dc output of a rectifier, and is defined as

$$r = \frac{\text{rms value of the components of wave}}{\text{average or dc value}} \tag{4.17}$$

The ***rectification efficiency*** tells us what percentage of total input ac power is converted into useful dc output power. Thus, rectification efficiency is defined as

$$\eta = \frac{\text{dc power delivered to load}}{\text{ac input power from transformer secondary}}$$

or
$$\eta = \frac{P_{dc}}{P_{ac}} \tag{4.18}$$

Here, P_{ac} is the power that would be indicated by a wattmeter connected in the rectifying circuit with its voltage terminals placed across the secondary winding and P_{dc} is the dc output power.

We shall now analyse half-wave and full-wave rectifiers to find their ripple factor and rectification efficiency.

4.7.1 Performance of Half-Wave Rectifier

The half-wave rectified current wave is plotted in Fig. 4.21 and is described mathematically as

$$i_L = I_m \sin \omega t; \quad \text{for} \quad 0 \leqslant \omega t \leqslant \pi \tag{4.19}$$

and
$$i_L = 0; \quad \text{for} \quad \pi \leqslant \omega t \leqslant 2\pi \tag{4.20}$$

For determining the ripple factor or rectification efficiency, we first find the rms value of the current.

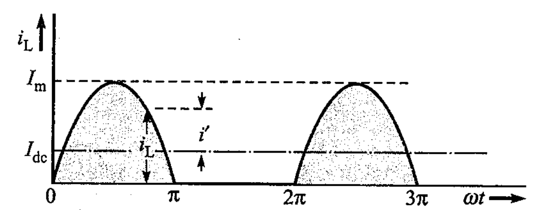


Fig. 4.21 *Half-wave rectified current wave (The instantaneous ac component of current is the difference between instantaneous total current and dc current, i.e., $i' = i_L - I_{dc}$)*

RMS value of current The rms or effective value of the current flowing through the load is given as

$$I_{rms} = \sqrt{\frac{1}{2\pi}\int_0^{2\pi} i_L^2 \, d(\omega t)}$$

where current i_L is described by Eqs. (4.19) and (4.20). Therefore,

$$I_{rms} = \sqrt{\frac{1}{2\pi}\left[\int_0^{\pi} I_m^2 \sin^2 \omega t \, d(\omega t) + \int_{\pi}^{2\pi} 0 \, d(\omega t)\right]}$$

$$= \sqrt{\frac{I_m^2}{2\pi}\int_0^{\pi} \frac{(1-\cos 2\omega t)}{2} d(\omega t)}$$

$$= \sqrt{\frac{I_m^2}{2\pi \times 2}\left|\omega t - \frac{\sin 2\omega t}{2}\right|_0^{\pi}}$$

or
$$I_{rms} = \frac{I_m}{2} \tag{4.21}$$

This is the rms value of the total current (dc value and ac components). As can be seen from Fig. 4.21, the instantaneous value of ac fluctuation is the difference of the instantaneous total value and the dc value. That is, the instantaneous ac value is given as

$$i' = i_L - I_{dc}$$

Therefore, the rms value of ac components is given as

$$I'_{rms} = \sqrt{\frac{1}{2\pi}\int_0^{2\pi} (i_L - I_{dc})^2 d(\omega t)}$$

$$= \sqrt{\frac{1}{2\pi}\int_0^{2\pi} (i_L^2 + I_{dc}^2 - 2 i_L I_{dc}) d(\omega t)} = \sqrt{I_{rms}^2 + I_{dc}^2 - 2I_{dc}^2}$$

or
$$I'_{rms} = \sqrt{I_{rms}^2 - I_{dc}^2} \tag{4.22}$$

Ripple factor From Eq. (4.17), the ripple factor is given as

$$r = \frac{I'_{rms}}{I_{dc}} = \sqrt{\frac{I_{rms}^2 - I_{dc}^2}{I_{dc}}} = \sqrt{\left(\frac{I_{rms}}{I_{dc}}\right)^2 - 1} \tag{4.23}$$

Using Eqs. (4.8) and (4.21), for half-wave rectifier, the ratio

$$\frac{I_{rms}}{I_{dc}} = \frac{I_m/2}{I_m/\pi} = 1.57$$

Therefore, the ripple factor is given as

$$r = \sqrt{(1.57)^2 - 1} = 1.21 \tag{4.24}$$

Thus, we see that the ripple current (or voltage) exceeds the dc current (or voltage). This shows that the half-wave rectifier is a poor converter of ac into dc.

Rectification efficiency For a half-wave rectifier, the dc power delivered to the load is

$$P_{dc} = I_{dc}^2 R_L = \left(\frac{I_m}{\pi}\right)^2 R_L$$

and the total input ac power is

$$P_{ac} = I_{rms}^2 (r_d + R_L) = \left(\frac{I_m}{2}\right)^2 (r_d + R_L)$$

Therefore, the rectification efficiency is

$$\eta = \frac{P_{dc}}{P_{ac}} = \frac{(I_m/\pi)^2 R_L}{(I_m/2)^2 (r_d + R_L)} \times 100\,\%$$

$$= \frac{40.6}{1 + r_d/R_L}\,\% \tag{4.25}$$

If $r_d \ll R_L$, $\eta \to 40.6$ per cent. It means that under the best conditions (i.e., no diode loss), only 40.6 % of the ac input power is converted into dc power. The rest remains as ac power in the load.

4.7.2 Performance of Full-Wave Rectifier

Figure 4.22 shows a full-wave rectified current wave. Its period may be seen to be π. The wave repeats itself after each π. Therefore, while computing the average or rms values, we should take the integration between the limits 0 to π, instead of 0 to 2π. The waveshape between 0 to π is described as

$$i_L = I_m \sin \omega t \tag{4.26}$$

where $\omega\,(= 2\pi f)$ is the angular frequency of the input ac voltage.

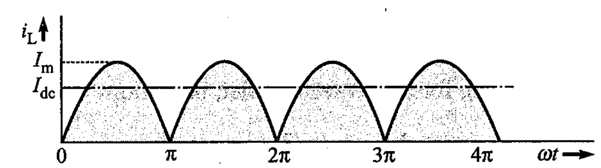


Fig. 4.22 *Full-wave rectified current wave*

RMS value of current Effective or rms value of current is given as

$$I_{rms} = \sqrt{\frac{1}{\pi}\int_0^{\pi} i_L^2\, d(\omega t)} = \sqrt{\frac{1}{\pi}\int_0^{\pi} I_m^2 \sin^2 \omega t\, d(\omega t)}$$

$$= \sqrt{\frac{I_m^2}{\pi}\int_0^{\pi}\left(\frac{1-\cos 2\omega t}{2}\right) d(\omega t)} = \sqrt{\frac{I_m^2}{\pi}\left|\frac{\omega t}{2} - \frac{\sin 2\omega t}{4}\right|_0^{\pi}}$$

$$= \sqrt{\frac{I_m^2}{\pi} \times \frac{\pi}{2}}$$

or $$I_{rms} = \frac{I_m}{\sqrt{2}} \tag{4.27}$$

Note that this is the same as the rms value of the full sinusoidal ac wave.

The dc or average value of the current is

$$I_{dc} = \frac{1}{\pi}\int_0^{\pi} i_L\, d(\omega t) = \frac{1}{\pi}\int_0^{\pi} I_m \sin \omega t\, d(\omega t)$$

$$= \frac{2 I_m}{\pi} \tag{4.28}$$

This current, as it should be, is double the dc current of a half-wave rectifier.

Ripple factor Equation (4.22) is valid for a full-wave rectifier too. We can therefore use Eq. (4.23) to calculate the ripple factor of a full-wave rectifier.

$$r = \sqrt{\left(\frac{I_{rms}}{I_{dc}}\right)^2 - 1} = \sqrt{\left(\frac{I_m/\sqrt{2}}{2I_m/\pi}\right)^2 - 1}$$

$$= 0.482 \tag{4.29}$$

Rectification efficiency For a full-wave rectifier, the dc power delivered to the load is

$$P_{dc} = I_{dc}^2 R_L = \left(\frac{2I_m}{\pi}\right)^2 R_L$$

and the total input ac power is

$$P_{ac} = I_{rms}^2 (r_d + R_L) = \left(\frac{I_m}{\sqrt{2}}\right)^2 (r_d + R_L)$$

Therefore, the rectification efficiency is

$$\eta = \frac{P_{dc}}{P_{ac}} = \frac{(2I_m/\pi)^2}{(I_m\sqrt{2})^2 (r_d + R_L)} \times 100\,\%$$

$$= \frac{81.2}{1 + r_d/R_L}\,\% \tag{4.30}$$

This shows that the rectification efficiency of a full-wave rectifier is twice that of a half-wave rectifier under identical conditions. The maximum possible efficiency can be 81.2 % (when $r_d \ll R_L$).

Example 4.3 In a centre-tap full-wave rectifier, the load resistance $R_L = 1\ k\Omega$. Each diode has a forward-bias dynamic resistance r_d of 10 Ω. The voltage across half the secondary winding is 220 sin 314t volts. Find (*a*) the peak value of current, (*b*) the dc or average value of current, (*c*) the rms value of current, (*d*) the ripple factor, and (*e*) the rectification efficiency.

Solution: The voltage across half the secondary winding is given as

$$v = 220 \sin 314t \text{ V}$$

(*a*) The peak value of voltage is

$$V_m = 220 \text{ V}$$

Therefore, peak value of current is

$$I_m = \frac{V_m}{r_d + R_L} = \frac{220}{10 + 1000} = \mathbf{0.2178\ A}$$
$$= \mathbf{217.8\ mA}$$

(*b*) The dc or average value of current is

$$I_{dc} = \frac{2I_m}{\pi} = \frac{2 \times 217.8}{\pi} = \mathbf{138.66\ mA}$$

(*c*) The rms value of current is

$$I_{rms} = \frac{I_m}{\sqrt{2}} = \mathbf{154\ mA}$$

(*d*) The ripple factor is given as

$$r = \sqrt{\left(\frac{I_{rms}}{I_{dc}}\right)^2 - 1} = \sqrt{\left(\frac{154}{138.66}\right)^2 - 1} = \mathbf{0.482}$$

(*e*) The rectification efficiency is given as

$$\eta = \frac{P_{dc}}{P_{ac}}$$

But $P_{dc} = I_{dc}^2 R_L = (138.66 \times 10^{-3})^2 \times 1000 = 19.2265 \text{ W}$

and $P_{ac} = I_{rms}^2 (r_d + R_L) = (154 \times 10^{-3})^2 \times (10 + 1000) = 23.953 \text{ W}$

$$\therefore \quad \eta = \frac{P_{dc}}{P_{ac}} = \frac{19.2265}{23.953} = 0.8026 = 0.8026 \times 100\ \% = \mathbf{80.26\ \%}$$

A full-wave rectifier is preferred to a half-wave rectifier, because its rectification efficiency is double and its ripple factor is low. Table 4.1 gives the comparison between different rectifiers discussed so far. Unless otherwise indicated, all rectifiers discussed from now on are full-wave rectifiers (either centre tap or bridge).

Table 4.1 *Comparison between different rectifiers*

	Half-wave	*Full-wave*	
		Centre-tap	*Bridge*
Number of diodes	1	2	4
Transformer necessary	No†	Yes	No†
Peak secondary voltage	V_m	V_m^*	V_m
Peak inverse voltage	V_m^{**}	$2V_m$	V_m
Peak load current, I_m	$V_m/(r_d + R_L)$	$V_m/(r_d + R_L)$	$V_m/(2r_d + R_L)$
RMS current, I_{rms}	$I_m/2$	$I_m/\sqrt{2}$	$I_m/\sqrt{2}$
DC current, I_{dc}	I_m/π	$2I_m/\pi$	$2I_m/\pi$
Ripple factor, r	1.21	0.482	0.482
Rectification efficiency (max)	40.6 %	81.2 %	81.2 %
Lowest ripple frequency, f_r	f_i	$2f_i$	$2f_i$

* It is the voltage between centre tap and one of the terminals.

** With a capacitor-input filter, the PIV of a half-wave circuit becomes $2V_m$, as we shall see later.

† Transformer may be used for isolation even if not required for stepping up (or down) the input ac.

4.8 HOW TO GET A BETTER DC

The object of rectification is to provide a steady dc voltage, similar to the voltage from a battery. We have seen that a full-wave rectifier provides a better dc than a half-wave rectifier. But, even a full-wave rectifier does not provide ripple-free dc voltage. These rectifiers provide what we may call "a pulsating dc". We can *filter* or *smooth* out the ac variations from the rectified voltage. For this we use a filter or smoothing circuit (see Fig. 4.11). In this section, we shall discuss different types of filter circuits.

4.8.1 Shunt Capacitor Filter

This is the simplest and the cheapest filter. You just have to connect a large value capacitor C in shunt with the load resistor R_L as shown in Fig. 4.23*a*. The capacitance offers a low-reactance path to the ac components of current. To dc (with zero frequency), this is an open circuit. All the dc current passes through the load. Only a small part of the ac component passes through the load, producing a small ripple voltage.

The capacitor changes the conditions under which the diodes (of the rectifier) conduct. When the rectifier output voltage is increasing, the capacitor charges to the peak voltage V_m. Just past the positive peak, the rectifier output voltage tries to fall (see the dotted curve in Fig. 4.23*b*). But at point B, the capacitor has $+V_m$ volts across it. Since the source voltage becomes slightly less than V_m, the capacitor will try to send current back through the diode (of the rectifier). This reverse-biases the diode, i.e., it becomes open-circuited.

The diode (open circuit) disconnects or separates the source from the load. The capacitor starts to discharge through the load. This prevents the load voltage from falling to zero. The capacitor continues to discharge until the source voltage (the dotted curve) becomes more than the capacitor voltage (at point C). The diode again starts conducting and the capacitor is again charged to peak value V_m. During the time the capacitor is charging (from point C to point D) the rectifier supplies the charging current i_C through the capacitor branch as well as the load current i_L. When the capacitor discharges (from point B to point C), the rectifier does not supply any current; the capacitor sends current i_L through the load. The current is maintained through the load all the time.

The rate at which the capacitor discharges between points B and C (in Fig. 4.23*b*) depends upon the time constant CR_L. The longer this time constant is, the steadier is the output voltage. If the load current is fairly small (i.e., R_L is sufficiently large) the capacitor does not discharge very much, and the average load voltage V_{dc} is slightly less than the peak value V_m (see Fig. 4.23*b*).

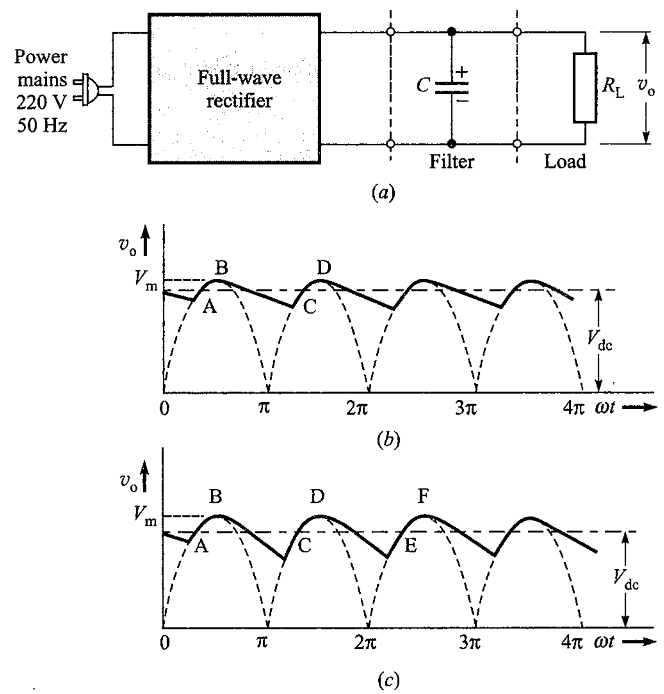


Fig. 4.23 *Full-wave rectifier with shunt capacitance filter*

An increase in the load current (i.e., decrease in the value of R_L) makes the time constant of the discharge path smaller. The capacitor then discharges more rapidly, and the load voltage is not constant (see Fig. 4.23*c*). The ripple increase with increase in load current. Also, the dc output voltage, V_{dc} decreases.

A much more steadier load voltage can be obtained if a capacitor of too large a value is used. But, the maximum value of the capacitance that can be employed is limited by another factor. The larger the capacitance value, the shorter is the period of charging the capacitor (from point C to D), and hence the greater is the current required to charge the capacitor to a given voltage. The maximum current that can be safely handled by a diode is limited by a figure quoted by the manufacturer. This puts a limit on the maximum value of the capacitance used in the shunt capacitor filter.

4.8.2 Series Inductor Filter

An inductor has the fundamental property of opposing any change in current flowing through it. This property is used in the series inductor filter of Fig. 4.24. Whenever the current through an inductor tends to change, a 'back emf' is induced in the inductor. This induced back emf prevents the current from changing its value. Any sudden change in current that might have occurred in the circuit without an inductor is smoothed out by the presence of the inductor.

Since the reactance of the inductor increases with frequency, better filtering of the higher harmonic ripples takes place. The output voltage waveform will therefore consist principally of the second harmonic frequency (the lowest ripple frequency), as shown in Fig. 4.24*b*. It shows a large dc component and a small ac component.

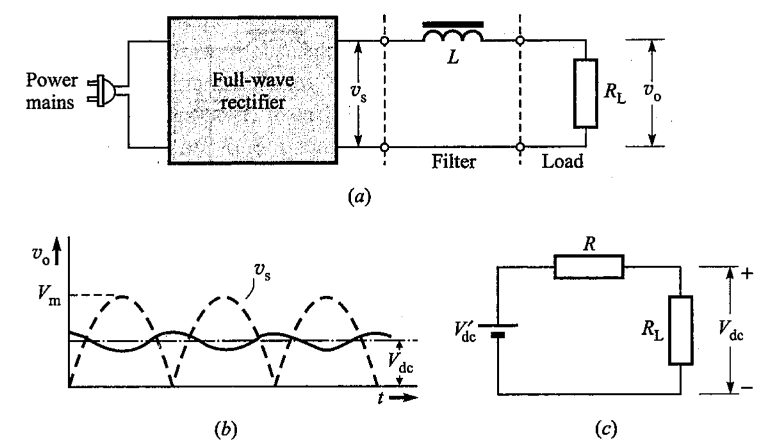


Fig. 4.24 *Full-wave rectifier with series inductor filter*

For dc (zero frequency), the choke resistance R in series with the load resistance R_L forms a voltage divider as shown in Fig. 4.24c. If V'_{dc} is the dc voltage from a full-wave rectifier, the dc voltage V_{dc} across the load is given as

$$V_{dc} = \frac{R_L}{R + R_L} V'_{dc} \tag{4.31}$$

Usually, R is much smaller than R_L; therefore, almost all of the dc voltage reaches the load.

The operation of a series inductor filter depends upon the current through it. Therefore, this filter (and also the choke-input LC filter discussed in the next section) can only be used together with a full-wave rectifier (since it requires current to flow at all times). Furthermore, the higher the current flowing through it, the better is its filtering action. Therefore, an increase in load current results in reduced ripple.

4.8.3 Choke-Input LC Filter

Figure 4.25 shows a choke-input filter using an inductor L in series and capacitor C in shunt with load. We have seen that a series inductor filter has the feature of decreasing the ripples when the load current is increased. Reverse is the case with a shunt capacitor filter. In this case, as the load current is increased, the ripples also increase. An LC filter combines the features of both the series inductor filter and shunt capacitor filter. Therefore, the ripples remain fairly the same even when the load current changes.

The choke (iron-core inductor) allows the dc component to pass through easily because its dc resistance R is very small. For dc, the capacitor appears as open circuit and all the dc current passes through the load resistance R_L. Therefore, the circuit acts like a dc voltage divider of Fig. 4.25*c*, and the output dc voltage is given by Eq. (4.31).

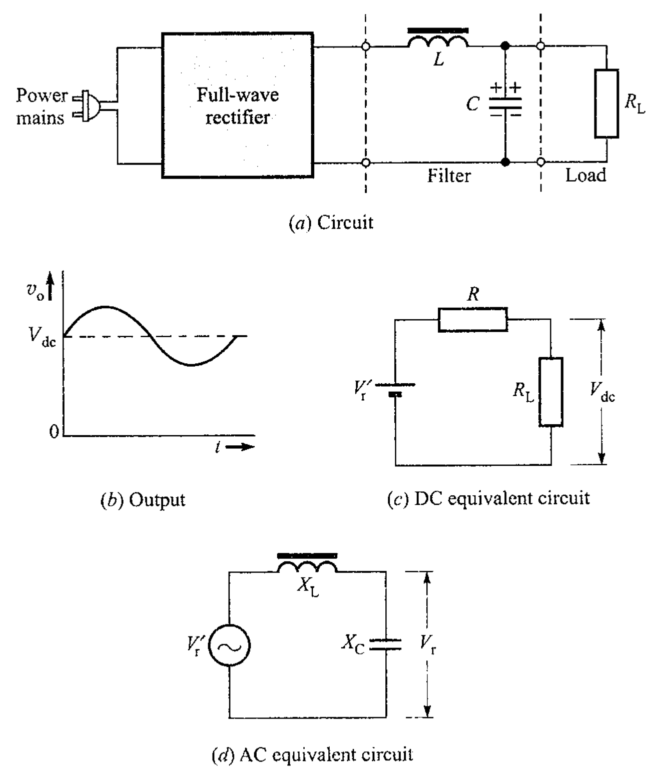


Fig. 4.25 *Full-wave rectifier with choke-input filter*

The fundamental frequency of the ac component in the output of the rectifier is 100 Hz (twice the line frequency). For this ac, the reactance $X_L(= 2\pi fL)$ is high. The ac current has difficulty in passing through the inductor. Even if some ac current manages to pass through the choke, it flows through the low reactance $X_C(= 1/2\pi fC)$ rather than through load resistance R_L. The ripples are reduced very effectively because X_L is much greater than X_C, and X_C is much smaller than R_L. The circuit works like the ac voltage divider of Fig. 4.25*d*. If V_r' is the rms value of the ripple voltage from the full-wave rectifier, then the rms value of the output ripple is given as

$$V_r \cong \frac{X_C}{X_L} V_r' \tag{4.32}$$

The reactances X_C and X_L are computed at 100 Hz. Typical values for L are 5 to 30 H and for C, 5 to 40 µF.

In the capacitor-input filter, the current flows through the transformer in a series of pulses. But in the choke-input filter, the current flows continuously. This means, the transformer is utilized more efficiently. A further advantage of the choke-input filter is that the ripple content at the output is not only low but is also less dependent on the load current.

Bleeder resistor Since an inductor depends upon current for its operation, it functions best under large current demands. For optimum functioning, the inductor should have a minimum current flowing at all times. If the current through the inductor falls below this minimum value, the output voltage rises sharply. The voltage regulation becomes poor. In order to provide this minimum current through the choke, a *bleeder resistor* R_b is usually included in the circuit. Figure 4.26 shows a bridge rectifier with a choke-input filter using a bleeder resistor.

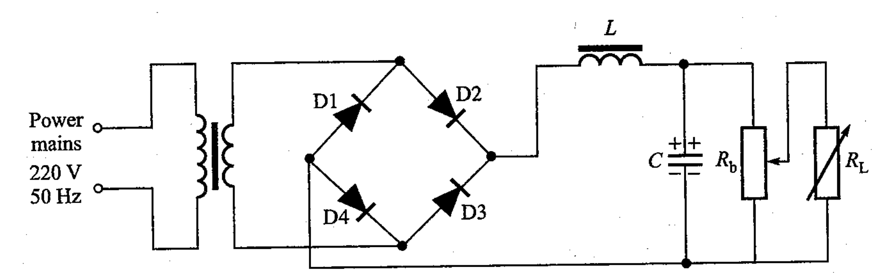


Fig. 4.26 *Bridge rectifier with a choke-input filter, using a bleeder resistor (The bleeder resistor maintains minimum current in L, discharges C and is used for varying the output)*

In Fig. 4.26, even if load resistance R_L becomes open circuit, the bleeder resistor R_b maintains the minimum current necessary for optimum inductor-operation. The bleeder resistor can serve a number of other functions as well. For example, it can be used as a voltage divider for providing a variable output voltage. It can also serve as a discharge path for the capacitor, so that voltage does not remain across the output terminals after the load has been disconnected, and the circuit de-energised.

This reduces the hazard of electrical shock when the load is connected to the output terminals next time.

4.8.4 π Filter

Very often, in addition to the *LC* filter, we use an additional capacitor C_1 for providing smoother output voltage. This filter is called π filter (its shape is like the Greek letter π). Such a filter is shown in Fig. 4.27. The rectifier now feeds directly into the capacitor C_1. Therefore, the filter is also called *capacitor-input filter*.

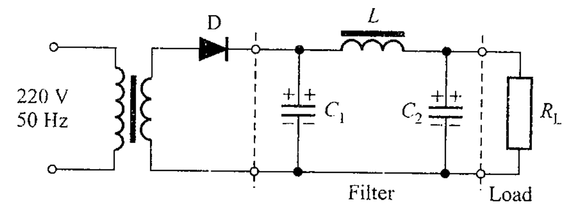


Fig. 4.27 *A half-wave rectifier feeding into a π-filter*

Since the rectifier feeds into the capacitor C_1, this type of filter can be used together with a half-wave rectifier. (The choke-input filter cannot be used with a half-wave rectifier.) Typical values for C_1 and C_2 for a half-wave rectifier are 32 μF each; and for *L*, 30 H. The half-wave rectifier ripple frequency being 50 Hz, these components have reactances of $X_L = 100\ \Omega$ and $X_C = 9492\ \Omega$ approximately. The reactances of *L* and C_2 act as an ac potential divider. This reduces the ripple voltage to approximately 100/9426 times its original value.

In the full-wave rectifier, the ripple frequency is 100 Hz. It means that a filter using the same component values would be more efficient in reducing the ripple. In other words, for a given amount of ripple smaller components can be used. Typical values are $C_1 = C_2 = 8$ μF and $L = 15$ H. Electrolytic capacitors have fairly large capacitances values and yet occupy minimum space. Usually both capacitors C_1 and C_2 are made inside one metal container. The metal container serves as the common ground for the two capacitors.

The disadvantages of the capacitor-input *LC* filter are the cost, weight, size and external field produced by the series inductor. These disadvantages can be overcome by replacing the series inductor with a series resistor of 100 to 200 Ω. It is then called capacitor-input *RC* filter. But this has the disadvantage of increasing the dc voltage drop in the filter. The voltage regulation becomes poorer. It also requires adequate ventilation to conduct away the heat produced in the resistor. As a result, it is only used to supply dc power to equipments taking only a small current.

4.9 TYPES OF DIODES

The important characteristics of semiconductor diodes are:

1. Maximum forward current

2. PIV rating
3. Forward and reverse ac resistances
4. Junction capacitance
5. Behaviour in breakdown region

One or more of these characteristics may be of prime importance depending upon the intended application of the diode. The main types of diodes used in electronic circuitry are:

1. Signal diodes
2. Power diodes
3. Zener diodes
4. Varactor diodes
5. Light-emitting diodes (LEDs)

4.9.1 Signal Diodes

These diodes are not required to handle large currents and/or voltages. The usual requirements are a large reverse-resistance/forward-resistance ratio and a minimum of junction capacitance. Some of the commercially available signal diodes are listed in the data book as general-purpose diodes. Some are best suited to a particular type of circuit application, such as a radio waves detector or as an electronic switch in logic circuitry. The maximum reverse voltage or peak inverse voltage, that the diode may be required to handle is usually not very high and neither is the maximum forward current. Most types of diodes have a PIV rating in the range 30 V to 150 V. The maximum forward current range may be somewhere between 40 mA and 250 mA.

4.9.2 Power Diodes

Power diodes are mostly used in rectifiers. The important parameters of a power diode are the peak inverse voltage, the maximum forward current, and the reverse-resistance/forward-resistance ratio. The peak inverse voltage rating is likely to be somewhere between 50 V and 1000 V. The maximum forward current may be perhaps 30 A or even more. As semiconductor technology advances, diodes capable of handling larger and larger power are being made available. Power diodes are usually silicon diodes. A power diode must have a forward resistance as low as possible. This helps in reducing the voltage drop across the diode when a large forward current flows. The forward resistance is usually not very much more than an ohm or two. The reverse resistance of a power diode must be as high as possible. Almost no current should flow through the diode when reverse biased.

4.9.3 Zener Diodes

Zener diodes are designed to operate in the breakdown region without damage. By varying the doping level, it is possible to produce zener diodes with breakdown voltages from about 2 V to 200 V.

As discussed in Section 4.2.3, the large current at breakdown is brought about by two factors, known as the zener and avalanche effects. When a diode is heavily doped, the depletion layer is very narrow. When the voltage across the diode is increased (in reverse bias) the electric field across the depletion layer becomes very intense. When this field is about 3×10^7 V/m, electrons are pulled from the covalent bonds. A large number of electron-hole pairs are thus produced and the reverse current sharply increases. This is known as the *zener effect*.

Avalanche effect occurs because of a cumulative action. The external applied voltage accelerates the minority carriers in the depletion region. They attain sufficient kinetic energy to ionise atoms by collision. This creates new electrons which are again accelerated to high-enough velocities to ionise more atoms. This way, an avalanche of free electrons is obtained. The reverse current sharply increases.

The zener effect is predominant for breakdown voltages less than about 4 V. The avalanche breakdown is predominant for voltages greater than 6 V. Between 4 and 6 V, both effects are present. It is the zener effect that was first discovered and the term 'zener diode' is in wide use for a *break-down diode* whether it uses zener effect or avalanche effect or both. If the applied reverse voltage exceeds the breakdown voltage, a zener diode acts like a *constant-voltage* source. For this reason, a zener diode is also called *voltage reference diode*.

The circuit symbol of a zener diode is shown in Fig. 4.28. A zener diode is specified by its breakdown voltage and the maximum power dissipation. The most common application of a zener diode is in the voltage stabilising or regulator circuits.

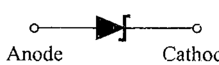


Fig. 4.28 *Circuit symbol of a zener diode*

Example 4.4 A zener diode is specified as having a breakdown voltage of 9.1 V, with a maximum power dissipation of 364 mW. What is the maximum current the diode can handle?

Solution: The maximum permissible current is

$$I_{Z\max} = \frac{P}{V_Z} = \frac{364\times10^{-3}}{9.1} = 40 \text{ mA}$$

Zener diode voltage regulator After the ripples have been smoothed or filtered from the rectifier output, we get a sufficiently steady dc output. But for many applications, even this sort of power supply may not serve the purpose. First, this supply does not have a good enough voltage regulation. That is, the output voltage reduces as the load (current) connected to it is increased. Secondly, the dc output voltage varies with the change in the ac input voltage. To improve the constancy of the dc output voltage as the load and/or the ac input voltage vary, a voltage-regulator circuit is used. The stabilizer circuit is connected between the output of the filter and the load (see Fig. 4.11).

The simplest regulator circuit consists merely of a resistor R_S connected in series with the input voltage, and a zener diode connected in parallel with the load (Fig. 4.29). The voltage from an unregulated power supply is used as the input voltage V_I to the *regulator* circuit. As long as the voltage across R_L is less than the zener breakdown voltage V_Z, the zener diode does not conduct. If the zener diode does not conduct, the resistors R_S and R_L make a potential divider across V_I. At an increased V_I, the voltage across R_L becomes greater than the zener breakdown voltage. It then operates in its breakdown region. The resistor R_S limits the zener current from exceeding its rate maximum $I_{Z\max}$.

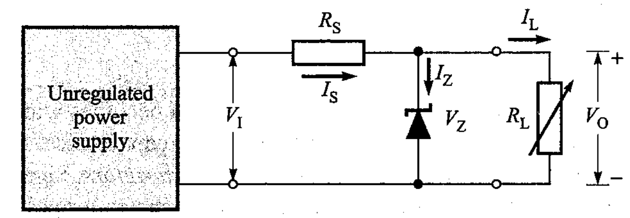


Fig. 4.29 *The zener-diode voltage regulator*

The current from the unregulated power supply splits at the junction of the zener diode and the load resistor. Therefore,

$$I_S = I_Z + I_L \tag{4.33}$$

When the zener diode operates in breakdown region, the voltage V_Z across it remains fairly constant even though the current I_Z flowing through it may vary considerably. If the load current I_L should increase (because of the reduction in load resistance), the current I_Z through the zener diode falls by the same percentage in order to maintain constant current I_S. This keeps the voltage group across R_S constant. Hence, the output voltage V_O remains constant. If, on the other hand, the load current should decrease, the zener diode passes an extra current I_Z such that the current I_S is kept constant. The output voltage of the circuit is thus stabilised.

Let us examine the other cause of the output voltage variation. If the input voltage V_I should increase, the zener diode passes a larger current so that extra voltage is dropped across R_S. Conversely, if V_I should fall, the current I_Z also falls, and the voltage drop across R_S is reduced. Because of the self-adjusting voltage drop across R_S, the output voltage V_O fluctuates to a much lesser extent than does the input voltage V_I.

4.9.4 Varactor Diodes

A reversed-biased PN-junction can be compared to a charged capacitor. The P and N regions (away from the space charge region) are essentially low resistance areas due to high concentration of majority carriers. The space-charge region, which is depleted of majority carriers, serves as an effective insulation between the P and N regions. The P and N regions act as the plates of the capacitor while the space-charge region acts as the insulating dielectric. The reverse-biased PN-junction thus has an effective capacitance, whose value is given as

$$C = \frac{\varepsilon A}{W} \tag{4.34}$$

where ε (the Greek letter "epsilon") is the permittivity of the semiconductor material, A is the area of the junction, and W is the width of the space-charge region. The width W of the space-charge region is approximately proportional to the square root of the reverse bias voltage V. The area A and permittivity ε being constant, we can write Eq. (4.34) as

$$C = \frac{K}{\sqrt{V}} \tag{4.35}$$

As the reverse bias increases, the space-charge region becomes wider, thus effectively increasing the plate separation and decreasing the capacitance. Silicon diodes optimised for this variable-capacitance effect are called *varactors*. Figure 4.30 shows the symbol used to represent a varactor diode. It also shows graphically how the capacitance of a varactor diode varies with the reverse-bias voltage. Typically, the capacitance variation might be 2-12 pF, or 20-28 pF, or perhaps 28-76 pF.

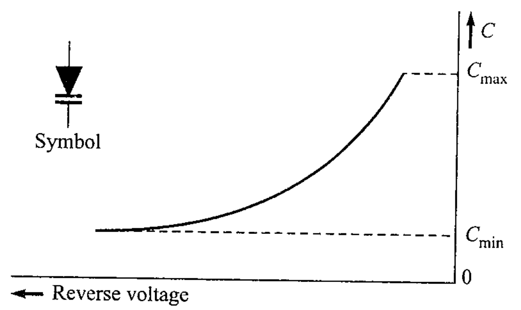


Fig. 4.30 *Varactor diode characteristic and its symbol*

Varactor diodes are replacing mechanically tuned capacitors in many applications. A varactor diode in parallel with an inductor gives a resonant tank circuit. The resonant frequency of this tank circuit can easily be changed by varying the reverse voltage across the diode.

Example 4.5 A varactor diode has a capacitance of 18 pF when the reverse bias voltage applied across it is 4 V. Determine the capacitance if the diode bias voltage is increased to 8 V.

Solution: The capacitance of a varactor diode is inversely proportional of the square root of the bias voltage, i.e.,

$$C = \frac{K}{\sqrt{V}}$$

Here, $V = 4$ V; $C = 18$ pF $= 18 \times 10^{-12}$ F.

$$\therefore \quad 18 \times 10^{-12} = \frac{K}{\sqrt{4}}$$

or

$$K = 36 \times 10^{-12}$$

Hence, when the voltage has increased to 8 V, the capacitance becomes

$$C = \frac{K}{\sqrt{V}} = \frac{36 \times 10^{-12}}{\sqrt{8}} = 12.728 \times 10^{-12} \text{ F}$$
$$= 12.728 \text{ pF}$$

4.9.5 Light-Emitting Diodes (LEDs)

When a PN-junction diode is forward-biased, the potential barrier is lowered. The majority carriers start crossing the junction. The conduction-band electrons from the N region cross the barrier and enter the P region. Immediately on entering the P region, each electron falls into a hole and recombination takes place. Also, some holes may cross the junction from the P region into the N region. A conduction-band electron in the N region may fall into a hole even before it crosses the junction. In either case, recombination takes place around the junction.

Each recombination radiates energy. In an ordinary diode (power diode or signal diode), the radiated energy is in the form of heat. In the *light-emitting diode* (LED), the radiated energy is in the form of light (or photons),

Germanium and silicon diodes have less probabilities of radiating light. By using materials such as gallium arsenide phosphide (GaAsP) and gallium phosphide (GaP), a manufacturer can produce LEDs that radiate red, green or orange lights. Infrared LEDs use gallium arsenide (GaAs), and they emit invisible (infrared) radiation. These find applications in burglar-alarm systems and other areas requiring invisible radiation.

Figure 4.31*a* shows the schematic symbol of an LED. The LEDs that emit visible light find applications in instrument displays, panel indicators, digital watches, calculators, multimeters, intercoms, telephone switch boards, etc. A *seven-segment display* unit as shown in Fig. 4.31*b* is made by using a number of LEDs. By activating suitable combination of LEDs in this unit, any digit from 0 to 9 can be displayed by it.

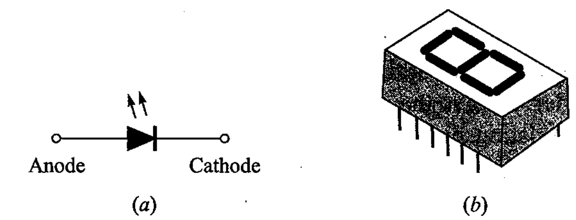


Fig. 4.31 *(a) Schematic symbol of an LED; (b) A seven-segment display using LEDs*

LEDs have a number of advantages over ordinary incandescent lamps. They work on low voltages (1 or 2 V) and currents (5 to 10 mA) and thus consume less power. They require no heating, no warm-up time, and hence are very fast in action. They are small in size and light in weight. They are not affected by mechanical vibrations and have long life (more than 20 years).

Review Questions

1. (*a*) What causes majority carriers to flow at the moment when a P region and an N region are brought together? (*b*) Why does this flow not continue until all the carriers have recombined?
2. Explain the formation of the 'depletion region' in an open circuited PN-junction.
3. State what you understand by barrier potential across a PN-junction. Also explain its significance.
4. The barrier potential developed across an open-circuited PN-junction aids the flow of minority carriers. Explain how this flow of charge carriers is counterbalanced.
5. Explain why the peak inverse voltage of a semiconductor diode is an important parameter.
6. Sketch, on the same axes, typical static characteristics for germanium and silicon diodes. Label clearly the values of forward voltage drop and reverse saturation current.
7. What do you understand by 'an ideal diode'? Draw its *V-I* characteristics.
8. What limits the number of reverse current carriers?
9. Why is the reverse current in a silicon diode much smaller than that in a comparable germanium diode ?
10. Explain how the process of avalanche breakdown occurs in a PN-junction diode. How is it different from zener breakdown?
11. Explain why a PN-junction possesses capacitance.
12. Which carriers conduct forward current in a diode? Draw the symbol for a PN-junction diode showing the direction of forward current.
13. Roughly how much forward voltage is needed to cause current to flow in (*a*) a silicon diode; (*b*) a germanium diode?
14. Draw the circuit diagram of a half-wave rectifier. Explain its working. What is the minimum frequency of ripple in its output?
15. If V_m is the peak value of the voltage across the secondary winding in a half-wave rectifier, what is the value of the dc component in its output voltage? Derive this relationship.
16. Draw the circuit of a half-wave rectifier with a capacitor-input filter. Give typical component values and describe the operation of the circuit. What is the peak inverse voltage across the diode : (*a*) without the capacitor connected, and (*b*) with the capacitor connected?

17. Draw the circuit diagram of a full-wave rectifier using (*a*) centre-tap connection, and (*b*) bridge connection. Explain the working of each. What is the PIV in each case?
18. Explain why a bridge rectifier is preferred over a centre-tap rectifier. Is there any application where a centre-tap rectifier is preferred over a bridge rectifier ?
19. Prove that the ripple factor of a half-wave rectifier is 1.21 and that of a full-wave rectifier is 0.482.
20. Show that the maximum rectification efficiency of a half-wave rectifier is 40.6 %.
21. Derive an expression for the rectification efficiency of a full-wave rectifier.
22. Draw the output voltage waveform of a half-wave rectifier and then show the effect, on this waveform, of connecting a capacitor across the load resistance.
23. Repeat Q. 4.22 for a full-wave rectifier.
24. Explain the need of using smoothing circuits in a power supply.
25. Explain why a series-inductor filter cannot be used with a half-wave rectifier.
26. Explain the working of a choke-input filter.
27. What is a bleeder resistor? What functions can it serve in a power supply?
28. Draw the circuit diagram, including typical component values, for a 12-V, 2-A power supply using bridge rectifier and a π-filter.
29. What are the important specifications of semiconductor diodes?
30. State the difference between the specifications of a signal diode and a power diode.
31. Name the two types of reverse breakdowns which can occur in a PN-junction diode. Which type occurs at lower voltages?
32. Explain why is it necessary to use a voltage regulator circuit in a power supply.
33. Draw the block diagram of a regulated power supply. Explain in brief the functioning of each block.
34. Draw the circuit diagram of a voltage regulator circuit using a zener diode. Explain its working. Is there any limitation on the value of the series resistor used in this circuit?
35. In what respect is an LED different from an ordinary PN-junction diode? State applications of LEDs. Why should you prefer LEDs over conventional incandescent lamps ?

Objective-Type Questions

Below are some incomplete statements. Four alternatives are provided for each. Choose the alternative that completes the statement correctly.

1. The potential barrier at a PN-junction is due to the charges on either side of the junction. These charges are
   (*a*) minority carriers
   (*b*) majority carriers

(*c*) both majority and minority carriers
(*d*) fixed donor and acceptor ions

2. In an unbiased PN-junction, the junction current at equilibrium is
(*a*) due to diffusion of minority carriers only
(*b*) due to diffusion of majority carriers only
(*c*) zero, because equal but opposite carriers are crossing the junction
(*d*) zero, because no charges are crossing the junction

3. In a PN-junction diode, holes diffuse from the P region to the N region because
(*a*) the free electrons in the N region attract them
(*b*) they are swept across the junction by the potential difference
(*c*) there is greater concentration of holes in the P region as compared to N region
(*d*) None of the above

4. In a PN-junction diode, if the junction current is zero, this means that
(*a*) the potential barrier has disappeared
(*b*) there are no carriers crossing the junction
(*c*) the number of majority carriers crossing the junction equals the number of minority carriers crossing the junction
(*d*) the number of holes diffusing from the P region equals the number of electrons diffusing from the N region

5. In a semiconductor diode, the barrier potential offers opposition to only
(*a*) majority carriers in both regions
(*b*) minority carriers in both regions
(*c*) free electrons in the N region
(*d*) holes in the P region

6. When we apply reverse bias to a junction diode, it
(*a*) lowers the potential barrier
(*b*) raises the potential barrier
(*c*) greatly increases the minority-carrier current
(*d*) greatly increases the majority-carrier current

7. The number of minority carriers crossing the junction of a diode depends primarily on the
(*a*) concentration of doping impurities
(*b*) magnitude of the potential barrier
(*c*) magnitude of the forward-bias voltage
(*d*) rate of thermal generation of electron-hole pairs

8. The reverse saturation current in a junction diode is the current that flows when
(*a*) only majority carriers are crossing the junction

(*b*) only minority carriers are crossing the junction
(*c*) the junction is unbiased
(*d*) the potential barrier is zero

9. When forward bias is applied to a junction diode, it
(*a*) increases the potential barrier
(*b*) decreases the potential barrier
(*c*) reduces the majority-carrier current to zero
(*d*) reduces the minority-carrier current to zero

10. The depletion or space-charge region in a junction diode contains charges that are
(*a*) mostly majority carriers
(*b*) mostly minority carriers
(*c*) mobile donor and acceptor ions
(*d*) fixed donor and acceptor ions

11. Avalanche breakdown in a semiconductor diode occurs when
(*a*) forward current exceeds a certain value
(*b*) reverse bias exceeds a certain value
(*c*) forward bias exceeds a certain value
(*d*) the potential barrier is reduced to zero

12. When a PN-junction is biased in the forward direction
(*a*) only holes in the P region are injected into the N region
(*b*) only electrons in the N region are injected into the P region
(*c*) majority carriers in each region are injected into the other region
(*d*) no carriers move

13. The forward bias applied to a PN-junction diode is increased from zero to higher values. Rapid increase in the current flow for a relatively small increase in voltage occurs
(*a*) immediately
(*b*) only after the forward bias exceeds the potential barrier
(*c*) when the flow of minority carriers is sufficient to cause an avalanche breakdown
(*d*) when the depletion area becomes larger than the space-charge area

14. The capacitance of a reverse-biased PN-junction
(*a*) increases as the reverse bias is decreased
(*b*) increases as the reverse bias is increased
(*c*) depends mainly on the reverse saturation current
(*d*) makes the PN-junction more effective at high frequencies

15. In a half-wave rectifier, the load current flows for
(*a*) the complete cycle of the input signal
(*b*) only for the positive half-cycle of the input signal
(*c*) less than half-cycle of the input signal
(*d*) more than half-cycle but less than the complete cycle of the input signal

16. In a full-wave rectifier, the current in each of the diodes flows for
(*a*) the complete cycle of the input signal
(*b*) half-cycle of the input signal
(*c*) less than half-cycle of the input signal
(*d*) zero time

17. In a half-wave rectifier, the peak value of the ac voltage across the secondary of the transformer is $20\sqrt{2}$ V. If no filter circuit is used, the maximum dc voltage across the load will be
(*a*) 28.28 V (*b*) 14.14 V
(*c*) 20 V (*d*) 9 V

18. If V_m is the peak voltage across the secondary of the transformer in a half-wave rectifier (without any filter circuit), then the maximum voltage on the reverse-biased diode is
(*a*) V_m (*b*) $\frac{1}{2}V_m$
(*c*) $2V_m$ (*d*) none of these

19. In the above question, if we use a shunt capacitor filter, the maximum voltage that occurs on the reverse-biased diode is
(*a*) V_m (*b*) $\frac{1}{2}V_m$
(*c*) $2V_m$ (*d*) none of these

20. In a centre-tap full-wave rectifier, V_m is the peak voltage between the centre-tap and one end of the secondary. The maximum voltage across the reverse-biased diode is
(*a*) V_m (*b*) $\frac{1}{2}V_m$
(*c*) $2V_m$ (*d*) none of these

21. A zener diode
(*a*) has a high forward-voltage rating
(*b*) has a sharp breakdown at low reverse voltage
(*c*) is useful as an amplifier
(*d*) has a negative resistance

22. The light-emitting diode (LED)
(*a*) is usually made from silicon
(*b*) uses a reverse-biased junction

(*c*) gives a light output which increases with increase in temperature
(*d*) depends on the recombination of holes and electrons

Answers

1. (*d*)	2. (*c*)	3. (*c*)	4. (*c*)	5. (*a*)	6. (*b*)
7. (*d*)	8. (*b*)	9. (*b*)	10. (*d*)	11. (*b*)	12. (*c*)
13. (*b*)	14. (*a*)	15. (*b*)	16. (*b*)	17 (*d*)	18. (*a*)
19. (*c*)	20. (*c*)	21. (*b*)	22. (*d*)		

• Tutorial Sheet 4.1 •

1. Calculate the maximum dc voltage available from a half-wave rectifier shown in Fig. T. 4.1.1. Also find the reading of the milliammeter. [**Ans.** 4.95 V; 4.95 mA]

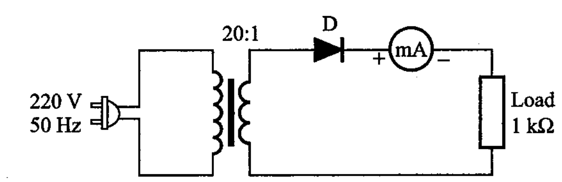


Fig. T. 4.1.1

2. Calculate the PIV rating of the diodes used in the full-wave rectifier shown in Fig. T. 4.1.2. Also find the maximum dc voltage that can be obtained from this circuit. Mark the polarity on the milliammeter and determine how much dc current it will indicate. [**Ans.** 9.9 V; 9.9 mA]

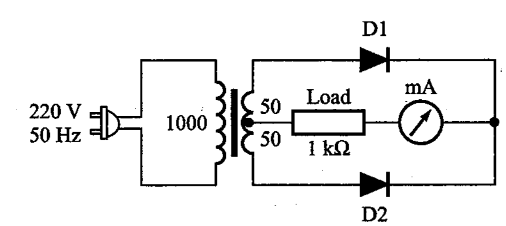


Fig. T. 4.1.2

3. It is desired to obtain a maximum of 15 V (dc) from a bridge rectifier circuit. It uses silicon diodes (the voltage drop across each diode is 0.7 V). It is energised from an ac mains supply (220 V, 50 Hz) through a step-down transformer. Find the turns ratio of this transformer. [**Ans.** 12 : 1]

4. In Fig. T. 4.1.3, calculate the load current I_L and zener diode current I_Z. Breakdown voltage of the zener diode may be assumed to be 5 V.

[**Ans.** 4.16 mA; 10.84 mA]

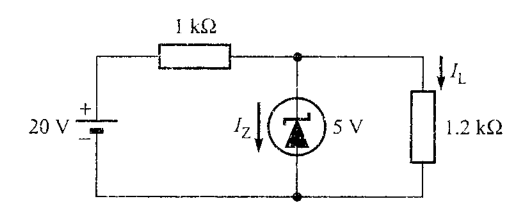


Fig. T. 4.1.3

5. The silicon diode shown in Fig. T. 4.1.4 is rated for a maximum current of 100 mA. Calculate the minimum value of the resistor R_L. Assume the forward voltage drop across the diode to be 0.7 V.

[**Ans.** 93 Ω – we can take a safer value of 100 Ω]

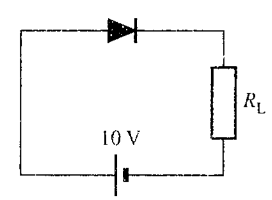


Fig. T. 4.1.4

Experimental Exercise 4.1

Title Semiconductor (or crystal) diode characteristics.

Objectives To

1. trace the circuit meant to draw the diode-characteristics;
2. measure the current through the diode for a particular value of forward voltage;
3. plot the forward characteristics of a germanium and a silicon diode;
4. compare the forward characteristic of a Ge diode with that of a Si diode;
5. calculate the forward static and dynamic resistance of the diode at a particular operating point.

Apparatus Required Experimental board, regulated power supply, milliammeter, electronic multimeter.

Circuit Diagram The circuit diagram is given in Fig. E. 4.1.1.

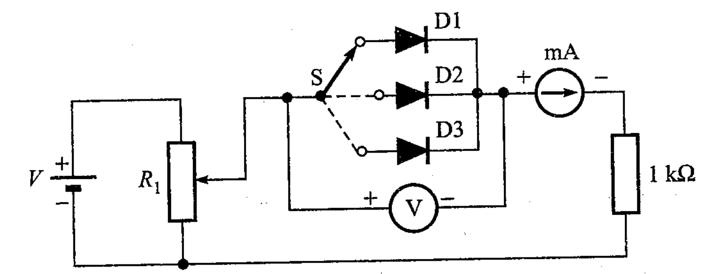


Fig. E. 4.1.1

Brief Theory A diode conducts in forward bias (i.e., when its anode is at higher potential than its cathode). It does not conduct in reverse bias. When the diode is forward biased, the barrier potential at junction reduces. The majority carriers then diffuse across the junction. This causes current to flow through the diode. In reverse bias, the barrier potential increases, and almost no current can flow through the diode.

The external battery is connected so that its positive terminal goes to the anode and its negative terminal goes to cathode. The diode is then forward-biased. The amount of forward bias can be varied by changing the externally applied voltage. As shown in Fig. E. 4.1.1, the external voltage applied across the diode can be varied by the potentiometer R_1. A series resistor (say, 1 kΩ) is connected in the circuit so that excessive current does not flow through the diode. We can note down different value of the current through the diode for various values of the voltage across it. A plot between this voltage and current give the *diode forward characteristics*.

At a given operating point we can determine the static resistance (R_d) and dynamic resistance (r_d) of the diode from its characteristic. The static resistance is defined as the ratio of the dc voltage to dc current, i.e.,

$$R_d = \frac{V}{I}$$

The dynamic resistance is the ratio of a small change in voltage to a small change in current, i.e.,

$$r_d = \frac{\Delta V}{\Delta I}$$

Procedure

1. Find the type number of the diodes connected in the experimental board.
2. Trace the circuit and identify different components used in the circuit. Read the value of the resistor using the colour code.
3. Connect the milliammeter and voltmeter of suitable ranges, say 0 to 25 mA for ammeter and 0 to 1.5 V for voltmeter.

4. Switch on the power supply. With the help of the potentiometer R_1, increase the voltage slowly.
5. Note the milliammeter and voltmeter readings for each setting of the potentiometer. Tabulate the observations.
6. Draw the graph between voltage and current.
7. At a suitable operating point, calculate the static and dynamic resistance of the diode, as illustrated in Fig. E. 4.1.2.
8. Bring the other diode in the circuit and repeat the above.

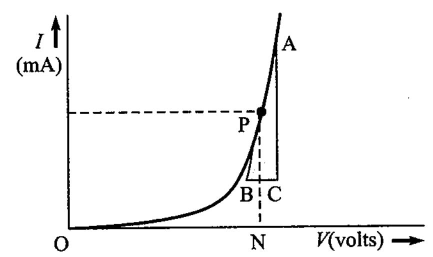


Fig. E. 4.1.2

Observations

1. *Type number of the diode* = ______
2. *Information from the data book*:
   (*a*) Maximum forward current rating = ______mA
   (*b*) Maximum peak inverse voltage rating = ______V
3. *Characteristics*:

Sr. No.	*Type No.*____		*Type No.*____	
	Voltage (*in* V)	*Current* (*in* mA)	*Voltage* (*in* V)	*Current* (*in* mA)
1. 2. 3.				

Calculations

1. Static resistance, $R_d = \frac{V}{I} = \text{\_\_\_\_\_} = \text{\_\_\_\_}\ \Omega$

2. Dynamic resistance, $r_d = \frac{\Delta V}{\Delta I} = \text{\_\_\_\_\_} = \text{\_\_\_\_\_}\ \Omega$

Results

1. The *V-I* characteristics of the diodes are shown in the graph.

2. The values of static and dynamic resistance of the two diodes are as given below:

	Diode type No. ___	*Diode type No.* ___
R_d		
r_d		

• Experimental Exercise 4.2 •

Title Zener diode characteristics.

Objectives To

1. trace the circuit;
2. plot the *V-I* characteristic of a zener diode under reverse-biased condition;
3. calculate the dynamic resistance of the diode under reverse-biased condition (when conducting).

Apparatus Required Experimental board, milliammeter, electronic multimeter, regulated power supply.

Circuit Diagram The circuit diagram is shown in Fig. E. 4.2.1.

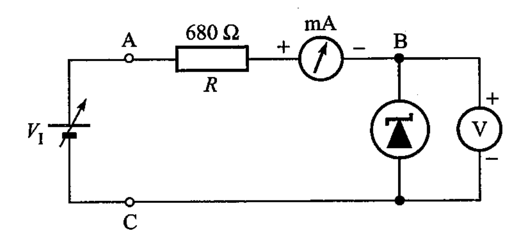


Fig. E. 4.2.1

Brief Theory A PN-junction diode normally does not conduct when reverse biased. But if the reverse bias is increased, at a particular voltage it starts conducting heavily. This voltage is called *breakdown voltage*. High current through the diode can permanently damage it. To avoid high current, we connect a resistor in series with it. Once the diode starts conducting, it maintains almost constant voltage across its terminals whatever may be the current through it. That is, it has very low dynamic resistance. A zener diode is a PN-junction diode, specially made to work in the breakdown region. It is used in voltage regulators.

Procedure

1. Note the type number of the zener diode. Find breakdown voltage, wattage and maximum current ratings of the diode from the data book.
2. Trace the circuit. Note the value of the current-limiting resistor.
3. Connect milliammeter and electronic voltmeter of suitable range. (The information obtained from the data book will help in choosing suitable range of the meters).
4. Connect the negative lead of the voltmeter to point C (Fig. E. 4.2.1). By connecting positive lead to point A, you can read the input dc voltage V_I. By connecting positive lead to point B, you get the voltage V_Z across the zener diode.
5. Switch on the power supply. Increase slowly the supply voltage in steps. Measure the voltages V_I and V_Z, and current I_Z. Once break-down occurs, V_Z remains fairly constant even though I_Z increases.
6. Plot graph between V_Z and I_Z. This is the *V-I* characteristic of the zener diode.
7. Calculate the dynamic resistance of zener diode in breakdown region, as illustrated in Fig. E. 4.2.2.

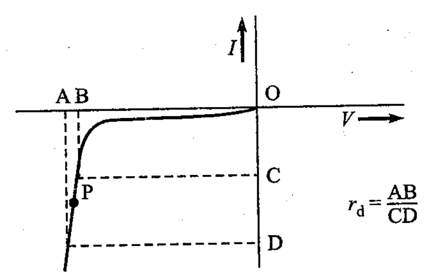


Fig. E. 4.2.2

Observations

1. *Type number of the zener diode* = ______
2. *Information from the data book*:
   (*a*) Breakdown voltage = ______V
   (*b*) Maximum current rating = ______mA
   (*c*) Maximum wattage rating = ______W
3. *V-I characteristics*:

S. No.	V_I (*in* V)	V_Z (*in* V)	I_Z (*in* mA)
1.			
2.			
3.			
4.			

Calculations Dynamic resistance, $r_d = \frac{\Delta V_Z}{\Delta I_Z} = \frac{AB}{CD} = \_\_\_\_ = \_\_\_\_\ \Omega$

Results

1. The *V-I* characteristic of the zener diode is shown in the graph.
2. The dynamic resistance of the diode, r_d = ______ Ω.

Experimental Exercise 4.3

Title Half-wave rectifier.

Objectives To

1. trace the circuit of half-wave rectifier;
2. draw the waveshape of the electrical signal at the input and output points (after observing it in CRO) of the half-wave rectifier:
3. measure the following voltages:
   (*a*) AC voltage at the input of the rectifier,
   (*b*) AC voltage at the output points,
   (*c*) DC voltage at the output points;
4. verify the formula $V_{dc} = \frac{V_m}{\pi}$
5. verify that ripple factor for a half-wave rectifier is 1.21.

Apparatus Required Half-wave rectifier circuit, a CRO, an electronic (or ordinary) multimeter.

Circuit Diagram As shown in Fig. E. 4.3.1.

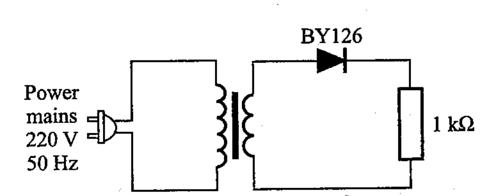


Fig. E. 4.3.1

Brief Theory A diode is a unidirectional conducting device. It conducts only when its anode is at a higher voltage with respect to its cathode. In a half-wave rectifier circuit, during positive half-cycle of the input, the diode gets forward biased and it conducts. Current flows through the load resistor R_L and voltage is developed across it. During negative half-cycle of the input, the diode gets reverse biased. Now no current (except the leakage current which is very small) flows. The voltage across the load resistance during this period of input cycle is zero. Thus a pure ac signal is converted into a unidirectional signal. It can be shown that

1. $V_{dc} = \dfrac{V_m}{\pi}$

   where, V_{dc} is the output dc voltage and V_m is peak ac voltage at the input of rectifier circuit

2. Ripple factor $= \dfrac{\text{ac voltage at the output}}{\text{dc voltage at the output}} = 1.21$

Procedure

1. Look at the given circuit of the half-wave rectifier. Trace the circuit. Note the type number of the diode. Also note the value of the load resistor used in the circuit.
2. Connect the primary side of the transformer to the ac mains. Connect the CRO probe to the output points. Adjust different knobs of the CRO so that a good and stable waveshape is visible on its screen. Plot this waveform in your record book. Take the CRO probes at the input points of the rectifier. Note the waveshape of the signal. Compare them.
3. Now, use a multimeter to measure the ac voltage at the secondary terminals of the transformer. This gives the rms value. Also measure the ac and dc voltages at the output points.
4. Multiply this rms value by $\sqrt{2}$ to get the peak value. Calculate the theoretical value of dc voltage using formula:

$$V_{dc} = \frac{V_m}{\pi}$$

   Compare this value with the practically measured value of output dc voltage.
5. Using the measured values of dc and ac output voltages, calculate ripple factor. This value should be about 1.21.

Observations

1. *Code number or type of diode* = ______
2. *Information from data book*:
   (*a*) Maximum forward dc current = _______ mA
   (*b*) Peak inverse voltage (PIV) = _______ V
3. *Waveforms from CRO*:
4. *Measurement of different voltages*:
   (*a*) AC voltage at the input = _______ V
   (*b*) DC voltage at the output = _______ V
   (*c*) AC voltage at the output = _______ V

5. *Verification of theoretical formula*:

Quantity	*Theoretical value*	*Practical value*
1. Output dc voltage		
2. Ripple factor	1.21	

Results

1. Input and output waveshapes are seen on CRO.
2. Practical value of dc voltage is little less than the theoretical value. Difference is only _______ V.
3. The practical value of ripple factor is more than its theoretical value. The difference is _______ .

• Experimental Exercise 4.4 •

Title Full-wave rectifier (centre-tap type).

Objectives To

1. trace the circuit of a full-wave rectifier;
2. draw the waveshape of the electrical signal at the input and output points (after observing it in CRO) of the full-wave rectifier;
3. measure the following voltages:
   (*a*) AC voltage at the input points,
   (*b*) AC voltage at the output points,
   (*c*) DC voltage at the output points,
4. verify the formula:

$$V_{dc} = \frac{2V_m}{\pi}$$

5. verify that the ripple factor for a full-wave rectifier is 0.482.

Apparatus Required Full-wave rectifier circuit, CRO, an electronic (or ordinary) multimeter.

Circuit Diagram As shown in Fig. E. 4.4.1.

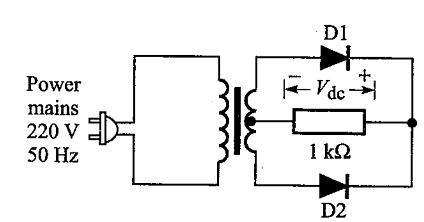


Fig. E. 4.4.1

Brief Theory In a full-wave rectifier circuit there are two diodes, a transformer and a load resistor. The transformer has a centre tap in its secondary winding. It provides out-of-phase voltages to the two diodes. During the positive half-cycle of the input, the diode D2 is reverse biased and it does not conduct. But diode D1 is forward biased and it conducts. The current flowing through D1 also passes through the load resistor, and a voltage is developed across it. During the negative half-cycle, the diode D2 is forward biased and D1 is reverse biased. Now, current flows through D2 and load resistor. The current flowing through load resistor R_L passes in the same direction in both the half-cycles. The dc voltage obtained at the output is given as

$$V_{dc} = \frac{2V_m}{\pi}$$

where V_m is the peak value of the ac voltage between the centre-tap point and one of the diodes. It can be proved that the ripple factor of a full-wave rectifier is 0.482.

Procedure

1. Trace the circuit. Note the value of the load resistor and the type number of the two diodes.
2. Connect the mains voltage to the primary of the centre-tapped transformer. Connect the output terminals to the vertical plates of the CRO. Adjust different knobs of the CRO and obtain a stationary pattern on its screen. Now touch the CRO probes at the centre tap and one of the diodes. Observe the waveshape on the CRO. Plot both the waveshapes in your record book. Compare the two voltage waveshapes.
3. Measure ac voltage at the input (centre tap and one of the diodes) and output points. Also measure the dc voltage across the load resistor.
4. From the measured ac voltage, calculate the dc voltage. Compare it with the measured value of dc output voltage. Now calculate the ripple factor by dividing ac voltage (at the output) by dc voltage at the output. How much does it differ from the theoretical value of 0.482?

Observations

1. *Type numbers of the diodes* = _____
2. *Information from data book*:
   (*a*) Maximum forward-current rating = ______ mA
   (*b*) Peak inverse voltage (PIV) rating = ____ V
3. *Waveshape at the input and output points*:
4. *Measurements of voltages*:
   (*a*) AC voltages at the input points (between centre tap and one of the diodes)
   (*b*) AC voltage at the output points = _______V
   (*c*) DC voltage at the output points = _______V

5. *Verification of the formula*:
   (*a*) Output dc voltage

Quantity	*Theoretical value*	*Practical value*
1. Output dc voltage	$\frac{2V_m}{\pi} =$	
2. Ripple factor	0.482	$\frac{V_{ac}}{V_{dc}} =$

Results

1. The waveshapes at input and output are observed on the CRO and they are plotted.
2. The output dc voltage is a little less than the theoretical value.
3. There is a little difference between the theoretical value and measured value of ripple factor.

• Experimental Exercise 4.5 •

Title Bridge rectifier circuit.

Objectives To

1. trace the given circuit of bridge rectifier;
2. draw the electrical waveshape at the input and output points after observing it on CRO;
3. measure the following voltages:
   (*a*) AC voltage at the input of rectifier,
   (*b*) AC voltage at the output points,
   (*c*) DC voltage at the output points;
4. verify the following formula

$$V_{dc} = \frac{2V_m}{\pi}$$

5. verify that the ripple factor for bridge-rectifier circuit is 0.482.

Apparatus Required Bridge-rectifier circuit, a CRO and an electronic multimeter.

Circuit Diagram As shown in Fig. E. 4.5.1.

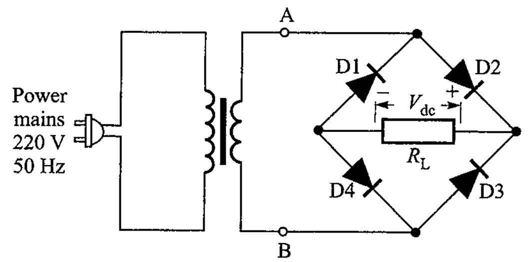


Fig. E. 4.5.1

Brief Theory In a bridge rectifier circuit there are four diodes, a transformer and a load resistor. When the input voltage is positive at point A diodes D2 and D4 conduct. The current passes through the load resistor R_L. During the other half of the input signal, the point A is negative with respect to point B. The diodes D1 and D3 conduct. The current passes through the load resistor in the same direction as during the positive half cycle. DC voltage is developed across the load. It can be proved that the output dc voltage is given by

$$V_{dc} = \frac{2V_m}{\pi}$$

where V_m is the peak ac voltage at the input of the rectifier. Also we can show that its ripple factor is 0.482.

Procedure

1. Find the type number of the diodes connected in the circuit. Trace the circuit and note down the value of the load resistor.
2. Energise the rectifier with the ac mains. Connect the output of the rectifier to the CRO. Adjust different knobs of CRO till you get a stable pattern on the screen. Similarly observe the voltage waveshape at the input of the rectifier. Compare the two waveshapes.
3. Now, measure the ac voltage at the secondary of the transformer. Also measure ac and dc voltage at the output points.
4. Using the theoretical formula

$$V_{dc} = \frac{2V_m}{\pi}$$

calculate the dc voltage at the output. Compare this value with the measured dc voltage.

5. Use the measured values of ac and dc voltage at the output points to calculate the ripple factor. Compare this value with the theoretical value, which is 0.482.

Observations

1. *Type numbers of the diodes* = ____

2. *Information from data book*:
   (*a*) Maximum forward current rating = ______ mA
   (*b*) Peak inverse voltage (PIV) rating = ______ V
3. *Waveshape at the input and output points*:
4. *Measurement of voltages*:
   (*a*) AC voltage at the input points (across the secondary winding terminals) = ____ V
   (*b*) AC voltage at the output points = ______ V
   (*c*) DC voltage at the output points = ______ V
5. *Verification of the formula*:
   (*a*) Output dc voltage

Quantity	*Theoretical value*	*Practical value*
1. Output dc voltage	$\frac{2V_m}{\pi} =$	
2. Ripple factor	0.482	$\frac{V_{ac}}{V_{dc}} =$

Results

1. The waveshapes at input and output are observed on CRO and are plotted.
2. The output dc voltage is a little less than the theoretical value. Why? We had not taken the voltage drop across the diodes into consideration while deriving the formula $V_{dc} = 2V_m/\pi$.

• Experimental Exercise 4.6 •

Title Different filter circuits.

Objectives To

1. plot the waveshape of the electrical signal at the output point with and without shunt capacitor filter in a half- and full-wave rectifier;
2. plot the waveshape of the electrical signal at the output points, with and without series inductor filter in a half- and full-wave rectifier;
3. plot the waveshape of the electrical signal at the output points, with and without π filter in a half- and full-wave rectifier;
4. measure the output dc voltage, with and without shunt capacitor filter in a half-wave and full-wave rectifier circuit;
5. measure the output dc voltage, with and without series inductor filter in a half- and full-wave rectifier circuit;

6. measure the output dc voltage with and without π filter in a half- and full-wave rectifier circuit;
7. verify that dc voltage at the output is approximately equal to the peak value of the input ac signal when shunt capacitor (and π) filter is used in a half- and full-wave rectifier.

Apparatus Required Rectifier circuit with different filters, a CRO and an electronic (or an ordinary) multimeter.

Circuit Diagram As shown in Fig. E. 4.6.1.

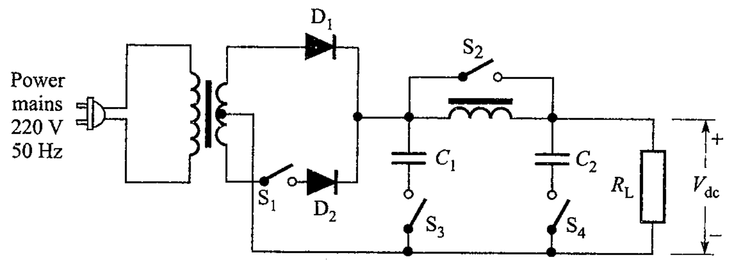


Fig. E. 4.6.1

Brief Theory The output of a half-wave or full-wave rectifier contains an appreciable amount of ac voltage in addition to dc voltage. But, what we desire is pure dc without any ac voltage in it. The ac variations can be *filtered* out or *smoothed* out from the rectified voltage. This is done by *filter circuits*.

In a shunt capacitor filter, we put a high-value capacitor in shunt with the load. The capacitor offers a low impedance path to the ac components of current. Most of the ac current passes through the shunt capacitor. All the dc current passes through the load resistor. The capacitor tries to maintain the output voltage constant at V_m. This is shown in Fig. E. 4.6.2, for half-wave rectifier.

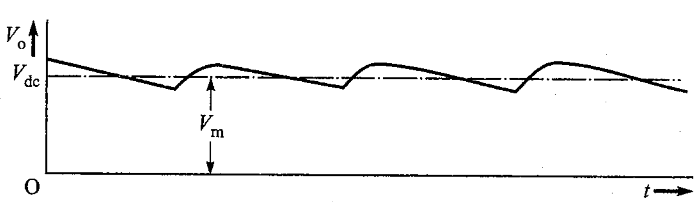


Fig. E. 4.6.2

In a series inductor filter, an inductor is used in series with the load. The inductor offers high impedance to ac variations of current and low impedance to dc. As a result, the output across the load has very low ac content. The output becomes a much better dc.

A π filter utilises the filtering properties of both the inductor and capacitor. It uses two capacitors (in shunt) and one inductor (in series). With this type of filter, the rectified output becomes almost free from ac.

Procedure

1. Trace the given rectifier circuit with different filter components. Identify every component in the circuit. Note down their values. Identify the switches S_1, S_2, S_3 and S_4.
2. With switch S_1 on, diode D_2 is in the circuit. It behaves as a full-wave rectifier. When switch S_1 is open, it becomes a half-wave rectifier. By closing switches S_3 and S_4, the capacitors C_1 and C_2 respectively can be brought into the circuit. If the switch S_2 is closed, the inductor L becomes out of circuit (the whole of the current passes through the closed switch S_2). When S_2 is open, the inductor comes in series with the load resistor R_L.
3. Keep switch S_1 open. The circuit becomes a half-wave rectifier. Open the switches S_3 and S_4, and close the switch S_2. Observe output voltage waveshape on CRO and plot it. Measure the output voltages (ac as well as dc). To obtain a shunt capacitor filter, switch on S_3. Observe and plot output-voltage waveshape. Again measure output ac and dc voltages. To have larger values of shunt capacitor, switch on S_4 also (capacitors C_1 and C_2 are in parallel). Again observe the output-voltage waveshape. Measure ac and dc voltages.
4. Switch on S_1 (to make it full-wave rectifier) and repeat the above.
5. Switch off S_1. Also switch off S_2, S_3 and S_4. It becomes a half-wave rectifier with series inductor filter. Observe and plot the output-voltage waveshape. Measure output dc and ac voltages.
6. Switch on S_1 and repeat the above.
7. Switch off S_1 and switch on S_3 and S_4 (switch S_2 is in off position). It becomes a half-wave rectifier with π filter. Observe and plot the output-voltage waveshape. Measure output voltage (ac as well as dc).
8. Switch on S_1 and repeat the above.
9. Measure the ac voltage between the centre tap and one of the end-terminals of the secondary of the transformer. From this, calculate the peak value V_m of the input voltage. Now, keeping the switch S_1 open, make a shunt capacitor filter by switching on S_3. (Switch S_4 is open and switch S_2 is closed.) Measure the output dc voltage. Compare it with V_m. Now switch on S_4. Again measure the output dc voltage. It becomes nearer to V_m.
10. Switch on S_1 and repeat the above.

Observations

1. Filters

Rectifier type	*Filter type*	V_{ac} (volts)	V_{dc} (volts)
Half-wave	1. No filter 2. Shunt capacitor filter 3. Series inductor filter 4. π filter		
Full-wave	1. No filter 2. Shunt capacitor filter 3. Series inductor filter 4. π filter		

2. Input ac voltage, V_m = _____ V (rms)

   Peak value, V_m = _____ $\times \sqrt{2}$ = ____ V

   Output dc voltage when shunt capacitor filter is used in half-wave rectifier circuit = _____ V

   Output dc voltage when shunt capacitor filter is used in full-wave rectifier circuit = ____ V

Results

1. With the use of shunt capacitor filter in half-wave and full-wave rectifier circuits, ripple voltages are very much reduced.
2. When a π filter is used, output of half-wave and full-wave rectifier is almost a pure dc.

5 UNIT
BIPOLAR JUNCTION TRANSISTORS (BJTs)

"The new electronic independence re-creates the world in the image of a global village."

Marshall McLuhan (1911–1980)
Canadian Communications Theorist, Educator, Writer and Social Reformer

After completing this unit, students will be able to:

- explain the construction of a transistor
- explain the action of transistor on the basis of current flow due to the movement of electrons and holes
- explain the flow of leakage current in a transistor in CB configuration
- draw the symbols of NPN and PNP transistors
- mark the direction of different currents in the symbols of NPN and PNP transistors
- explain the effect of temperature on leakage current
- connect the external batteries and ac input signal to a transistor in its three configurations (CB, CE and CC)
- explain the meaning of α (alpha), β (beta), I_{CBO}, I_{CEO}, input dynamic resistance and output dynamic resistance
- draw the input and output characteristics of a transistor in CB and CE configurations
- calculate transistor parameters from its characteristics.
- derive the relationship between alpha and beta of a transistor
- compare the CB and CE configurations
- explain the superiority of CE configuration over CB configuration in amplifier circuits

- write code numbers (or type numbers) of at least five commonly used Ge and Si transistors manufactured in India
- draw the circuit diagram of a basic transistor amplifier in CE configuration
- draw the dc loadline on the output characteristics, for the given amplifier circuit
- calculate the current gain, voltage gain, and the power gain for a simple amplifier circuit, by using the dc load line
- explain the phase reversal of the signal when it is amplified by CE amplifier
- explain the basic construction of alloy junction transistors and silicon planar transistors
- refer to the transistor data sheet for a given transistor
- explain the phenomenon of thermal runaway of a transistor
- explain the use of heat sinks in power transistors

5.1 INTRODUCTION

The transistor was invented in 1948 by John Bardeen, Walter Brattain and William Shockley at Bell Laboratory in America. They were awarded the Nobel Prize in recognition of their contributions to Physics. This invention completely revolutionised the electronics industry. Since then, there has been a rapidly expanding effort to utilise and develop many types of semiconductor devices such as FET, MOSFET, UJT, SCR, etc.

5.2 JUNCTION TRANSISTOR STRUCTURE

A transistor is basically a silicon or germanium crystal containing three separate regions. It can either be NPN-type or PNP-type. Figure 5.1*a* shows an NPN transistor. It has three regions. The middle region is called the *base* and the two outer regions are called the *emitter* and the *collector*. Although the two outer regions are of the same type (N-type), their functions cannot be interchanged. The two regions have different physical and electrical properties. In most transistors, the collector region is made physically larger than the emitter region since it is required to dissipate more heat. The base is very *lightly doped* and is very *thin*. The emitter is heavily doped. The doping of the collector is between the heavy doping of the emitter and the light doping of the base. The function of the emitter is to *emit* or inject electrons (holes in case of a PNP transistor) into the base. The base passes most of these electrons (holes in case of PNP) onto the collector. The collector has the job of *collecting* or gathering these electrons (holes in case of a PNP) from the base.

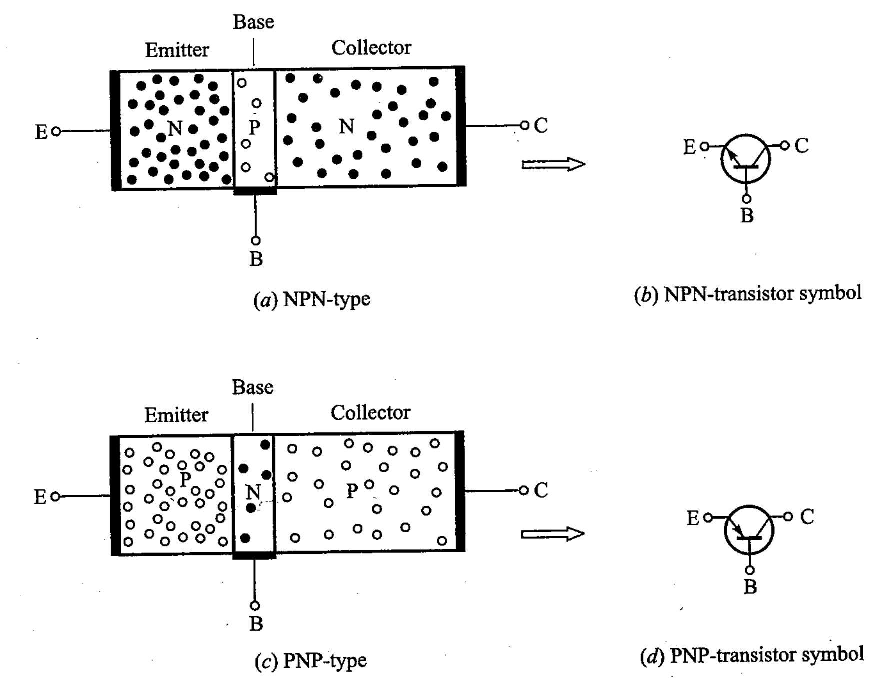


Fig. 5.1 *Junction transistor*

A transistor has two PN-junctions. One junction is between the emitter and the base and is called the emitter-base junction or simply the *emitter junction*. The other junction is between the base and the collector and is called collector-base junction or simply the *collector junction*. Thus, a transistor is like two PN-junction diodes connected back-to-back. The PN-junction theory, learnt in the last chapter, will be used to discuss the action of a transistor.

Figure 5.1*b* shows the symbol for an NPN transistor. Note that in the symbol, the emitter (not the collector) has an arrowhead. The arrowhead points in the direction of the conventional emitter current (from *P* region to *N* region).

Figure 5.1*c* shows the structure of a PNP transistor and Fig. 5.1*d* shows its symbol. Note the direction of the arrowhead in the emitter. In a PNP transistor, the conventional emitter current will flow from the emitter to the base. That is why the direction of arrowhead is inward (from *P* region to *N* region).

Both types (PNP and NPN) of transistors are widely used; sometimes together in same circuit. We shall study both the types. However, to avoid confusion, the discussion in this chapter will concentrate on the NPN type. Since a PNP transistor is the *complement* of an NPN transistor, it is merely necessary to read hole for electron, electron for hole, negative for positive, and positive for negative, for the corresponding operation of a PNP transistor. The choice of an NPN transistor is found to be more suitable because the major part of the current is transported by electrons which have higher mobility compared to holes.

5.3 THE SURPRISING ACTION OF A TRANSISTOR

A transistor has two junctions—emitter junction and a collection junction. There are four possible ways of biasing these two junctions (see Table 5.1). Of these four possible combinations, only one interests us at the moment. It is condition I, where emitter junction is forward biased and collector junction is reverse biased. This condition is often described as forward reverse (FR).

Table 5.1 *Four ways of biasing a junction transistor*

Condition		*Emitter junction*	*Collector junction*	*Region of operation*
I.	FR	Forward biased	Reverse biased	Active
II.	FF	Forward biased	Forward biased	Saturation
III.	RR	Reverse biased	Reverse biased	Cutoff
IV.	RF	Reverse biased	Forward biased	Inverted

Let us connect a junction transistor in the circuit shown in Fig. 5.2. For the sake of clarity, the base region has been shown very wide. (Remember, the base is actually made very narrow.) The battery V_{EE} acts to forward bias the emitter junction, and the battery V_{CC} acts to reverse bias the collector junction. Switches S_1 and S_2 have been provided in the emitter and collector circuits. When the two switches are open, the two junctions are unbiased. We thus have depletion or space-charge regions at the two junctions.

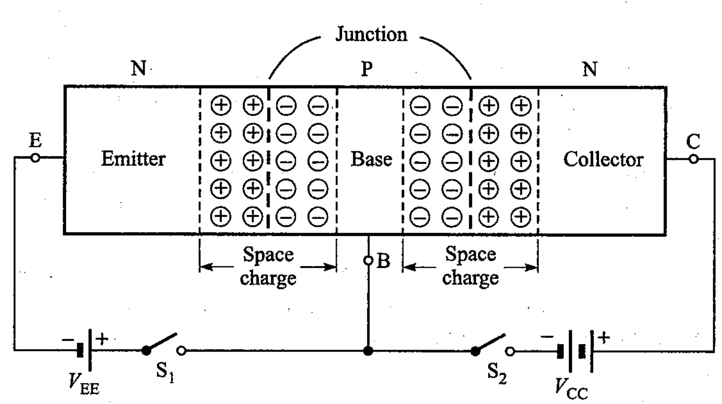


Fig. 5.2 *Biasing an NPN transistor for active operation*

If we close the switch S_1 and keep the switch S_2 open, the emitter junction will be forward biased as shown in Fig. 5.3. The barrier at the emitter junction is reduced. Since the emitter and base regions are just like those in a PN diode, we can expect a large current due to forward biasing. This current consists of majority carriers diffusing across the junction. Electrons diffuse from the emitter to the base and holes from the base to the emitter. *The total current flowing across the junction is the sum of the electron diffusion current and the hole diffusion current.* In a transistor, the base region is deliberately doped very lightly compared to the emitter region. Because of this, there are very few holes in the base region. As a result, over 99 % of the total current is carried by the electrons (diffusing from the emitter to the base). The emitter current I_E and the base current I_B in Fig. 5.3 are quite large. The two currents must be equal ($I_E = I_B$). The collector current I_C is zero.

4.68
4.50
5.1

4.61
4.31
5.30

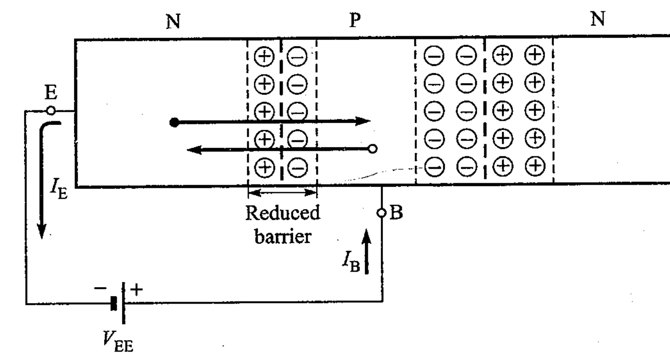


Fig. 5.3 *Only emitter junction is forward-biased—a large current flows*

Next, we close switch S_2 and keep the switch S_1 open in Fig. 5.2. This situation is shown in Fig. 5.4. The collector junction is reverse biased. Very small current flows across this reverse-biased junction. The reverse leakage current is due to the movement of minority carriers. These carriers are accelerated by the potential barrier. Just as in the PN-junction diode, this leakage current is very much temperature dependent. The current flows into the collector lead and out of the base lead. There is no emitter current ($I_E = 0$). The small collector current is called the collector leakage current. It is given a special symbol, I_{CBO}. The subscript CBO in this symbol signifies that it is a current between Collector and Base, when the third terminal (i.e., emitter) is Open.

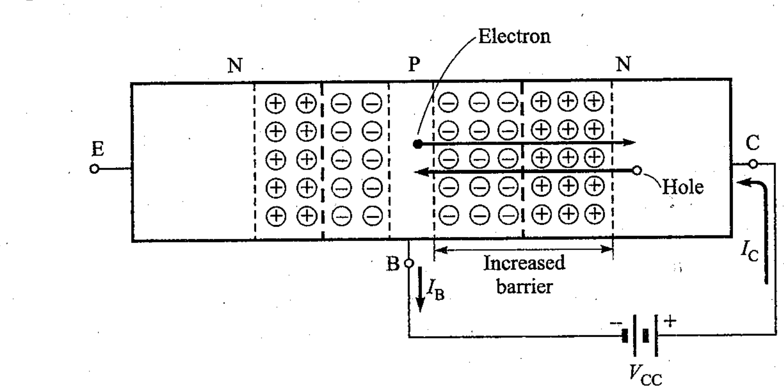


Fig. 5.4 *Only collector junction is reverse biased—a small leakage current flows*

Refer again to Fig. 5.2. What should we expect if both switches S_1 and S_2 are closed? As discussed above, we would expect both I_E and I_B to be large and I_C to be very small. However, the result of closing both switches turns out to be very surprising. The emitter current I_E is large, as expected. But I_B turns out to be a very small current, and I_C turns out to be a large current. It is entirely unexpected. It is because of this unexpected result that the transistor is such a great invention. In the next section, we shall investigate the reason for I_C being large and I_B being small.

5.4 THE WORKING OF A TRANSISTOR

Let us consider an NPN transistor biased for active operation. As shown in Fig. 5.5, the emitter-base junction is forward biased by V_{EE}, and the collector-base junction is reverse biased by V_{CC}. The directions of various currents that flow in the transistor are also indicated in Fig. 5.5. As is the usual convention, the direction of current flow has been taken opposite to the direction of electron movement. To understand the action of the transistor, we have numbered some of the electrons and holes. This will simplify the description.

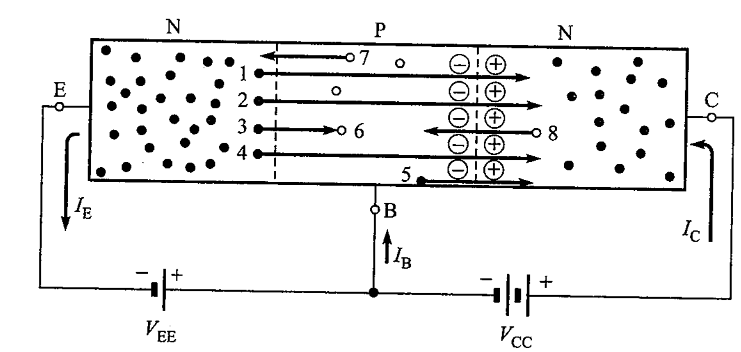


Fig. 5.5 *An NPN transistor biased for active operation*

The emitter junction is forward biased (may be, by a few tenths of a volt). The barrier potential is reduced. The space-charge region at the junction also becomes narrow. As such, majority charge carriers diffuse across the junction. The resulting current consists of electrons travelling from the emitter to the base and holes passing from the base to the emitter. As will soon be evident, only the electron current is useful in the action of the transistor. Therefore, the electron current is made much larger than the hole current. This is done by doping the base region more lightly than the emitter region. In Fig. 5.5, we have shown electrons 1, 2, 3 and 4 crossing from the emitter to the base, and hole 7 from the base to the emitter. The total sum of these charge-carrier movements constitutes the emitter current I_E. Only a portion of this current is due to the movement of electrons 1, 2, 3 and 4. These are the electrons injected by the emitter into the base. The ratio of the electron current to the total emitter current is known as *emitter injection ratio,* or the *emitter efficiency.* This ratio is denoted by symbol γ (greek letter gamma). Typically, γ is equal to 0.995. This means

that only 0.5 % of the emitter current consists of the holes passing from the base to the emitter.

Once the electrons are injected by the emitter into the base, they become minority carriers (in the base region). These electrons do not have separate identities from those which are thermally generated in the base region itself. (Note that these electrons are emitted by the emitter, and are in addition to the thermally generated minority carriers in the base region.) The central idea in transistor action is that the base is made very narrow (about 25 μm) and is very lightly doped. Because of this, most of the minority carriers (electrons) travelling from the emitter end of the base region to its collector end do not recombine with holes in this journey. Only a few electrons (like 3) may recombine with holes (like 6). The ratio of the number of electrons arriving at collector to the number of emitted electrons is known as the *base transportation factor*. It is designated by symbol* β'. Typically, $\beta' = 0.995$.

Refer to Fig. 5.5. Movement of hole 8 from the collector region and electron 5 from the base region constitutes leakage current, I_{CBO}. Movement of electron 3 and hole 7 constitute a part of emitter current I_E. These two currents are not equal. Actually, the number of electrons (like 3) and holes (like 7) crossing the emitter-base junction is much more than the number of electrons (like 5) and holes (like 8) crossing the collector-base junction. The difference of these two currents in the base region makes the base current I_B.

The collector current is less than the emitter current. There are two reasons for this. First, a part of the emitter current consists of holes that do not contribute to the collector current. Secondly, not all the electrons injected into the base are successful in reaching the collector. The first factor is represented by the emitter injection ratio γ, and the second, by the base transport factor β'. Hence the ratio of the collector current to the emitter current is equal to $\beta'\gamma$. This ratio is called dc *alpha* (α_{dc}) of the transistor. Typically, $\alpha_{dc} = 0.99$.

5.4.1 Relations between Different Currents in a Transistor

Let us now examine the role played by the batteries V_{EE} and V_{CC} in Fig. 5.5. These batteries help in maintaining the current flow in the transistor. To understand this, see Fig. 5.6.

We have seen that the emitter region emits a large number of electrons into the base region. Also, some holes diffuse from base to the emitter region. These holes recombine with electrons available in the emitter region. This way, the emitter region becomes short of electrons temporarily. This shortage is immediately made up by the battery V_{EE}. The negative terminal of this battery supplies electrons to the emitter region. After all, the battery is a storehouse of charge; they can supply as much charge as needed. To make matters simple, let us assume that 100 electrons are supplied by the negative terminal of the battery V_{EE}. (In actual practice, the electrons that flow are

* We are using the symbol β' to represent base transport factor, so as not to confuse it with the β of the transistor. The β of a transistor stands for its short-circuit current gain in CE mode.

very large in number.) These 100 electrons enter the emitter region and constitute the current I_E in the emitter terminal. The conventional current (flow of positive charge or holes) flows in a direction opposite to that of the electron flow. The current I_E is shown coming out of the emitter terminal. This is why the symbol of an NPN transistor has an arrow in the emitter lead, pointing outward (Fig. 5.1*b*).

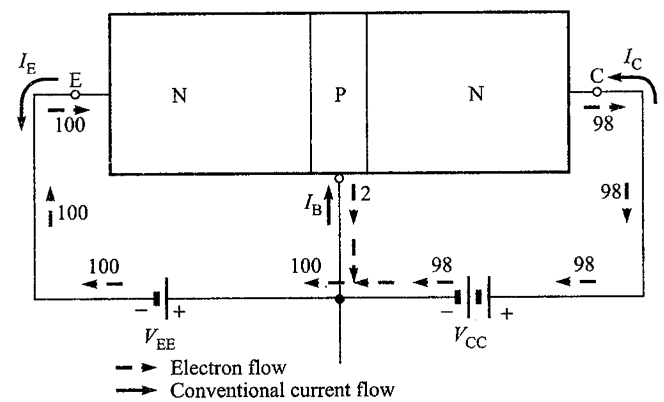


Fig. 5.6 *Relationship between different transistor currents*

What happens to the 100 electrons that enter the emitter region? The majority of these electrons (say 99 electrons) are injected into the base region. One electron is lost in the emitter region because of the recombination with a hole that has diffused from the base region. Out of the 99 electrons injected into the base, say, only one recombines with a hole; the rest of them (98 electrons) reach the collector region. This happens because of the special properties of the base region (it is lightly doped and very thin). In this manner, the base region loses only two holes (one diffuses to the emitter region and the other is lost in the base region itself, due to recombination.) The loss of these two holes is made up by creation of two fresh holes in the crystal near the base terminal. In the process of creation of holes, two electrons are generated. These two electrons flow out of the base terminal and constitute the base current. The conventional base current I_B flows into the transistor and is very small in magnitude.

The 98 electrons reaching the collector region experience an attractive force due to the battery V_{CC}. They travel out of the collector terminal and reach the positive terminal of the battery V_{CC}. The conventional collector current I_C (due to the flow of 98 electrons) flows into the transistor. The current I_C is almost equal to, but slightly less than the emitter current I_E.

The negative terminal of the battery V_{CC} gives out as many electrons as are received by its positive terminal. These 98 electrons from the negative terminal of V_{CC} and the 2 electrons from the base terminal combine together (at the junction) to make up a total of 100 electrons. These 100 electrons reach the positive terminal of the battery V_{EE}. The circuit is thus complete. The battery V_{EE} had given out 100 electrons from its negative terminal.

The total current flowing into the transistor must be equal to the total current flowing out of it. Hence, *the emitter current is equal to the sum of the collector and base currents*. That is,

$$I_E = I_C + I_B \tag{5.1}$$

This equation is a simple statement of what we have discussed up to now; the emitter current distributes itself into the collector current and base current.

There is another point. From the discussions above we can state that the collector current is made up of two parts: (*i*) The fraction of emitter current which reaches the collector; and (*ii*) The normal reverse leakage current I_{CO}. In equation form, we can write

$$I_C = \alpha_{dc} I_E + I_{CO} \tag{5.2}$$

where α_{dc} is the fraction of the emitter current I_E that reaches the collector.

5.4.2 DC Alpha (α_{dc})

We can solve Eq. (5.2) for α_{dc} and write

$$\alpha_{dc} = \frac{I_C - I_{CO}}{I_E}$$

Usually, the reverse leakage current I_{CO} is very small compared to the total collector current. Neglecting this current, the above equation can be written as

$$\alpha_{dc} = \frac{I_C}{I_E} \tag{5.3}$$

Here it is simply given as the ratio of the dc collector to dc emitter current in the transistor. If, for instance, in a transistor, we measure an I_C of 4.9 mA and an I_E of 5 mA, its dc alpha will be

$$\alpha_{dc} = \frac{4.9}{5} = 0.98$$

The thinner and more lightly doped the base is, the greater is the value of α_{dc}. But dc alpha of a transistor can never exceed unity. Many transistors have α_{dc} greater than 0.99, and almost all have α_{dc} greater than 0.95.

Example 5.1 A certain transistor has α_{dc} of 0.98 and a collector leakage current I_{CO} of 1 μA. Calculate the collector and the base currents, when I_E = 1 mA.

Solution: With I_E = 1 mA, we can use Eq, (5.2) to calculate the collector current.

$$\begin{aligned} I_C &= \alpha_{dc} I_E + I_{CO} \\ &= 0.98 \times 1 \times 10^{-3} + 1 \times 10^{-6} = 0.981 \times 10^{-3} \\ &= \mathbf{0.981\ mA} \end{aligned}$$

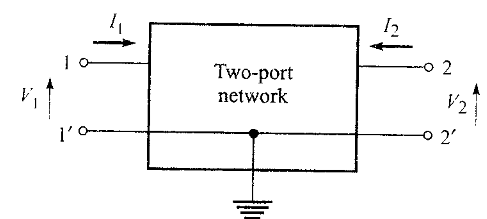
Now, using Eq. (5.1), the base current can be calculated as

$$I_B = I_E - I_C$$
$$= 1 \times 10^{-3} - 0.981 \times 10^{-3} = 0.019 \times 10^{-3} = 0.019 \text{ mA}$$
$$= \mathbf{19\ \mu A}$$

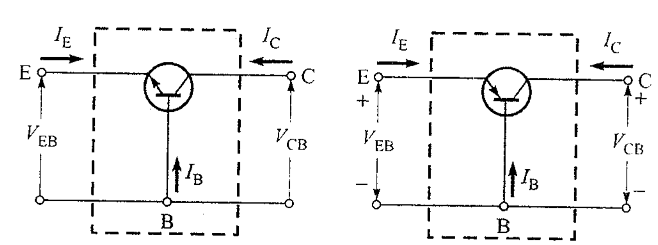
Note that I_C and I_E are almost equal and I_B is very small.

5.4.3 Sign Conventions

The sign convention for the currents and voltages in a transistor is the same as followed in a two-port network. Figure 5.7*a* shows a general two-port network. The port on the left (with terminals 1 1′) is the input port and the one on the right (with terminals 2 2′) is the output port. Usually, one terminal is made common to the input and the output, and is often grounded. Figure 5.7*a* also shows the reference directions of input and output currents as well as voltages.



(*a*) Two-port network



(*b*) NPN transistor (*c*) PNP transistor

Fig. 5.7 *Sign convention for currents and voltages*

A transistor is a three-terminal device. If one of the terminals is considered common to input and output, a transistor becomes a two-port device. Figure 5.7*b* shows an NPN transistor as a two-port network in which base is made common to the input and the output. Similarly, Fig. 5.7*c* shows a PNP transistor.

As a standard convention, *all the currents entering into the transistor are taken to be positive.* A current flowing out is negative. In other words, if the actual conventional current flows in the outward direction, a negative sign is included along with the magnitude. Hence, for an NPN transistor (see Fig. 5.7*b*), the emitter current I_E

is negative, whereas both the base current and the collector current are positive. In a PNP transistor (Fig. 5.7*c*) the emitter current is positive, but the base and collector currents are negative. In many textbooks, however, to avoid confusion, the actual direction of current flow is indicated in the diagrams.

In Figs. 5.7*b* and *c*, the transistors are connected in common-base configuration. The base is common to the input and the output. The potential (or voltages) of the emitter and collector terminals are written with reference to the common terminal (here, base). Thus, voltage V_{EB} is the voltage of emitter with respect to base. The reference direction of the voltage is indicated by a single-ended arrow (as in Fig. 5.7*b*), or by a double-ended arrow with a plus and a minus sign (as in Fig. 5.7*c*). The voltage V_{CB} represents the voltage of the collector with respect to the base. In case, the common (reference) terminal is at higher potential, the voltage is given negative sign. For an NPN transistor, biased to operate in active region (as is done in Fig. 5.6), the voltage V_{EB} is negative and voltage V_{CB} is positive (since the battery V_{EE} sets the emitter at lower potential and the battery V_{CC} sets the collector at higher potential with respect to the base).

5.4.4 Other Conditions of Operation

To understand the operation of a transistor completely, we should briefly discuss other conditions of operation given in Table 5.1. Condition II has FF bias (both the junctions are forward biased). The transistor is in *saturation*. The collector current becomes almost independent of the base current. The transistor acts like a closed switch.

Condition III has RR bias, and it represents *cut-off* operation. In this condition, both junctions are reverse biased. The emitter does not emit carriers into the base. There are no carriers to be collected by the collector. The collector current is thus zero (except a little current because of thermally generated minority carriers). The transistor acts like an open switch in this condition. We will talk more about saturation and cut-off when we discuss transistor characteristics.

Condition IV has RF bias, and it leads to *inverted* operation. This operation is quite different from the normal operation (condition I, active). Since the emitter and the collector are not doped to the same extent, they cannot be interchanged. The RF bias will result in very poor transistor action and is rarely used.

5.5 TRANSISTOR AMPLIFYING ACTION

Though a transistor can perform a number of other functions, its main use lies in amplifying electrical signals. Figure 5.8 shows a basic transistor amplifier. Here, the transistor (NPN) is connected in common-base configuration. The emitter is the input terminal and the collector is the output terminal. The transistor is biased to operate in the active region. That is, the battery V_{EE} forward biases the emitter-base junction, and the battery V_{CC} reverse biases the collector-base junction. A signal source V_s is connected in the input circuit. A load resistance R_L of 5 kΩ is connected in the output circuit. An output signal voltage V_o is developed across this resistor.

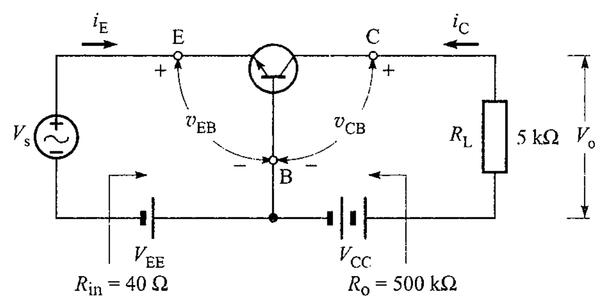


Fig. 5.8 *A basic transistor amplifier in common-base configuration*

When the signal V_s is superimposed on the dc voltage V_{EE}, the emitter voltage v_{EB} varies with time. As a result, the emitter current i_E also varies with time. Since the collector current is a function of the emitter current, a similar variation occurs in the collector current i_C. This varying current passes through the load resistor R_L and a varying voltage is developed across it. This varying voltage is the output voltage V_o.

The output signal V_o is many times greater than the input signal voltage V_s. To understand how the signal voltage is magnified (or amplified), let us consider how the transistor responds to the ac signal. Since the emitter-base junction is forward biased, it offers very low impedance to the signal source V_s. In the common-base configuration, the input resistance typically varies from 20 Ω to 100 Ω. The output junction (the collector-base junction) being reverse-biased, offers high resistance. Typically, the output resistance may vary from 100 kΩ to 1 MΩ. Assume that the input signal voltage is 20 mV (rms or effective value). Using an average value of 40 Ω for the input resistance, we get the effective value of the emitter-current variation as

$$I_e = \frac{20 \times 10^{-3}}{40} = 0.5 \text{ mA}$$

Since the collector current is almost the same as the emitter current (in fact it is slightly less), the effective value of the collector current variation is

$$I_c \cong I_e = 0.5 \text{ mA}$$

Now, the output resistance of the transistor is very high (say, 500 kΩ) and the load resistance is comparatively low (5 kΩ). The output side of the transistor acts like a constant current source; almost all the current I_e passes through the load resistance R_L.

Therefore, the effective value of output signal voltage is

$$\begin{aligned} V_o &= I_c R_L \\ &= (0.5 \times 10^{-3}) \times (5 \times 10^3) = 2.5 \text{ V} \end{aligned}$$

The ratio of the output voltage V_o to the input voltage V_s is known as the *voltage amplification* or *voltage gain* A_v of the amplifier. For the amplifier in Fig. 5.8,

$$\begin{aligned} A_v &= \frac{V_o}{V_s} \\ &= \frac{2.5}{20 \times 10^{-3}} = 125 \end{aligned}$$

The transistor's amplifying action is basically due to its capability of transferring its signal current from a *low resistance* circuit to *high resistance* circuit. Contracting the two terms *transfer* and *resistor* results in the name *transistor*; that is,

*trans*fer + resistor → *transistor*

5.5.1 Standard Notation for Symbols

When a transistor is used in a circuit, we talk of various quantities to explain its working. A standard notation of symbols to denote these quantities has been adopted by the Institution of Electrical and Electronics Engineers (IEEE). The notation is summerised as follows:

1. Instantaneous values of quantities which vary with time are represented by italic lower case (small) letters (for example, i for current, v for voltage, and p for power).
2. Italic upper case (capital) letters are used to indicate either the dc values or the effective (rms) values of ac.
3. Average (or dc) values and instantaneous total values are indicated by the roman (upright) capital subscripts of the proper electrode symbol (E for emitter, C for collector, and B for base; and S for source, D for drain, and G for gate in case of FET).
4. Time varying components (ac components) are indicated by roman lower case (small) letter subscripts of the proper electrode symbol.
5. The current reference direction is indicated by an arrow. The voltage reference polarity is indicated by plus and minus signs or by an arrow that points from the negative to the positive terminal. For example, in Fig. 5.8, instantaneous total value of the emitter-to-base voltage is written as v_{EB}, but if the base terminal is understood to be common and grounded, we may shorten the symbol v_{EB} to simply v_E. Here, the voltage v_E is the voltage of emitter (with respect to the common terminal base) and is negative.
6. The conventional current flow into an electrode from the external circuit is taken as positive.
7. The magnitude of dc supply is indicated by using roman capital double subscripts of the proper electrode symbol. For instance, in Fig. 5.8, V_{CC} represents the magnitude (the sign is taken care of separately) of the dc supply in the collector circuit.

For better understanding of the above rules, let us examine the input circuit of Fig. 5.8. Before the signal voltage V_s is connected to the input circuit, the emitter-to-base voltage is V_{EB} (or simply V_E). This voltage is the same as the dc supply voltage

V_{EE} with a negative sign. That is, $V_{EB} = V_E = -V_{EE}$. See Fig. 5.9*a*. The variation of signal voltage v_s with time is shown in Fig. 5.9*b*. When this signal voltage is connected, the total instantaneous value of the emitter voltage (with respect to the base) v_E varies with time as shown in Fig. 5.9*c*, since

$$v_E = -V_{EE} + v_s$$

At any instant t_1, the instantaneous value of the ac component (v_e) of the voltage (v_E) is also shown. In the figure, note that the voltage v_e is the same as signal voltage v_s.

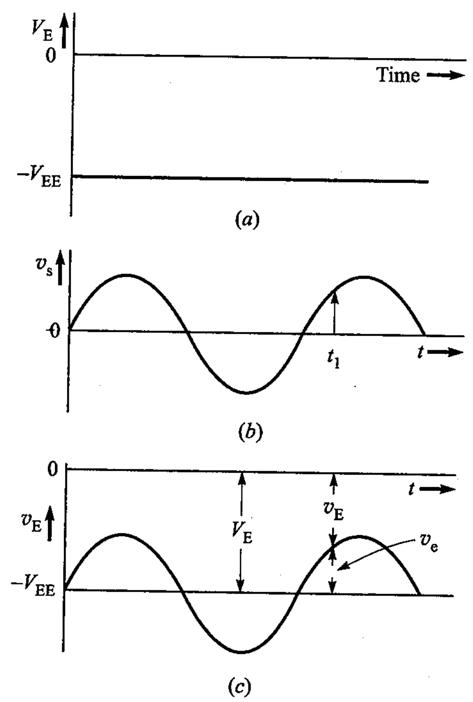


Fig. 5.9 *When signal voltage v_s is connected in the input circuit, the instantaneous value of the emitter voltage changes with time*

5.6 THREE CONFIGURATIONS

In the previous section, we have seen how a transistor amplifies ac signals when connected in common-base configuration. Is common-base (CB) the only configuration in which a transistor can work as an amplifier? No. In fact, a transistor can be used as an amplifier in any one of the three configurations. Any of its three electrodes can be made common to input and output. (This common terminal is usually grounded or connected to the chassis.) The connection is then described in terms of the common

electrode. For example, in the circuit of Fig. 5.8, the base terminal has been made common to both input and output. This connection is called common-base connection. The input signal is fed between the emitter and the base. The output signal is developed between the collector and the base. By making the emitter or the collector common, we can have what are known as common-emitter (CE) or common-collector (CC) configurations, respectively. In all the configurations, the *emitter-base junction is always forward-biased and the collector-base junction is always reverse-biased.*

Figure 5.10 shows three configurations from the ac (signal) point of view. None of the configurations shows dc biasing. But it is understood that in all the three configurations, the transistor is working in the active region (i.e., it has FR bias). In common-emitter configuration (see Fig. 5.10*b*) the base is the input terminal and the collector is the output terminal. The input signal is connected between the base and the emitter and the load resistor is connected between the collector and the emitter. The output appears across this load resistor.

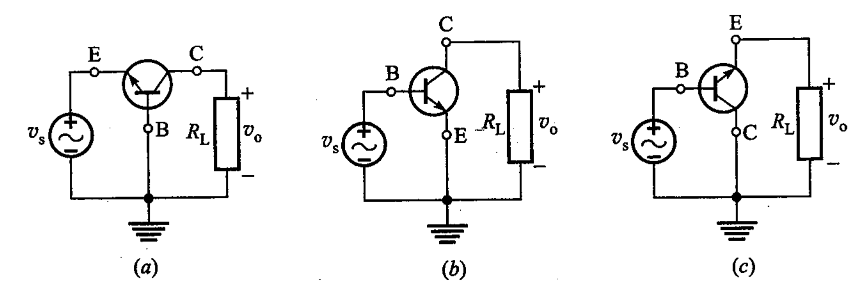


Fig. 5.10 *Three configurations in which a transistor may be connected*

Figure 5.10*c* shows common-collector (CC) configuration. Here, the input signal is connected between the base and the collector. The output appears between the emitter and the collector. This circuit is popularly known as *emitter follower.* The voltage gain of this amplifier is poor (it never exceeds unity). But it has got an important characteristics of having very high input resistance and very low output resistance. This property of the emitter follower makes it very useful in certain applications.

5.7 TRANSISTOR CHARACTERISTICS

Knowing α_{dc} (dc alpha) of a transistor does not describe its behaviour, many more details about a transistor can be studied with the help of curves that relate transistor currents and voltages. These curves are known as *static characteristic curves*. Though many sets of characteristic curves can be plotted for a given configuration, two of them are most important. In fact, these two sets of characteristics completely describle the static operation of the transistor. One is the *input characteristic* and the other is the *output characteristic*. Each curve of the input characteristic relates the input current with the input voltage, for a given output voltage. The output characteristic curve relates the output current with the output voltage, for a given input current.

Although, it is possible to draw the CC characteristics of a transistor, usually they are not needed. The common-collector configuration can be treated as a special case of common-emitter configuration (with feedback applied). The details (or the design) of a CC amplifier can be known from the CE characteristics. The static characteristics of a transistor in CC configuration are therefore not required. This is why the CC characteristics are not discussed in this book.

5.7.1 Common-Base (CB) Configuration

The circuit arrangement for determining CB characteristics of a transistor (here, we have taken PNP type) is shown in Fig. 5.11. The emitter-to-base voltage v_{EB} can be varied with the help of a ponentiometer R_1. Since the voltage v_{EB} is quite low (less than one volt) we include a series resistor R_S (say, 1 kΩ) in the emitter circuit. This helps in limiting the emitter current i_E to a low value; without this resistor, the current i_E may change by large amount even if the potentiometer (R_1) setting is moved slightly.

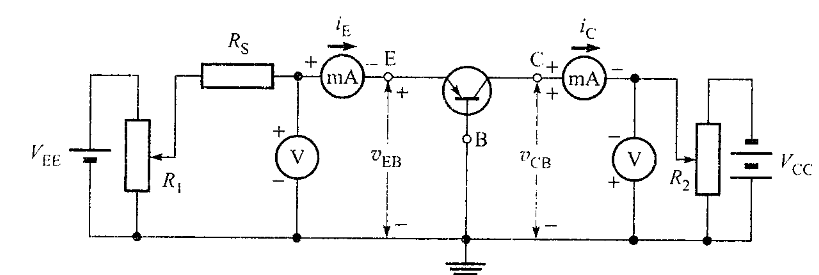


Fig. 5.11 *Circuit arrangement for determining the static characteristics of a PNP transistor in CB configuration*

The collector voltage can be varied by adjusting the potentiometer R_2. The required currents and voltages for a particular setting of the potentiometers can be read from the milliammeters and voltmeters connected in the circuit.

Input CB characteristics The common-base input characteristics are plotted between emitter current i_E and the emitter-base voltage v_{EB}, for different values of collector–base voltage V_{CB}. Figure 5.12 shows a typical input characteristics for a PNP transistor in common-base configuration.

For a given value of V_{CB}, the curve is just like the diode characteristic in forward-bias region. Here, the emitter-base is the PN-junction diode which is forward biased. This junction becomes a better diode as V_{CB} increases. That is, there will be a greater i_E for a given v_{EB} as V_{CB} increases, although the effect is very small.

For a diode, we had seen that its dynamic resistance is calculated from the slope of its forward characteristic curve. In a similar way, from the slope of the input characteristic we can get the *dynamic input resistance* of the transistor:

$$r_i = \left.\frac{\Delta v_{EB}}{\Delta i_E}\right|_{V_{CB}=\text{const.}} \tag{5.4}$$

The dynamic input resistance r_i is very low (20 to 100 Ω). Since the curve is not linear, the value of r_i varies with the point of measurement. As the emitter-base voltage increases, the curve tends to become more vertical. As a result, r_i decreases.

The input characteristics of an NPN transistor are similar to those in Fig. 5.12, differing only in that both i_E and v_{EB} would be negative and V_{CB} would be positive.

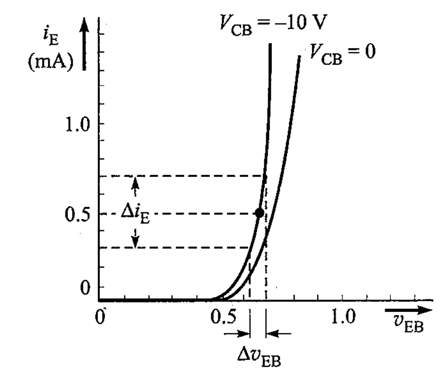


Fig. 5.12 *Common-base input characteristics for a typical PNP silicon transistor*

Example 5.2 The input characteristics of a PNP transistor in common-base configuration are given in Fig. 5.12. Determine the dynamic input resistance of the transistor at a point where $I_E = 0.5$ mA and $V_{CB} = -10$ V.

Solution: Around $I_E = 0.5$ mA, we take a small change Δi_E. Let

$$\Delta i_E = 0.7 - 0.3 = 0.4 \text{ mA}$$

From the curve for $V_{CB} = -10$ V (see Fig. 5.12), the corresponding change Δv_{EB} in emitter-base voltage is

$$\Delta v_{EB} = 0.70 - 0.62 = 0.08 \text{ V}$$

The dynamic input resistance is

$$r_i = \left.\frac{\Delta v_{EB}}{\Delta i_E}\right|_{(V_{CB}=-10\text{ V})} = \frac{0.08}{0.4 \times 10^{-3}}$$
$$= \mathbf{200\ \Omega}$$

This value of the input resistance is somewhat higher than what is expected. When the transistor is operated as an amplifier, the emitter current may be a few milliamperes. For higher values of emitter currents, the input characteristics curve becomes steeper. The input resistance r_i decreases to a very low value (say, 20 Ω).

Output CB characteristics For the same PNP transistor in CB configuration, a set of output characteristics are shown in Fig. 5.13. The output characteristic curve indicates the way in which the collector current i_C varies with change in collector-base voltage v_{CB}, with the emitter current I_E kept constant. As per standard convention, a current entering into a transistor is positive. For a PNP transistor, current i_C is flowing out of the transistor and is negative. Since the collector junction is reverse biased, the voltage v_{CB} is negative. The emitter is entering into the transistor and is taken as positive.

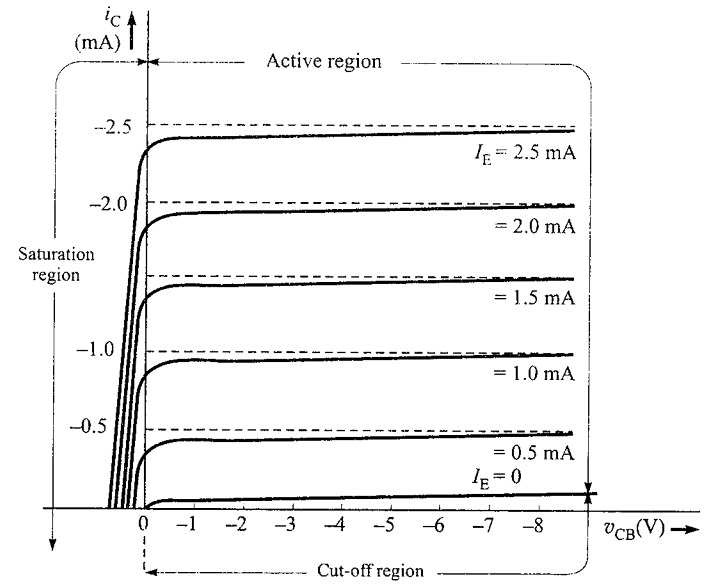


Fig. 5.13 *Common-base output characteristics for a PNP transistor*

A close look at the output characteristics of Fig. 5.13 reveals the following interesting points:

1. The collector current I_C is approximately equal to the emitter current I_E. This is true only in the *active* region, where collector-base junction is reverse biased.
2. In the active region, the curves are almost flat. This indicates that i_C (for a given I_E) increases only slightly as v_{CB} increases. Is it not what happens in a constant current source? The transistor characteristic (in CB configuration) is similar to that of a current source. It means that the transistor should have high output resistor (r_o).
3. As v_{CB} becomes positive (the collector-base junction becomes forward biased), the collector current i_C (for a given I_E) sharply decreases. This is the *saturation* region. In this region, the collector current does not depend much upon the emitter current.
4. The collector current is not zero when $I_E = 0$. It has a very small value. This is the reverse leakage current I_{CO}. The conditions that exist when $I_E = 0$ for CB configuration is shown in Fig. 5.14. The notation most frequently used for I_{CO} is I_{CBO}, as indicated in the figure. In this notation, the subscript CBO means that it is the current between the collector and base when the third terminal (the emitter) is open. Mind you, the current I_{CBO} is like the reverse saturation current for a diode. This too is temperature sensitive. At room temperature, typical values of I_{CBO} ranges from 2 μA to 5 μA for germanium transistors and 0.1 μA to 1 μA for silicon transistors.

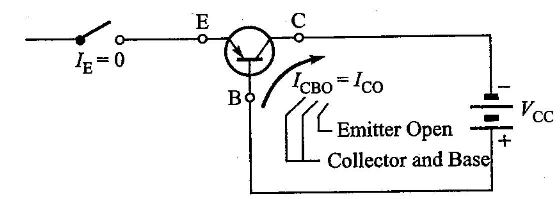


Fig. 5.14 *Reverse leakage current in CB configuration*

From the output characteristics of Fig. 5.13, we can determine a number of important transistor parameters, such as dynamic output resistance (r_o), dc current gain (α_{dc}), and ac current gain (α). The dynamic output resistance is defined as

$$r_o = \left.\frac{\Delta v_{CB}}{\Delta i_C}\right|_{I_E = \text{const.}} \tag{5.5}$$

where Δv_{CB} and Δi_C are small changes in collector voltage and collector current around a given point on the characteristic curve (for given I_E). Since the output curves are very flat, for a given Δv_{CB}, the Δi_C is very small. It means the output resistance is very high (of the order of 1 MΩ).

The dc alpha of the transistor is defined as

$$\alpha_{dc} = \frac{I_C}{I_E} \tag{5.6}$$

The characteristics tell us that at any point (in the active region) on the curve, I_C is less than I_E and the difference is very small. The value of α_{dc} is less than, but very close to unity. A typical value is 0.98.

A transistor is used as an amplifier. The amplifier handles ac (varying) signals. Under such a condition, we are more interested in the small changes in the voltages and currents rather than their absolute (dc) values. Specifically, we would like to know what change occurs in collector current for a given change in emitter current. This information is given by a parameter called ac current gain or ac alpha (α or h_{fb})*. It is defined as

$$h_{fb} \text{ or } \alpha = \left.\frac{\Delta i_C}{\Delta i_E}\right|_{V_{CB} = \text{const.}} \tag{5.7}$$

* The symbol h_{fb} comes from the analysis based on *h*-parameters or *hybrid* parameters. The letter *f* in the subscript stands for *forward*. And letter *b* indicates a *common-base* connection. For details about *h*-parameters, the reader is advised to see Unit 8 on 'Small-Signal Amplifiers' (Sec. 8.4.2).

In the above definition of h_{fb} (or α), we have stated that V_{CB} = constant (see Fig. 5.8). When you apply an ac signal to the input, the current will change, and so will the collector voltage. The only way to keep V_{CB} constant (even when the ac signal is applied to the input) is to short circuit the load resistor R_L. Therefore, the current gain h_{fb} should have been more appropriately called *short-circuit current gain*. However, very often, h_{fb} is simply referred to as *current gain*, with the understanding that it is defined under short-circuit condition. The value of h_{fb} is in the range from 0.95 to 0.995.

Summarising the common-base configuration, we can say that the current gain h_{fb} (or α) is less than unity (typical value is 0.98), dynamic input resistance r_i is very low (typical value is 20 Ω), and dynamic output resistance is very high (typical value is 1 MΩ). The leakage current I_{CBO} is quite low (typically, 4 μA for germanium and 20 nA for silicon transistors). This current is temperature dependent.

Example 5.3 In a certain transistor, a change in emitter current of 1 mA produces a change in collector current of 0.99 mA. Determine the short-circuit current gain of the transistor.

Solution: The short-circuit current gain of the transistor is given as

$$\alpha \text{ or } h_{fb} = \frac{\Delta i_C}{\Delta i_E}$$

$$= \frac{0.99 \times 10^{-3}}{1 \times 10^{-3}} = \mathbf{0.99}$$

5.7.2 Common-Emitter (CE) Configuration

In CE configuration, the emitter is made common to the input and the output. The signal is applied between the base and emitter and the output is developed between the collector and emitter. Whether the transistor works in CB or CE configuration, it is to be ensured that it works in the active region. It means that the emitter-base junction is forward biased and the collector-base junction is reverse biased. Such biasing (FR biasing) is achieved in CE configuration by connecting the batteries V_{BB} and V_{CC} as shown in Fig. 5.15*a*. Here an NPN transistor is used. The emitter-base junction is forward biased by the battery V_{BB}. This forward biasing needs a very small voltage (say, 0.6 V). The battery V_{CC} (say, 9 V) is connected between emitter and collector. Since the base is at $+V_{BB}$ potential with respect to the emitter, and the collector is at $+V_{CC}$ potential with respect to the emitter, the net potential of the collector with respect to the base is $V_{CC} - V_{BB}$. The collector-base junction is reverse biased by this potential. Since V_{CC} is much larger than V_{BB}, the reserse-bias voltage may be taken as merely V_{CC}.

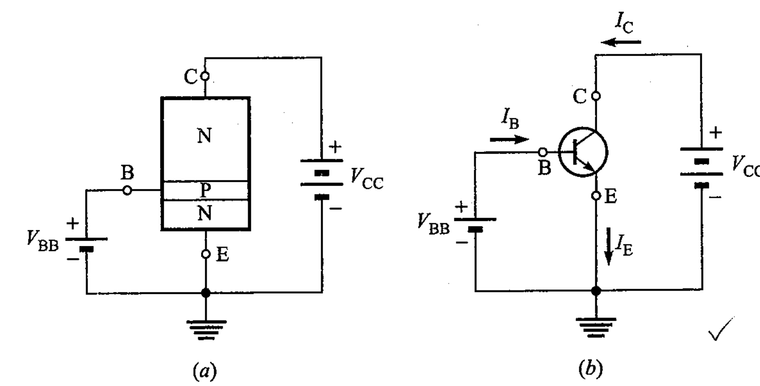


Fig. 5.15 *FR biasing of an NPN transistor in common emitter (CE) configuration*

In Fig. 5.15*b*, the transistor is replaced by its symbol. The directions of actual currents are also marked in the figure.

Current relations in CE configuration We have seen that in CB configuration I_E is the input current and I_C is the output current. These currents are related through Eqs. (5.1) and (5.2) (rewritten below for convenience):

$$I_E = I_C + I_B \tag{5.1}$$

$$I_C = \alpha_{dc} I_E + I_{CBO} \tag{5.2}$$

In CE configuration, I_B becomes the input current and the I_C is the output current. We are interested in knowing how the output current I_C is related with the input current I_B. That is, we should find a relation such as

$$I_C = f(I_B) \tag{5.8}$$

To obtain this relation, we simply substitute the expression of I_E from Eq. (5.1) into Eq. (5.2), so that

$$I_C = \alpha_{dc}(I_C + I_B) + I_{CBO}$$

or $$(1 - \alpha_{dc})I_C = \alpha_{dc} I_B + I_{CBO}$$

or $$I_C = \frac{\alpha_{dc}}{1-\alpha_{dc}} I_B + \frac{1}{1-\alpha_{dc}} I_{CBO} \tag{5.9}$$

In this equation, I_C is given in terms of I_B. The equation can be simplified somewhat by defining

$$\beta_{dc} = \frac{\alpha_{dc}}{1-\alpha_{dc}} \tag{5.10}$$

and $$I_{CEO} = \frac{I_{CBO}}{1-\alpha_{dc}} \tag{5.11}$$

Thus, Eq. (5.9) becomes

$$I_C = \beta_{dc} I_B + I_{CEO} \tag{5.12}$$

This equation states that I_C is equal to β_{dc} multiplied by the input current I_B, plus a leakage current I_{CEO}. This leakage current is the current which would flow between the collector and the emitter, if the third terminal (base) were open. This is illustrated in Fig. 5.16. The magnitude of I_{CEO} is much larger than that of I_{CBO} as indicated by Eq. (5.11). For example, if $\alpha_{dc} = 0.98$, the value of I_{CEO} is fifty times that of I_{CBO}. For silicon transistors, I_{CEO} would typically be a few microamperes, but it may be a few hundred microamperes for germanium transistors.

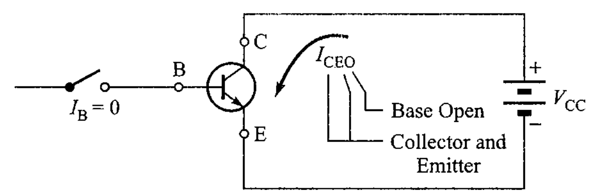


Fig. 5.16 *Reverse leakage current in CE configuration*

The factor β_{dc} is called the common-emitter dc current gain. It relates the dc output current I_C to the input current I_B. Equation (5.10) indicates that β_{dc} can be very large. For example, if $\alpha_{dc} = 0.98$, the value of β_{dc} is

$$\beta_{dc} = \frac{0.98}{1-0.98} = 49$$

Typically β_{dc} can have values in the range from 20 to 300.

If we solve Eq. (5.12) for β_{dc}, we obtain

$$\beta_{dc} = \frac{I_C - I_{CEO}}{I_B} \tag{5.13}$$

If I_{CEO} is very small compared to I_C (as is the case usually) then

$$\beta_{dc} = \frac{I_C}{I_B} \tag{5.14}$$

Thus, β_{dc} is the ratio of dc collector current to dc base current.

How beta of a transistor is related to its alpha The dc current gain of a transistor when connected in CE configuration, is β_{dc}. It is defined by Eq. (5.14). The same transistor connected in CB configuration gives a dc current gain of α_{dc}. Therefore, there is nothing surprising if beta (β_{dc}) of a transistor is related to its alpha (α_{dc}). This relation is given by Eq. (5.10). If the value of α_{dc} of a transistor is known, its β_{dc} can be calculated. Manipulating Eq. (5.10), we get

$$\beta_{dc} = \frac{\alpha_{dc}}{1-\alpha_{dc}}$$

or

$$\beta_{dc} - \beta_{dc}\alpha_{dc} = \alpha_{dc}$$

or

$$\beta_{dc} = \alpha_{dc}(1+\beta_{dc})$$

or

$$\alpha_{dc} = \frac{\beta_{dc}}{\beta_{dc}+1} \tag{5.15}$$

Thus, knowing the value of β_{dc}, we can calculate α_{dc} using the above equation.

When a transistor is used as an amplifier, we are more interested in knowing the ratio of *small changes* in the collector and base currents, rather than the ratio of their absolute values. This ratio is called *ac* or *dynamic beta* (β_{ac} or simply β). Thus, the ac beta is

$$\beta = \left.\frac{\Delta i_C}{\Delta i_B}\right|_{V_{CE}=\text{const.}} \tag{5.16}$$

To a very close approximation, the value β_{dc} is same as the ac beta (β). Like β_{dc}, the typical values of β vary from 20 to 300.

Just as β_{dc} is related to α_{dc}, so is β related to α. We can establish this relation by considering that

$$I_E = I_C + I_B$$

If we let I_C and I_B change by small amounts Δi_C and Δi_B, so that I_E changes by Δi_E, we would still have

$$\Delta i_E = \Delta i_C + \Delta i_B$$

Dividing the above equation by Δi_C and rearranging the terms, we get

$$\frac{\Delta i_E}{\Delta i_C} = 1 + \frac{\Delta i_B}{\Delta i_C}$$

or

$$\frac{1}{\alpha} = 1 + \frac{1}{\beta}$$

or

$$\beta = \frac{\alpha}{1-\alpha} \tag{5.17}$$

Example 5.4 When the emitter current of a transistor is changed by 1 mA, its collector current changes by 0.995 mA. Calculate (*a*) its common-base short circuit current gain α and (*b*) its common-emitter short circuit current gain β.

Solution:

(*a*) Common-base short circuit current gain is given by

$$\alpha = \frac{\Delta i_C}{\Delta i_E} = \frac{0.995 \times 10^{-3}}{1 \times 10^{-3}} = \mathbf{0.995}$$

(*b*) Common-emitter short circuit current gain is

$$\beta = \frac{\alpha}{1-\alpha} = \frac{0.995}{1-0.995} = \mathbf{199}$$

Example 5.5 The dc current gain of a transistor in common-emitter configuration is 100. Find its dc current gain in common-base configuration.

Solution: We can use Eq. (5.15) to calculate the dc current gain in common-base configuration

$$\alpha_{dc} = \frac{\beta_{dc}}{\beta_{dc}+1} = \frac{100}{100+1} = \mathbf{0.99}$$

Input CE characteristics In CE configuration, i_B and v_{BE} are the input variables. The output variables are i_C and v_{CE}. We can use the circuit arrangement of Fig. 5.17 to determine the input characteristics of a PNP transistor (for an NPN transistor, terminals of all the batteries, milliammeters and voltmeters will have to be reversed). Typical input characteristics are shown in Fig. 5.18. They relate i_B to v_{BE} for different values of V_{CE}. These curves are similar to those obtained for CB configuration (Fig. 5.12). Note that the change in output voltage V_{CE} does not result in a large deviation of the curves. In fact, for the commonly used dc voltages, the effect of changing V_{CE} on input characteristics may be ignored.

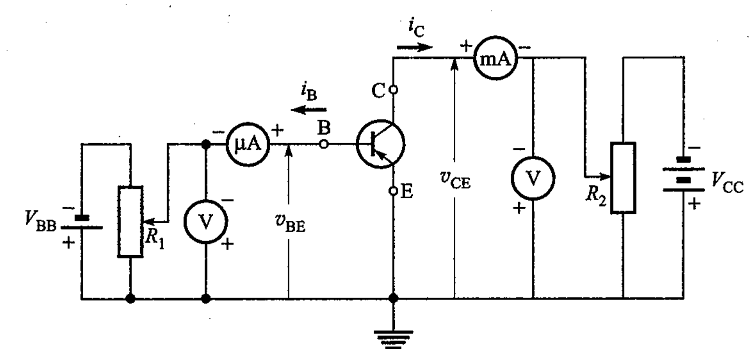


Fig. 5.17 *Circuit arrangements for determining the static characteristics of PNP transistor in CE configuration*

We can find the *dynamic input resistance* of the transistor at a given voltage V_{BE}, from Fig. 5.18. It is given by the reciprocal of the slope of the curve at the point. That is,

$$r_i = \left.\frac{\Delta v_{BE}}{\Delta i_B}\right|_{V_{CE}=\text{const.}} \tag{5.18}$$

For example, the input resistance of the transistor at the point

$$V_{BE} = -0.75 \text{ V, and } V_{CE} = -2 \text{ V}$$

is calculated from Fig. 5.18, as follows:

$$r_i = \left.\frac{\Delta v_{BE}}{\Delta i_B}\right|_{V_{CE}=-2\text{ V}} = \frac{0.78-0.72}{(68-48)\times 10^{-6}} = \frac{0.06}{20\times 10^{-6}} = 3 \text{ k}\Omega$$

The value of r_i is typically 1 kΩ, but can range from 800 Ω to 3 kΩ.

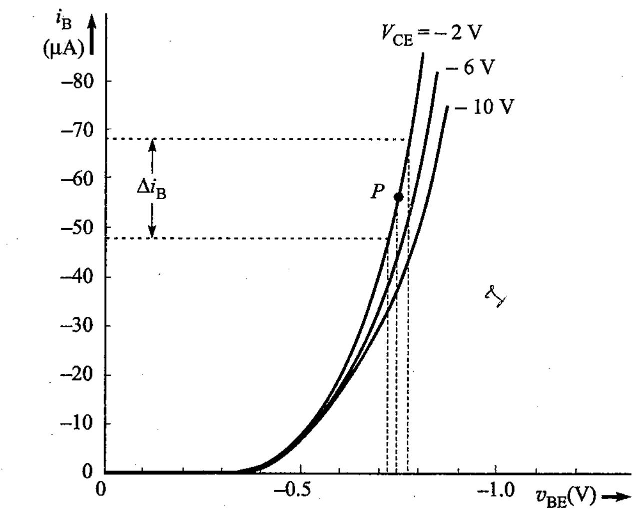


Fig. 5.18 *Common-emitter input characteristics of a PNP transistor*

Output CE characteristics From the circuit arrangement of Fig. 5.17, we can also determine the output characteristics of a PNP transistor. Figure 5.19 shows typical output characteristics of a PNP transistor. They relate the output current i_C, to the voltage between collector and emitter, v_{CE}, for various values of input current, I_B. Note that the quantities v_{CE}, i_C and I_B are all negative for a PNP transistor. If the transistor is NPN type, we reverse the terminals of the batteries V_{CC} and V_{BB}, so that v_{CE}, i_C and I_B become positive.

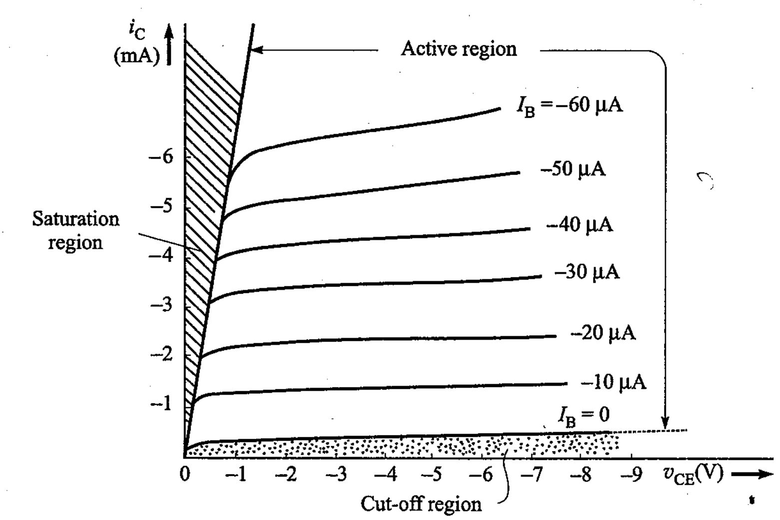


Fig. 5.19 *Common-emitter output characteristics of a PNP transistor*

A study of these output characteristics reveals following interesting points:

1. In the active region, i_C increases slowly as v_{CE} increases. The slope of these curves is somewhat greater than the CB output characteristics (see Fig. 5.13). We know that β_{dc} is equal to the ratio I_C/I_B. For each curve of Fig. 5.19, the input current I_B is constant, but current i_C increases with v_{CE}. This indicates that β_{dc} increases with v_{CE}.
2. When v_{CE} falls below a few tenths of a volt, i_C decreases rapidly as v_{CE} decreases. This occurs as v_{CE} drops below the value of V_{BE}; the collector-base junction then becomes forward biased. In this condition, both junctions of the transistor are forward biased. The transistor is working in the *saturation region*. It is called saturation region, because the current I_C no longer depends upon the input current I_B.
3. In the active region, the collector current is β_{dc} times greater than the base current. Thus, small input current I_B produces a large output current I_C.
4. The collector current is not zero when I_B is zero. It has a value of I_{CEO}, the reverse leakage current. The current I_{CEO} is related with I_{CBO} by Eq. (5.11). Using Eq. (5.10), we can write Eq. (5.11) in another form:

$$I_{CEO} = \frac{1}{1-\alpha_{dc}} I_{CBO}$$

$$= \frac{1}{1-\dfrac{\beta_{dc}}{1+\beta_{dc}}} I_{CBO}$$

or

$$I_{CEO} = (1+\beta_{dc}) I_{CBO} \tag{5.19}$$

For a germanium transistor, I_{CEO} may typically have a value of 500 μA. For silicon transistor, it is only about 20 μA.

From the output characteristics of Fig. 5.19, we can determine the dynamic output resistance r_o, the dc current gain β_{dc}, and the ac current gain β as follows:

$$r_o = \left.\frac{\Delta v_{CE}}{\Delta i_C}\right|_{I_B=\text{const.}} \tag{5.20}$$

$$\beta_{dc} = \left.\frac{I_C}{I_B}\right|_{V_{CE}=\text{const.}} \tag{5.21}$$

$$\beta = \left.\frac{\Delta i_C}{\Delta i_B}\right|_{V_{CE}=\text{const.}} \tag{5.22}$$

Example 5.6 Figure 5.20 gives the output characteristics of an NPN transistor in CE configuration. Determine, for this transistor, the dynamic output resistance, the dc current gain and the ac current gain, at an operating point V_{CE} = 10 V, when I_B = 30 μA.

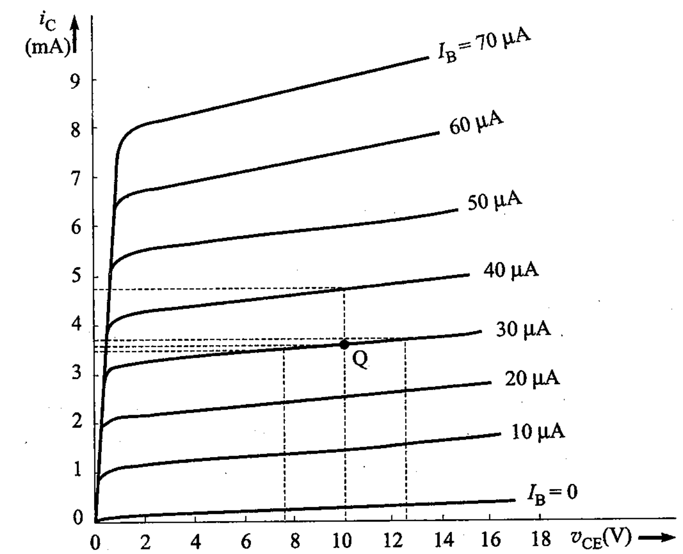


Fig. 5.20 *Determination of dynamic output resistance, dc beta, and ac beta, of an NPN transistor in CE mode*

Solution: Let us first mark the given operating point on the given characteristics. We draw a vertical line at V_{CE} = 10 V. The point of intersection of this line with the characteristic curve for I_B = 30 μA, gives the operating point Q. The collector current at this point is 3.6 mA.

To determine the dynamic output resistance of the transistor, we take a small change of collector voltage around the operating point. Let the voltage v_{CE} change from 7.5 V to 12.5 V. For a constant base current of 30 μA, the corresponding change in collector current may be seen to be from 3.5 mA to 3.7 mA. Therefore, the dynamic output resistance is given as

$$r_o = \left.\frac{\Delta v_{CE}}{\Delta i_C}\right|_{I_B=30\,\mu A} = \frac{12.5-7.5}{(3.7-3.5)\times 10^{-3}} = \frac{5}{0.2\times 10^{-3}}$$

$$= \mathbf{25\ k\Omega}$$

To find β_{dc} we should know the value of dc collector current corresponding to I_B = 30 μA. From the characteristics it can be seen that I_C = 3.6 mA at this point. Therefore,

$$\beta_{dc} = \frac{I_C}{I_B} = \frac{3.6\text{ mA}}{30\ \mu A} = \mathbf{120}$$

In order to calculate ac current gain (β), a vertical line corresponding to V_{CE} = 10 V is drawn. From the given characteristics it is clear that when base current changes from 30 μA to 40 μA, the collector current changes from 3.6 mA to 4.7 mA. Therefore, the ac current gain is given as

$$\beta = \left.\frac{\Delta i_C}{\Delta i_B}\right|_{V_{CE}=10\,\text{V}} = \frac{4.7\text{ mA} - 3.6\text{ mA}}{40\,\mu\text{A} - 30\,\mu\text{A}} = \frac{1.1\times10^{-3}}{10\times10^{-6}}$$
$$= \mathbf{110}$$

5.7.3 Common-Collector (CC) Configuration

In CC configuration, we make the collector common to the input and the output. This is shown in Fig. 5.21*a*. The same circuit can be drawn in a different way (Fig. 5.21*b*). Here the transistor is shown in the conventional manner (the collector terminal at the upper end and the emitter terminal at the lower end). Now, do you see some similarity between this circuit and that of CE configuration (Fig. 5.10*b*). The two circuits look alike, except for the fact that in the CC configuration the output is taken at the emitter rather than the collector. Also, the load resistance R_L is connected between the emitter and the ground.

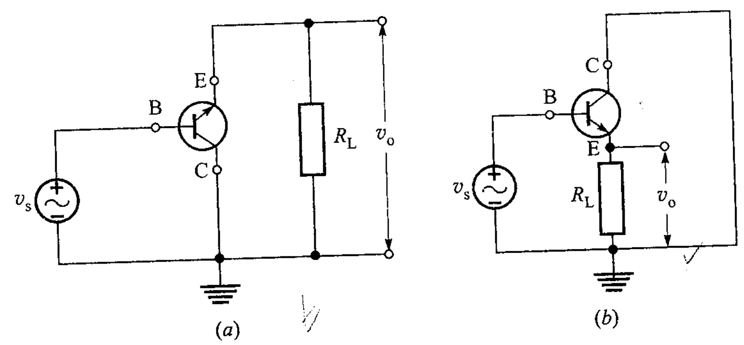


Fig. 5.21 *Transistor connected in CC configuration*

The biasing arrangement for a CC configuration is shown in Fig. 5.22. The battery V_{BB} forward biases the base-emitter junction. The battery V_{CC} has large voltage so that the collector-junction is reverse biased. If a PNP transistor is used in place of the NPN, the polarities of the batteries V_{BB} and V_{CC} are reversed. Note that the load resistor is connected to the emitter terminal. Quite often, we name it R_L. You will see later that this circuit is also called *emitter follower.*

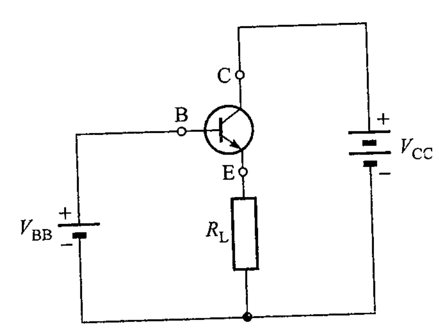


Fig. 5.22 *Biasing arrangement for an NPN transistor connected in common-collector configuration*

Current relations in CC configuration In CC configuration, the base current is the input current, and the emitter current is the output current. The output current is dependent on the input current. That is,

$$I_E = f(I_B) \tag{5.23}$$

To find this functional relationship, we start with the basic current relations of a transistor (see Eqs. (5.1) and (5.2)):

$$I_E = I_B + I_C$$

and $$I_C = \alpha_{dc} I_E + I_{CBO}$$

Since the collector is the common terminal, we are not interested in the value of collector current. We, therefore, eliminate the collector current I_C from the above two equations. Substituting the expression for collector current from the second equation into the first equation, we get

$$I_E = I_B + \alpha_{dc} I_E + I_{CBO}$$

or $$(1-\alpha_{dc})I_E = I_B + I_{CBO}$$

or $$I_E = \frac{1}{1-\alpha_{dc}} I_B + \frac{1}{1-\alpha_{dc}} I_{CBO}$$

Since

$$\frac{1}{1-\alpha_{dc}} = \beta_{dc} + 1$$

Therefore,

$$I_E = (\beta_{dc}+1)I_B + (\beta_{dc}+1)I_{CBO} \tag{5.24}$$

If we neglect the leakage current I_{CBO}, then

$$I_E = (\beta_{dc}+1)I_B$$

or $$\frac{I_E}{I_B} = (\beta_{dc}+1) \tag{5.25}$$

Equation (5.25) shows that the dc current gain (sometimes designated as γ_{dc}) of this configuration is maximum. It is equal to $(\beta_{dc} + 1)$. The leakage current in this configuration is as high as in CE configuration.

5.8 COMPARISON BETWEEN THE THREE CONFIGURATIONS

We have seen that a transistor can be connected in any one of the three configurations. It behaves differently in different configurations. In which configuration should we connect a transistor? This depends upon the particular application we desire. A configuration may be suitable for some application, whereas it may not be suitable for the other. What are the important parameters that govern the suitability of the configuration? We should know the input dynamic resistance, output dynamic resistance, dc current gain, ac voltage gain and leakage current of the transistor in a given configuration.

Out of the three configurations, the common-collector configuration has maximum input dynamic resistance. So we use this configuration where high input resistance is of prime importance, even though its voltage gain is less than unity. The decreased voltage gain can be compensated by subsequently using the CE configuration. We do not study the CC configuration separately as an independent circuit. It is usual practice to consider the CC configuration as a special case of the CE configuration*. We shall therefore consider and compare only the CB and CE configurations. Table 5.2 gives the typical values of the important parameters in the two configurations.

Table 5.2 *Comparison between CB and CE configurations*

Parameters	*Common-base configuration*	*Common-emitter configuration*
1. Input dynamic resistance	Very low (20 Ω)	Low (1 kΩ)
2. Output dynamic resistance	Very high (1 MΩ)	High (10 kΩ)
3. Current gain	Less than unity (0.98)	High (100)
4. Leakage current	Very small (5 μA for Ge, 1 μA for Si)	Very large (500 μA for Ge, 20 μA for Si)

5.8.1 Input Dynamic Resistance

The input dynamic resistance of the CB configuration is much lower than that of the CE configuration. This fact can be seen from the definition of the input dynamic resistance of the two configurations:

1. r_i for CB configuration $= \left.\dfrac{\Delta v_{EB}}{\Delta i_E}\right|_{V_{CB}=\text{const.}}$

2. r_i for CE configuration $= \left.\dfrac{\Delta v_{BE}}{\Delta i_B}\right|_{V_{CE}=\text{const.}}$

The numerators of the above two expressions are the same. But, the denominator of the first is of the order of a few milliamperes, whereas that of the second is of the order of a few microamperes. Hence, r_i for the CB configuration is much lower than that for the CE configuration.

5.8.2 Output Dynamic Resistance

Let us look at the output static characteristics of the two configurations (see Figs. 5.13 and 5.19). We note that the output characteristics of the CB configuration (Fig. 5.13)

* This circuit, also called emitter follower, is discussed in Unit 12 on "Feedback in Amplifier".

are almost horizontal. There is hardly any change in the collector current for a given variation in collector-to-base voltage. This means that the output resistance

$$r_o = \left.\frac{\Delta v_{CB}}{\Delta i_C}\right|_{I_E=\text{const.}}$$

is very high (since Δi_C is very small for a certain value of Δv_{CB}). Now see Fig. 5.19. These curves are not so horizontal. As we increase v_{CE}, the collector current is seen to increase by an appreciable amount. This shows that the output dynamic resistance

$$r_o = \left.\frac{\Delta v_{CE}}{\Delta i_C}\right|_{I_B=\text{const.}}$$

of the common-emitter configuration is not very high. It is of the order of 40 kΩ. Note that the slope of the output characteristic curve is not the same everywhere. It is for this reason that the value of the output dynamic resistance of the transistor depends upon the point around which the variations are taken.

5.8.3 Current Gain

The current gain (both dc as well as ac) of CB configuration is less than unity. It is typically 0.98. The closer its value to unity, the better is the transistor. A transistor having a low value of alpha (say, less than 0.95) will not make a good amplifier. Such transistors are rejected during manufacture.

The current gain of the CE configuration is quite high. It is typically 100, and it may be as high as 250. Such high current gain in the CE configuration makes it possible to have quite high voltage gain as well as high power gain.

5.8.4 Leakage Current

The leakage current in the CB configuration is very low (of the order of only a few μA). In the CE configuration, it is quite high (a few hundred μA). The leakage current in a transistor is due to the flow of minority carriers. The concentration of these minority carriers is very much dependent on temperature. Thus, the leakage current is temperature dependent. As the temperature rises, the leakage current rises. This may lead to what is known as thermal runaway of the transistor. The high value of the leakage current (and its rapid increase with temperature) in CE configuration is its great disadvantage. Since, silicon transistors have much less leakage current as compared to germanium transistors, we prefer to use silicon transistors.

5.9 WHY IS CE CONFIGURATION WIDELY USED IN AMPLIFIER CIRCUITS?

The main utility of a transistor lies in its ability to amplify weak signals. The transistor alone cannot perform this function. We have to connect some passive components (such as resistors and capacitors) and a biasing battery. Such a circuit is then called an *amplifier.* Thus, an amplifier is an electronic circuit that is capable of amplifying (or increasing the level of) signals.

Very often, a single transistor amplifying stage is not sufficient. In almost all applications we use a number of amplifier stages, connected one after the other. The signal to be amplified is fed to the input of the first stage. The output of the first stage is connected to the input of the second stage. The second stage feeds the third stage, and so on. Ultimately, the output appears across the load connected to the output of the final stage. Such a connection of amplifier stages is known as *cascaded amplifier.*

Figure 5.23 shows a cascaded amplifier having two stages. The first stage is energised by a signal source having voltage v_s and internal resistance R_s. The load is connected to the output of the second stage at terminals A_3B_3. If this cascaded amplifier is to work properly, certain conditions must be satisfied. The working of one stage should not adversely affect the performance of the other.

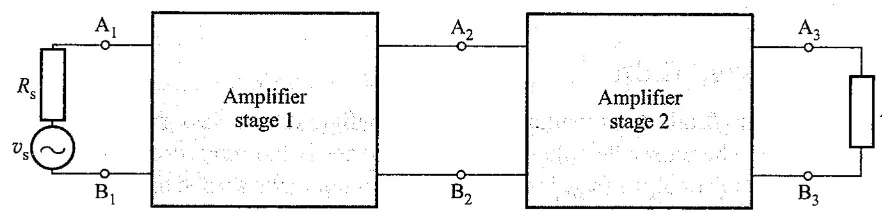


Fig. 5.23 *Two amplifier stages cascaded to increase amplifying action*

First of all, we would want the whole of (if not whole, then atleast most of) the signal voltage v_s to reach the input of the first stage. This can happen only when the input resistance of this stage at terminals A_1B_1 is high (compared to source resistance R_s). Recall that a source works as a good voltage source when the load resistance is much greater than the source resistance. Here, the input resistance of the first stage acts as the load resistance to the source. Secondly, it is desirable that the performance of first stage is not disturbed when we connect the second stage at terminals A_2B_2. For this, the output resistance of the first stage should be low. Also, the input resistance of the second stage (which comes in parallel with the load resistance of the first stage) should be high. You may recollect that connecting a high resistance in parallel with a low resistance element of a circuit does not much affect the working of the circuit. The resistance of the parallel combination will almost be the same as the low resistance itself.

Moreover, the first stage serves as the voltage source for the second stage. The input resistance of second stage acts as the load resistance for the voltage source (i.e., the first stage). The output resistance of the first stage is the internal resistance of the voltage source. The internal resistance of the source must be low compared to load resistance. Again, the second amplifier stage will deliver more power to the load R_L (this load may be a loudspeaker) only if its output resistance is low. Thus, we find that *a good amplifier stage is one which has high input resistance and low output resistance.*

A transistor in CB configuration has a very low input resistance ($\simeq$ 20 Ω) and a very high output resistance ($\simeq$ 1 MΩ). It is just the reverse of what we desire (high input resistance and low output resistance). That is why the CB configuration is unpopular. Comparatively, the CE configuration is much better, as regards its input and output resistances. Its input resistance is about 1 kΩ and output resistance about 10 kΩ. A transistor in the CE configuration makes a much better amplifier. Furthermore, the current gain, voltage gain and power gain of CE is much greater than those of CB.

From the point of view of cascading of amplifier stages, the CC configuration would have been the best. Its input resistance is very high ($\simeq$ 150 kΩ) and output resistance is quite low ($\simeq$ 800 Ω). However, unfortunately the voltage gain of the CC amplifier is low (less than unity). Therefore, we use CC amplifier only in such applications where the requirement of high input resistance is of prime importance.

Thus we see that CE configuration is best suited for most of the amplifier circuits. We shall study this circuit in some detail.

5.10 BASIC CE AMPLIFIER CIRCUIT

Figure 5.24 shows a basic CE amplifier circuit*. Here, we have used an NPN transistor. The battery V_{BB} forward biases the emitter junction. The series resistance R_B is meant to limit the base current within certain specified values. The battery V_{CC} is a relatively high-voltage battery (9 V). It reverse biases the collector junction. The resistor R_C in the collector circuit is the load resistance. The amplified ac voltage appears across it.

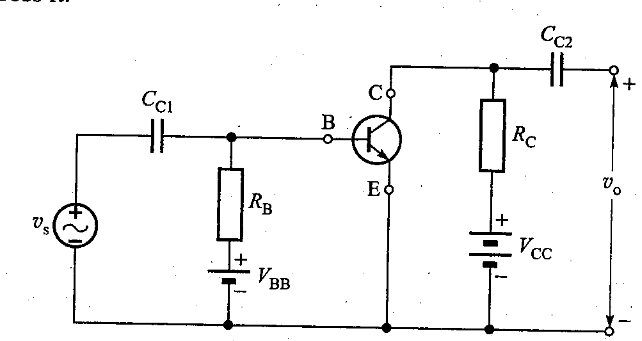


Fig. 5.24 *Basic CE amplifier circuit*

The signal to be amplified is represented by voltage source v_s. The signal is applied to the base through the coupling capacitor C_{C1}. The capacitor permits only ac to pass through. It blocks dc voltage. The dc base current flows only through resistor R_B, and not through the voltage source v_s. Similarly, the coupling capacitor C_{C2} blocks dc from reaching the output terminals. Only ac signal voltage appears at the output v_o.

* This is not a practical circuit. In practice, we use only one battery (say, V_{CC}) for biasing both the collector junction as well as the emitter junction. We shall study such practical circuits later.

To observe the performance of the amplifier circuit, we take the help of a dc load line.

5.10.1 DC Load Line

Let us consider the amplifier circuit of Fig. 5.24, when no signal is applied to its input. This condition (of having no input signal) is described as a *quiescent condition.* The circuit then reduces to the one shown in Fig. 5.25. The battery V_{CC} sends current I_C through the load resistor R_C and the transistor. There is some voltage drop across the load resistor R_C due to the flow of current I_C. The polarity of this voltage drop $I_C R_C$ is shown in the figure. The remaining voltage drops across the transistor. This voltage is written as V_{CE}. Applying Kirchhoff's voltage law to the collector circuit, we get

$$V_{CC} = I_C R_C + V_{CE} \tag{5.26}$$

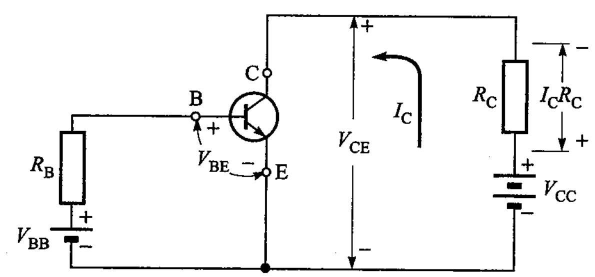


Fig. 5.25 *CE amplifier in quiescent condition*

We can rearrange the terms of the above equation and put it as

$$I_C = \left(-\frac{1}{R_C}\right) V_{CE} + \frac{V_{CC}}{R_C} \tag{5.27}$$

We have rewritten Eq. (5.26) in above form, because we wanted to put it in the form

$$y = mx + c \tag{5.28}$$

which is the equation of a straight line. If Eq. (5.27) is plotted on the transistor's output characteristics (i.e., the curves between v_{CE} and i_C), we get a straight line. Comparison of Eq. (5.27) with Eq. (5.28) indicates that the slope of this line is

$$m = -\frac{1}{R_C} \tag{5.29}$$

and its intercept on the i_C axis is

$$c = \frac{V_{CC}}{R_C} \tag{5.30}$$

The straight line represented by Eq. (5.27) is called the ***dc load line.***

Plotting of the dc load line on collector characteristics is easy. Find any two points satisfying Eq. (5.27), and then join these points. The simplest way, then, is to take one point on the v_{CE} axis and the other on i_C axis. On the v_{CE} axis, the current I_C must be zero. Hence, from Eq. (5.27), we should have $V_{CE} = V_{CC}$. When $V_{CE} = 0$, Eq. (5.27) gives $I_C = V_{CC}/R_C$. Thus, the two points on the dc load line are

1. $V_{CE} = V_{CC}$; $I_C = 0$ (Point A in Fig. 5.26)
2. $V_{CE} = 0$; $I_C = \dfrac{V_{CC}}{R_C}$ (Point B in Fig. 5.26)

These two points can be located on the collector characteristics. See Fig. 5.26. Join these two points. This is the dc load line. The slope of this line is $(-1/R_C)$ and is decided by the value of resistor R_C. Since this resistance is the dc load* of the amplifier, we call this line as dc load line.

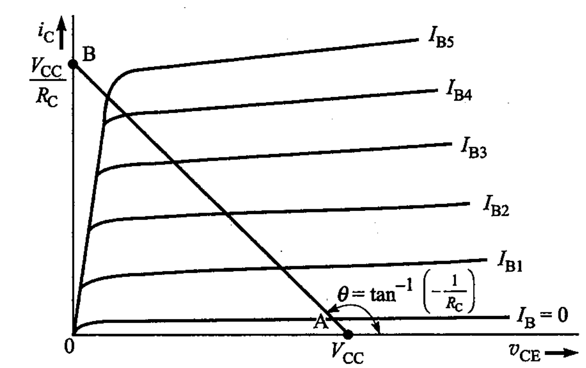


Fig. 5.26 *Plotting of dc load line on collector characteristics*

In an amplifier circuit, the operating conditions of the transistor are described by the values of its V_{CE} and I_C. These values fix up the *operating point* of the transistor. The operating point is decided not only by the characteristics of the transistor itself, but also by a number of other factors. These factors are V_{CC}, R_C, R_B, V_{BE} and V_{BB}. First, we fix the values of V_{CC} and R_C in an amplifier circuit. This ensures that the operating point of the transistor must lie on the dc load line? Now, where exactly does the operating point lie on the dc load line. This is decided by the value of the base current I_B. And, in turn, base current I_B is decided by the values of V_{BE} (of the transistor), R_B and V_{BB}. Applying Kirchhoff's voltage law to the base circuit of the transistor, we get

$$V_{BB} = I_B R_B + V_{BE}$$

or

$$I_B = \frac{V_{BB} - V_{BE}}{R_B} \cong \frac{V_{BB}}{R_B} \tag{5.31}$$

Knowing the values of V_{BB}, R_B and V_{BE} (value of V_{BE} is 0.7 V for Si transistors and 0.3 V for Ge transistors), the above equation gives the value of base current I_B.

* Later we shall learn that the ac load of an amplifier may be different from its dc load.

Corresponding to this base current, there will be a collector characteristic curve. If by chance this curve is not present on the characteristics, we can plot the curve (see Example 5.7). The exact operating point will lie at the intersection of this curve and the dc load line. This point is called *quiescent operating point* or simply Q point.

Example 5.7 A silicon transistor is used in the circuit of Fig. 5.25, with $V_{CC} = 12$ V, $R_C = 1$ kΩ, $V_{BB} = 10.7$ V, and $R_B = 200$ kΩ. The collector characteristics of the transistor are given in Fig. 5.27. Determine the Q point.

Solution: First we plot the dc load line on the output characteristic curves. Two points on the dc load line are (V_{CC}, 0) and (0, V_{CC}/R_C). Here, $V_{CC} = 12$ V and $R_C = 1$ kΩ. Therefore, the two points are (12 V, 0) and (0, 12 mA). In other words, the load line cuts the v_{CE} axis at 12 V and the i_C axis at 12 mA. We join these two points to get the dc load line.

The operating point may lie anywhere on this dc load line. To fix the Q point, we will determine the base current I_B. Applying Kirchhoff's voltage law (KVL) to the input circuit gives

$$V_{BB} = I_B R_B + V_{BE}$$

or

$$I_B = \frac{V_{BB} - V_{BE}}{R_B}$$

Here, $V_{BB} = 10.7$ V, and $R_B = 200$ kΩ. For a silicon transistor, $V_{BE} = 0.7$ V. Therefore, the base current is

$$I_B = \frac{10.7 - 0.7}{200 \times 10^3} = 50\ \mu\text{A}$$

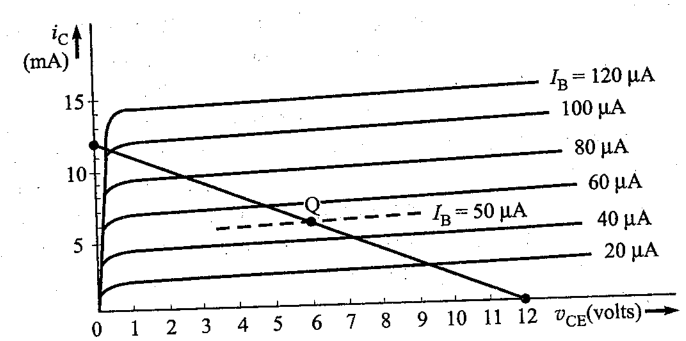


Fig. 5.27 *Fixing the Q point of the transistor on its output characteristics*

However, it is seen from Fig. 5.27 that the curve for $I_B = 50$ μA is not given. We draw this curve between the curves for $I_B = 40$ μA, and $I_B = 60$ μA. This curve is shown

dotted. The point of intersection of this curve and the dc load line gives the Q point. At this point

$$V_{CE} = \mathbf{6\ V}$$

and

$$I_C = \mathbf{6\ mA}$$

5.10.2 Amplifier Analysis using DC Load Line

A transistor can amplify ac signals only after its dc operating point is suitably fixed. We have seen in the last section how to fix the Q point on the output characteristics. The Q point should preferably lie in the middle portion of the active region of the characteristics. This helps the transistor to amplify ac signals faithfully, i.e., without distorting its waveshape.

Under quiescent condition, the base current has a constant dc value. It is determined from the Q point. Now, we apply the ac signal to the input of the amplifier circuit (see Fig. 5.24). The base voltage varies as per the signal voltage v_s. As a result, the base current will also vary. As the base current varies, the instantaneous operating point of the transistor moves along the dc load line. Thus, the instantaneous values of collector current and voltage also vary according to the input signal. The variation in collector voltage is many times larger than the variation of the input signal. The collector-voltage variation reaches the output terminals through capacitor C_{C2}. The output is therefore many times larger than the input.

Let us take an illustrative example. See the amplifier circuit in Fig. 5.24. As in Example 5.7, let us assume that $V_{CC} = 12$ V, $R_C = 1$ kΩ, $V_{BB} = 10.7$ V and $R_B = 200$ kΩ. In this circuit, the dc base current is found to be 50 μA. The collector dc current and dc voltage are 6 mA and 6 V, respectively. Let us now apply a small ac signal voltage, say, 7 mV to the input. This voltage will have about 20 mV peak-to-peak variation. If the input dynamic resistance r_i (or h_{ie}) of the transistor is assumed to be 1 kΩ, the input voltage will produce a peak-to-peak variation of 20 μA in base current, since

$$\Delta i_B = \frac{20\ \text{mV}}{1\ \text{k}\Omega} = 20\ \mu\text{A}$$

This variation in base current takes place around its quiescent value of 50 μA. As the base current varies, the instantaneous operating point moves along the dc load line between the points A($I_B = 60$ μA) and B($I_B = 40$ μA). To show the variation in I_B on the collector characteristics of the transistor, we draw a line perpendicular to the dc load line and passing through the Q point. This line is taken as ωt axis, and then, variation in I_B (assumed sinusoidal) is plotted (Fig. 5.28).

As the instantaneous operating point moves along dc load line between the point A and B, both the collector current and collector voltage vary. The current i_C varies between the points A_1($I_C = 7.3$ mA) and B_1($I_C = 4.8$ mA). This variation is shown on the left side of the characteristics. The voltage v_{CE} varies between points A_2 ($V_{CE} = 4.9$ V) and B_2 ($V_{CE} = 7.1$ V). The collector-voltage variation is shown at the bottom of the characteristics.

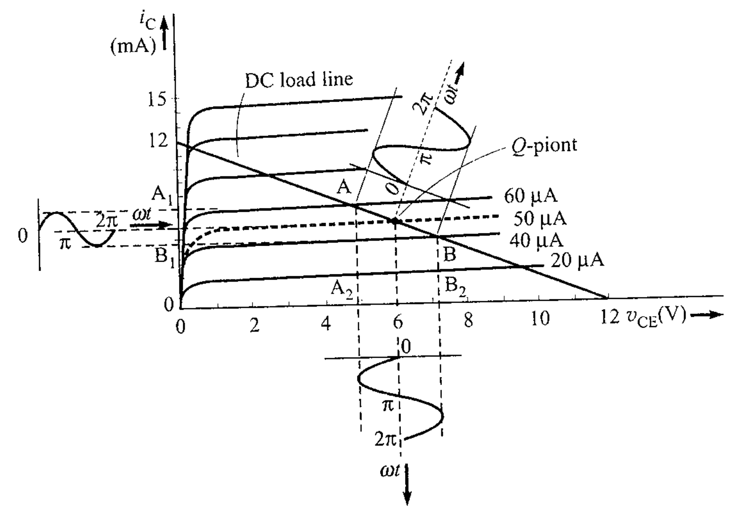


Fig. 5.28 *Variation in base current produces variation in collector current and voltage in a CE amplifier*

The current gain, voltage gain and the power gain of the amplifier can now be computed. We shall make the calculations on the basis of peak-to-peak variation:

1. Current gain, $A_i = \dfrac{\Delta i_C}{\Delta i_B} = \dfrac{(7.3 - 4.8)10^{-3}}{(60 - 40) \times 10^{-6}} = \mathbf{125}$

2. Voltage gain, $A_v = \dfrac{\Delta v_{CE}}{\Delta v_{BE}} = \dfrac{7.1 - 4.9}{20 \times 10^{-3}} = \mathbf{110}$

3. Power gain $= \dfrac{\text{output ac power}}{\text{input ac power}} = \dfrac{I_c V_{ce}}{I_b V_{be}}$

$$= A_i \times A_v = 125 \times 110 = \mathbf{13\,750}$$

We find that the CE amplifier has sufficiently large values of current gain, voltage gain, and power gain.

5.11 CONSTRUCTION OF TRANSISTORS

In recent years, the construction of transistors has undergone a great many changes and improvements. A number of different methods of manufacturing transistors have been developed since the invention of transistor in 1948. A description of all the methods is outside the scope of this book. The most commonly used types of transistors are alloy junction transistor and silicon planar transistor. We shall discuss these two types here.

5.11.1 Alloy Junction Transistor

The alloy junction transistor is one of the earliest types of transistor that is still in use. It is relatively inexpensive and provides high current gain. It can be constructed to operate at high current and power levels.

The construction of a germanium alloy junction transistor is illustrated in Fig. 5.29. We start with a very thin (of the order of 250 μm) N-type germanium crystal wafer. This wafer is lightly doped and serves as the base of the transistor. On the two sides of this wafer, indium dots (P-type impurity) are placed. It is then heated to a temperature above the melting point of indium and below the melting point of germanium. The indium melts and dissolves the germanium. A liquid solution of germanium in indium is obtained. The wafer is then slowly cooled. During cooling, a region of P-type germanium is produced and an alloy of germanium and indium (mainly indium) is deposited on the wafer. The emitter and collector leads are connected to this alloy (on the two sides of the wafer). The process results in a PNP transistor.

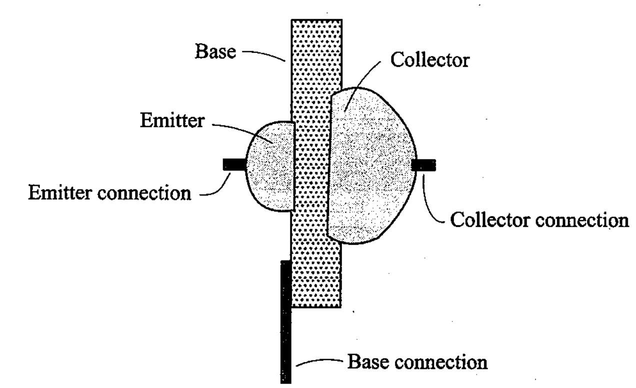


Fig. 5.29 *Construction of an alloy junction transistor*

5.11.2 Silicon Planar Transistor

The construction of a silicon planar transistor is shown in Fig. 5.30. The important feature of this type of transistor is that the PN-junctions are buried under a layer of silicon dioxide. This layer protects the PN-junctions from impurities.

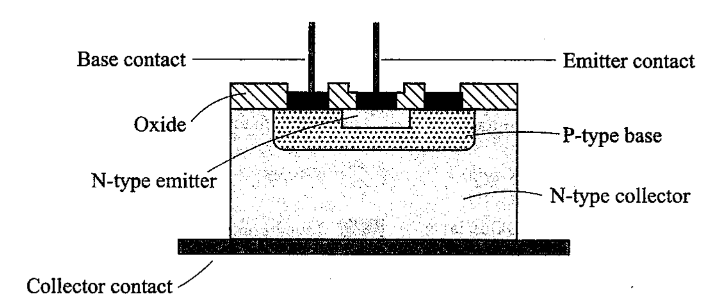


Fig. 5.30 *Construction of a silicon planar transistor*

The steps involved in the manufacture of a silicon planar transistor are illustrated in Fig. 5.31. To make an NPN transistor, we start with an N-type silicon wafer, which would ultimately make the collector. The top surface of this wafer is oxidised to a depth of approximately 1 μm (Fig. 3.31*a*). SiO_2 is an insulating material which cannot be penetrated by impurities. To make the base region, we diffuse acceptor-type impurity (e.g., boron) into the wafer. However, because the SiO_2 film checks impurity diffusion, we must remove the film from those areas on the wafer where the base is to be diffused. This is done by etching away the SiO_2 from that area, with a masked photo-resist process (Fig. 5.31*b*). The wafer is now exposed to a vapour of boron (P-type impurity) and the impurity is allowed to diffuse into the wafer to a predecided depth. Now, another layer of SiO_2 is grown over the entire wafer (Fig. 5.31*c*). A part of the SiO_2 film is again etched away by the photo-resist process using another mask (Fig. 5.31*d*). The wafer is now exposed to a vapour of donor-type impurity (e.g., phosphorus) and is also reoxidised again (Fig. 5.31*e*). The wafer now contains a layer of P-type material that makes the base of the transistor and a layer of N-type material that makes the emitter. SiO_2 is again etched away from the surface of

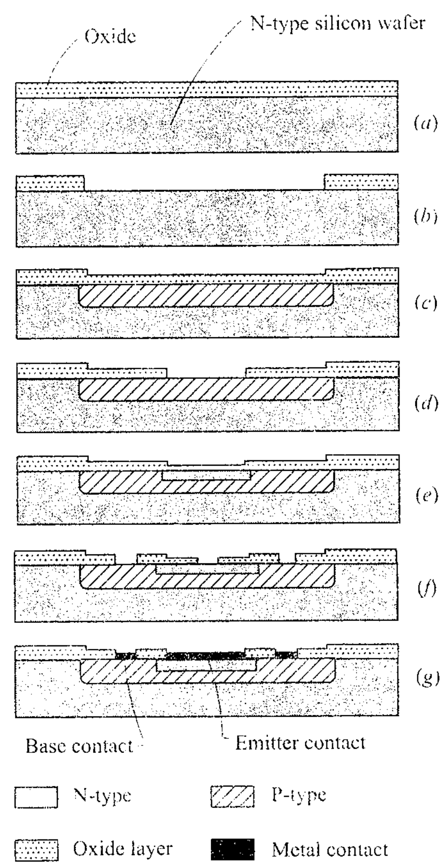


Fig. 5.31 *Various stages in the manufacture of a silicon planar NPN transistor*

the wafer to separate the base and emitter regions (Fig. 5.31*f*). Finally, metal contacts are made onto the etched areas (Fig. 5.31*g*). The wafer is now cut to the required size. It is mounted on a suitable collector contact. Leads are then connected to the base and emitter contacts.

5.12 TRANSISTOR DATA SHEETS

To analyse or to design a transistor circuit, one must have sufficient information about the transistor. This information is obtained from the manufacturer's data sheets. These sheets describe the transistor. Sometimes the outline and dimensions are also given. The lead orientation is also identified here. Commonly, the lead orientation is as shown in Fig. 5.32*a*. A red dot is placed near one of the terminals. This represents the collector lead. Now put the transistor such that the leads are facing you as in Fig. 5.32*b*. The central terminal is the base. The third one is the emitter. However, a word of caution is necessary. The convention described above for recognising the three leads is not a standard one. Different manufacturers use different conventions for this purpose.

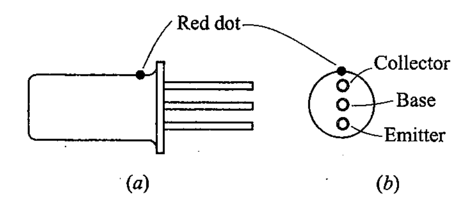


Fig. 5.32 *Orientation of leads in a transistor. A red dot is placed on the body of the transistor near the collector lead*

The important set of data, from a user's point of view are as follows:

1. The maximum power dissipation in the transistor at 25 °C.
2. The maximum allowable collector-base voltage.
3. The current gain β or h_{fe}.
4. The transition frequency f_T of the transistor.

Generally, one or the other of these four factors is of prime importance, depending upon the application. In no case should the maximum ratings given in items 1 and 2 above, be exceeded. Otherwise the transistor may be damaged.

If a transistor is required for a small-signal audio frequency amplifier, the most important factor in selecting a transistor is its current gain. In some cases, it may be necessary to see the collector-base voltage. Since the power involved will be small enough, and the transistor is not required to handle high frequencies, it is not necessary to consider the factors at 1 and 4.

5.12.1 Transistor Testing

Today, the market is flooded with transistors of all sorts and makes. Very often, we come across a transistor whose specifications are not known. Sometimes, the transistor type number may be obliterated from its body. Even if the transistor type is known, the reference data-book may not be readily available. In these circumstances, it becomes necessary to test a transistor. In this section we shall see how to conduct the test to determine whether the transistor is NPN or PNP. Also, a test is given to identify the transistor terminals.

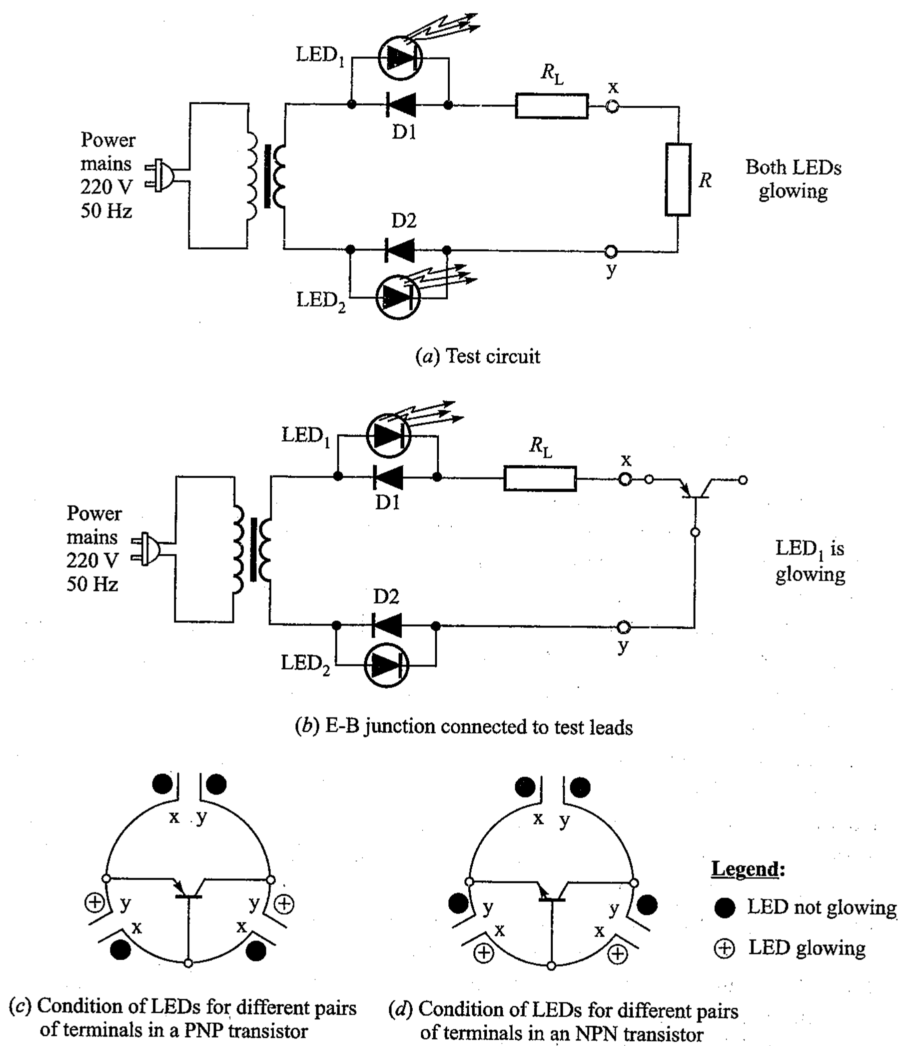
Test to distinguish between PNP and NPN transistors Figure 5.33*a* shows a simple circuit for testing a transistor for its nature (PNP and NPN). In this circuit, two germanium rectifier diodes and two LEDs (light-emitting diodes) are used. A resistor R_L is also placed in series so as to prevent a heavy current from flowing in the circuit. The two leads of the tester are marked x and y. If a resistor R is placed between these terminals, the current passes for both the halves of the input wave. In the positive half, current flows through LED_1, R_L, R, and D2. The diode D1 will not conduct during this half. The forward voltage drop of 0.3 V across D2 will prevent LED_2 from glowing in this half. During the negative half-cycle, the current flows through LED_2, R, R_L and D1. During this period, LED_2 will glow, while LED_1 will not. Thus both the LEDs will glow alternatively. As the frequency of supply is 50 Hz (quite high) we shall observe both the LEDs glowing continuously.

Now, consider the case when one of the junctions (say E-B junction) of a transistor (say, PNP-type) is connected across the test leads x and y. This is shown in Fig. 5.33*b*. In this case, current cannot flow for those half-cycles when the E-B junction is reverse biased. However, current flows in the direction from emitter to base (from P to N) during those half-cycles when the E-B junction is forward biased. As such, only LED_1 will glow. This test indicates that the terminal connected to the lead x is P-type (and that connected to lead y is N-type). Thus, this simple circuit identifies P- and N-type terminals of a PN-junction.

None of the LEDs will glow when test leads are connected to the terminals of the same type (emitter and collector) of the transistor. Under this condition, the base is open circuited. No current (except a very small leakage current) flows through the transistor.

As seen earlier, only one LED glows when the test leads are connected across an E-B junction or across a C-B junction. That is, between those two pair of terminals, the common terminal must be the base terminal. The remaining two must be the emitter and collector terminals. In case the common terminal (the base) is P-type, the transistor is obviously NPN. In the other case, when the common terminal is N-type, it is a PNP transistor. The conditions of the two LEDs when the test leads are connected to different pairs of terminals of a PNP transistor are shown in Fig. 5.33*c*. Figure 5.33*d* shows the same procedure for an NPN transistor.

It is obvious that the glowing of both the LEDs indicate a short circuited pair of terminals, i.e., the transistor is faulty. This test cannot distinguish between emitter and collector terminals. For this, we conduct another test described as follows.



(*a*) Test circuit

(*b*) E-B junction connected to test leads

(*c*) Condition of LEDs for different pairs of terminals in a PNP transistor

(*d*) Condition of LEDs for different pairs of terminals in an NPN transistor

Fig. 5.33 *PNP/NPN check for a transistor*

Identification of emitter and collector terminals Once the transistor type (PNP or NPN) is known and the base terminal is identified, we can use the arrangement shown in Fig. 5.34 to identify the emitter and collector terminals. We use an ohmmeter to measure the resistance offered by the E-B junction and the C-B junction, when forward biased. If it is a PNP transistor (as shown in Fig. 5.34) connecting the positive (or red) lead of ohmmeter to the emitter (or collector), and the negative (or black) lead to the base, then the junction is forward biased. We measure the resistance of one junction, and then using the same ohmmeter we measure the resistance of the other junction. *The measurement that results in the higher resistance* *reading indicates the emitter terminal.* The other terminal is obviously the collector terminal.

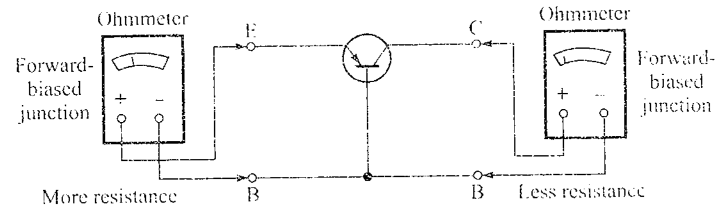


Fig. 5.34 *Test arrangement for the identification of emitter and collector terminals*

5.13 THERMAL RUNAWAY AND HEAT SINK

If the temperature of the collector-base junction increases, the collector leakage current I_{CBO} increases. Because of this, collector current increases. The increase in collector current produces an increase in the power dissipated at the collector junction. This, in turn, further increases the temperature of the junction and so gives further increase in collector current. The process is cumulative. It may lead to the eventual destruction of the transistor. This is described as the *thermal runaway* of the transistor. In practice, thermal runaway is prevented in a well-designed circuit by the use of stabilisation circuitry.

For transistors handling small signals, the power dissipated at the collector is small. Such transistors, have little chances of thermal runaway. However in power transistors, the power dissipated at the collector junction is larger. This may cause the junction temperature to rise to a dangerous level. We can increase the power handling capacity of a transistor if we make a suitable provision for rapid conduction of heat away from the junction. This is achieved by using a sheet of metal called *heat sink*. As the power dissipated within a transistor is predominantly the power dissipated at its collector-base junction, sometimes the collector of the power transistor is connected to its metallic case. The case of the transistor is then bolted on to a sheet of metal as shown in Fig. 5.35*a*. This sheet serves as the heat sink.

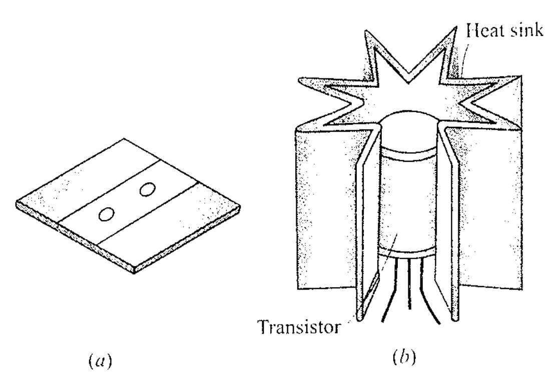


Fig. 5.35 *Two kinds of heat sinks used with power transistors*

Connecting a heat sink to a transistor increases the area from which heat is to be transferred to the atmosphere. Heat moves from the transistor to the heat sink by conduction and then it is removed from the sink to the ambient by convection and radiation.

Another type of heat sink is shown in Fig. 5.35*b*. It consists of a pushfit clip. This clip is pushed on to the transistor. To increase the surface area of the heat sink, it is usually given a ribbed structure. Because of this structure, the heat sink does not occupy much space within the equipment.

For maximum efficiency, a heat sink should (*i*) be in good thermal contact with the transistor case, (*ii*) have the largest possible surface area, (*iii*) be painted black, and (*iv*) be mounted in a position such that free air can flow past it.

Review Questions

1. Draw a sketch showing the structure of an NPN-junction transistor. Label the emitter, base and collector regions. Also label the emitter-base and collector-base junctions.
2. Repeat the above for a PNP-junction transistor.
3. Draw the circuit symbol of an NPN transistor and indicate the reference directions, according to standard convention, for the three currents.
4. Explain why an ordinary junction transistor is called bipolar.
5. Show the biasing arrangement for a PNP transistor in CB configuration so that it works in active region.
6. Explain the function of the emitter in the operation of a junction transistor.
7. What is done to the base region of a transistor to improve its operation?
8. Though the collector-base junction of a transistor operating in active region is reverse-biased, the collector current is still quite large. Explain briefly, say, within 10 lines.
9. What do you understand by the collector reverse-saturation current? In which configuration (CB or CE) does it have a greater value?
10. Besides the active region of operation of a transistor, what are the other possible conditions of operation of a transistor? Give the biasing conditions of each.
11. What causes collector current to flow when the emitter current is zero? What is this collector current called?
12. Explain the reason why the base current in a transistor is usually much smaller than I_E or I_C in active operation.
13. For a PNP transistor in the active region, what is the sign (positive or negative) of I_E, I_B, I_C, V_{EB} and V_{CE}?
14. Draw an NPN transistor in the CB configuration biased for operation in active region.
15. What is considered the input terminal and what is the output terminal in the CB configuration?

16. Sketch typical CB input characteristic curves for an NPN transistor. Label all variables. Explain how you will calculate the input dynamic resistance of the transistor from these curves.
17. Sketch typical output characteristic curves for a PNP transistor in CB configuration. Label all variables and indicate active, cut-off and saturation regions.
18. What are the input and output terminals in the CE configuration?
19. An NPN transistor is to be used in common-emitter configuration. Show how you will connect the external batteries so that the transistor works in the active region.
20. Sketch typical CE input characteristics for an NPN transistor. Label all variables. Outline the procedure of calculating the input dynamic resistance of the transistor at a given point from these curves.
21. Sketch typical CE output characteristic curves for an NPN transistor. Label all variables. Explain in brief how you will compute the beta of the transistor from these characteristic curves?
22. Derive the relationship between the beta and alpha of a transistor.
23. Compare the relative values of input and output resistances for the common-base and common-emitter configurations. Give their typical values.
24. Explain why CE configuration is most popular in amplifier circuits.
25. Draw the circuit diagram of a simple transistor amplifier in CE configuration. Write the equation of a dc line.
26. Explain how you will determine the voltage gain of the CE amplifier by plotting the dc load line on the output characteristics of the transistor.
27. Describe briefly the procedure for manufacturing alloy junction transistors.
28. Explain the important steps in making a silicon planar transistor. Why has this technology of manufacturing transistors become so popular recently?
29. You are given a transistor. Somehow the red dot on its body has been obliterated. Explain how you will determine its three terminals.
30. What is meant by thermal runaway in a transistor? Explain.
31. What is a heat sink? Draw a typical heat sink. List the factors which determine its efficiency.
32. State the order of magnitude of the collector reverse saturation current I_{CBO} for (*a*) germanium transistor, and (*b*) silicon transistor. How does it vary with temperature in each case?

Objective-Type Questions

I. Below are some incomplete statements. Four alternatives are provided for each. Choose the alternative that completes the statement correctly.

1. In a PNP transistor with normal bias
   (*a*) only holes cross the collector junction
   (*b*) only majority carriers cross the collector junction

   (*c*) the collector junction has a low resistance
   (*d*) the emitter-base junction is forward biased and the collector-base junction is reverse biased.
2. In a transistor with normal bias, the emitter junction
   (*a*) has a high resistance
   (*b*) has a low resistance
   (*c*) is reverse biased
   (*d*) emits such carriers into the base region which are in majority (in the base)
3. For transistor action
   (*a*) the collector must be more heavily doped than the emitter region
   (*b*) the collector-base junction must be forward-biased
   (*c*) the base region must be very narrow
   (*d*) the base region must be N-type material
4. The symbol I_{CBO} signifies the current that flows when some dc voltage is applied
   (*a*) in the reverse direction to the collector junction with the emitter open circuited
   (*b*) in the forward direction to the collector junction with the emitter open circuited
   (*c*) in the reverse direction to the emitter junction with the collector open circuited
   (*d*) in the forward direction to the emitter junction with the collector open circuited
5. The current I_{CBO}
   (*a*) is generally greater in silicon than in germanium transistors
   (*b*) depends largely on the emitter-base junction bias
   (*c*) depends largely on the emitter doping
   (*d*) increases with an increase in temperature
6. The main current crossing the collector junction in a normally biased NPN transistor is
   (*a*) a diffusion current
   (*b*) a drift current
   (*c*) a hole current
   (*d*) equal to the base current
7. In a PNP transistor, electrons flow
   (*a*) out of the transistor at the collector and base leads
   (*b*) into the transistor at the emitter and base leads
   (*c*) into the transistor at the collector and base leads
   (*d*) out of the transistor at the emitter and base leads
8. The current I_{CBO} flows in
   (*a*) the emitter, base, and collector leads
   (*b*) the emitter and base leads

(*c*) the emitter and collector leads
(*d*) the collector and base leads

9. The emitter region in the PNP junction transistor is more heavily doped than the base region so that
(*a*) the flow across the base region will be mainly because of electrons
(*b*) the flow across the base region will be mainly because of holes
(*c*) recombination will be increased in the base region
(*d*) base current will be high

10. For a given emitter current, the collector current will be higher if
(*a*) the recombination rate in the base region were decreased
(*b*) the emitter region were more lightly doped
(*c*) the minority-carrier mobility in the base region were reduced
(*d*) the base region were made wider

11. The arrowhead on the transistor symbol always points in the direction of
(*a*) hole flow in the emitter region
(*b*) electron flow in the emitter region
(*c*) minority-carrier flow in the emitter region
(*d*) majority-carrier flow in the emitter region

12. A small increase in the collector reverse bias will cause
(*a*) a large increase in emitter current
(*b*) a large increase in collector current
(*c*) a large decrease in collector current
(*d*) very small change in collector reverse saturation current

13. One way in which the operation of an NPN transistor differs from that of a PNP transistor is that
(*a*) the emitter junction is reverse biased in the NPN
(*b*) the emitter injects minority carriers into the base region of the PNP and majority carriers in the base region of the NPN
(*c*) the emitter injects holes into the base region of the PNP and electrons into the base region of the NPN
(*d*) the emitter injects electrons into the base region of the PNP and holes into the base region of the NPN

14. The emitter current in a junction transistor with normal bias
(*a*) may be greatly increased by a small change in collector bias
(*b*) is equal to the sum of the base current and collector current
(*c*) is approximately equal to the base current
(*d*) is designated as I_{CO}

15. In CB configuration, the output volt-ampere characteristics of the transistor may be shown by plots of
(*a*) v_{CB} versus i_C for constant values of I_E
(*b*) v_{CB} versus i_B for constant values of I_E
(*c*) v_{CE} versus i_E for constant values of I_B
(*d*) v_{CE} versus i_C for constant values of I_B

16. A transistor-terminal current is considered positive if
(*a*) the electrons flow out of the transistor at the terminal
(*b*) the current is due to the flow of holes only
(*c*) the current is due to the flow of electrons only
(*d*) the electrons flow into the transistor at the terminal

17. A transistor-terminal voltage is considered positive if
(*a*) the terminal is more negative than the common terminal
(*b*) the terminal is more positive than the common terminal
(*c*) the terminal is the output terminal
(*d*) the terminal is connected to P-type material

18. The current I_{CEO} is
(*a*) the collector current in the common-emitter connected transistor with zero base current
(*b*) the emitter current in the common-collector connected transistor with zero base current
(*c*) the collector current in the common-emitter connected transistor with zero emitter current
(*d*) the same as I_{CBO}

19. The common-emitter input volt-ampere characteristics may be shown by plots of
(*a*) v_{CB} versus i_C for constant values of I_E
(*b*) v_{CE} versus i_C for constant values of I_B
(*c*) v_{CE} versus i_E for constant values of V_{EB}
(*d*) v_{BE} versus i_B for constant values of V_{CE}

20. In CE configuration, the output volt-ampere characteristics may by shown by plots of
(*a*) v_{CB} versus i_C for constant values of I_E
(*b*) v_{CE} versus i_C for constant values of I_B
(*c*) v_{CE} versus i_E for constant values of V_{EB}
(*d*) v_{BE} versus i_B for constant values of V_{CE}

21. The beta (β) of a transistor may be determined directly from the plots of
(*a*) v_{CB} versus i_C for constant values of I_E
(*b*) v_{EC} versus i_E for constant values of I_B
(*c*) v_{CE} versus i_C for constant values of I_B
(*d*) v_{BE} versus i_B for constant values of V_{CE}

22. The most noticeable effect of a small increase in temperature in the common emitter connected transistor is
(*a*) the increase in the ac current gain
(*b*) the decrease in the ac current gain
(*c*) the increase in output resistance
(*d*) the increase in I_{CEO}

23. When determining the common-emitter current gain by making small changes in direct currents, the collector voltage is held constant so that
    (*a*) the output resistance will be high
    (*b*) the transistor will not burn out
    (*c*) the change in emitter current will be due to a change in collector current
    (*d*) the change in collector current will be due to a change in base current
24. The high resistance of the reverse-biased collector junction is due to the fact that
    (*a*) a small change in collector bias voltage causes a large change in collector current
    (*b*) a large change in collector bias voltage causes very little change in collector current
    (*c*) a small change in emitter current causes an almost equal change in collector current
    (*d*) a small change in emitter bias voltage causes a large change in collector current
25. A transistor connected in common-base configuration has
    (*a*) a low input resistance and high output resistance
    (*b*) a high input resistance and a low output resistance
    (*c*) a low input resistance and a low output resistance
    (*d*) a high input resistance and a high output resistance
26. Compared to a CB amplifier, the CE amplifier has
    (*a*) lower input resistance
    (*b*) higher output resistance
    (*c*) lower current amplification
    (*d*) higher current amplification
27. A transistor, when connected in common-emitter mode, has
    (*a*) a high input resistance and a low output resistance
    (*b*) a medium input resistance and a high output resistance
    (*c*) very low input resistance and a low output resistance
    (*d*) a high input resistance and a high output resistance
28. The input and output signals of a common-emitter amplifier are
    (*a*) always equal
    (*b*) out of phase
    (*c*) always negative
    (*d*) in phase
29. A transistor is said to be in a quiescent state when
    (*a*) no signal is applied to the input
    (*b*) it is unbiased
    (*c*) no currents are flowing
    (*d*) emitter-junction bias is just equal to collector-junction bias

30. When a positive voltage signal is applied to the base of a normally biased NPN common-emitter transistor amplifier
    (*a*) the emitter current decreases
    (*b*) the collector voltage becomes less positive
    (*c*) the base current decreases
    (*d*) the collector current decreases

II. Indicate which of the following statements pertain to NPN transistors and which pertain to PNP transistor:

1. The emitter injects holes into the base region.
2. When biased in the active region, current flows into the emitter terminal.
3. The electrons are the minority carriers in the base region.
4. The collector is biased negatively related to the base for active operation.
5. The principal current carriers are electrons.
6. The E-B junction is forward biased for active operation.
7. The base is made by doping the intrinsic semiconductor with indium.

Answers

I.	1. (*d*)	2. (*b*)	3. (*c*)	4. (*a*)	5. (*d*)	6. (*b*)
	7. (*c*)	8. (*d*)	9. (*b*)	10. (*a*)	11. (*a*)	12. (*d*)
	13. (*c*)	14. (*b*)	15. (*a*)	16. (*a*)	17. (*b*)	18. (*a*)
	19. (*d*)	20. (*b*)	21. (*c*)	22. (*d*)	23. (*d*)	24. (*b*)
	25. (*a*)	26. (*d*)	27. (*b*)	28. (*b*)	29. (*a*)	30. (*b*)
II.	1. PNP	2. PNP		3. NPN	4. PNP	
	5. NPN	6. Both PNP and NPN		7. NPN		

Tutorial Sheet 5.1

1. For a certain transistor $\alpha_{dc} = 0.98$ and emitter current $I_E = 2$ mA. Calculate the values of collector current I_C and base current I_B.
   [Ans. $I_C = 1.96$ mA, $I_B = 40$ μA]
2. The collector current $I_C = 2.9$ mA in a certain transistor circuit. If the base current $I_B = 100$ μA, calculate α_{dc} of the transistor. **[Ans.** $\alpha_{dc} = 0.97$]
3. The emitter current I_E in a transistor is 2 mA. If the leakage current I_{CBO} is 5 μA and $\alpha_{dc} = 0.985$, calculate the collector and base currents.
   [Ans. $I_C = 1.975$ mA, $I_B = 25$ μA]
4. In an NPN silicon transistor, $\alpha_{dc} = 0.995$, $I_E = 10$ mA, leakage current $I_{CO} = 0.5$ μA. Determine I_C, I_B, β_{dc} and I_{CEO}.
   [Ans. $I_C = 9.9505$ mA, $I_B = 49.5$ μA, $\beta_{dc} = 199$, $I_{CEO} = 100$ μA]

5. A transistor is supplied with dc voltages so that $I_B = 40\ \mu A$. If $\beta_{dc} = 80$ and the leakage current is $5\ \mu A$, what is the value of emitter current I_E?
[**Ans.** $I_E = 3.645$ mA]

Tutorial Sheet 5.2

1. In a transistor circuit, $I_E = 5$ mA, $I_C = 4.95$ mA and $I_{CEO} = 200\ \mu A$. Calculate β_{dc} and the leakage current I_{CBO}. [**Ans.** $\beta_{dc} = 99$, $I_{CBO} = 2\ \mu A$]
2. Collector current in a BC107 transistor is 5 mA. If $\beta_{dc} = 140$ and base current is $35\ \mu A$. Calculate the leakage current I_{CO}. [**Ans.** $I_{CO} = 0.71\ \mu A$]
3. A transistor is connected in CB configuration. When the emitter voltage is changed by 200 mV, the emitter current changes by 5 mA. During this variation, collector-to-base voltage is kept fixed. Calculate the dynamic input resistance of the transistor. [**Ans.** $r_i = 40\ \Omega$]
4. A variation of $5\ \mu A$ in the base current produces a change of 1.2 mA in the collector current. Collector-to-emitter voltage remains fixed during this variation. Calculate the current amplification factor β_{dc}. [**Ans.** $\beta_{dc} = 240$]

Tutorial Sheet 5.3

1. Table T. 5.3.1 gives values of the collector current and collector voltage for a series of base current values in a transistor in the CE configuration. Plot these characteristics and hence find (*a*) the current gain when the collector voltage is 6 V, (*b*) the output resistance for a base current of $45\ \mu A$.
[**Ans.** (*a*) 40.25; (*b*) 13.33 kΩ]

Table T. 5.3.1

V_{CE} (V)	*Collector current* (mA)			
	$I_B = 25\ \mu A$	$I_B = 45\ \mu A$	$I_B = 65\ \mu A$	$I_B = 85\ \mu A$
3	0.91	1.59	2.25	3.00
5	0.92	1.69	2.45	3.20
7	0.96	1.84	2.65	3.50
9	0.99	2.04	2.95	4.00

2. Table T. 5.3.2 gives the data of a transistor which is used in a common-emitter amplifier. Plot the output characteristics assuming them to be linear between the values indicated. The collector supply voltage is 10 V, and the collector load resistance is 1.2 kΩ. Draw the load line and choose a suitable operating point. Use this load line to calculate (*a*) the voltage gain and (*b*) the current gain, when a $12\ \mu A$ peak signal is applied at the base. Assume the dynamic input resistance of the transistor to be 1.8 kΩ. [**Ans.** (*a*) 41.61; (*b*) 62.5]

Table T. 5.3.2

V_{CE} (V)	I_C (mA)		
	$I_B = 40\ \mu A$	$I_B = 60\ \mu A$	$I_B = 80\ \mu A$
1	3	4.5	6.0
3	3.4	5.0	6.5
5	3.8	5.5	7.0
9	4.2	6.0	7.6
11	4.6	6.5	8.2

3. In a basic transistor amplifier shown in Fig. T. 5.3.1*a* an NPN transistor is used. The output characteristics of this transistor are shown in Fig. T. 5.3.1*b*. Draw the dc load line on the characteristics and locate the *Q* point. (*a*) Write the coordinates of the *Q* point. (*b*) Determine the current gain of this amplifier.
[**Ans.** (*a*) 7.0 V, 2.7 mA; (*b*) 16]

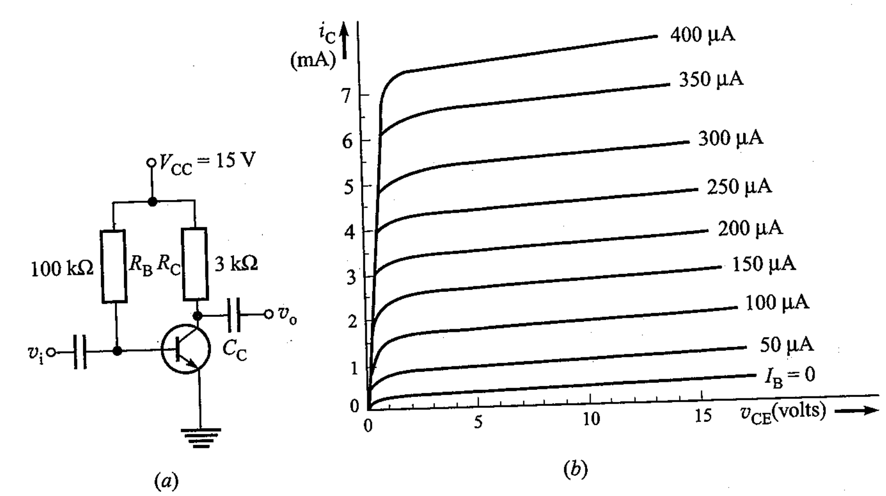


Fig. T. 5.3.1

Experimental Exercise 5.1

Title Common-base transistor characteristics.

Objectives To

1. trace the given circuit;
2. measure emitter current for different values of emitter-base voltage keeping collector-base voltage constant;

3. calculate the input dynamic resistance from the input characteristic at a given operating point;
4. plot the output characteristics (graph between the collector current and collector-to-base voltage, keeping emitter current fixed) for the given transistor;
5. calculate the output dynamic resistance r_o, α_{dc} and α at a given operating point.

Apparatus Required Experimental board, transistor (or IC) power supply, two milliammeters (0 to 50 mA), two electronic multimeters.

Circuit Diagram The circuit diagram is shown in Fig. E. 5.1.1.

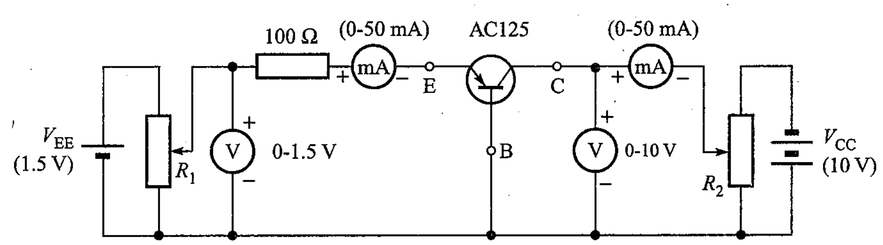


Fig. E. 5.1.1

Brief Theory A transistor is a three-terminal active device. The three terminals are emitter, base and collector. In common-base configuration, we make the base common to both input and output. For normal operation, the emitter-base junction is forward-biased and the collector-base junction is reverse biased.

The input characteristic is a plot between i_E and v_{EB} keeping voltage V_{CB} constant. This characteristic is very similar to that of a forward-biased diode. The input dynamic resistance is calculated using the formula,

$$r_i = \left.\frac{\Delta v_{EB}}{\Delta i_E}\right|_{V_{CB}=\text{const.}}$$

The output characteristic curves are plotted between i_C and v_{CB}, keeping I_E constant. These curves are almost horizontal. This shows that the output dynamic resistance, defined below is very high.

$$r_o = \left.\frac{\Delta v_{CB}}{\Delta i_C}\right|_{I_E=\text{const.}}$$

The collector current I_C is less than, but almost equal to the emitter current. The current I_E divides into I_C and I_B. That is,

$$I_E = I_C + I_B$$

When the input side is open (i.e., $I_E = 0$), the collector current is not zero, but has a small (a few μA) value. This value of collector current is called collector reverse saturation current, I_{CBO}.

At a given operating point, we define the dc and ac current gains (alpha) as follows:

$$\text{dc current gain, } \alpha_{dc} = \frac{I_C}{I_E}$$

$$\text{ac current gain, } \alpha = \left.\frac{\Delta i_C}{\Delta i_E}\right|_{V_{CB}=\text{const.}}$$

Procedure

1. From the experimental board, note down the type number of the transistor. Note the important specifications of the transistor from the data book. Identify the terminals of the transistor. Trace the circuit.
2. Make the circuit connections as shown in Fig. E. 5.1.1. Use milliammeters of proper range.
3. For input characteristics, first fix the voltage V_{CB}, say, at 6 V. Now vary the voltage v_{EB} slowly (say, in steps of 0.1 V) and note the current i_E for each value of v_{EB}.
4. Repeat the above for another value of V_{CB} say, 10 V.
5. For output characteristics, first fix the collector voltage, say, at 4 V. Open the input circuit. Note the collector current by using a microammeter. Vary the collector voltage in steps and note collector current for each value of collector voltage. This will give the curve for reverse saturation current. Now, close the input circuit. Adjust the emitter current I_E to, say, 1 mA with the help of potentiometer R_1. Again vary the voltage V_{CB} in steps. Note current I_C for each. Repeat this process for three to four different values of emitter current (say, 2 mA, 3 mA, 4 mA, etc). See to it that you do not exceed the maximum ratings of the transistor.
6. Plot the input and output characteristics by using the readings taken above.
7. Select a suitable operating point well within the active region (say, $V_{CB} = 6$ V, $I = 3$ mA). At this operating point, draw a tangent to the curve of input characteristic (you should have the curve for the selected value of V_{CB}). The slope of this curve will give the input dynamic resistance. Similarly, by drawing tangent to the output characteristic curve gives the output dynamic resistance.
8. To determine dc alpha, simply divide the dc collector current (at the selected operating point) by the dc emitter current.
9. To determine ac alpha, draw a vertical line through the selected operating point on the output characteristics. Take a small change in i_E (say, 1 mA) around the operating point and read from the graph, the corresponding change in i_C. Divide the change in i_C by the change in i_E to get ac alpha.

Observations

1. *Type number of the transistor* = ________
2. *Information from data book*:
   (*a*) Maximum collector current rating = ________ mA
   (*b*) Maximum collector-to-emitter voltage rating = ______ V
   (*c*) Maximum collector dissipation power = ________ W
3. *Input characteristics*:

S. No.	$V_{CB} = 6$ V		$V_{CB} = 10$ V	
	v_{EB} *in* mV	I_E *in* mA	v_{EB} *in* V	I_E *in* mA
1. 2. 3.				

4. *Output characteristics*:

S. No.	$I_E = 0$		$I_E = 1$ mA		$I_E = 2$ mA		$I_E = 3$ mA	
	v_{CB}(V)	I_C(mA)	v_{CB}(V)	I_C(mA)	v_{CB}(V)	I_C(mA)	v_{CB}(V)	I_C(mA)
1. 2. 3.								

Calculations

1. Input dynamic resistance,

$$r_i = \left.\frac{\Delta v_{EB}}{\Delta i_E}\right|_{V_{CB}=\_\_\_ \text{V}} = \_\_\_\_\_\_ = \_\_\_\_\_\_ \ \Omega$$

2. Output dynamic resistance,

$$r_o = \left.\frac{\Delta v_{CB}}{\Delta i_C}\right|_{I_E=\_\_\_ \text{mA}} = \_\_\_\_\_\_ = \_\_\_\_\_\_ \ \text{k}\Omega$$

3. DC current gain, $\alpha_{dc} = \dfrac{I_C}{I_E} = \_\_\_\_\_\_ = \_\_\_\_\_\_$

4. AC current gain, $\alpha = \left.\dfrac{\Delta i_C}{\Delta i_E}\right|_{V_{CB}=\_\_\_ \text{V}} = \_\_\_\_\_\_ = \_\_\_\_\_\_$

Results

1. Input and output characteristics are plotted on the graph.

2. The transistor parameters are given below:

Parameter	*Value determined*
1. r_i	______ Ω
2. r_o	______ kΩ
3. α_{dc}	______
4. α	______

Experimental Exercise 5.2

Title Transistor characteristics in common-emitter configuration.

Objectives To

1. trace the given circuit;
2. plot the input characteristics (graph between the base current i_B and base-to-emitter voltage v_{BE}, keeping collector-to-emitter voltage V_{CE} constant);
3. calculate the input dynamic resistance from the input characteristic at a given operating point;
4. plot the output characteristics (graph between i_C and v_{CE}, for fixed values of I_B);
5. calculate the output ac resistance (r_o), the dc beta (β_{dc}), and ac beta at a given operating point.

Apparatus Required Experimental board, transistor (or IC) power supply, one milliammeter, (0-50 mA), one microammeter (0-50 µA), two electronic multimeters.

Circuit Diagram The circuit diagram is shown in Fig. E. 5.2.1.

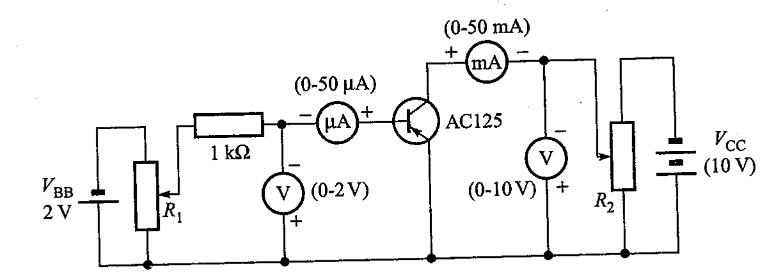


Fig. E. 5.2.1

Brief Theory In Experimental Exercise 5.1, we had drawn the transistor characteristics in CB configuration. In CE configuration, we make the emitter terminal common to the input and output. Whether the transistor is connected in CB or CE, the E-B junction is forward biased and the C-B junction is reverse biased.

For CE configuration, we defined the important parameters as follows:

1. Input dynamic resistance, $r_i = \left.\dfrac{\Delta v_{BE}}{\Delta i_B}\right|_{V_{CE}=\text{const.}}$
2. Output ac resistance, $r_o = \left.\dfrac{\Delta v_{CE}}{\Delta i_C}\right|_{I_B=\text{const.}}$
3. DC current gain, $\beta_{dc} = \dfrac{I_C}{I_B}$
4. AC current gain, $\beta = \left.\dfrac{\Delta i_C}{\Delta i_B}\right|_{V_{CE}=\text{const.}}$

Procedure

1. Note down the type number of the transistor used in the experimental board. Find the important specifications of the transistor from the data book. Identify the terminals of the transistor and trace the circuit.
2. Mark the circuit connections as shown in Fig. E. 5.2.1. Use meters with proper range.
3. For input characteristic, first fix the voltage V_{CE}, say, at 9 V. Vary the voltage v_{BE} slowly, in steps. Note the value of current i_B at each step.
4. For output characteristics, first open the input circuit. Vary the collector voltage v_{CE} in steps and note the collector current. This current is the reverse saturation current I_{CEO}, and its magnitude will be small. Now close the input circuit and fix the base current I_B at, say, 10 μA. For this you can use the potentiometer R_1. Vary the voltage v_{CE} with the help of potentiometer R_2 in steps. Note current i_C for each step. Repeat the process for other values of I_B (say, 20 μA, 30 μA, 40 μA, etc.) Be careful not to go beyond the maximum ratings of the transistor.
5. Plot the input and output characteristics by using the readings taken above.
6. Select a suitable operating point in the linear portion of the characteristics. Determine the slope of the input characteristic curve at this operating point. This gives the input dynamic resistance. Similarly, using the definition given above (in brief theory), calculate the output ac resistance r_o, dc beta and ac beta.

Observations

1. *Type number of the transistor* = ______
2. *Information from the data book*:
   (a) Maximum collector current rating = ______ mA
   (b) Maximum collector voltage rating = ______ V
   (c) Maximum collector dissipation power rating = ______ W

3. *Input characteristics*:

S. No.	V_{CE} = ______ V		V_{CE} = ______ V	
	v_{BE} (in V)	i_B (in μA)	v_{BE} (in V)	i_B (in μA)
1.				
2.				
3.				

4. *Output characteristics*:

S. No.	$I_B = 0$		$I_B = 10$ μA		$I_B = 20$ μA		$I_B = 30$ μA	
	v_{CE}(V)	i_C(mA)	v_{CE}(V)	i_C(mA)	v_{CE}(V)	i_C(mA)	v_{CB}(V)	i_C(mA)
1.								
2.								
3.								

Calculations

1. Input dynamic resistance,

$$r_i = \left.\frac{\Delta v_{BE}}{\Delta i_B}\right|_{V_{CE}=\_\_\_\text{V}} = \_\_\_\_\_\_ = \_\_\_\_\_\_ \text{ k}\Omega$$

2. Output ac resistance,

$$r_o = \left.\frac{\Delta v_{CE}}{\Delta I_C}\right|_{I_B=\_\_\_\text{mA}} = \_\_\_\_\_\_ = \_\_\_\_\_\_ \text{ k}\Omega$$

3. DC current gain,

$$\beta_{dc} = \left.\frac{I_C}{I_B}\right|_{V_{CE}=\_\_\_\text{V}} = \_\_\_\_\_\_ = \_\_\_\_\_\_$$

4. AC current gain,

$$\beta = \left.\frac{\Delta I_C}{\Delta I_B}\right|_{V_{CE}=\_\_\_\text{V}} = \_\_\_\_\_\_ = \_\_\_\_\_\_$$

Results

1. Input and output characteristics are plotted on the graph.
2. The parameters of the transistor in CE mode are given below:

Parameters	*Value determined*
1. r_i	______ Ω
2. r_o	______ kΩ
3. β_{dc}	______
4. β	______

6 UNIT

FIELD EFFECT TRANSISTORS (FETs)

"History always changes and everything that we owned 10 years ago is becoming obsolete electronically."

Chris Watts (1965–present)

American Inventor, Businessman, Award Winning Filmmaker and Visual Effects Supervisor

After completing this unit, students will be able to:

- explain the structure and working of a junction field-effect transistor JFET
- draw the circuit symbols of an N-channel and P-channel JFET
- draw the output characteristics of a JFET
- define JFET parameters
- explain the structure and working of two types of MOSFETs, depletion-type and enhancement type
- draw the circuit symbols of an N-channel and P-channel DE MOSFET and EN MOSFET
- draw the output characteristics of the two types of MOSFETs
- compare the three types of transistors—JFET, MOSFET and BJT
- explain how MOSFETs are better than JFETs
- explain the structure of complementary MOS (CMOS) transistor
- state applications of FETs

6.1 INTRODUCTION

The FET was developed in the early 1960s. The FET operates under principles which are completely different from those of the BJT. The name *'field effect'* is derived from the fact that the current flow in the device is controlled by an electric field set up by an externally applied voltage.

There are two main types of FETs:

1. Junction field-effect transistor (JFET)
2. Metal-oxide-semiconductor field-effect transistor (MOSFET)

Both types are fabricated as discrete components and as components of integrated circuits (ICs). Compared to BJTs, MOS transistors can be made quite small (that is, occupying a small silicon area on the IC chip). Furthermore, digital logic memory functions can be implemented with circuits that exclusively use MOSFETs (that is, no resistors or diodes are needed). For these reasons, most of the very-large-scale-integrated (VLSI) circuits are made at present using MOS technology.

MOSFETs are again of two types:

1. Depletion-type MOSFET (DE MOSFET)
2. Enhancement-type MOSFET (EN MOSFET)

In many respects, a DE MOSFET is similar to a JFET. Whatever be the construction of an FET (whether JFET, or DE MOSFET, or EN MOSFET), it can have either (*i*) N-type channel, or (*ii*) P-type channel.

In a BJT, current is conducted by charge carriers of both the polarity (i.e., electrons and holes). That is the reason why the conventional transistor is called ***bipolar***. In contrast, the current in an FET is conducted by the majority charge carriers in the channel (i.e., by electrons in N-channel, and by hole in P-channel). Since the conduction is performed by charge carriers of only one polarity, FETs are called ***unipolar transistors.***

6.2 JUNCTION FIELD-EFFECT TRANSISTOR (JFET)

6.2.1 Structure of a Junction Field-Effect Transistor (JFET)

A JFET can be of N-*channel* type or of P-*channel* type. (The meaning of *channel* will be made clear later in the section.) We shall describe the structure of an N-channel JFET. The structure of a P-channel JFET is similar to that of an N-channel JFET, except that in its structure, N-type is replaced by P-type and vice versa.

In its simplest form, the structure of an N-channel JFET starts with nothing more than a bar of N-type silicon. This bar behaves like a resistor between its two terminals, called *source* and *drain* (Fig. 6.1*a*). We introduce heavily doped P-type regions on either side of the bar. These P regions are called *gates* (Fig. 6.1*b*). Usually, the two gates are connected together (Fig. 6.1*c*). The gate terminal is analogous to the base of a BJT. This is used to control the current flow from source to drain. Thus, source and drain terminals are analogous to emitter and collector terminals respectively, of a BJT.

In Fig. 6.1*d*, the bar of the JFET has been placed vertically. The circuit symbol of N-channel JFET is shown in Fig. 6.1*e*. Note that the arrow is put in the gate terminal (and not in the source terminal, though source is analogous to emitter in a BJT). The gate arrow points into the JFET. (In a P-channel JFET, the gate arrow would point out of the JFET).

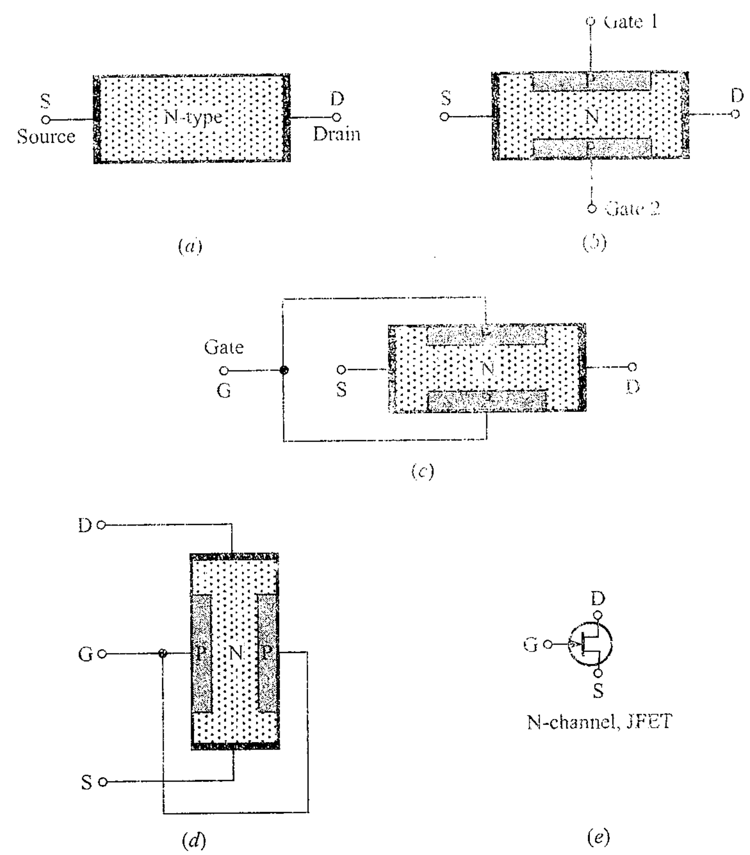


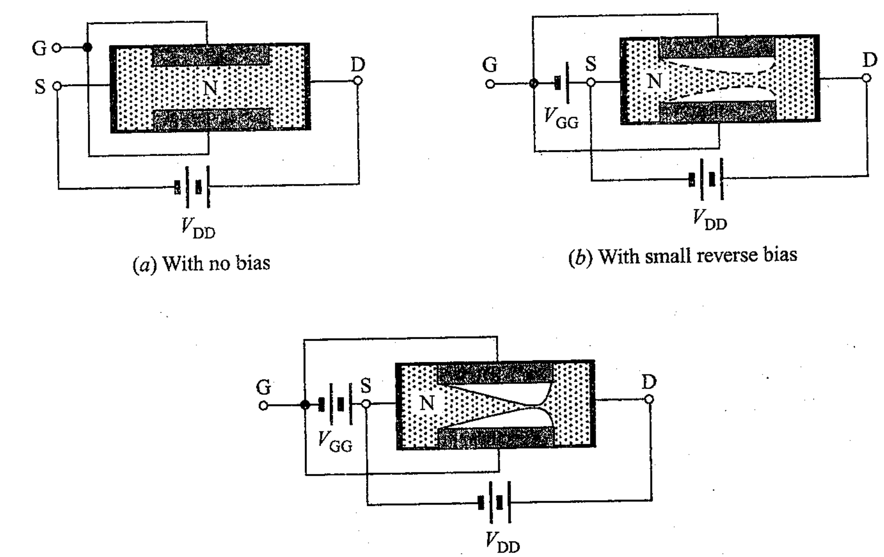
Fig. 6.1 *Junction field-effect transistor (N-channel type)*

Let us now see why the N-type bar is called a channel. Normally, we operate an N-channel JFET by applying positive voltage to the drain with respect to the source (Fig. 6.2*a*). Due to this voltage, the majority carriers in the bar (electrons in this case) start flowing from the source to the drain. This flow of electrons makes the drain current I_D. The current I_D is analogous to the collector current I_C in a BJT. The electrons in the bar have to pass through the space between the two P regions. As we shall see, the width of this space between the P regions can be controlled by varying the gate voltage. That is why this space is called a channel.

To see how the width of the channel changes by varying the gate voltage, let us consider Fig. 6.2*b*. Here we have applied a small reverse bias to the gate. Because of the reverse bias, the width of the depletion increases. Since the N-type bar is lightly doped compared to the P regions, the depletion region extends more into the N-type bar. This reduces the width of the channel. Recall that the depletion regions do not contain any charge carriers. The electrons have to pass through the channel of

reduced width. Reduction in the width of the channel (the conductive portion of the bar) increases its resistance. This reduces the drain current I_D.

See Fig. 6.2*b* carefully. There is one important point about the channel shape. It is narrower at the drain end. This happens because the amount of reverse bias is not same throughout the length of the PN-junction. When current flows through the bar, a potential drop occurs across its length. As a result, the reverse bias between the gate and the drain end of the bar is more than that between the gate and the source end of the bar. The width of the depletion region is more at the drain end than at the source end. As a result, the channel becomes narrower at the drain end.



(*a*) With no bias

(*b*) With small reverse bias

(*c*) Pinch-off occurs at large reverse bias

Fig. 6.2 *Effect of gate-source voltage on the channel*

Let us see what happens if the reverse gate-bias is increased further. The channel becomes narrower at the drain end and the drain current further reduces. If the reverse bias is made sufficiently large, the depletion regions tends to extend into the channel and meet. This *pinches off* the current flow (Fig. 6.2*c*). The gate-source voltage at which pinch-off occurs is called *pinch-off voltage* V_P.

You may think that the channel completely closes at the drain end when the gate-source voltage reaches the pinch-off value. But in practice it does not happen, simply because it cannot happen. Suppose, if it were possible, the channel completely closes at the drain end. The drain current would then reduce to zero. As a result, there would be no voltage drop along the length of the channel. The amount of reverse bias would become uniformly same throughout the length. The wedge shaped depletion region would try to become straight (rectangular shaped). The channel would then open at the drain end. The drain current flows.

When the gate-source voltage reaches the pinch-off value, the channel width reduces to a constant minimum value. The drain current flows through this constricted channel.

Some important terminology regarding a JFET:

1. **Source** The source is the terminal through which the majority carriers (electrons in case of N-channel JFET, and holes in case of P-channel JFET) enter the bar.
2. **Drain** The drain is the terminal through which the majority carriers leave the bar.
3. **Gate** On both sides of the N-type bar, heavily doped P regions are formed. These regions are called *gates*. Usually, the two gates are joined together to form a single gate.
4. **Channel** The region between the source and drain, sandwiched between the two gates, is called *channel*. The majority carriers move from source to drain through this channel.

6.2.2 JFET Characteristics

As a BJT has static collector characteristics, so does a JFET have static drain characteristics. Such characteristics are shown in Fig. 6.3. For each curve, the gate-to-source voltage V_{GS} is constant. Each curve shows the variation of drain current i_D versus drain-to-source voltage v_{DS}.

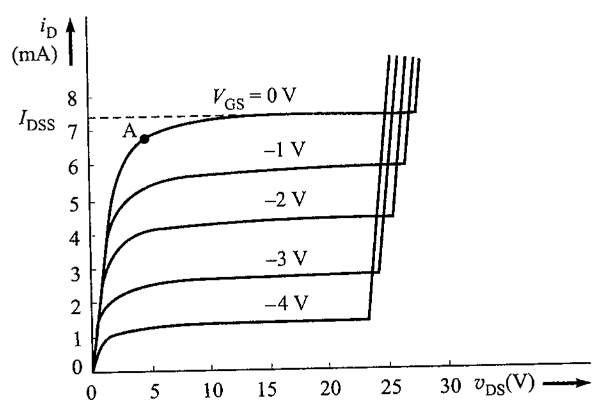


Fig. 6.3 *Typical drain characteristics of an N-channel JFET*

Let us consider first, the curve for zero gate bias. For this curve, $V_{GS} = 0$. When v_{DS} is zero, the channel is entirely open. But the drain current is zero, because the drain terminal does not have any attractive force for the majority carriers. For small applied voltage v_{DS}, the bar acts as a simple resistor. Current i_D increases linearly with voltage v_{DS}. This region (to the left of point A) of the curve is called *ohmic region*, because the bar acts as an ohmic resistor.

Ohmic voltage drop is caused in the bar due to the flow of current i_D. This voltage drop along the length of the channel reverse biases the gate junction. The reserve biasing of the gate junction is not uniform throughout. The reverse bias is more at the drain end than at the source end of the channel. So, as we start increasing v_{DS}, the channel starts constricting more at the drain end. The channel is eventually *pinched off*. The current i_D no longer increases with the increase in v_{DS}. It approaches a constant saturation value. The voltage v_{DS} at which the channel is "pinched off" (that is, all the free charges from the channel are removed), is called *pinch-off voltage*, V_P. Note that the voltage V_P is not sharply defined on the curve. The region of the curve to the right of point A is called *pinch-off region*.

A special significance is attached to the drain current in the pinch-off region when $V_{GS} = 0$. It is given the symbol I_{DSS}. It signifies the drain source current at pinch-off, when the gate is shorted to the source. It is measured well into the pinch-off region. In this case, $I_{DSS} = 7.4$ mA.

Further increase in voltage v_{DS} increases the reverse bias across the gate junction. Eventually, at high v_{DS} breakdown of the gate junction occurs. The drain current I_D shoots to a high value. Of course, when we use a JFET in a circuit, we avoid the gate junction breakdown.

If the gate reverse bias is increased (say, $V_{GS} = -1$ V), the curve shifts downward. The pinch-off occurs for smaller value of v_{DS}. The maximum saturation drain current is also smaller, because the conducting channel now becomes narrower.

For an increased reverse bias at the gate, the avalanche breakdown of the gate junction occurs at lower value of v_{DS}. This happens because the effective bias at the gate junction (at the drain end) is the voltage V_{GS} plus voltage V_{DS}. The greater the value of V_{GS}, the lower the value of V_{DS} required for the junction to breakdown.

6.2.3 JFET Parameters

An important parameter of a JFET is the current I_{DSS}. It signifies the drain saturation current when $V_{GS} = 0$. It is specified by the manufacturer. Besides this, there are the following three important parameters of a JFET.

1. **Dynamic drain resistance (r_d)** Dynamic drain resistance at an operating point is defined as the ratio of small change in drain voltage to the small change in drain current, keeping the gate voltage constant. That is

$$r_d = \left.\frac{\Delta v_{DS}}{\Delta i_D}\right|_{V_{GS}=\text{const.}} \tag{6.1}$$

Typically, r_d is about 400 kΩ.

2. **Mutual conductance or transconductance (g_m)** The mutual conductance, at an operating point, is defined as the ratio of small change in drain current to the small change in gate voltage, keeping the drain voltage constant. That is

$$g_m = \left.\frac{\Delta i_D}{\Delta v_{GS}}\right|_{V_{DS}=\text{const.}} \tag{6.2}$$

It is measured in siemens (S). Typically, its value ranges from 150 μS to 250 μS.

3. **Amplification factor (μ)** It is defined as the ratio of small change in drain voltage to the small change in gate voltage, when current I_D is kept constant. That is

$$\mu = \left. \frac{\Delta v_{DS}}{\Delta v_{GS}} \right|_{I_D = \text{const.}} \tag{6.3}$$

Since μ is a ratio of two voltages, it does not have any units. The amplification factor of a JFET can be as high as 100.

The above three parameters of a JFET are related as

$$\mu = r_d g_m \tag{6.4}$$

Thus, if any two parameters are known, the third can be computed.

Example 6.1 For a JFET type BFW10 (made by BEL, Bangalore), the typical values of amplification factor and transconductance are specified as 80 and 200 μS, respectively. Calculate the dynamic drain resistance of this JFET.

Solution: The three parameters of a JFET are related by the formula,

$$\mu = r_d g_m$$

Here, $\mu = 80$, and $g_m = 200\ \mu\text{S} = 200 \times 10^{-6}$ S. Therefore, the dynamic drain resistance is given as

$$r_d = \frac{\mu}{g_m} = \frac{80}{200 \times 10^{-6}} = 4 \times 10^5\ \Omega = \mathbf{400\ k\Omega}$$

6.3 METAL-OXIDE SEMICONDUCTOR FET (MOSFET)

The Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET) is similar to the JFET in many ways. A MOSFET too has *drain* (D), *source* (S) and *gate* (G) terminals. Like the JFET, the channel conductivity of a MOSFET is also controlled by gate-to-source voltage, v_{GS}.

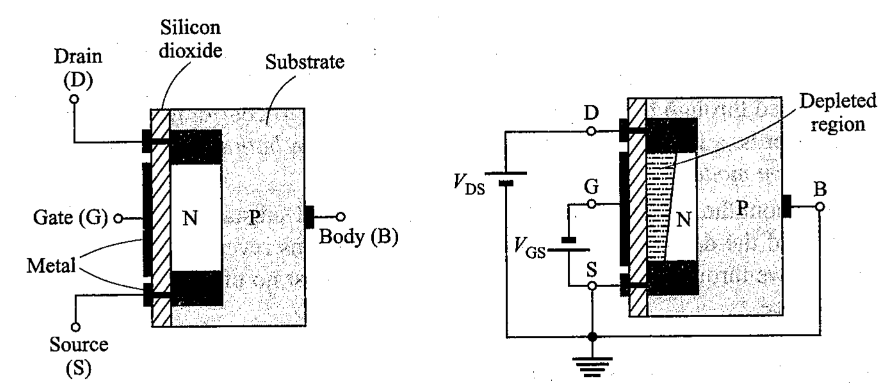
A MOSFET differs from a JFET in the sense that its gate terminal is electrically *insulated* from its channel region. For this reason, a MOSFET is also called Insulated-Gate Field-Effect Transistor (IGFET). It is because of this reason that the gate current in a MOSFET is extremely small ($\approx 10^{-15}$ A).

There are two types of MOSFETs: *depletion* type (DE MOSFET) and *enhancement* type (EN MOSFET). These names are derived from the two different ways in which the conductivity of the channel is changed by varying v_{GS}.

6.4 DEPLETION-TYPE MOSFET (DE MOSFET)

6.4.1 Structure of DE MOSFET

Figure 6.4*a* shows the structure of an N-channel depletion-type MOSFET. A block of high-resistance, P-type silicon forms the *substrate* or the *body* (B). It provides physical support to the device. Two heavily doped N-type wells (or pockets) are created on the surface of the block. These are labelled as N^+ in the figure. In between these two wells, there is a lightly doped N region which makes the channel. A thin layer of an insulating material—silicon oxide (SiO_2)—is deposited along the surface. Two metal contacts penetrate the silicon oxide layer to reach the two N^+ wells. These make the *source* (S) and the *drain* (D) terminals of the device.



(*a*) Structure (*b*) Working principle

Fig. 6.4 *N-channel depletion-type MOSFET*

Unlike a JFET, there is no PN-junction formed between the gate and the channel. Here, the silicon oxide layer insulates the gate from the channel. Also note that while going form gate to channel, we come across ***M***etal, ***O***xide and ***S***emiconductor, in that sequence. Hence the name MOSFET.

It is obvious that a P-channel DE MOSFET is made from a lightly doped N substrate. The drain and source are heavily doped P^+ wells. In between these well, there is a lightly doped P region which makes the channel.

6.4.2 Working Principle of DE MOSFET

As shown in Fig. 6.4*b*, a voltage source V_{DS} is connected between the drain and source, making the drain (D) positive with respect to the source (S). The body (B) is usually connected to the source (S), as shown in the figure. We make the gate negative with respect to the source by connecting a battery V_{GS}.

The gate being at negative voltage with respect to the body, an electric field (having direction from channel to the gate) is created in the channel. This field repels the electrons away from the portion of the channel near the SiO_2 layer. This portion is therefore *depleted* of the carriers. The channel width is effectively reduced. The narrower the channel, the greater is its resistance, and the smaller is the current from drain to source. It means that by varying the gate voltage V_{GS}, the drain current I_D can be controlled. This is quite similar to an N-channel JFET. What is the difference between the two? The main difference is that in a MOSFET the channel width is controlled by the action of the electric field, whereas in a JFET the channel width is controlled by the size of the depletion region of the reverse-biased PN-junction.

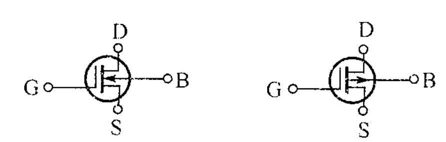
Note that a DE MOSFET has no PN-junction at the gate. Therefore, there is no risk of making the gate current I_G large due to PN-junction becoming forward biased, if voltage V_{GS} is made positive. That is, the gate current remains negligibly small, even if voltage V_{GS} is made positive.

When V_{GS} is made positive, the field thus produced attracts more electrons into the channel from N^+ region. This increases or *enhances* the conductivity of the channel. As a result, the drain current increases. Thus, in a DE MOSFET the gate voltage can be varied through both negative and positive values to control the drain current. In other words, a DE MOSFET is capable of *working in both the depletion and the enhancement mode.*

What about the PN-junction between the channel and substrate? Will it effect the operation of the device? As this junction always remains reverse biased, very little current flows through the substrate. Hence, it has almost no effect on the operation of the device.

6.4.3 Circuit Symbol of DE MOSFET

Figure 6.5 shows the circuit symbols of two types of DE MOSFET. The thick vertical line represents the channel. An arrow is drawn on the body terminal. It points from P to N. Thus, for N-channel the arrow is inward; for P-channel it is outward. Note that the gate terminal is connected to a line which is separated from the solid thick line representing the channel. This emphasises that the gate is insulated from the channel.



(*a*) N-channel (*b*) P-channel

Fig. 6.5 *Circuit symbols of DE MOSFETs*

6.4.4 Output Characteristics of DE MOSFET

Figure 6.6 shows the output characteristics of an N-channel DE MOSFET. These characteristics appear similar to those of an N-channel JFET shown in Fig. 6.3. This is as expected, since the working of the two devices is quite similar. The only difference is that here the voltage V_{GS} is shown both negative as well as positive. When V_{GS}

is negative, the device works in *depletion mode*; when positive, it works in *enhancement mode*.

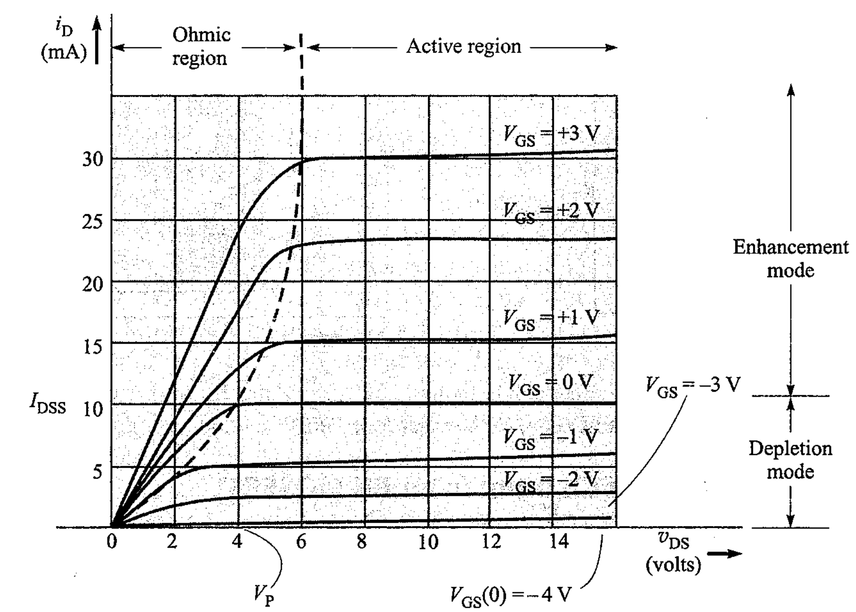


Fig. 6.6 *Output characteristics of an N-channel DE MOSFET*

Consider the characteristic curve for $V_{GS} = 0$. For small value of v_{DS}, the N material between the drain and source acts as a simple resistance. The current i_D increases linearly with v_{DS}, as per Ohm's law. As v_{DS} is increased further, the channel becomes narrower and the current i_D begins to level off. When v_{DS} becomes equal to *pinch off voltage* (V_p), the channel is pinched-off. The current saturates at I_{DSS}. In Fig. 6.6, $V_p = 4$ V, and $I_{DSS} = 10$ mA.

When V_{GS} is made negative, the pinch-off condition occurs at a lower value of v_{DS}. The drain current saturates at a lower value. If V_{GS} is made sufficiently negative to deplete the entire channel, the drain current is completely cut off.

When V_{GS} is made positive, the device works in enhancement mode. The resistance of the channel reduces, and the drain current i_D increases. The pinch-off occurs at a larger value of v_{DS}. Also, the current i_D saturates at a larger value than I_{DSS}. The parabolic dashed line passes through the pinch-off points for different curves.

6.5 ENHANCEMENT-TYPE MOSFET (EN MOSFET)

6.5.1 Structure of EN MOSFET

The structure of EN MOSFET is similar to that of DE MOSFET. The only difference is that there is no N-type material between the drain and source. Instead the

P-type substrate extends all the way to the SiO_2 layer adjacent to the gate as shown in Fig. 6.7*a*.

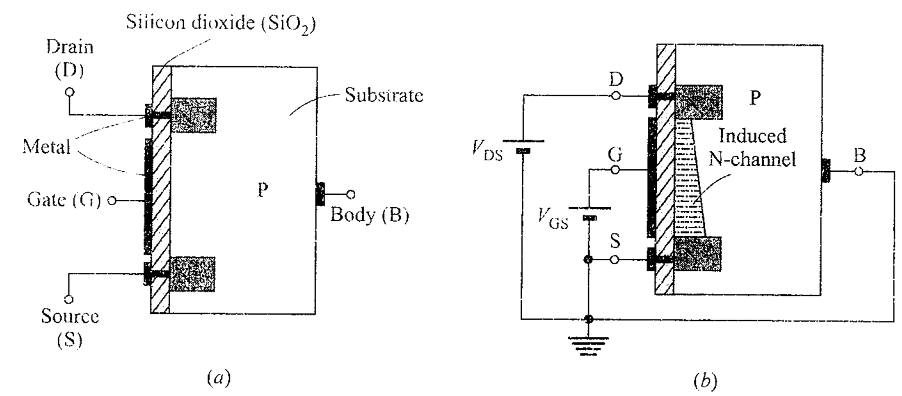


Fig. 6.7 *N-channel EN MOSFET*

6.5.2 Formation of Channel in an EN MOSFET

Figure 6.7*b* shows the normal connections for an N-channel EN MOSFET. As in the DE MOSFET, the substrate or the body (B) is connected to the source. The voltage V_{GS} is connected so that *gate is positive with respect to the source.* The gate repels holes from the region under it. This leaves behind a depletion region containing negative immobile ions.

In addition, the positive gate attracts electrons from the N^+ drain and N^+ source where they are available in plenty. This, in effect, creates an N region near the surface of the substrate under the gate. The drain and source are connected by this N-region. If a positive voltage V_{DS} is applied a current flows from drain to source through this N-region. This *induced* N-region thus acts as a *channel* for the current flow. The MOSFET of Fig. 6.7 is called **N-channel MOSFET** or simply **NMOS transistor**. Note that the NMOS transistor is formed in a P-type substrate. As seen above, the channel is created by inverting the substrate from P-type to N-type. Hence, the induced channel is called the ***inversion layer.***

The induced N-channel in Fig. 6.7*b* does not become sufficiently conductive to allow the drain current to flow until V_{GS} reaches a certain value. This minimum value of V_{GS} at which sufficient number of electrons accumulates in the channel is called ***threshold voltage*** V_T. The value of V_T is typically in the range of 1 to 3 V.

6.5.3 Working Principle of EN MOSFET

Once the channel is formed, on applying a small voltage v_{DS} (Fig. 6.7*b*) causes free electrons to travel from source to drain, constituting a small current i_D. The magnitude of i_D depends on the density of electrons in the channel, which in turn depends on V_{GS}. At $v_{DS} = V_T$, the channel is just induced. The current i_D is negligibly small. As v_{DS} increases beyond V_T, the conductivity of the channel increases and as a result

the current I_D also increases. In fact, the conductivity of the channel is proportional to the ***excess gate voltage*** ($v_{GS} - V_T$).

Suppose that $V_T = 2$ V and V_{GS} is set at 10 V. As v_{DS} is gradually increased above zero volt, the current i_D also increases gradually. The channel behaves like a resistor. The current i_D increases linearly with voltage v_{DS} (Ohm's law), as shown in Fig. 6.8.

As we continue to increase v_{DS}, the channel starts becoming narrower at the drain end, as shown in Fig. 6.7*b*. This happens because the gate-to-drain voltage V_{GD} reduces as v_{DS} increases. This results in reduced field at the drain end. For example, if $V_{GS} = 10$ V and $v_{DS} = 3$ V, the voltage v_{GD} becomes $10 - 3 = 7$ V. But when V_{DS} is increased to 4 V, the voltage V_{GD} reduces to $10 - 4 = 6$ V. Thus, the field at drain gets reduced. The channel width decreases. As a result, the resistance of the channel begins to increase and the drain current i_D begins to level off. Eventually, when v_{DS} reaches 8 V, the voltage v_{GD} becomes $10 - 8 = 2$ V (same as V_T). The channel width at the drain end reduces to almost zero. That is, the channel is pinched off. If we increase v_{DS} beyond this point, the channel shape remains almost same and the current through the channel remains constant. The MOSFET enters into ***saturation*** or ***active region***.

6.5.4 Output Characteristics of EN MOSFET

Figure 6.8 shows the output characteristics of an N-channel EN MOSFET. These characteristics are similar to those of an N-channel JFET. The only difference is that in EN MOSFET we keep V_{GS} positive; whereas in JFET, V_{GS} is kept negative. Also, note that all values of V_{GS} are positive. It means that an EN MOSFET can be operated only in the enhancement mode.

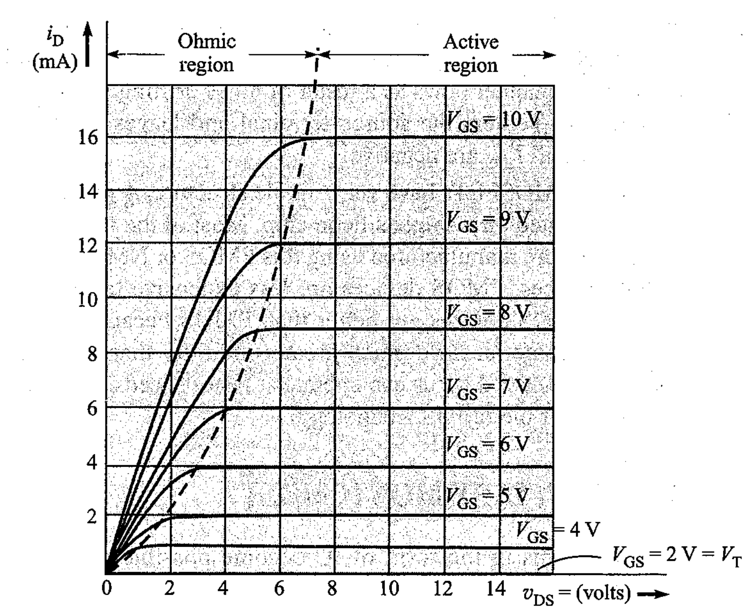


Fig. 6.8 *Output characteristics of an N-channel EN MOSFET*

Note that when V_{GS} is reduced to threshold voltage $V_T = 2$ V, the drain current i_D reduces to zero for all values of v_{DS}. The dashed parabolic line joins the saturation voltages for different curves. The region to the left of this line is called the *voltage-controlled-resistance region* or *ohmic region*. The region to the right is called *saturation* or *active region*.

6.5.5 Circuit Symbol of EN MOSFET

Figure 6.9 shows the circuit symbols of N-channel and P-channel EN MOSFET. The symbol is very descriptive. The vertical solid line denotes the gate electrode. The vertical broken thick line denotes channel. It is shown broken to indicate that the channel is induced rather than being an inherent part of the structure.

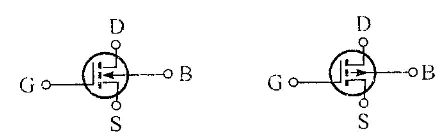


Fig. 6.9 *Circuit symbols of EN MOSFETs*

The spacing between the gate line and the channel represents the insulating SiO_2 layer. The arrowhead on the line representing substrate (body) points from P to N. It indicates whether the channel is N-type or P-type.

6.5.6 The P-Channel MOSFET

A P-channel MOSFET (also called **PMOS transistor)** is complement of NMOS transistor. It is fabricated on an N-type substrate with P^+ regions for the drain and source. The current in the channel flows as a result of holes drifting from the source to the drain. The current I_D enters the source terminal and leaves from the drain terminal. Here, both V_{GS} and V_{DS} are negative.

The enhancement-type MOSFETs have very simple structure. A great number of such devices can be fabricated on a single silicon chip. Most of the ***very-large-scale integrated*** (VLSI) circuits are manufactured using this PMOS or NMOS technology. Compared to NMOS devices, PMOS devices are less expensive to produce. But, performance-wise the NMOS devices are better than PMOS because the majority carriers (electrons) in N material have much greater mobility than the holes in P material. As a result, the NMOS devices can operate at faster speeds. Hence, NMOS technology is preferred over the PMOS technology.

6.6 COMPLEMENTARY MOS (CMOS)

Power-saving circuit-designs in the form of IC become possible, if we use both NMOS and PMOS together embedded in the same substrate. Such a complementary

MOS (or CMOS) technology is most widely used nowadays. Both analogue and digital circuits are manufactured in the form of ICs using CMOS technology.

Figure 6.10 shows a cross section of a CMOS chip illustrating how NMOS and PMOS transistors are fabricated. The NMOS transistor is implemented directly in P-type substrate. To fabricated PMOS transistor, first an N region is specially created. This region is called N ***well*** or ***tub***. The two devices are isolated from each other by a thick region of insulating SiO_2.

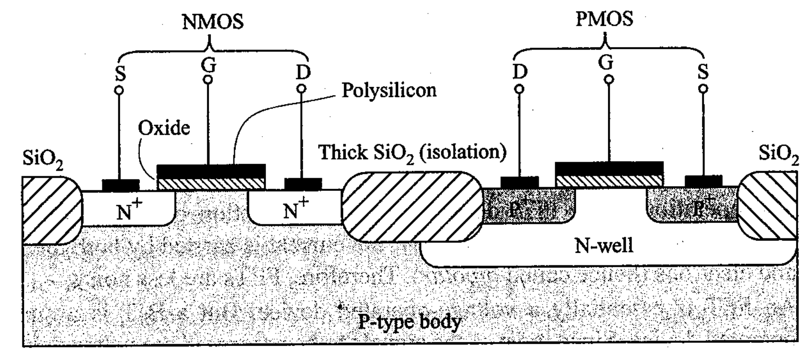


Fig. 6.10 *Cross-section of a CMOS integrated circuit*

Note that the arrangement shown in Fig. 6.10 could also be fabricated in an N-type substrate. The NMOS transistor is then made in a P well.

6.6.1 VMOS Technology

This technology is a little modification of MOS technology. The V-shaped groove penetrates alternate N and P layers, as shown in Fig. 6.11 for an N-channel device. The term VMOS is derived from the V appearance of the cross-sectional view. The length of the induced N channel is determined by the thickness of the P layer. This technique saves space on the chip surface, thereby enabling many more devices to be fabricated on a single chip. The channel can be made longer by simply making P layer thicker.

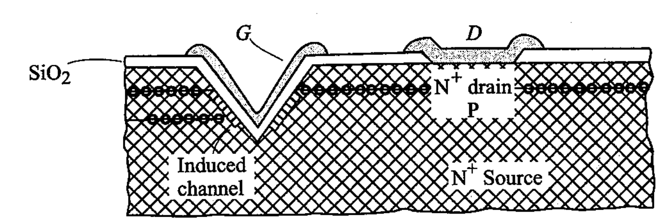


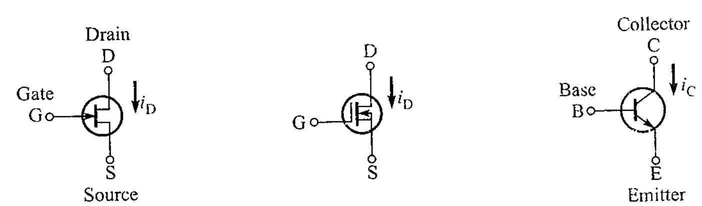
Fig. 6.11 *VMOS structure for an N-channel transistor*

VMOS transistors have greater current-handling capabilities. Hence, these transistors find use in power-amplifier applications.

6.7 COMPARISON OF JFET, MOSFET AND BJT

As shown in Fig. 6.12, an FET is quite analogous to a BJT. Any circuit, whether analogue or digital, whether discrete or integrated, can be designed using either FETs or BJTs or their combination. However, in many ways the FETs prove advantageous over BJTs.

1. Since the PN-junction in the input circuit of a JFET (i.e., gate to source) is reverse biased, this device has high input impedance of the order of 10 MΩ. MOSFETs are even better. These devices have no PN-junction in the input circuit. The gate is insulated from the channel by a SiO_2 layer. This offers very high input impedance (more than 10^{12} Ω). On the contrary, the PN-junction in the input circuit of a BJT (i.e., emitter to base) is forward biased, giving quite low input impedance (of the order of 1 kΩ).
2. The operation of an FET depends on the conduction of majority carriers (hence called *unipolar*). But in a BJT, the current is carried by both the holes and electrons (hence called *bipolar*). Therefore, FETs are less noisy.
3. An FET is essentially a voltage-operated device. But a BJT is a current-operated device. Since most of the signals to be processed are voltage signals, FETs are better than BJTs.
4. Power gain of an FET is much higher than that of a BJT. Hence, there is no need of a driver stage if an FET is used as power amplifier.
5. In digital circuits, power requirement is an important criterion. FET digital circuits need much less power compared to BJT circuits.
6. There is no risk of thermal runaway in FET circuits, as the drain current of an FET decreases with increase in temperature.
7. MOSFETs are simpler and less expensive to fabricate than the BJTs.
8. A MOS device requires much smaller area on the silicon chip than a BJT. This allows a greater number of devices to be accommodated on a single silicon chip in an IC.



(*a*) N-channel JFET (*b*) N-channel DE MOSFET (*c*) NPN BJT

Fig. 6.12 *Analogy between FET and BJT*

6.7.1 How MOSFETs are Better than JFETs

A JFET is a depletion-type device. Its characteristics are similar to those of the depletion-type MOSFET. However, there are two important differences because of which the MOSFETs are more widely used than JFETs.

1. A DE MOSFET can be operated in the enhancement mode (by simply applying a positive voltage V_{GS}, if the device is N-channel). But it is impossible to operate a JFET in enhancement mode. If we attempt to apply positive voltage V_{GS} (in case of N-channel JFET), the gate-channel junction becomes forward biased. As a result, the gate ceases to control the channel.
2. Even if a JFET is operated with a reverse bias, the gate current (caused by the flow of minority carriers) is much larger than that in a comparable MOSFET.

6.7.2 Applications of FETs

1. FETs can be used in circuits of amplifier, oscillator, etc.
2. FETs can be used as switches in digital circuits.
3. FETs can be used as analogue switches in circuits such as sample and hold, amplitude modulation, ADC (analogue-to-digital) or DAC (digital-to-analogue) converters.
4. CMOS is an excellent device for use in an inverter circuit, as it needs almost no power.
5. FETs, when operated in ohmic region, can be used as voltage variable resistors (VVR).

• Review Questions •

1. Sketch the basic structure of N-channel JFET.
2. Draw the circuit symbols of (*a*) an N-channel JFET and (*b*) a P-channel JFET.
3. Show the biasing arrangement of an N-channel JFET.
4. Draw typical drain characteristics curves of a JFET. Explain the shape of these curves qualitatively.
5. What do you understand by the term 'channel' in a JFET?
6. Define all the important parameters of a JFET.
7. How does a MOSFET differ from a JFET?
8. What is the basic difference between an FET and a BJT?
9. What are the two types of MOSFETs? How do they differ in their structure?
10. Briefly explain the working of a depletion-type MOSFET.
11. Briefly explain the working of an enhancement-type MOSFET.
12. Draw the circuit symbols of (*a*) a P-channel EN MOSFET and (*b*) an N-channel DE MOSFET. Highlight the differences between the two.
13. Why is the NMOS technology considered better than the PMOS?
14. Briefly explain the structure of a CMOS.
15. How does the VMOS structure save space on the chip?
16. Why are the FETs considered better than the BJTs?
17. Why are the MOSFETs considered a better choice than the JFETs or BJTs in making ICs?

Objective-Type Questions

I. Below are some incomplete statements. Four alternatives are provided for each. Choose the alternative that completes the statement correctly.

1. A junction field-effect transistor (JFET)
   (a) has three PN-junctions
   (b) incorporates a forward-biased junction
   (c) depends on the variation of a magnetic field for its operation
   (d) depends on the variation of the depletion-layer width with reverse v_{GS} in its operation
2. The operation of a JFET involves
   (a) a flow of minority carriers
   (b) a flow of majority carriers
   (c) recombination
   (d) negative resistance
3. A field-effect transistor (FET)
   (a) uses a high-concentration emitter junction
   (b) uses a forward-biased PN-junction
   (c) has a very high input resistance
   (d) depends on minority-carrier flow
4. Which one of the following has the highest input impedance?
   (a) NPN transistor in CB mode
   (b) NPN transistor in CE mode
   (c) N-channel JFET
   (d) P-channel enhancement-type MOSFET
5. Which one of the following is a unipolar device?
   (a) N-channel MOSFET
   (b) PNP Si transistor
   (c) NPN Ge transistor
   (d) PN junction diode
6. Which one of the following is the most likely value of threshold voltage of an N-channel EN MOSFETY?
   (a) $V_T = -3$ V
   (b) $V_T = -1.5$ V
   (c) $V_T = +2$ V
   (d) $V_T = +10$ V

II. Some statements are written below. Write whether they are TRUE or FALSE, in the space provided against each.

1. The input resistance of a JFET is much greater than that of a DE MOSFET. (True/False)

2. A depletion-type MOSFET can work in both the depletion mode as well as in the enhancement mode. (True/False)
3. FETs are nowadays widely used in both analogue and digital applications. (True/False)
4. MOSFETs are also known as IGFETs. (True/False)
5. Various forms of FETs can be produced in both N-channel and P-channel versions. (True/False)
6. For an enhancement-type NMOS transistor, the gate voltage is always kept negative in its normal range of operation. (True/False)
7. The spacing between the gate line and the channel line in the circuit symbol of a MOSFET represents a silicon dioxide insulation layer. (True/False)
8. In an N-channel JFET, for $V_{GS} = 0$ V, if V_{DS} is gradually increased from zero, at a particular value of V_{DS} the channel is pinched off and the current I_D reduces to zero. (True/False)
9. For a JFET, the shorted-gate drain current I_{DSS} is the drain current when the gate is shorted to the drain and V_{DS} is equal to or more than the pinch-off voltage. (True/False)

Answers

I.	1. (d)	2. (b)	3. (c)	4. (d)	5. (a)	6. (c)
II.	1. F	2. T	3. T	4. T	5. T	6. F
	7. T	8. F	9. F			

Experimental Exercise 6.1

Title JFET characteristics.

Objectives To

1. trace the given circuit of JFET;
2. plot the static drain characteristics of JFET;
3. calculate the JFET parameters (drain dynamic resistance r_d, mutual conductance g_m, and amplification factor μ) at a given operating point.

Apparatus Required Experimental board, transistor (or IC) power supply, milliammeter (0 to 25 mA), two electronic multimeters.

Circuit Diagram The circuit diagram is shown in Fig. E. 6.1.1

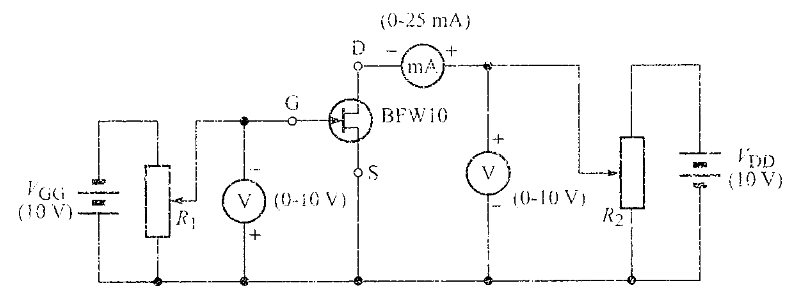


Fig. E. 6.1.1

Brief Theory Like an ordinary junction transistor, a field-effect transistor is also a three terminal device. It is a unipolar device, because its function depends only upon one type of carrier. (The ordinary transistor is bipolar, hence it is called bipolar-junction transistor). Unlike a BJT, a JFET has high input impedance. This is a great advantage.

A field-effect transistor can be either a JFET or MOSFET. Again, a JFET can either have N-channel or P-channel. An N-channel JFET, has an N-type semiconductor bar; the two ends of which make the drain and source terminals. On the two sides of the bar, PN-junctions are made. These P regions make gates. Usually, these two gates are connected together to form a single gate. The gate is given a negative bias with respect to the source. The drain is given positive potential with respect to the source. In case of a P-channel JFET, the terminals of all the batteries are reversed.

The important parameters of a JFET are defined below:

1. *Drain dynamic resistance,* $r_d = \left.\dfrac{\Delta v_{DS}}{\Delta I_D}\right|_{V_{GS}=\text{const.}}$
2. *Mutual conductance,* $g_m = \left.\dfrac{\Delta i_D}{\Delta v_{GS}}\right|_{V_{DS}=\text{const.}}$
3. *Amplification factor,* $\mu = \left.\dfrac{\Delta v_{DS}}{\Delta v_{GS}}\right|_{I_D=\text{const.}}$

These parameters are related by the equation,

$$\mu = r_d g_m$$

Procedure

1. Note the type number of JFET connected in the experimental board. See its specifications from the data book. Identify its terminals. Trace the circuit.
2. Make the circuit connections as shown in E. 6.1.1. Use milliammeter and electronic voltmeter in suitable range.
3. First, fix V_{GS} at some value, say 0 V. Increase the drain voltage v_{DS} slowly in steps. Note drain current i_D for each step. Now, change V_{GS} to another value and repeat the above. This way, take readings for 3 to 4 gate-voltage values.

4. Plot the drain characteristics (graph between i_D and v_{DS} for fixed values of V_{GS}).
5. Use the definitions given in brief theory to calculate the JFET parameters, from the characteristics.

Observations

1. *Type number of the JFET* = __________
2. *Information from the data book:*
   (*a*) Maximum drain current rating = __________ mA
   (*b*) Maximum drain voltage rating = __________ V
3. *Drain characteristics:*

S. No.	v_{DS} (*in* V)	*Drain current* i_D (*in* mA)				
		V_{GS} = 0 V	V_{GS} = −1 V	V_{GS} = −2 V	V_{GS} = −3 V	V_{GS} = −4 V
1.						
2.						
3.						

Calculations A suitable operating point is selected, say at V_{DS} = 8 V, V_{GS} = −3 V. At this operating point, the parameters are calculated as follows:

1. $r_d = \left.\dfrac{\Delta v_{DS}}{\Delta i_D}\right|_{V_{GS}=-3\,V}$ = __________ = __________ kΩ
2. $g_m = \left.\dfrac{\Delta i_D}{\Delta V_{GS}}\right|_{V_{DS}=8\,V}$ = __________ = __________ mS
3. $\mu = \left.\dfrac{\Delta v_{DS}}{\Delta v_{GS}}\right|_{I_D=\_\_\_mA}$ = __________ = __________

Results

1. The drain characteristics of the JFET are plotted on the graph.
2. The parameters of JFET determined from the drain characteristics are given below.

	Parameter	*Value determined*
1.	r_d	__________ kΩ
2.	g_m	__________ mS
3.	μ	__________

7 UNIT

TRANSISTOR BIASING AND STABILISATION OF OPERATING POINT

"There is a young and impressionable mind out there that is hungry for information. It has latched on to an electronic tube as its main source of nourishment."

Joan Ganz Cooney (1929–present)
American Television Producer

After completing this unit, students will be able to:

- draw different biasing arrangements in transistor circuits.
- explain with the help of simple equations as to how the operating point is obtained in different biasing circuits.
- calculate the operating point current and voltage in different biasing circuits.
- explain the effect of change in temperature on the operating point in different biasing circuits.
- explain the effect of change in transistor parameters on the operating point in different biasing circuits.
- explain with the help of simple equations as to why the potential divider biasing circuit is the most widely used circuit.

7.1 INTRODUCTION

Transistors are used in different kinds of circuits. These circuits are meant to serve a specific purpose. For example, a circuit may be used to increase the voltage or power level of an electrical signal; such a circuit is called an *amplifier*. There is another class of circuits which generates sine or square wave; such circuits are called *oscillators*. It is very difficult to study all the circuits in which transistors are used.

However, the study of some fundamental aspects of transistor circuits may be helpful, because the knowledge gained during such a study can help us to understand other difficult circuits too. In this chapter, some basic concepts dealing with the dc biasing of transistors are discussed.

7.2 WHY BIAS A TRANSISTOR?

The purpose of dc biasing of a transistor is to obtain a certain dc collector current at a certain dc collector voltage. These values of current and voltage are expressed by the term *operating point* (or *quiescent point*). To obtain the operating point, we make use of some circuits; and these circuits are called "biasing circuits". Of course, while fixing the operating point, it has to be seen that it provides proper dc conditions so that the specific function of the circuit is achieved. The suitability of an operating point for the specific application of the circuit should be seen on the transistor characteristics. In this chapter, we shall discuss the suitability of the operating point in amplifier circuits.

7.3 SELECTION OF OPERATING POINT

In order that the circuit amplifies the signal properly, a judicious selection of the operating point is very necessary. The biasing arrangement should be such as to make the emitter-base junction forward biased and the collector-base junction reverse biased. Under such biasing, the transistor is said to operate in the *active region* of its characteristics. Various transistor ratings are to be kept in view while designing the biasing circuit. These ratings—specified by the manufacturer—limit the range of useful operation of the transistor. $I_{C(max)}$ is the maximum current that can flow through the device and $V_{CE(max)}$ is the maximum voltage that can be applied across it safely. In no case should these current and voltage limits be crossed.

If a transistor is to work as an amplifier, a load resistance R_C must be connected in the collector circuit. Only then the output ac signal voltage can develop across it. The dc load line corresponding to this resistance R_C and a given collector supply V_{CC} is shown in Fig. 7.1. The operating point will necessarily lie somewhere on this load line. Depending upon the base current, the operating point could be either at point A, B, or C. Let us now consider which one of these is the most suitable operating point.

After the dc (or static) conditions are established in the circuit, an ac signal voltage is applied to the input. Due to this voltage, the base current varies from instant to instant. As a result of this, the collector current and the collector voltage also vary with time. That is how an amplified ac signal is available at the output. The variations in collector current and collector voltage corresponding to a given variation (which may be assumed sinusoidal) of base current can be seen on the output characteristics of the transistor.

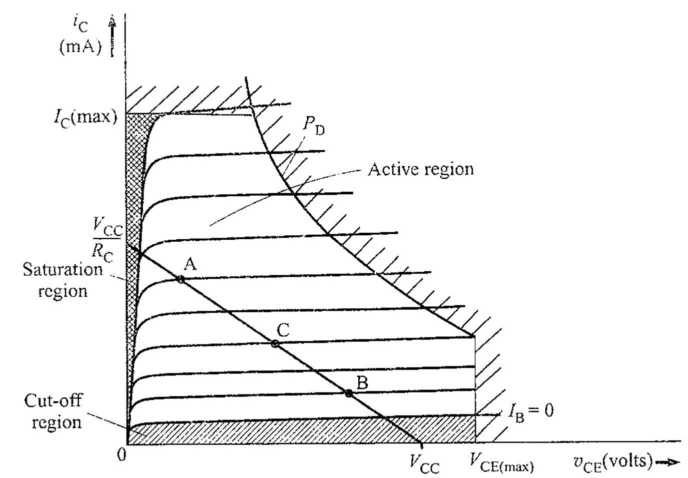


Fig. 7.1 *Output characteristics of a transistor in common-emitter configuration. Maximum current, voltage and power ratings are indicated*

These variations are shown in Figs. 7.2, 7.3 and 7.4 for the operating point A, B and C, respectively. In Fig. 7.2 point A is very near to the saturation region. Even though the base current is varying sinusoidally, the output current (and also output voltage) is seen to be clipped at the positive peaks. This results in *distortion* of the signal. At the positive peaks, the base current varies, but collector current remains constant at saturation value. Thus we see that point A is not a suitable operating point.

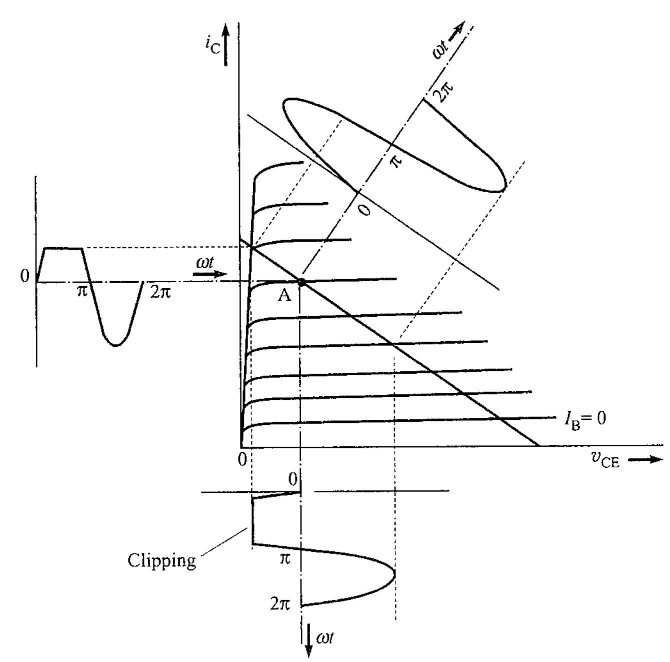


Fig. 7.2 *Operating point near saturation region gives clipping at the positive peaks*

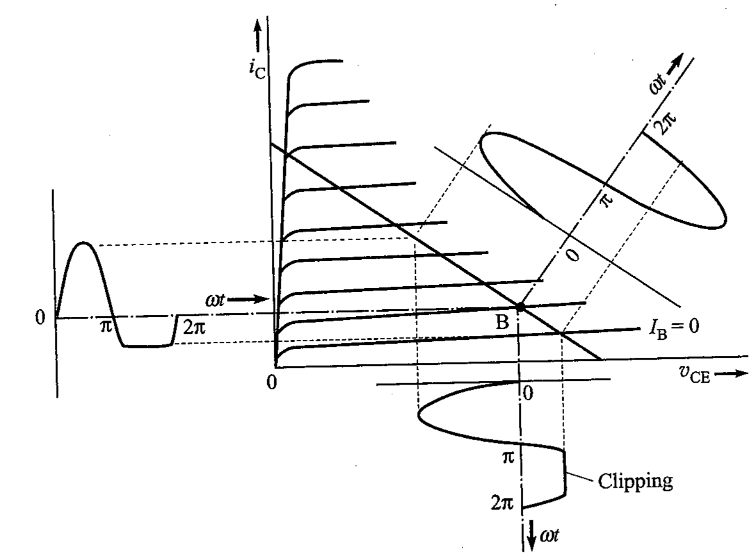


Fig. 7.3 *Operating point near cut-off region gives clipping at the negative peaks*

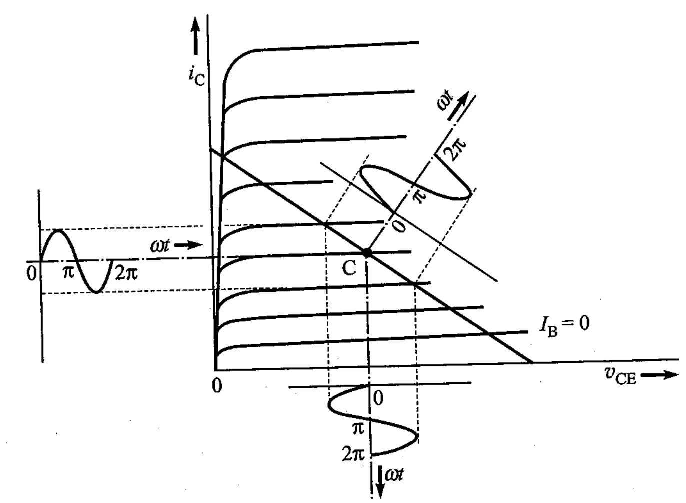


Fig. 7.4 *Operating point at the centre of active region is most suitable*

In Fig. 7.3, the point B is very near to the cut-off region. The output signal is now clipped at the negative peaks. Hence, this too is not a suitable operating point.

It is clear from Fig. 7.4 that the output signal is not at all distorted if point C is chosen as the operating point. A good amplifier amplifies signals without introducing distortion, as much as possible. Thus, point C is the most suitable operating point.

Even for the operating point C, distortion can occur in the amplifier if the input signal is large. As shown in Fig. 7.5, the output current and output voltage is clipped at both the positive and the negative peaks. Thus, the maximum signal that can be handled by an amplifier is decided by the choice of the operating point.

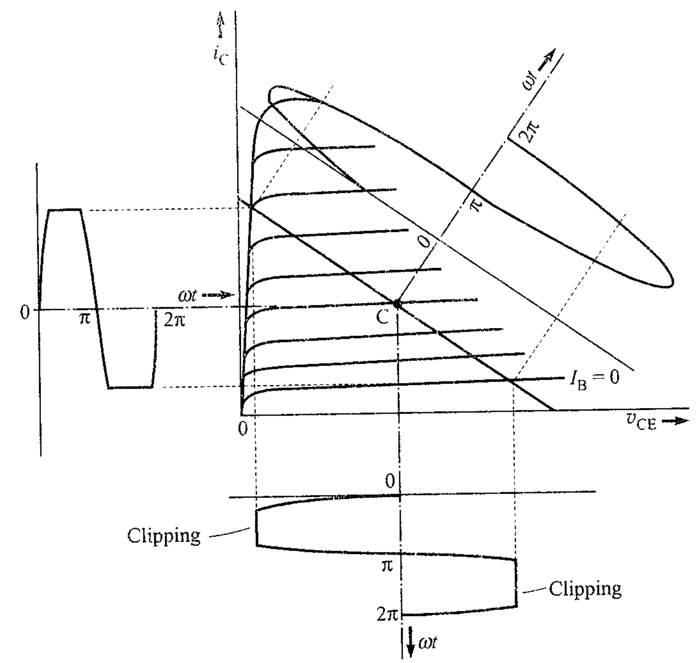


Fig. 7.5 *Distortion may result because of too large an input signal*

7.4 NEED FOR BIAS STABILISATION

Only the fixing of a suitable operating point is not sufficient. It must also be ensured that it remains where it was fixed. It is unfortunate that in the transistor circuits the operating point shifts with the use of the circuit. Such a shift of operating point may drive the transistor into an undesirable region. The amplifier then becomes useless.

There are two reasons for the operating point to shift. First, the transistor parameters are temperature dependent. Secondly, the parameters (such as β) change from unit to unit. In spite of tremendous advancement in semiconductor technology, the transistor parameters vary between wide limits even among different units of the same type. Thus, when a transistor is replaced by another of the same type, the operating point may shift.

Flow of current in the collector circuit produces heat at the collector junction. This increases the temperature. More minority carriers are generated in base-collector region (since more bonds are broken). The leakage current I_{CBO} increases. Since

$$I_{CEO} = (1 + \beta)\, I_{CBO} \tag{7.1}$$

and

$$I_C = \beta I_B + I_{CEO} \tag{7.2}$$

the increase in I_{CBO} will cause I_{CEO} to increase, which in turn increases the collector current I_C. This further raises the temperature of the collector-base junction, and the whole cycle repeats again. Such a cumulative increase in I_C will ultimately shift the operating point into the saturation region. This situation may prove to be very dangerous. The excess heat produced at the junction may even burn the transistor. Such a situation is described by the term *thermal runaway*.

The sequence of events resulting in thermal runaway of the transistor can be summarised as shown in Fig. 7.6. Here an upward arrow indicates an increase in the quantity written with it. Thus, the diagram means: "As temperature T increases, the leakage current I_{CBO} also increases; as I_{CBO} increases, the leakage current in common-emitter configuration I_{CEO} also increases; as I_{CEO} increases, ... and so on."

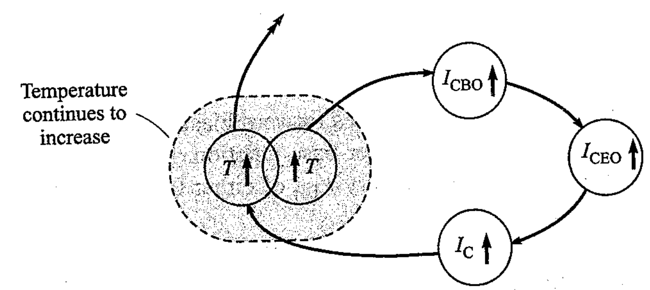


Fig. 7.6 *Increase in I_{CEO} with temperature leads to thermal runaway*

7.5 REQUIREMENTS OF A BIASING CIRCUIT

The discussion in the above sections may be summarised by stating that the biasing circuit should fulfill the following requirements:

1. Establish the operating point in the centre of the active region of the characteristics, so that on applying the input signal the instantaneous operating point does not move either to the saturation region or to the cut-off region, even at the extreme values of the input signal.
2. Stabilise the collector current against temperature variations.
3. Make the operating point independent of the transistor parameters so that it does not shift when the transistor is replaced by another of the same type in the circuit.

7.6 DIFFERENT BIASING CIRCUITS

The simplest biasing circuit could be the one shown in Fig. 7.7. The emitter-base junction is forward biased by the battery V_{BB} and the collector-base junction is reverse biased by the battery V_{CC}. The voltage V_{BE} across the forward-biased junction is very low (for a germanium transistor, $V_{BE} = 0.3$ V; and for silicon transistor, $V_{BE} = 0.7$ V). This requires that the battery voltage V_{BB} must also be of the same order. The voltage V_{CC} should be of a much larger value than the voltage V_{BB}; only then is the collector-base junction reverse biased.

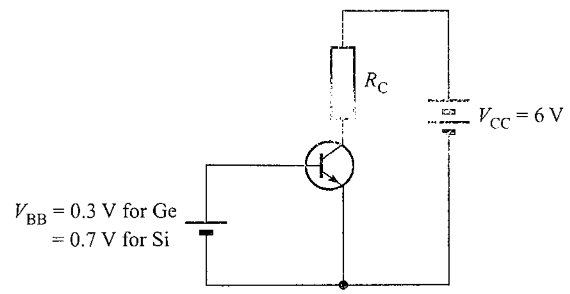


Fig. 7.7 *Simplest biasing circuit*

Though the circuit of Fig. 7.7 achieves forward biasing of the emitter base junction and reverse biasing of the collector-base junction, it is not a practical circuit. It is extremely difficult to have a battery V_{BB} of either 0.3 V or 0.7 V. This circuit, therefore, is never used.

A modified circuit of Fig. 7.7 is shown in Fig. 7.8*a*. This circuit is commonly called a *fixed-bias circuit*.

7.6.1 Fixed-bias Circuit

In the circuit shown in Fig. 7.8*a*, the battery V_{BB} need not be of low value. When current I_B flows through the series resistance R_B, a major portion of the voltage is dropped across it. The supply V_{BB} can now be of 1.5 V. Such a supply is easily available. It is interesting to note that the same circuit can also be drawn in a different way as in Fig. 7.8*b*.

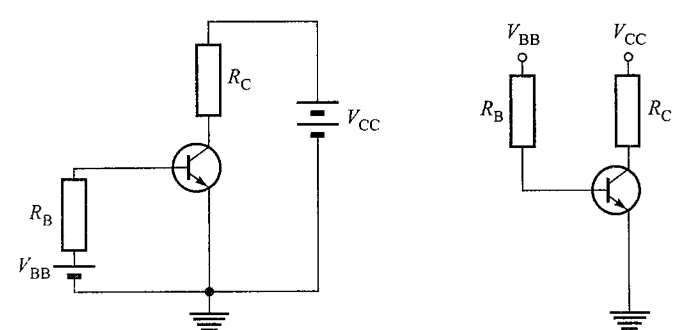


Fig. 7.8 *Fixed-bias circuit*

The circuit in Fig. 7.8 uses two batteries, V_{CC} and V_{BB}. The positive terminals of both the batteries are connected to the collector and base resistors. We can use a single battery instead of the two, as shown in Fig. 7.9. The value of R_B is then suitably modified. From a practical point of view, it is always preferable to have an electronic circuit that works on a single battery. Given a fixed-bias circuit, can we determine its operating point? We shall now develop a step-by-step procedure for doing this.

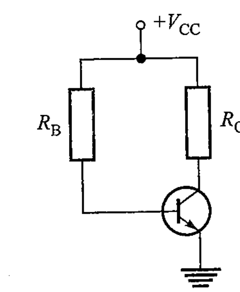


Fig. 7.9 *Fixed-bias circuit using a single battery*

Input Section: Let us first consider only the input section of the circuit, as shown in Fig. 7.10*a*. We can apply Kirchhoff's voltage law (KVL) to this base-emitter loop and get

$$V_{CC} = I_B R_B + V_{BE}$$

from which the base current is given as

$$I_B = \frac{V_{CC} - V_{BE}}{R_B} \tag{7.3}$$

Since V_{BE} is very small compared to V_{CC}, not much error will be committed if it is neglected. The base current is then given by the simple expression,

$$I_B \simeq \frac{V_{CC}}{R_B} \tag{7.4}$$

The supply voltage V_{CC} being of fixed value—once the resistance R_B is selected—the base current I_B is also fixed. Hence the name *fixed-bias circuit*.

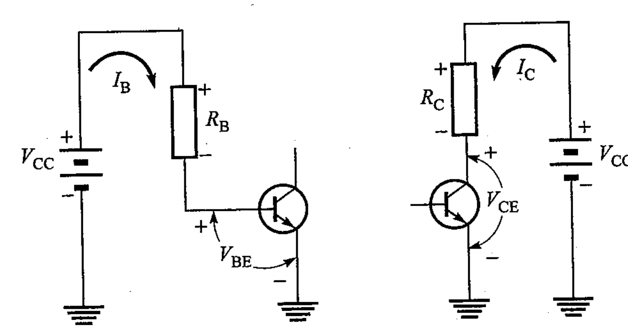


Fig. 7.10 *(a) Input section of the fixed-bias circuit (b) Output section of the fixed-bias circuit*

Output Section: We now consider the output section of the circuit, as shown in Fig. 7.10*b*. The collector current I_C that flows through the resistor R_C is given as

$$I_C = \beta I_B + I_{CEO} \tag{7.5}$$

In this equation, βI_B is the portion of current transferred from the input side. The current I_{CEO} is the leakage current in the CE configuration. Though the current I_{CEO} is not as small as the leakage current in the CB configuration I_{CBO}, yet it is very small compared to the usual values of I_C. A very small error will be introduced if we neglect the current I_{CEO} in our calculations. Therefore, to a good approximation, the collector current I_C is given as

$$I_C = \beta I_B \tag{7.6}$$

Applying KVL to the output section of Fig. 7.10*b*, we get

$$V_{CC} = I_C R_C + V_{CE} \tag{7.7}$$

One word of caution: It is clear from the above equation that the supply voltage V_{CC} provides the voltages across the resistor R_C and also across the collector-emitter terminals. Obviously, the voltage drop $I_C R_C$ can never be more than the supply voltage, V_{CC}, or

$$I_C \leqslant \frac{V_{CC}}{R_C} \tag{7.8}$$

If the value of I_C turns out to be the greater than the maximum value given by Eq. (7.8), it is certainly wrong. It is so, because the operating point is lying in the saturation region of the characteristics. Here, the collector current I_C is limited due to saturation, and its value remains at its maximum (given by Eq. (7.8)) whatever the value of base current I_B is. Equation (7.6) then becomes invalid.

Having taken care of the above caution, we are to find V_{CE} now, to determine the operating point. Equation (7.7) can be written as

$$V_{CE} = V_{CC} - I_C R_C \tag{7.9}$$

Of course, when the transistor is in saturation, the voltage V_{CE} is almost zero (actually a few tenths of a volt), and collector saturation current

$$I_{C(sat)} = \frac{V_{CC}}{R_C}$$

To summarise: The operating point in the fixed-bias circuit can be calculated in the following three steps:

1. Calculate base current I_B using Eq. (7.4). In case V_{BE} is known, use Eq. (7.3) to obtain more accurate results.
2. Calculate collector current I_C from Eq. (7.6). Make sure that this value is not greater than the one calculated from Eq. (7.8).
3. Calculate collector-emitter voltage V_{CE} using Eq. (7.9).

Example 7.1 Calculate the collector current and the collector-to-emitter voltage for the circuit given in Fig. 7.11.

Solution:

(*a*) The base current I_B is given as

$$I_B = \frac{(V_{CC} - V_{BE})}{R_B} \simeq \frac{V_{CC}}{R_B} = \frac{9}{300 \times 10^3}$$

$$= 3 \times 10^{-5}\ \text{A} = 30\ \mu\text{A}$$

(*b*) The collector current I_C is given as

$$I_C = \beta I_B = 50 \times 30 \times 10^{-6}\ \text{A} = \mathbf{1.5\ mA}$$

Let us check if this current is less than the collector saturation current.

$$I_{C(sat)} = \frac{V_{CC}}{R_C} = \frac{9}{2 \times 10^3}$$

$$= 4.5 \times 10^{-3}\ \text{A} = 4.5\ \text{mA}$$

Thus, the transistor is not in saturation.

(*c*) The collector-to-emitter voltage

$$V_{CE} = V_{CC} - I_C R_C = 9 - 1.5 \times 10^{-3} \times 2 \times 10^3 = \mathbf{6\ V}$$

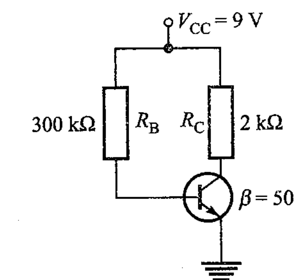


Fig. 7.11

Example 7.2 Calculate the coordinates of the operating point as fixed in the circuit in Fig. 7.12*a*. Given: $R_C = 1\ \text{k}\Omega$, $R_B = 100\ \text{k}\Omega$.

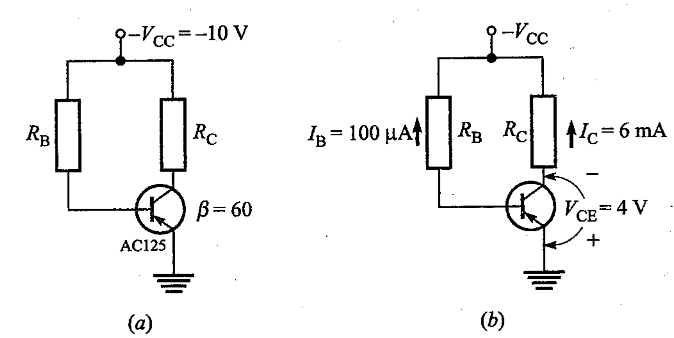


Fig. 7.12

Solution:

(*a*) The base current is

$$I_B = \frac{(V_{CC} - V_{BE})}{R_B} \simeq \frac{V_{CC}}{R_B} = \frac{10}{100 \times 10^3}\ \text{A} = 100\ \mu\text{A}$$

(*b*) The collector current is

$$I_C = \beta I_B = 60 \times 100 \times 10^{-6}\ \text{A} = 6\ \text{mA}$$

We shall now check if this current is less than the collector saturation current

$$I_{C(sat)} = \frac{V_{CC}}{R_C} = \frac{10}{1 \times 10^3}\ \text{A} = 10\ \text{mA}$$

Therefore, the transistor is not in saturation.

(*c*) The voltage between the collector and emitter terminals is

$$V_{CE} = V_{CC} - I_C R_C = 10 - 6 \times 10^{-3} \times 10^3 = 4\ \text{V}$$

Figure 7.12*b* shows the value and the direction of base current I_B, collector current I_C and collector-emitter voltage V_{CE}.

Example 7.3 In the circuit shown in Fig. 7.12*a*, the transistor is replaced by another unit of AC125. This new transistor has $\beta = 150$ instead of 60. Determine the quiescent operating point.

Solution:

(*a*) The base current remains the same, i.e., 100 μA.

(*b*) The collector current is

$$I_C = \beta I_B = 150 \times 100 \times 10^{-6}\ \text{A} = 15\ \text{mA}$$

The collector saturation current was 10 mA in the last example. Here also, this current remains the same. But the calculated current I_C is seen to be greater than $I_{C(sat)}$. Hence, the transistor is now in saturation. In this case, the operating point is specified as

$$I_C = I_{C(sat)} = \mathbf{10\ mA}$$
$$V_{CE} = \mathbf{0\ V}$$

Example 7.4 In the biasing circuit shown in Fig. 7.13, a supply of 6 V and a load resistance of 1 kΩ is used. (*a*) Find the value of resistance R_B so that a germanium transistor with $\beta = 20$ and $I_{CBO} = 2$ μA draw an I_C of 1 mA. (*b*) What I_C is drawn if the transistor parameters change to $\beta = 25$ and $I_{CBO} = 10$ μA due to rise in temperature?

Solution:

(*a*) We know that

$$I_C = \beta I_B + (\beta + 1) I_{CBO}$$

or

$$I_B = \frac{I_C - (\beta + 1) I_{CBO}}{\beta}$$

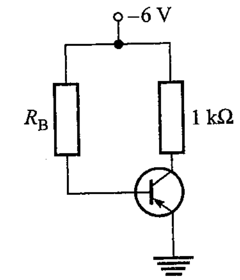


Fig. 7.13

Here, $I_C = 1$ mA; $\beta = 20$; and $I_{CBO} = 2$ μA

$$\therefore \quad I_B = \frac{1 \times 10^{-3} - (20 + 1) \times 2 \times 10^{-6}}{20}\ \text{A}$$
$$= 47.9\ \mu\text{A}$$

Writing loop equation for the input section, we get

$$6 = I_B R_B + V_{BE}$$

Since it is stated that the transistor used in the circuit is a germanium transistor, V_{BE} can be assumed to be 0.3 V. Thus, from the above equation, we have

$$R_B = \frac{6 - 0.3}{I_B} = \frac{5.7}{47.9 \times 10^{-6}}\ \Omega = 118.998\ \text{k}\Omega$$

It is worthwhile to see how much error is committed if we neglect I_{CBO} and V_{BE} from the above calculations. Since,

$$I_C = \beta I_B$$
$$I_B = \frac{I_C}{\beta} = \frac{1 \times 10^{-3}}{20}\ \text{A} = 50\ \mu\text{A}$$

The input loop equation now becomes

$$6 = I_B R_B$$

Therefore, the resistance R_B is given as

$$R_B = \frac{6}{50 \times 10^{-6}}\ \Omega = \mathbf{120\ k\Omega}$$

$$\text{The percentage error} = \frac{120 - 118.998}{118.998} \times 100 = 0.842\ \%$$

Comment: This error is too small to bother about. Moreover the resistors available in the market ordinarily have a tolerance of ±10 %. It is, therefore, not very *incorrect* to neglect V_{BE} and I_{CBO} while making calculations.

(*b*) Here, due to rise in temperature, the transistor parameters have changed to $\beta = 25$ and $I_{CBO} = 10$ μA. The collector current is now given as

$$I_C = \beta I_B + (\beta + 1) I_{CBO}$$
$$= 25 \times 47.9 \times 10^{-6} + (25 + 1) \times 10 \times 10^{-6}\ \text{A}$$
$$= \mathbf{1.46\ mA}$$

Comment: It may be noted that the collector current has increased by almost 50 % due to rise in temperature.

Why fixed-bias circuit is seldom used The fixed-bias circuit in Fig. 7.9 is a simple circuit. It uses very few components (only two resistors and one battery). It is very easy to fix the quiescent operating point anywhere in the active region of the characteristics by simply changing the value of resistor R_B. It provides maximum flexibility in the design. In spite of all this, it is seldom used in practice.

This circuit meets the first requirement stated in Section 7.5 very well. However, it miserably fails to meet the second and third requirements. With the rise in temperature, a cumulative action takes place and the collector current goes on increasing. The circuit provides no check on the increase in collector current. The operating point is not stable. This situation can be shown as in Fig. 7.14.

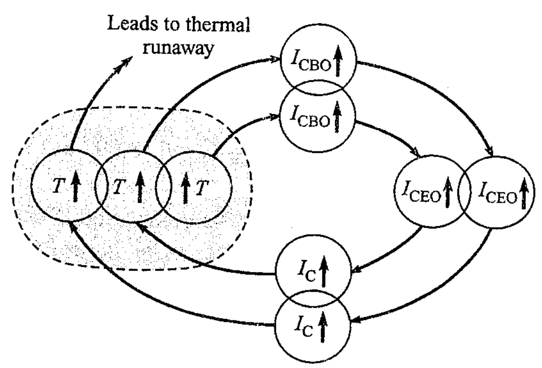


Fig. 7.14 *Fixed-bias circuit leads to thermal runaway of the transistor*

As regard the third requirement—i.e., the Q point should not shift on replacing the transistor with another of same type—the circuit utterly fails. Since, $I_C = \beta I_B$, and the base current I_B is already fixed, the current I_C is solely dependent on β. When the transistor is replaced by another with different value of β, the operating point will shift. The stabilisation of the operating point is very poor. Therefore, the biasing circuit needs some modification.

7.6.2 Collector-to-Base Bias Circuit

Figure 7.15 shows a modified biasing circuit. Here, the base resistor R_B is connected to the collector instead of connecting it to the battery V_{CC}. Writing the loop equation for the input circuit, we get

$$V_{CC} = R_C(I_C + I_B) + I_B R_B + V_{BE}$$

or $$V_{CC} = R_C I_C + (R_C + R_B) I_B + V_{BE} \tag{7.10}$$

or $$I_B = \frac{(V_{CC} - I_C R_C) - V_{BE}}{R_C + R_B} \tag{7.11}$$

From the output section of the circuit, we have

$$V_{CE} = V_{CC} - (I_C + I_B) R_C$$

or $$V_{CE} \simeq V_{CC} - I_C R_C \quad (\text{since } I_B \ll I_C) \tag{7.12}$$

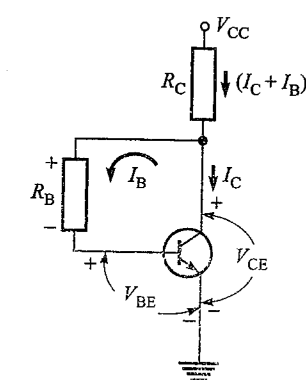


Fig. 7.15 *Collector-to-base bias circuit*

Substituting Eq. (7.12) in Eq. (7.11), we get

$$I_B = \frac{V_{CE} - V_{BE}}{R_C + R_B} \tag{7.13}$$

Let us now see what happens when the temperature rises.

Suppose the temperature increases causing the leakage current (and also β) to increase. This increases the collector current (since $I_C = \beta I_B + I_{CEO}$). As the collector current increases, the voltage V_{CE} decreases (since $V_{CE} = V_{CC} - I_C R_C$). As can be seen from Eq. (7.13), the reduced V_{CE} causes decrease in base current I_B. The lowered base current in turn reduces the original increase in collector current. Thus, a mechanism exists in the circuit because of which the collector current is not allowed to increase rapidly. There is a tendency in the circuit to stabilise the operating point.

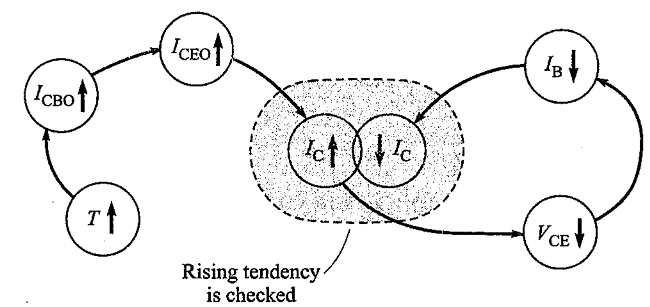


Fig. 7.16 *Collector-to-base bias circuit checks the rising tendency of collector current*

Note that the resistor R_B connects the collector (the output) with the base (the input). This means that a feedback exists in the circuit. The base current is dependent on the collector voltage. And this dependence is such as to nullify the changes in base current. That is why this circuit is also called a *voltage feedback* bias circuit.

Suppose the transistor in this circuit is replaced by another having different value of β. The shift in the operating point will not be as much as it occurs in case of fixed-bias circuit. This can be seen as follows.

For determining the operating point, we substitute βI_B for I_C in Eq. (7.10) to get

$$V_{CC} = R_C \beta I_B + (R_C + R_B) I_B + V_{BE}$$

or $$V_{CC} = V_{BE} + [R_B + (\beta + 1) R_C] I_B$$

or $$I_B = \frac{V_{CC} - V_{BE}}{R_B + (\beta + 1) R_C} \simeq \frac{V_{CC}}{R_B + \beta R_C} \tag{7.14}$$

Since $I_C = \beta I_B$, we can determine the collector current. To determine the collector voltage, we write the loop equation for the output section of the circuit (Fig. 7.15).

$$V_{CC} - (I_B + I_C) R_C - V_{CE} = 0$$

or $$V_{CE} = V_{CC} - (I_C + I_B)R_C \simeq V_{CC} - I_C R_C \tag{7.15}$$

Why collector-to-base bias circuit is seldom used This circuit has a tendency to stabilise the operating point against temperature and β variations. But the circuit is not used very much. The resistor R_B not only provides a dc feedback for the stabilisation of operating point, but it also causes an ac feedback. This reduces the voltage gain of the amplifier. It is not desirable. After all, the biasing of a transistor was needed so that it could amplify the ac signals properly. Because of this drawback, the circuit is not very commonly used.

Example 7.5 How much is the emitter current in the circuit in Fig. 7.17? Also calculate V_C.

Solution: From Eq. (7.14), the base current is given as

$$I_B = \frac{V_{CC}}{R_B + \beta R_C}$$

Here, $V_{CC} = 10$ V; $R_B = 500 \times 10^3\ \Omega$;

$R_C = 500\ \Omega$; $\beta = 100$.

Therefore,

$$I_B = \frac{10}{500 \times 10^3 + 100 \times 500}$$

$$= 18 \times 10^{-6}\ \text{A} = 18\ \mu\text{A}$$

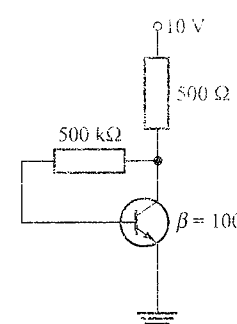


Fig. 7.17

The emitter current is then given as

$$I_E \cong I_C = \beta I_B = 100 \times 18 \times 10^{-6} = 1.8 \times 10^{-3}\ \text{A} = \mathbf{1.8\ mA}$$

The collector voltage

$$V_C = V_{CE} = V_{CC} - I_C R_C = 10 - 1.8 \times 10^{-3} \times 500 = \mathbf{9.1\ V}$$

Comment: Since the collector voltage $V_C (= V_{CE})$ is only slightly less than V_{CC}, the quiescent operating point is near the cut-off region.

Example 7.6 Calculate the minimum and maximum collector current in Fig. 7.18, if the β of the transistor varies within the limits indicated.

Solution: From Eq. (7.14), the base current is given as

$$I_B = \frac{V_{CC}}{R_B + \beta R_C}$$

(*a*) Let us first take the minimum value of β, so that $\beta = 50$;

$$V_{CC} = 20\ \text{V};\ R_B = 200\ \text{k}\Omega = 200 \times 10^3\ \Omega$$

$$R_C = 2\ \text{k}\Omega = 2 \times 10^3\ \Omega$$

Therefore,

$$I_B = \frac{20}{200 \times 10^3 + 50 \times 2 \times 10^3}$$

$$= 66.6 \times 10^{-6}\ \text{A}$$

The collector current is given as

$$I_C = \beta I_B = 50 \times 66.6 \times 10^{-6}$$

$$= 3.33 \times 10^{-3}\ \text{A} = \mathbf{3.33\ mA}$$

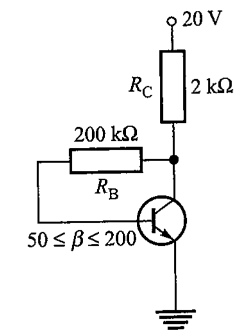


Fig. 7.18

(*b*) Now we take the maximum value of β, i.e., $\beta = 200$, so that,

$$I_B = \frac{20}{200 \times 10^3 + 200 \times 2 \times 10^3} = 33.33 \times 10^{-6}\ \text{A}$$

$$\therefore \quad I_C = \beta I_B = 200 \times 33.33 \times 10^{-6}\ \text{A}$$

$$= 6.66 \times 10^{-3}\ \text{A} = \mathbf{6.66\ mA}$$

Comment: It may be noted that in this circuit when β increases four times, the base current is halved and the collector current becomes double. However, in a fixed-bias circuit, if β had increased four times, the collector current would have also increased four times.

7.6.3 Bias Circuit with Emitter Resistor

We can modify the fixed-bias circuit of Fig. 7.9 by connecting a resistor to the emitter terminal. The modified circuit is shown in Fig. 7.19. In this circuit we have three resistors R_C, R_B and R_E and a battery V_{CC}.

We shall now see what happens to the Q point when the temperature increases. For this, we write the loop equation for the input section of the circuit,

$$V_{CC} = R_B I_B + V_{BE} + I_E R_E \tag{7.16}$$

or $$I_B = \frac{(V_{CC} - I_E R_E - V_{BE})}{R_B}$$

or $$I_B \simeq \frac{(V_{CC} - I_E R_E)}{R_B} \quad \text{(since } V_{BE} \text{ is very small)} \tag{7.17}$$

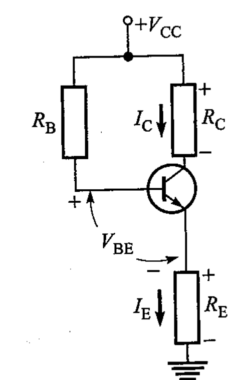


Fig. 7.19 *Bias circuit with emitter resistor*

As the temperature tends to increase, the sequence of events occuring is shown in Fig. 7.20.

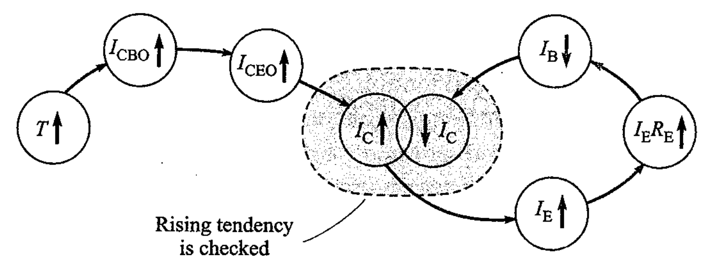


Fig. 7.20

Because of the temperature rise, the leakage current increases. This increases the collector current as well as the emitter current. As a result, the voltage drop across resistor R_E also increases. This reduces the numerator of Eq. (7.17) and hence the current I_B also decreases. This results in reduction of the collector current. Thus we see that the collector current is not allowed to increase to the extent it would have been in the absence of the resistor R_E.

In case the transistor is replaced by another of the same type (which may have different value of β), then also this circuit provides stabilisation of the Q point, as is shown in Fig. 7.21.

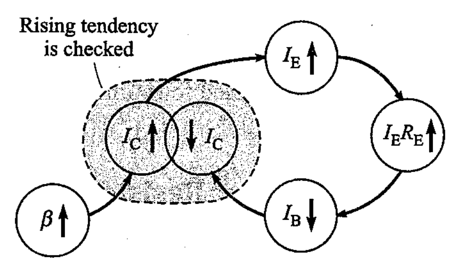


Fig. 7.21

Having seen that the operating point is stable in this circuit, let us determine the Q point. To do this, let us rewrite Eq. (7.16) as

$$V_{CC} = I_B R_B + V_{BE} + (\beta + 1)\, I_B R_E \qquad (\text{since } I_E = (\beta + 1) I_B)$$

or

$$I_B = \frac{V_{CC} - V_{BE}}{R_B + (\beta+1)R_E} = \simeq \frac{V_{CC}}{R_B + \beta R_E} \tag{7.18}$$

We can calculate the collector current easily, since

$$I_C = \beta I_B = \frac{\beta V_{CC}}{R_B + \beta R_E} = \frac{V_{CC}}{R_E + (R_B/\beta)} \tag{7.19}$$

To find V_{CE}, we write the loop equation for the output section,

$$V_{CC} = I_C R_C + V_{CE} + I_E R_E$$

$$V_{CE} = V_{CC} - (R_C + R_E) I_C \qquad (\text{since } I_C \simeq I_E) \tag{7.20}$$

Operating point is thus determined by Eqs. (7.19) and (7.20).

The resistor R_E is present in the output side as well as in the input side of the circuit. A feedback occurs through this resistor. The feedback voltage is proportional to the emitter current. Hence, this circuit is also called *current feedback* biasing circuit. While the dc feedback helps in the stabilisation of the Q point, the ac feedback reduces the voltage gain of the amplifier; again, an undesirable feature. Of course, this drawback can be remedied by putting a capacitor C_E across the resistor R_E, as shown in Fig. 7.22. The capacitor C_E offers very low impedance to the ac current. The emitter is effectively placed at ground potential for the ac signal. The circuit provides dc feedback for the stabilisation of the Q point, but does not give any ac feedback. The process of amplification of the ac signal remains unaffected.

Why this circuit is not used The circuit in Fig. 7.19 does provide some stabilisation of the Q point. But as you can see from Eq. (7.19), the denominator can be independent of β only if

$$R_E \gg \frac{R_B}{\beta}$$

This means we should either have a very high value of R_E or a very low value of R_B. A high value of R_E will cause a large dc drop across it. To obtain a particular operating point under this condition, it will require a high dc source V_{CC}. On the other hand if R_B is low, a separate low voltage supply has to be used in the base circuit. Both the alternatives are quite impractical. We should, therefore, look for some better circuit.

Example 7.7 Calculate the values of the three currents in Fig. 7.22.

Solution: From Eq. (7.18), the base current is given as

$$I_B = \frac{V_{CC}}{R_B + (\beta+1)R_E}$$

Here, $V_{CC} = 10$ V; $R_B = 1$ MΩ $= 1 \times 10^6$ Ω

$R_E = 1$ kΩ $= 1 \times 10^3$ Ω; $\beta = 100$

Therefore,
$$I_B = \frac{10}{1\times10^6 + (100+1)\times1\times10^3}$$
$$= 9.09 \times 10^{-6}\ \text{A} = \mathbf{9.09\ \mu A}$$

Now, the collector current

$$I_C = \beta I_B = 100 \times 9.09 \times 10^{-6}$$
$$= 0.909 \times 10^{-3}\ \text{A}$$
$$= \mathbf{0.909\ mA}$$

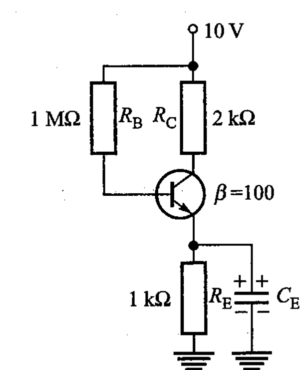


Fig. 7.22

The emitter current

$$I_E = I_C + I_B \simeq I_C = \mathbf{0.909\ mA}$$

Example 7.8 Calculate the minimum and maximum values of emitter current for the biasing circuit of Fig. 7.23. Also calculate the corresponding values of collector-to-emitter voltage. The transistor used in the circuit is a germanium transistor.

Solution: Since a germanium transistor is used in the circuit, $V_{BE} = 0.3$ V. The base current is given by Eq. (7.18) as

$$I_B = \frac{V_{CC} - V_{BE}}{R_B + (\beta+1)R_E}$$

Multiplying both sides of the above equation by $(\beta + 1)$, we get

$$(\beta+1)I_B = \frac{(V_{CC} - V_{BE})(\beta+1)}{R_B + (\beta+1)R_E}$$

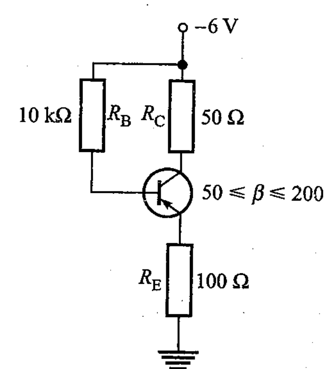


Fig. 7.23

Since $I_E = (\beta + 1)I_B$, the emitter current is given as

$$I_E = \frac{(V_{CC} - V_{BE})(\beta+1)}{R_B + (\beta+1)R_E} \qquad (7.21)$$

(*a*) Let us first take the minimum value of β, so that $\beta = 50$; $V_{CC} = 6$ V. $R_B = 10\ \text{k}\Omega = 10 \times 10^3\ \Omega$; $R_E = 100\ \Omega$. Therefore,

$$I_E = \frac{(6 - 0.3)(50+1)}{10 \times 10^3 + (50+1) \times 100} = \frac{5.7 \times 51}{15100}$$

$$= 19.25 \times 10^{-3}\ \text{A} = \mathbf{19.25\ mA}$$

The collector-to-emitter voltage is given by

$$V_{CE} = V_{CC} - (R_C + R_E)I_E \quad (\text{since } I_C \simeq I_E)$$
$$= 6 - (50 + 100) \times 19.25 \times 10^{-3} = \mathbf{3.1\ V}$$

(*b*) Let us now consider the maximum value of β, i.e., $\beta = 200$. The emitter current becomes

$$I_E = \frac{(6 - 0.3)(200+1)}{10 \times 10^3 + (200+1) \times 100} = \frac{5.7 \times 201}{(10 + 20) \times 10^3}$$
$$= 38.2 \times 10^{-3}\ \text{A} = \mathbf{38.2\ mA}$$

The collector-to-emitter voltage is

$$V_{CE} = V_{CC} - (R_C + R_E)\,I_E = 6 - (50 + 100)\ 38.2 \times 10^{-3}$$
$$= 6 - 5.7 = \mathbf{0.3\ V}$$

Comment: (*i*) Note that when β becomes four times, the emitter current becomes almost double. Had it been a fixed-bias circuit, the emitter current would have increased four times. Thus, the circuit does provide some stability of the Q point.

(*ii*) When β changes from 50 to 200, the Q point shifts from the active region to very near the saturation. With $\beta = 50$, $V_{CE} = 3.1$ V (almost half of V_{CC}). But, with $\beta = 200$, $V_{CE} = 0.3$ V (nearing zero volts).

Example 7.9 If the collector resistance R_C in Fig. 7.23 is changed to 1 kΩ, determine the new Q points for the minimum and maximum values of β.

Solution: Since the value of the emitter current does not depend upon the value of R_C (see Eq. (7.21)), the emitter current I_E remains the same as calculated in the previous example. That is

(*a*) For $\beta = 50$, $I_E = 19.25$ mA

(*b*) For $\beta = 200$, $I_E = 38.2$ mA

(*a*) The collector-to-emitter voltage is given by

$$V_{CE} = V_{CC} - (R_C + R_E)I_E = 6 - (1000 + 100) \times 19.25 \times 10^{-3}$$
$$= 6 - 21.17 = -15.17\ \text{V}$$

The above result is absurd! Sum of the voltage drops across R_C and R_E cannot be greater than the supply voltage V_{CC}. Is our calculation wrong? Certainly not. We face such difficulties when the transistor is in saturation. The maximum possible current that can be supplied by the battery V_{CC} to the output section is

$$I_{C(sat)} = \frac{V_{CC}}{R_C + R_E} = \frac{6}{1000 + 100}$$
$$= 5.45 \times 10^{-3}\ \text{A} = \mathbf{5.45\ mA}$$

Under saturation, the collector-to-emitter voltage is

$$V_{CE(sat)} = \mathbf{0\ V}$$

(*b*) We have seen that the transistor is in saturation when its $\beta = 50$. In case $\beta = 200$, there is all the more reason for the transistor to be in saturation. So, the Q point will be the same as calculated earlier, i.e.,

$$I_{C(sat)} = \mathbf{5.45\ mA},\ V_{C(sat)} = \mathbf{0\ V}$$

Example 7.10 Calculate the value of R_B in the biasing circuit of Fig. 7.24 so that the Q point is fixed at $I_C = 8$ mA and $V_{CE} = 3$ V.

Solution: The current I_B is given as

$$I_B = \frac{I_C}{\beta}$$

Here, $I_C = 8$ mA $= 8 \times 10^{-3}$ A and $\beta = 80$.

Therefore, $$I_B = \frac{8\times10^{-3}}{80} = 1\times10^{-4}\text{ A} = 100\ \mu\text{A}$$

From Eq. (7.18), we have

$$I_B R_B + (\beta+1) I_B R_E = V_{CC} - V_{BE} \simeq V_{CC}$$

or $$R_B = \frac{V_{CC} - (\beta+1) I_B R_E}{I_B}$$

Here, $V_{CC} = 9$ V; $\beta = 80$; $I_B = 1\times10^{-4}$ A, $R_E = 500\ \Omega$

Therefore, $$R_B = \frac{9-(80+1)\times1\times10^{-4}\times500}{1\times10^{-4}} = \frac{4.95}{1\times10^{-4}}\ \Omega$$

$$= \mathbf{49.5\ k\Omega}$$

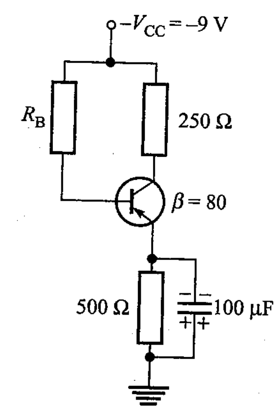


Fig. 7.24

7.6.4 Voltage Divider Biasing Circuit

This is the most widely used biasing circuit and is shown in Fig. 7.25. Compare this circuit with the one shown in Fig. 7.19. Here, an additional resistor R_2 is connected between base and ground. The name "voltage divider" comes from the voltage divider formed by the resistors R_1 and R_2. By suitably selecting this voltage divider network, the operating point of the transistor can be made almost independent of beta (β). This is why this circuit is also called "biasing circuit independent of beta".

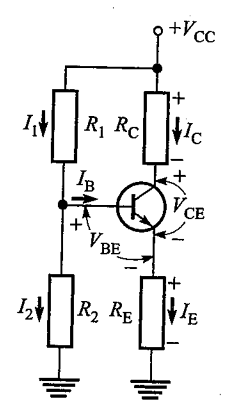


Fig. 7.25 *Voltage divider biasing circuit*

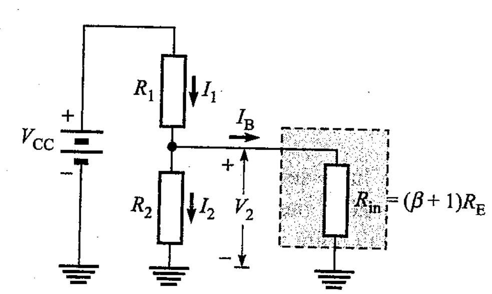


Fig. 7.26 *Input section of the voltage divider biasing circuit*

Approximate analysis To determine the operating point, we first consider the input of the circuit, redrawn in Fig. 7.26.

We make a basic assumption: The base current I_B is very small compared to the currents in R_1 and R_2. That is,

$$I_1 \simeq I_2 \gg I_B$$

The above assumption is valid because, in practice, the resistance seen looking into the base (R_{in}) is much larger than R_2. We can apply the voltage divider theorem to find the voltage across the resistor R_2 (same as base voltage V_B)

$$V_B = V_2 = \frac{R_2}{R_1+R_2}\times V_{CC} \tag{7.22}$$

The voltage across the emitter resistor R_E equals the voltage across R_2 minus the base-to-emitter voltage V_{BE}. That is,

$$V_E = V_2 - V_{BE}$$

The current in the emitter is then calculated as

$$I_E = \frac{V_E}{R_E} = \frac{V_2 - V_{BE}}{R_E} \tag{7.23}$$

The voltage at the collector (measured with respect to ground) V_C equals the supply voltage V_{CC} minus the voltage drop across R_C,

$$V_C = V_{CC} - I_C R_C$$

The collector-to-emitter voltage is then given as

$$V_{CE} = V_C - V_E = (V_{CC} - I_C R_C) - I_E R_E$$

or $$V_{CE} \simeq V_{CC} - (R_C + R_E) I_C \tag{7.24}$$

since I_C and I_E are approximately equal.

Note that in the above analysis, nowhere does β appear in any equation. It means that the operating point does not depend upon the value of β of the transistor. This is why the voltage divider circuit is most widely used. In the mass production of transistors, one of the main problems is the wide variation in β. It varies from transistor to transistor of the same type. For example, the transistor AC127 has a minimum β of 50 and maximum β of 150 for an I_C of 10 mA and a temperature of 25 °C. If this biasing circuit is used, no problem is faced on replacement of the transistor in the circuit. The operating point remains where it was fixed in the original design.

Example 7.11 Calculate the dc bias voltages and currents for the circuit of Fig. 7.27. Assume $V_{BE} = 0.3$ V and $\beta = 60$ for the transistor used.

Solution: From Eq. (7.22), the base voltage is

$$V_B = V_2 = \frac{R_2}{R_1+R_2}\times V_{CC}$$

Here, $R_2 = 5\ \text{k}\Omega = 5 \times 10^3\ \Omega$; $R_1 = 40\ \text{k}\Omega = 40 \times 10^3\ \Omega$; $V_{CC} = 12$ V

Therefore, $V_2 = \dfrac{5 \times 10^3}{(40+5) \times 10^3} \times 12 = 1.3$ V

The emitter voltage

$$V_E = V_2 - V_{BE} = 1.3 - 0.3 = 1.0 \text{ V}$$

Therefore, the emitter current

$$I_E = \frac{V_E}{R_E} = \frac{1.0}{1 \times 10^3} = 1.0 \times 10^{-3} \text{ A}$$
$$= 1.0 \text{ mA}$$

The collector current,

$$I_C \simeq I_E = \mathbf{1.0\ mA}$$

The collector voltage

$$V_C = V_{CC} - I_C R_C$$
$$= 12 - 1 \times 10^{-3} \times 5 \times 10^3 = \mathbf{7\ V}$$

Finally, the collector-to-emitter voltage

$$V_{CE} = V_C - V_E = 7 - 1 = \mathbf{6\ V}$$

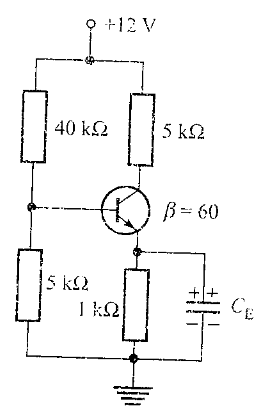


Fig. 7.27

Example 7.12 To set up 100 mA of emitter current in the power amplifier circuit of Fig. 7.28, calculate the value of the resistor R_E. Also calculate V_{CE}. The dc resistance of the primary of the output transformer is 20 Ω. Given that $R_1 = 200\ \Omega$; $R_2 = 100\ \Omega$; $R_C = 20\ \Omega$ and $V_{CC} = 15$ V.

Solution:

$$I_C \simeq I_E = 100 \text{ mA} = 0.1 \text{ A}$$

From Eq. (7.22), the base voltage is

$$V_B = \frac{R_2}{R_1 + R_2} \times V_{CC}$$
$$= \frac{100}{200 + 100} \times 15 = 5 \text{ V}$$

Neglecting V_{BE}, $V_E = V_B = 5$ V

From Eq. (7.23), the emitter resistance is given by

$$R_E = \frac{V_E}{I_E} = \frac{5}{0.1} = \mathbf{50\ \Omega}$$

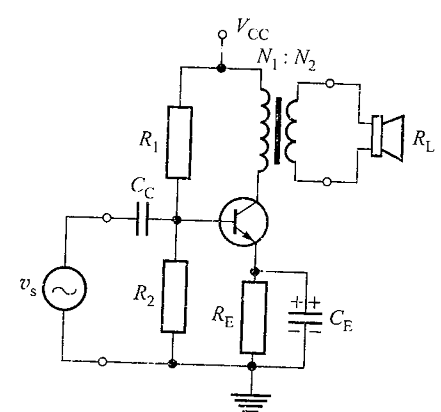


Fig. 7.28

The collector-to-emitter voltage is then calculated using Eq. (7.24),

$$V_{CE} = V_{CC} - (R_C + R_E) I_C$$
$$= 15 - (20 + 50) \times 0.1 = \mathbf{8\ V}$$

Accurate analysis You may be wondering why the operating point does not change when the transistor is replaced by another, in the circuit of Fig. 7.25. Is it really so? To verify this, let us try to analyse the circuit more accurately. For such an analysis, Thevenin's theorem is of great help. A brief review of this theorem for dc circuit is given below.

Thevenin's theorem Suppose we have a complicated network containing resistors and voltage sources (see Fig. 7.29*a*). A and B are two terminals in this network. Thevenin's theorem simply states that this circuit acts as if a voltage V_{TH} in series with a resistor R_{TH} is connected between this pair of terminals, as shown in Fig. 7.29*b*. Now, when we connect the resistor R_L across the terminals A and B, only one loop is seen. It becomes very easy to calculate the current in this resistor. The power of Thevenin's theorem lies in converting the complicated network into a single loop circuit.

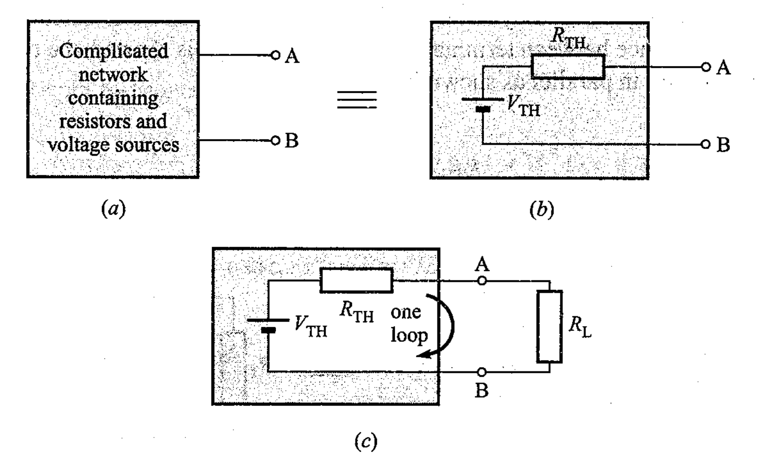
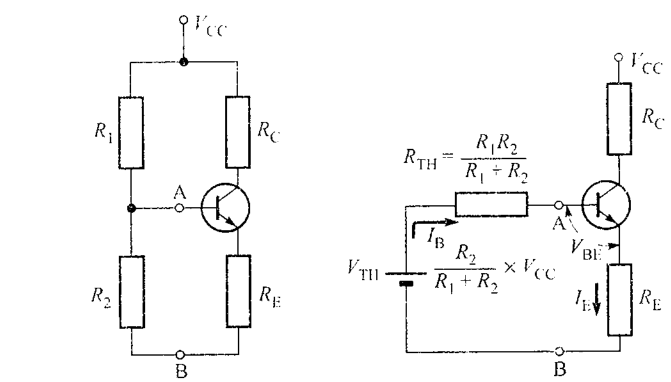


Fig. 7.29 *Thevenin's theorem*

Thevenin's equivalent voltage source V_{TH} is the open circuit voltage across terminals AB. Thevenin's resistor R_{TH} is the resistance from A to B when all the sources in the network are reduced to zero. After *Thevenising* the circuit at AB terminals (Thevenising means "finding the Thevenin's equivalent of") we may connect any resistor across AB and calculate the current flowing in it (see Fig. 7.29*c*).

The voltage-divider biasing circuit is drawn in Fig. 7.30*a*. Let us Thevenise the circuit on the left of the terminals AB. The result is shown in Fig. 7.30*b*. Here, V_{TH} is the Thevenin's voltage given as

$$V_{TH} = \frac{R_2}{R_1 + R_2} \times V_{CC} \tag{7.25}$$



(a) (b)

Fig 7.30 *Voltage divider biasing circuit*

The resistor R_{TH} is found by reducing the battery V_{CC} to zero and calculating the equivalent resistance between terminals A and B. When V_{CC} is shorted, the two resistors R_1 and R_2 are in parallel as shown in Fig. 7.31.

Thus,

$$R_{TH} = \frac{R_1 R_2}{R_1 + R_2} \tag{7.26}$$

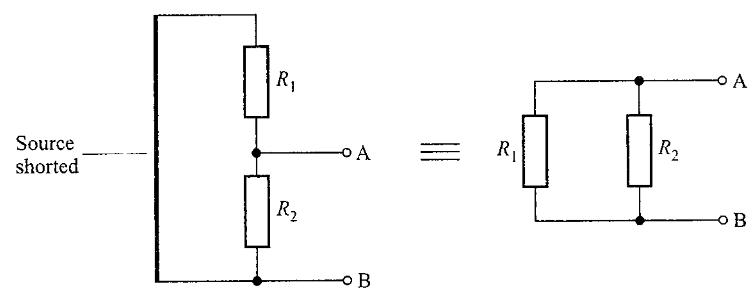


Fig. 7.31 *R_{TH} in Thevenin's equivalent*

We shall now analyse the circuit of Fig. 7.30*b* for calculation of the operating point. The loop equation for the input section can be written as

$$V_{TH} = I_B R_{TH} + V_{BE} + R_E I_E$$

or $$V_{TH} = I_B R_{TH} + V_{BE} + (\beta + 1) I_B R_E \quad (\text{since } I_E = (\beta + 1) I_B)$$

or $$I_B [R_{TH} + (\beta + 1) R_E] = V_{TH} - V_{BE}$$

or $$I_B = \frac{V_{TH} - V_{BE}}{R_{TH} + (\beta + 1) R_E} \simeq \frac{V_{TH}}{R_{TH} + \beta R_E} \tag{7.27}$$

Once the base current is fixed, the collector current can be calculated as

$$I_C = \beta I_B$$

The collector-to-base emitter voltage is then found by the familiar equation,

$$V_{CE} = V_{CC} - (R_C + R_E) I_C$$

It will be worthwhile to see to what extent the approximate analysis is valid. This is done in the next example.

Example 7.13 Make use of Thevenin's theorem to find accurate values of collector current and collector-to-emitter voltage in Fig. 7.27 (of Example 7.11).

Solution: The Thevenin voltage of the voltage-divider circuit is (see Eq. 7.25)

$$V_{TH} = \frac{R_2}{R_1 + R_2} \times V_{CC}$$

Here, $R_1 = 40\ \text{k}\Omega = 40 \times 10^3\ \Omega$; $R_2 = 5\ \text{k}\Omega = 5 \times 10^3\ \Omega$; $V_{CC} = 12$ V. Therefore,

$$V_{TH} = \frac{5 \times 10^3}{(40 + 5) \times 10^3} \times 12 = \mathbf{1.3\ V}$$

The Thevenin resistance, from Eq. (7.26) is

$$R_{TH} = \frac{R_1 R_2}{R_1 + R_2} = \frac{40 \times 10^3 \times 5 \times 10^3}{(40 + 5) \times 10^3} = 4.44 \times 10^3\ \Omega = \mathbf{4.44\ k\Omega}$$

Figure 7.32 gives the Thevenin's dc equivalent of the circuit of Fig. 7.27. We can now use Eq. (7.27) to determine base current,

$$I_B = \frac{V_{TH} - V_{BE}}{R_{TH} + \beta R_E}$$

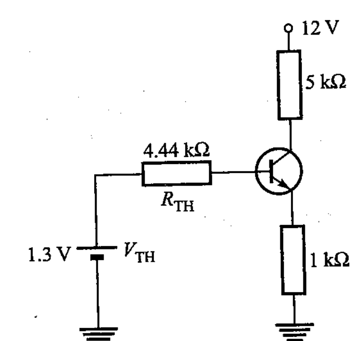


Fig. 7.32

Here, $V_{BE} = 0.3$ V; $\beta = 60$, $R_E = 1\ \text{k}\Omega = 1 \times 10^3\ \Omega$

Therefore, $$I_B = \frac{1.3 - 0.3}{4.44 \times 10^3 + 60 \times 1 \times 10^3}$$
$$= 15.52 \times 10^{-6}\ \text{A}$$

The collector current is then

$$I_C = \beta I_B = 60 \times 15.52 \times 10^{-6}$$
$$= 0.93 \times 10^{-3}\ \text{A} = \mathbf{0.93\ mA}$$

The collector-to-emitter voltage is

$$V_{CE} = V_{CC} - (R_C + R_E) I_C$$
$$= 12 - (5 \times 10^3 + 1 \times 10^3) \times 0.93 \times 10^{-3} = 12 - 5.58 = \mathbf{6.42\ V}$$

Comment: The values of I_C and V_{CE}, obtained above, may be compared with those obtained in Example 7.11. Note that the error committed is within 7 % only. Thus, we find that the approximations made are quite reasonable.

7.6.5 Emitter-Bias Circuit

Figure 7.33*a* shows an emitter bias circuit. The circuit gets this name because the negative supply V_{EE} is used to forward bias the emitter junction through resistor R_E. As usual, the V_{CC} supply reverse biases the collector junction. This circuit uses only three resistors and it provides almost as much stability of operating point as a voltage divider circuit does. However, the emitter biasing can be used only when two supplies —one positive and the other negative—are available. Figure 7.33*b* shows a simple way to draw this circuit using split supply.

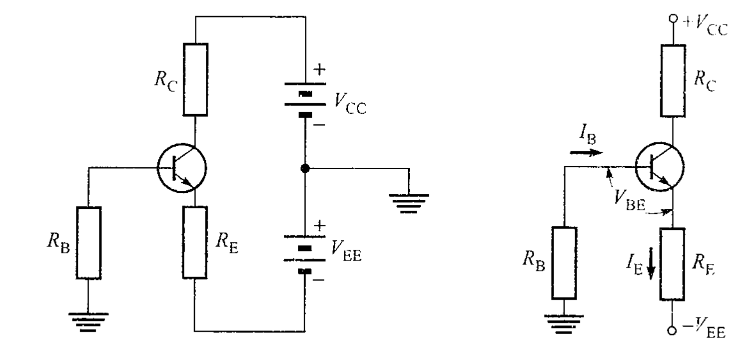


Fig. 7.33 *Emitter-bias circuit*

To determine the operating point, we apply Kirchhoff's voltage law to the emitter-base loop.

$$I_B R_B + V_{BE} + I_E R_E - V_{EE} = 0$$

Since $I_C \simeq I_E$ and $I_B = I_E/\beta$, we can rearrange the above equation to get

$$I_E = \frac{V_{EE} - V_{BE}}{R_E + R_B/\beta} \tag{7.28}$$

If we want the operating point to be independent of β, we should have

$$R_E \gg \frac{R_B}{\beta}$$

This condition can easily be met in practice. Since the supply V_{EE} is much greater than V_{BE}, we may approximate Eq. (7.28) to give

$$I_E \simeq \frac{V_{EE}}{R_E} \tag{7.29}$$

The above equation shows that the emitter is virtually at ground potential. All the V_{EE} supply voltage appears across R_E. If the emitter is at ground point, the collector-to-emitter voltage V_{CE} is simply given as

$$V_{CE} = V_{CC} - I_C R_C \tag{7.30}$$

Example 7.14 Calculate I_C and V_{CE} for the emitter-bias circuit of Fig. 7.33, where V_{CC} = !2 V; V_{EE} = 15 V, R_C = 5 kΩ, R_E = 10 kΩ, R_B = 10 kΩ, β = 100.

Solution: From Eq. (7.29), the emitter current is

$$I_E = \frac{V_{EE}}{R_E} = \frac{15}{10 \times 10^3} = 1.5 \times 10^{-3} \text{ A} = 1.5 \text{ mA}$$

The collector current is

$$I_C \cong I_E = \mathbf{1.5\ mA}$$

Using Eq. (7.30), the collector-to-emitter voltage V_{CE} is

$$V_{CE} = V_{CC} - I_C R_C = 12 - 1.5 \times 10^{-3} \times 5 \times 10^3$$
$$= 12 - 7.5 = \mathbf{4.5\ V}$$

7.7 PNP TRANSISTOR-BIASING CIRCUITS

You may be wondering why we have been using only NPN transistors? How will a biasing circuit change if a PNP transistor is used in place of an NPN transistor? To forward bias the emitter diode of a PNP transistor, V_{BE} must have the polarity as shown in Fig. 7.34. The collector junction is to be reverse biased and the polarity of V_{CE} is also to be reversed, as shown.

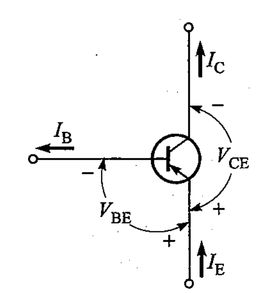


Fig. 7.34 *Notations for a PNP transistor*

In fact, the PNP transistor is the *complement* of NPN transistor. Here, the word complement signifies that all voltages and currents are opposite to those of the NPN transistor. Therefore, to find the complementary PNP circuit of a given NPN circuit, all you have to do is:

1. to replace the NPN transistor by a PNP transistor, and
2. to complement or reverse all voltages and currents.

Note that if you use *magnitudes* of voltages and currents, all formulae derived for NPN circuits apply to PNP circuits as well. For instance, in Example 7.2, use is made of the same formulae as in Example 7.1. After calculating the magnitudes of I_C and V_{CE}, the direction of I_C and the polarity of V_{CE} are properly marked in Fig. 7.12*b*, remembering that it is a PNP transistor.

7.8 BIASING THE FET

The problem of biasing an FET is not as serious as biasing a BJT. Unlike a BJT, where the emitter-base junction is always forward biased, we need reverse bias at the gate-source junction. This fact makes biasing of FETs much simpler.

Also, achieving bias stabilisation in case of FETs is much simpler than in BJTs.

7.8.1 Biasing JFET

A simple self-bias circuit (Fig. 7.35) can be used to provide reverse bias at the gate-source junction of a JFET. The design of the self-bias circuit is quite simple. For a specified value of drain current I_D, the corresponding gate-to-source voltage V_{GS} can be determined* from the transfer characteristics of the JFET. For this V_{GS}, the required value of source resistor R_S is to be calculated.

Since the gate-to-channel PN-junction is made reverse biased, the gate current is almost zero, i.e., $I_G = 0$. It has two effects. First, there is no voltage drop across resistor R_G. Hence the gate terminal and the point O (grounded) are at same potential. Thus, the biasing voltage V_{GS} is same as voltage V_{OS} (or same as voltage $-V_{SO}$). Secondly, since $I_G = 0$, source current I_S is same as drain current I_D. Therefore, the bias voltage is given as

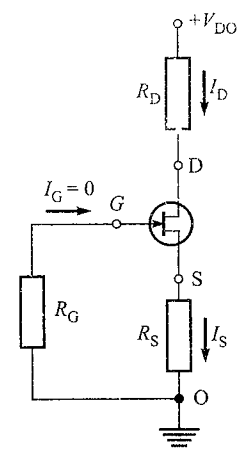


Fig. 7.35 *Self-bias circuit for an N-channel JFET*

$$V_{GS} = V_{OS} = -V_{SO} = -I_S R_S = -I_D R_S \tag{7.31}$$

Using above equation, the required value of R_S can easily be calculated.

There is no voltage drop across gate resistor R_G. This resistor is used just to provide an electrical connection between the gate terminal and ground. For this reason resistor R_G could have any value—low or high. However, we keep resistor R_G of very high value (say 10 MΩ). The reason is simple. When the circuit is used in an amplifier, the signal-voltage source at input faces very high input resistance of the amplifier—a very desirable feature.

7.8.2 Bias Stabilisation for JFET

The drain current is affected by rise in temperature in two ways:

1. The lattice ions vibrate more vigorously. Hence, the current carriers cannot move as freely in the channel. Their mobility decreases. Hence, the drain current I_D decreases.

* It is also possible to calculate V_{GS} for a given I_D by using Schockley's equation $I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_P}\right)^2$

2. The gate-to-channel depletion region decreases. The channel becomes wider, and hence I_D increases.

Figure 7.36 shows the transfer characteristics of a JFET for different temperatures ($T_1 > T_2 > T_3$). There exists a value V_{GS} for which I_D does not change with temperature. Hence, it is possible to bias a JFET for zero drain-current drift.

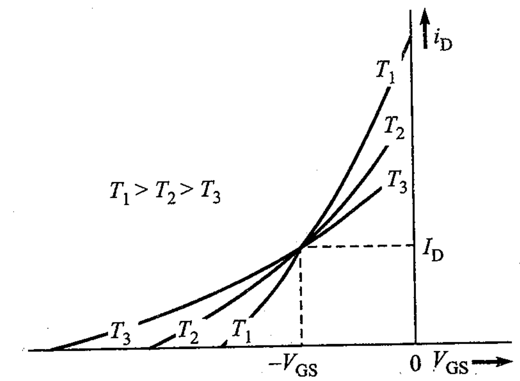


Fig. 7.36 *Transfer characteristics of an N-channel JFET for different temperatures*

7.8.3 Biasing DE MOSFET

A DE MOSFET is capable of working in both modes—depletion as well as enhancement. Therefore, the normal operating region for the DE MOSFET is around $V_{GS} = 0$. This makes biasing a DE MOSFET very simple. We have to just connect the gate to the source. But as shown in Fig. 7.37, instead of directly connecting the gate to the source (which is grounded), we do it through a high-value resistor R_G. This helps in keeping the input resistance of the amplifier quite high.

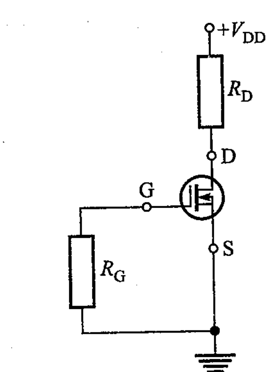


Fig. 7.37 *Bias circuit for circuit for a DE MOSFET*

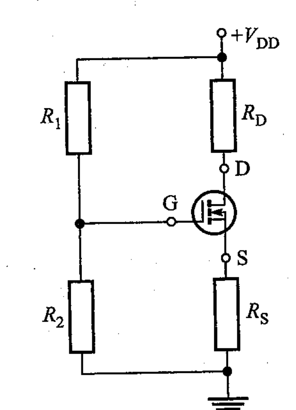


Fig. 7.38 *Voltage divider biasing an N-channel EN MOSFET*

7.8.4 Biasing EN MOSFET

Consider an N-channel EN MOSFET. Unlike a DE MOSFET, It can work only in enhancement mode. Therefore its normal operating region requires V_{GS} to be positive. This cannot be achieved by the self-bias circuit of Fig. 7.35. Instead, we use potential divider circuit to get the desired value of positive voltage V_{GS}, as shown in Fig. 7.38. The resistor R_S is included to provide dc feedback for bias stabilisation.

Review Questions

1. Explain the term "biasing".
2. Explain why a transistor should be biased.
3. Connect two external batteries between the two junctions of a transistor in its three configurations so that it works in the active region.
4. In case of the CE configuration, what are the approximate voltages of the dc batteries connected between base-emitter and collector-emitter terminals?
5. Explain why the battery connected between the emitter and base terminals requires a high resistance in series with it.
6. Draw transistor-biasing circuits using a 9-V battery and two resistors (1 kΩ and 150 kΩ) in two different ways. Point out the circuit in which bias stabilisation exists.
7. Draw a simple circuit in which only one battery is used and biasing is achieved by fixing the base current.
8. In the circuit given in Fig. R.7.1, derive the expression for I_C and V_{CE}.
9. Draw a biasing circuit using the following components:
   (*i*) two resistors of 1 kΩ each;
   (*ii*) one resistor of 100 kΩ
   (*iii*) one dc source of 6 V, and
   (*iv*) one PNP transistor (say AC126)
10. Draw a potential-divider biasing circuit making use of a 9-V battery. Mark the direction of current flowing through each resistor of the circuit.
11. For the circuit given in Fig. R.7.2, derive the following expression:

$$I_E = \frac{V_{CC} - V_{BE}}{R_B/(\beta+1) + R_E}$$

12. Prove mathematically that the operating point in a potential-divider biasing circuit is independent of β. Make relevant assumptions.
13. Explain why the fixed-bias circuit, in spite of its simplicity, is not much used in amplifiers.
14. Explain how stabilisation of operating point is achieved when one end of the base resistor R_B is connected to the collector terminal instead of the dc supply.
15. Explain the function of the emitter resistor R_E in the potential-divider biasing circuit.

16. Explain why operating point is fixed in the middle of the active region of transistor characteristics in a good voltage amplifier.
17. Explain why the operating point is not selected near the saturation region of the transistor characteristics in a voltage amplifier.
18. State the factors to be considered while designing a biasing circuit for a good transistor voltage amplifier.
19. Explain, how the circuits given in Fig. R.7.3 to R.7.6 respond to temperature and beta variations.
20. Derive the expressions for Q point in the circuit given in Fig. R.7.6, using Thevenin's theorem.

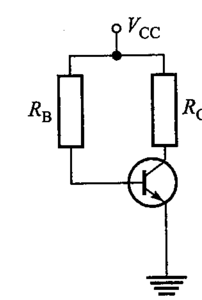


Fig. R.7.1

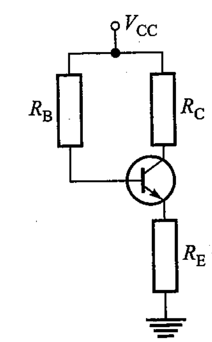


Fig. R.7.2

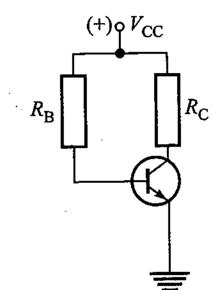


Fig. R.7.3

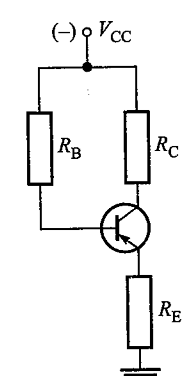


Fig. R.7.4

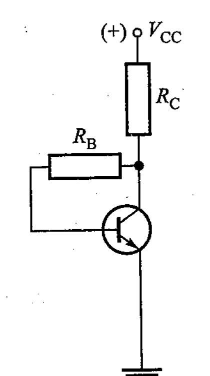


Fig. R.7.5

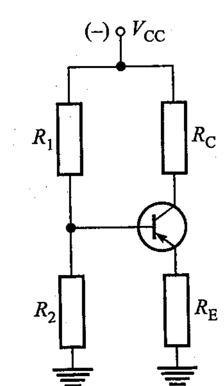


Fig. R.7.6

Objective-Type Questions

I. Below are some incomplete statements. Four alternatives are provided for each. Choose the alternative that completes the statement correctly.

1. The emitter resistor R_E bypassed by a capacitor
   (*a*) reduces the voltage gain
   (*b*) increases the voltage gain

(*c*) causes thermal runaway
(*d*) stabilises the Q point

2. The Q point in a voltage amplifier is selected in the middle of the active region because
(*a*) it gives a distortionless output
(*b*) the operating point then becomes very stable
(*c*) the circuit then requires less number of resistors
(*d*) it then requires a small dc voltage

3. The operating point of an NPN transistor amplifier should not be selected in the saturation region as
(*a*) it may drive the transistor to thermal runaway
(*b*) it may cause output to be clipped in the negative half of the input signal
(*c*) it may cause output to be clipped in the positive half of the input signal
(*d*) it may require high dc collector supply

4. The potential-divider method of biasing is used in amplifiers to
(*a*) limit the input ac signal going to the base
(*b*) make the operating point almost independent of β
(*c*) reduce the dc base current
(*d*) reduce the cost of the circuit

5. A transistor is operating in the active region. Under this condition
(*a*) both the junctions are forward biased
(*b*) both the junctions are reverse biased
(*c*) emitter-base junction is reverse biased, and collector-base junction is forward biased
(*d*) emitter-base junction is forward biased, and collector-base junction is reverse biased

6. The signal handling capacity of an amplifier is high if
(*a*) the operating point is selected near the cut-off region
(*b*) the operating point is selected near the saturation region
(*c*) the operating point is selected in the middle of the active region
(*d*) an NPN transistor of similar characteristics is used instead of PNP one

II. Some statements are given below. Write whether they are TRUE or FALSE.

1. The purpose of biasing a transistor is to obtain a certain dc collector current at a certain dc collector voltage.
2. In a certain biasing circuit, V_{CC} and V_{CE} are equal. This is because the transistor is heavily conducting.
3. A good biasing circuit should stabilise the collector current against temperature variations.
4. The emitter resistor R_E is bypassed by a capacitor so as to improve the stabilisation of Q point.

5. The dc collector current in a transistor circuit is limited by the junction capacitance.
6. Negative dc feedback through R_E is responsible for the stabilisation of the operating point in a potential-divider bias circuit.

III. Amplifier circuits shown in Fig. O. 7.1 may be either incomplete or wrongly drawn or both. If so, detect the same and redraw them correctly.

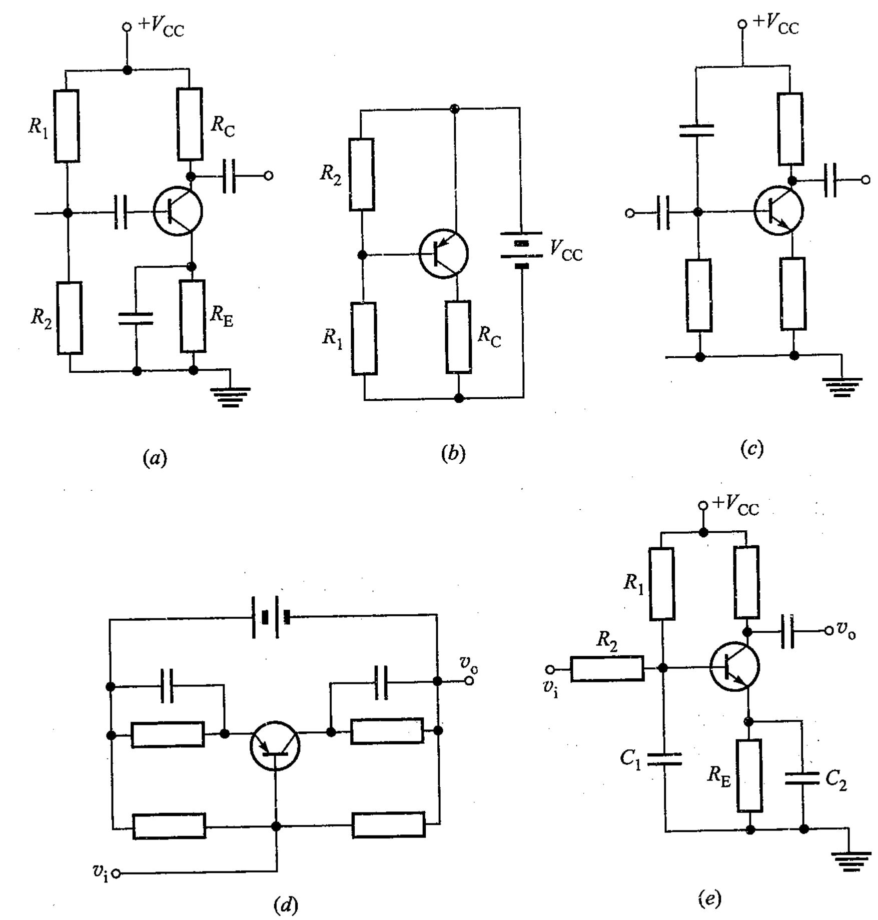


Fig. O. 7.1

Answers

I.	1. (*d*)	2. (*a*)	3. (*c*)	4. (*b*)	5. (*d*)	6. (*c*)
II.	1. T	2. F	3. T	4. F	5. F	6. T

Tutorial Sheet 7.1

1. Calculate the value of V_{CE} in a fixed-bias circuit given in Fig. T. 7.1.1. Assume $\beta_{dc} = 100$, $R_B = 200$ kΩ, $R_C = 1$ kΩ and $V_{CC} = 10$ V. **[Ans.** $V_{CE} = 5$ V]
2. Calculate the value of R_C and R_B if the dc operating point is to be fixed at $V_{CE} = 7$ V, $I_C = 5$ mA. Following data are given (refer to Fig. T. 7.1.1): $V_{CC} = 12$ V, $\beta_{dc} = 100$. **[Ans.** $R_B = 240$ kΩ, $R_C = 1$ kΩ]

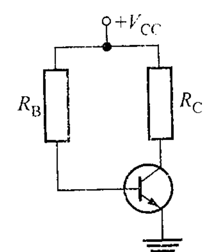


Fig. T. 7.1.1

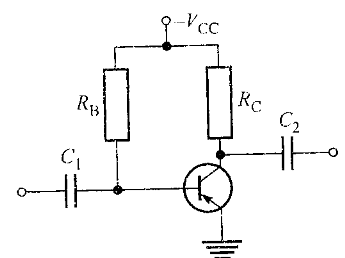


Fig. T. 7.1.2

3. A PNP transistor of $\beta = 200$ is used in the circuit given in Fig. T. 7.1.2. A dc supply of 9 V and R_C of 1.5 kΩ are used. The operating point is to be fixed at $I_C = 2$ mA. Calculate the value of R_B and the voltage V_{CE}. **[Ans.** $R_B = 0.9$ MΩ; $V_{CE} = 6$ V]
4. Design a simple fixed-biasing circuit for a PNP transistor having β such that $50 < \beta < 200$, if $V_{CC} = 12$ V and a load of 3 kΩ is used (refer to Fig. T. 7.1.2). Assume $V_{BE} = 0.3$ V. **[Ans.** $R_B < 585$ kΩ]
5. A PNP germanium transistor with $\beta = 100$ and $V_{BE} = 200$ mV is used in Fig. T. 7.1.2. Compute the Q point for the circuit conditions given below:
$$V_{CC} = 16 \text{ V}; R_C = 5 \text{ k}\Omega; R_B = 790 \text{ k}\Omega$$
[Ans. $V_{CE} = 6$ V; $I_C = 2$ mA]
6. Calculate the collector-to-emitter voltage for the PNP transistor connected in Fig. T. 7.1.3 neglecting V_{BE}. **[Ans.** $V_{CE} = -5.4$ V; $I_C = 3.6$ mA]
7. Calculate the highest value of R_C permissible in the circuit of Fig. T. 7.1.3. **[Ans.** $R_{C(max)} = 2.5$ kΩ]

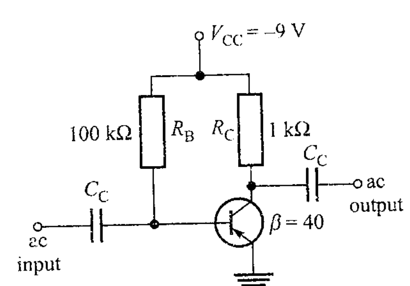


Fig. T. 7.1.3

Tutorial Sheet 7.2

1. Calculate the Q point for the dc-bias circuit in Fig. T. 7.2.1, given the following: $R_C = 3$ kΩ; $R_B = 60$ kΩ; $V_{CC} = 12$ V; $\beta = 60$; assume V_{BE} negligible. **[Ans.** $V_{CE} = 3$ V; $I_C = 3$ mA]

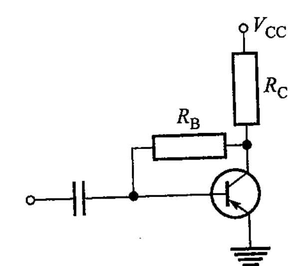


Fig. T. 7.2.1

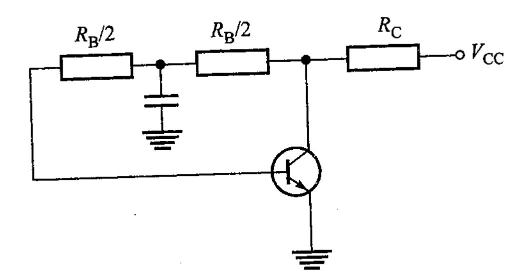


Fig. T. 7.2.2

2. Calculate all the dc currents and voltage V_{CE} in the transistor of Fig. T. 7.2.2 for the following given data:
$V_{CC} = 10$ V; $R_C = 3$ kΩ, $R_1 = 250$ kΩ; $\beta = 50$; neglect V_{BE}.
[Ans. $I_E = I_C = 1.25$ mA; $I_B = 25$ μA; $V_{CE} = 6.25$ V]
3. Select the value of R_1 to set up the biasing condition such that $V_{CE} = 0.5V_{CC}$, for the following circuit components (refer to Fig. T. 7.2.1).
$V_{CC} = 30$ V; $R_C = 5$ kΩ; $\beta = 40$. **[Ans.** 200 kΩ]
4. Calculate the biasing point of the transistor (refer Fig. T. 7.2.1) for the following data.
$R_C = 5$ kΩ; $V_{CC} = 15$ V; $\beta = 100$, $R_B = 215$ kΩ; $V_{BE} = 0.7$ V
[Ans. $I_C = 2$ mA; $V_{CE} = 5$ V]
5. Calculate the new operating point if the transistor of Problem 4 is replaced by the other silicon PNP transistor having $\beta = 300$.
[Ans. $I_C = 2.5$ mA; $V_{CE} = 2.5$ V]

Tutorial Sheet 7.3

1. Calculate V_{CE} and I_C in the circuit of Fig. T. 7.3.1 if $V_{CC} = 9$ V; $R_B = 50$ kΩ; $R_C = 250$ Ω; $R_E = 500$ Ω and $\beta = 80$. **[Ans.** $V_{CE} = 3$ V, $I_C = 8$ mA]
2. Compute the Q point of the transistor (refer to Fig. T. 7.3.1) if $R_B = 400$ kΩ; $R_C = 2$ kΩ; $R_E = 1$ kΩ; $\beta = 100$ and $V_{CC} = 20$ V, neglecting V_{BE}. Mark the direction of I_C. **[Ans.** $I_C = 4$ mA, $V_{CE} = 8$ V]

3. The NPN transistor in the circuit given in Fig. T. 7.3.2 has a $\beta = 56$. Calculate the Q point if the following circuit components are used:

$$V_{CC} = 18\text{ V};\ R_B = 50\text{ k}\Omega;\ R_E = 0.75\text{ k}\Omega \text{ and } R_C = 500\ \Omega.$$

Assume $V_{BE} = 0.7$ V. **[Ans.** $I_C = 10.53$ mA, $V_{CE} = 4.83$ V]

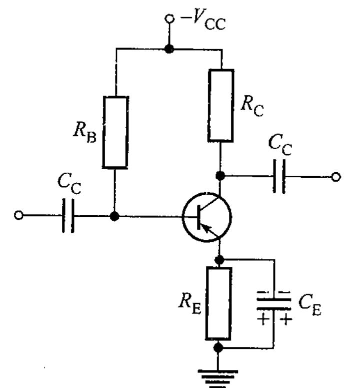


Fig. T. 7.3.1

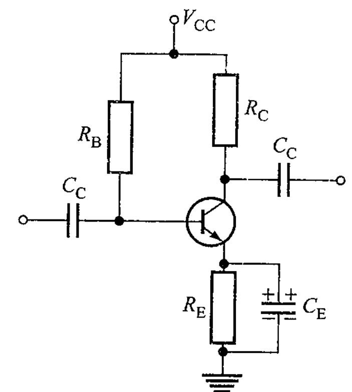


Fig. T. 7.3.2

4. Calculate the Q point for the transistor given in Fig. T. 7.3.2 for the given circuit parameters. $V_{CC} = 10$ V; $V_{BE} = 0.25$ V; $\beta = 80$; $R_B = 75$ kΩ; $R_C = 0.5$ kΩ and $R_E = 470\ \Omega$. **[Ans.** $I_C = 6.92$ mA, $V_{CE} = 3.29$ V]
5. A PNP transistor having a dc current gain in CE equal to 100 is to be biased at $I_C = 5$ mA and $V_{CE} = 3.8$ V. The collector load has a resistance of 500 Ω. If $V_{CC} = 10$ V and $V_{BE} = 0.3$ V, calculate the value of R_B and R_E (refer to Fig. T. 7.3.1). **[Ans.** $R_E = 740\ \Omega$, $R_B = 120$ kΩ]

Tutorial Sheet 7.4

Note

Use approximate method of solving biasing circuits unless specifically asked otherwise.

1. Calculate the collector current and collector-to-emitter voltage of the circuit given in Fig. T. 7.4.1 assuming the following circuit components and transistor specifications:

$R_1 = 40$ kΩ	$V_{BE} = 0.5$ V
$R_2 = 4$ kΩ	$\beta = 40$
$R_C = 10$ kΩ	$V_{CC} = 22$ V
$R_E = 1.5$ kΩ	

[Ans. $I_C \simeq I_E = 1$ mA, $V_{CE} = 10.5$ V]

2. Calculate the bias voltage, and currents for the PNP silicon transistor used in Fig. T. 7.4.2 assuming the following data:

$R_1 = 100$ kΩ	$V_{CC} = 12$ V
$R_2 = 27$ kΩ	$V_{BE} = 0.751$ V
$R_C = 2$ kΩ	$\beta = 75$
$R_E = 1$ kΩ	

[Ans. $I_C \simeq I_E = 1.8$ mA; $V_{CE} = 6.6$ V]

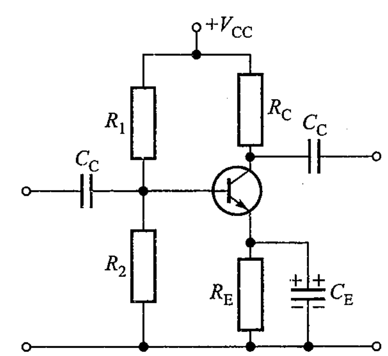


Fig. T. 7.4.1

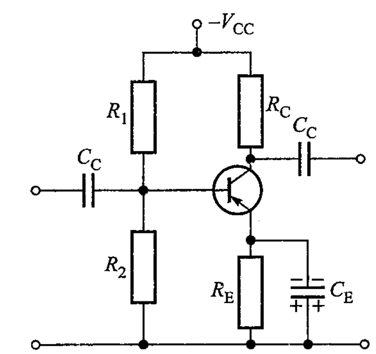


Fig. T. 7.4.2

3. Solve Problems 1 and 2 using an accurate method.
4. Calculate the value of resistors R_1 and R_L to place the Q point at $I_E = 2$ mA and $V_{CE} = 6$ V, in the two circuit of Fig. T. 7.4.3 and T. 7.4.4. In both circuits, a transistor of $\alpha = 0.985$, $I_{CBO} = 4$ μA and $V_{BE} = 200$ mV is used. The V_{CC} supply used is 16 V.

[Ans. (*a*) $R_1 = 5.54$ kΩ. $R_L = 3$ kΩ: (*b*) $R_1 = 56.3$ kΩ, $R_L = 4$ kΩ]

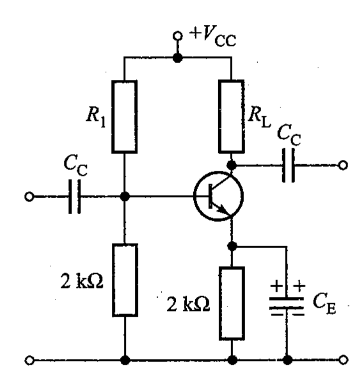


Fig. T. 7.4.3

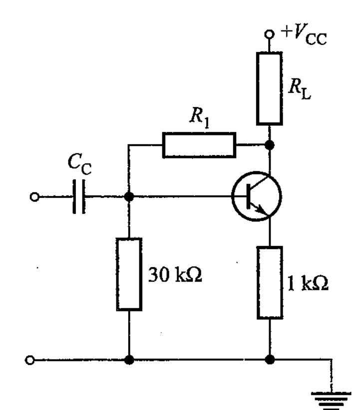


Fig. T. 7.4.4

• Experimental Exercise 7.1 •

Title Fixed-bias circuit with and without emitter resistor.

Objectives To

1. trace the given biasing circuit;
2. measure the Q point collector current and collector-to-emitter voltage with and without emitter resistor R_E;
3. note the variation of the Q point by increasing the temperature of the transistor in fixed-bias circuit with and without emitter resistor R_E;
4. note the variation in Q point by changing the base resistor in bias circuit, when emitter resistor is present and not present;

Apparatus Required Experimental board, electronic multimeter, milliammeter, power supply unit.

Circuit Diagram As given in Fig. E. 7.1.1. (typical values of components are also given).

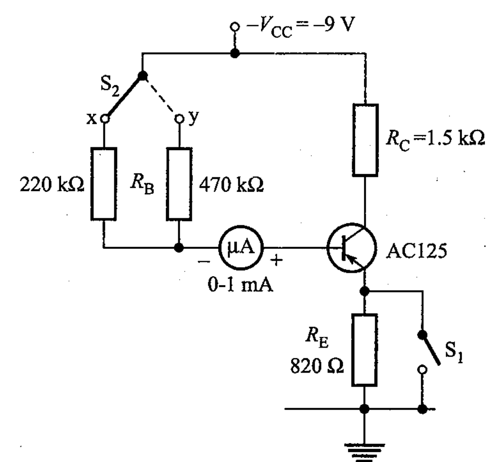


Fig. E. 7.1.1

Brief Theory The biasing circuit when the emitter resistance R_E is not present, is generally referred to as fixed-bias circuit (i.e., when switch S_1 is in closed position). In this circuit, we have

$$I_B = \frac{(V_{CC} - V_{BE})}{R_B} \cong \frac{V_{CC}}{R_B} \tag{1}$$

and

$$V_{CE} = V_{CC} - I_E R_C \tag{2}$$

When the temperature of the collector-to-base junction increases, the leakage current I_{CEO} increases. Since

$$I_C = \beta I_B + I_{CEO}$$

the operation point may go into the saturation region. Sometimes, thermal runaway may also take place. When the emitter resistor R_E is added in the circuit (i.e., when switch S_1 is in the open position) the operating point is given by

$$I_E = \frac{V_{CC} - V_{BE}}{R_E + R_B/\beta} \simeq \frac{V_{CC}}{R_E + R_B/\beta}$$

and

$$V_{CE} = V_{CC} - I_E R_C$$

If the temperature increases, the following sequence of events takes place.

$$T\uparrow \Rightarrow I_{CEO}\uparrow \Rightarrow I_C\uparrow \Rightarrow I_E\uparrow \Rightarrow R_E I_E\uparrow \Rightarrow V_E\uparrow \Rightarrow V_{BE}\downarrow \Rightarrow I_B\downarrow \Rightarrow I_C\downarrow$$

This shows that there is a tendency to make the operating point stable.

Procedure

1. Take the experimental board and identify the resistors R_B, R_C, and R_E. Also find the values of these resistors.
2. Apply $V_{CC} = 9$ V and close the switch S_1. Connect milliammeter and voltmeter. Note the values of I_C and V_{CE}.
3. Increase the temperature of the transistor (by rubbing your hands together and touching the transistor with one of the fingers; or by putting a lamp near it) and note the effect on collector current.
4. Now put off the switch S_1 so that the emitter resistor R_E comes in the circuit. Note the new operating emitter voltage.
5. Now increase the temperature and note the effect on the operating point.
6. Now change the switch S_2 in such a way that the base resistor R_B changes. With the new value of base resistor, repeat the above experiment.

Observations

1. When the switch S_1 is in closed position, i.e., R_E is not the circuit. Assume $\beta = 100$ for transistor.

S. No.	V_{CC}	R_B	I_C		V_{CE}	
			Theor.	Pract.	Theor.	Pract.
1.						
2.						

2. When the switch S_1 is in the open condition, i.e. when R_E is present in the circuit.

Circuit Diagram As shown in Fig. E. 7.3.1 (Typical values of the components are also shown.)

Brief Theory Potential-divider bias circuit is a widely used biasing circuit in amplifiers. The most significant advantage of this circuit is that the operating point in circuit is almost independent of β. The expression for emitter current (which is also equal to collector current) is given by

$$I_E = \frac{\frac{V_{CC} \times R_2}{R_1 + R_2} - V_{BE}}{R_E}$$

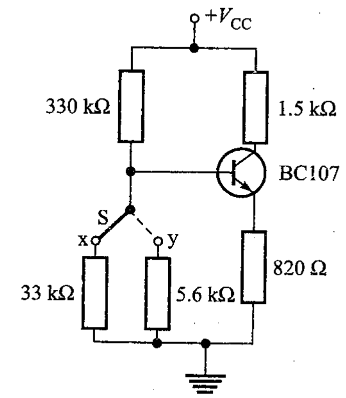


Fig. E. 7.3.1

The collector-to-emitter voltage is given by

$$V_{CE} = V_{CC} - (R_E + R_C)I_E$$

The operating point can be changed by changing one of the resistors of the potential-divider network. In the experiment, two values of the resistor R_2 are provided.

Procedure

1. From the given circuit, find out whether the transistor is PNP type or NPN type. Trace the circuit and note down the values of different resistors.
2. Connect the collector supply dc voltage in the circuit. Adjust the dc voltage to, say, 9 V.
3. Measure the collector supply voltage V_{CC}, the collector voltage V_C and collector-to-emitter voltage V_{CE}. Calculate the collector current by finding the voltage drop across the collector resistor R_C. This drop is $(V_{CC} - V_C)$ V.
4. Now put the switch S in the second position. Note the new Q point by measuring the collector current and collector-to-emitter voltage.

Observations

DC supply voltage $V_{CC} = 9$ V.

S. No	*(Base to ground resistor)* R_2	V_C	I_C	V_{CE}
1.				
2.				
3.				

Result When resistor R_2 connected between base and ground decreases, the collector current also decreases.

8 UNIT SMALL-SIGNAL AMPLIFIERS

"There's a basic principle about consumer electronics: it gets more powerful all the time and it gets cheaper all the time. That's true of all types of consumer electronics."

Trip Hawkins (1953–present)

Silicon Valley American entrepreneur and founder of Electronic Arts, The 3DO Company and Digital Chocolate

After completing this unit, students will be able to:

- draw a single-stage amplifier circuit (CE configuration)
- calculate the voltage gain of a single-stage amplifier when supplied with; (*a*) collector characteristics of the transistor; (*b*) values of different resistors used in the circuit; (*c*) value of the dc supply voltage; and (*d*) dynamic input resistance of transistor
- explain the phase relationship between the input and the output signal in a single-stage amplifier circuit
- calculate the voltage gain, input impedance and output impedance of a single-stage amplifier circuit when circuit parameters and transistor parameters, like β and r_{in} (dynamic input resistance), or the h-parameters, are given
- draw the circuit diagram of a single-stage field-effect transistor amplifier
- calculate the voltage gain of a single-stage field effect transistor amplifier by graphical as well as equivalent circuit method

8.1 INTRODUCTION

Almost no electronic system can work without an amplifier. Could the voice of a singer reach everybody in the audience in a hall, if the PA system (Public Address

System) fails? It is only because of the *enlargement* or the *amplification* of the signal picked up by microphone that we can enjoy a music orchestra. We are able to hear the news or the cricket commentary on our radio, simply because the amplifier in the radio *amplifies* the weak signals received by its antenna. The signal can only be of any use if it is amplified to give a suitable output (such as sound in radio, picture in TV, etc.).

In the previous chapter, the transistor circuits were analysed purely from the dc point of view. After a transistor is biased in the active region, it can work as an amplifier. We apply an ac voltage between the base and emitter terminals to produce fluctuations in the collector current. An amplified output signal is obtained when this fluctuating collector current flows through a collector resistor R_C. When the input signal is so weak as to produce *small* fluctuations in the collector current compared to its quiescent value, the amplifier is called *small-signal amplifier* (also "voltage amplifier"). Such an amplifier is used as the first stage of the amplifier used in receivers (radio and TV), CD players, stereos and measuring instruments.

8.2 SINGLE-STAGE TRANSISTOR AMPLIFIER

We have seen in the previous chapter that the voltage divider method of biasing is the best. The circuit is shown in Fig. 8.1*a*. Almost all amplifiers use this biasing circuit, because the design of the circuit is simple and it provides good stabilisation of the operating point. If this circuit is to amplify ac voltages, some more components must be added to it. The result is shown in Fig. 8.1*b*. We have added three capacitors.

The capacitors C_C are called the *coupling capacitors*. A coupling capacitor passes an ac signal from one side to the other. At the same time, it does not allow the dc voltage to pass through. Hence, it is also called a *blocking capacitor*. For instance, it is due to the capacitor C_C (connected between collector and output) in Fig. 8.1*b* that the output voltage v_o across the resistor R_o is free from the collector's dc voltage.

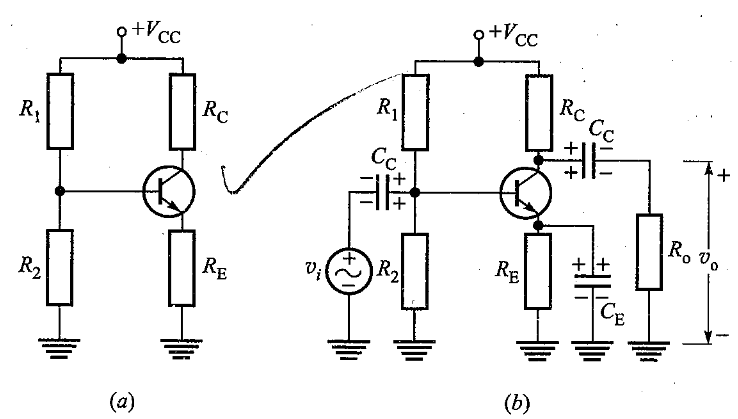


Fig. 8.1 (*a*) *Voltage-divider biasing circuit;* (*b*) *Same circuit converted into an amplifier*

The capacitor C_E works as a *bypass capacitor*. It bypasses all the ac currents from the emitter to the ground. If the capacitor C_E is not put in the circuit, the ac voltage developed across R_E will affect the input ac voltage. Such a feedback of ac signal is reduced by putting the capacitor C_E. If the capacitor C_E is good enough to provide an effective bypass to the lowest frequency of the signal, it will do so better to the higher frequencies. We, therefore, select such a value of capacitor C_E that gives quite a low impedance compared to R_E at the lowest frequency present in the input signal. As a practical guide, we make the reactance of the capacitor C_E at the lowest frequency, not more than one-tenth the value of R_E. That is

$$X_{CE} \leqslant \frac{R_E}{10} \tag{8.1}$$

The resistor R_o represents the resistance of whatever is connected at the output. Quite often, the amplification of the signal given by one amplifier may not suffice. More stages of amplifiers are then needed. The resistor R_o in Fig. 8.1*b* will then represent the input resistance of the next stage.

To what extent an amplifier enlarges signals is expressed in terms of its *voltage gain*. The voltage gain of an amplifier is given as

$$A_v = \frac{\text{Output ac voltage}}{\text{Input ac voltage}} = \frac{V_o}{V_i} \tag{8.2}$$

The other quantities of interest for a voltage amplifier are its current gain (A_i), input impedance (Z_i), and output impedance (Z_o). The amplifier can be analysed for its performance by the following two methods:

1. Graphical method
2. Equivalent circuit method

8.3 GRAPHICAL METHOD

For analysing an amplifier by this method, we need the output characteristics of the transistor. These are supplied by the manufacturer. When the ac voltage is applied to the input, the base current varies. The corresponding variations in collector current and collector voltage can be seen on the characteristics. This method involves no approximating assumptions. Hence, the results obtained by this method are more accurate than the equivalent circuit method. One can also visualise the maximum ac voltage that can be properly handled by this amplifier. In fact, for *large-signal amplifiers* (power amplifiers) this is the only suitable method.

8.3.1 Is dc Load Line Same as ac Load Line?

In the amplifier circuit of Fig. 8.1, the resistors R_1 and R_2 form a voltage divider arrangement for fixing a certain dc base voltage. This base voltage and the resistor R_E fix the emitter current. The collector current is almost the same as the emitter current. The resistor R_C then decides the value of V_{CE}. Writing the KVL equation for the output section of the circuit, we get

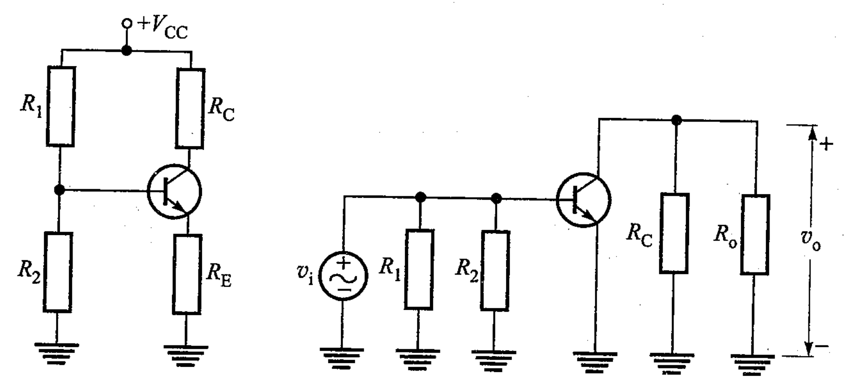
$$V_{CC} = I_C R_C + V_{CE} + I_E R_E$$

$$= V_{CE} + I_C(R_C + R_E) \quad [\text{since } I_C \simeq I_E]$$

or

$$I_C = \frac{-1}{(R_C + R_E)} V_{CE} + \frac{V_{CC}}{(R_C + R_E)} \tag{8.3}$$

This is the equation of the dc load line. By plotting this line on the output characteristics, the dc collector voltage and current can be determined for the given value of base current. As regards the dc currents and voltages, the amplifier circuit of Fig. 8.1*b* behaves like the circuit shown in Fig. 8.2*a*. This is obtained by opening all the capacitors in the original circuit. The capacitors are as good as open circuits for dc.



(*a*) DC behaviour (*b*) AC behaviour

Fig. 8.2 *Amplifier circuit of Fig. 8.1b*

Suppose we had changed the dc bias, giving a different value of base current. The collector current and collector voltage both will change. As a result, the Q point will shift on the dc load line. This is what roughly happens when we apply an input ac signal. But, in the ac signals, the variations occur very fast. The capacitors can no longer be considered as an open circuit. In fact, the variations in the currents and voltages occur so fast that the capacitors in the circuit may be treated as short circuits. Also while dealing with ac currents and voltages, we need not consider the dc supplies. If

we do this, the original circuit of Fig. 8.1*b* reduces to the one shown in Fig 8.2*b*. This circuit explains the behaviour of the amplifier from the ac point of view. You may now see that the resistor R_C comes in parallel with R_o and forms the ac load for the amplifier. The variation in the collector current and voltage are seen with the help of the ac load line corresponding to this ac load. How the ac load line is drawn is made clear in the next section.

8.3.2 Calculation of Gain

To understand how to calculate the current gain and voltage gain by the graphical method, we consider a typical amplifier circuit. One such circuit is shown in Fig. 8.3. The output characteristics of the transistor used in this circuit, are shown in Fig. 8.4.

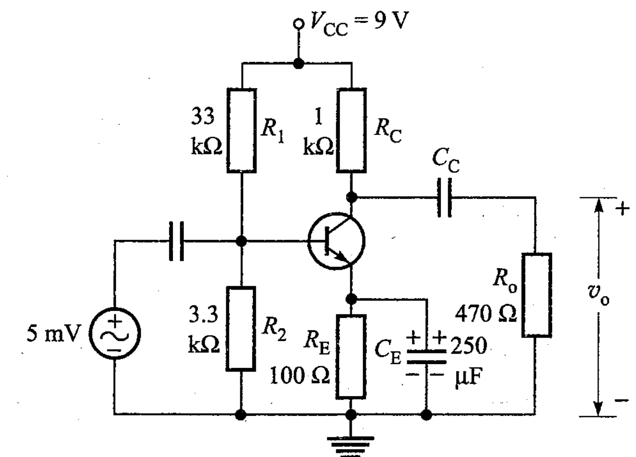


Fig. 8.3 *Typical transistor amplifier circuit*

We first plot the dc load line on the output characteristics. The equation of this dc load line is given by Eq. (8.3). This line is drawn by simply joining the points (V_{CC}, 0) and (0, V_{CC}/R_{dc}). It may be seen that the slope of this line is $-1/R_{dc}$, where the dc load, $R_{dc} = R_C + R_E$. In this case, $V_{CC} = 9$ V; and $R_{dc} = R_C + R_E = 1\ \text{k}\Omega + 0.1\ \text{k}\Omega = 1.1\ \text{k}\Omega$. Thus, the two points for plotting the dc load line are (9 V, 0) and (0, 8.2 mA). Let us assume that the biasing arrangement is such that the dc base current is 30 μA. The quiescent operating point Q is given by the intersection of the dc load line and the output characteristic corresponding to $I_B = 30$ μA. The dc collector current and collector-to-emitter voltage of the Q point may be seen to be 4 mA and 4.5 V, respectively.

When we apply an ac input signal V_i, the circuit behaves like the one shown in Fig. 8.2*b*. It is clear that for ac, the load resistance is R_C in parallel with R_o. This is the ac load R_{ac} for which the load line should be plotted. In our case, $R_{ac} = 1\ k\Omega \parallel 470\ \Omega = 320\ \Omega$. The ac load line will have a slope of $-1/R_{ac}$. Since the Q point describes the zero-signal conditions of the circuit, the ac load line should also pass through Q point. To draw such a line, we can first draw any line AB with the given slope. We can then draw the ac load line parallel to this line and passing through the Q point.

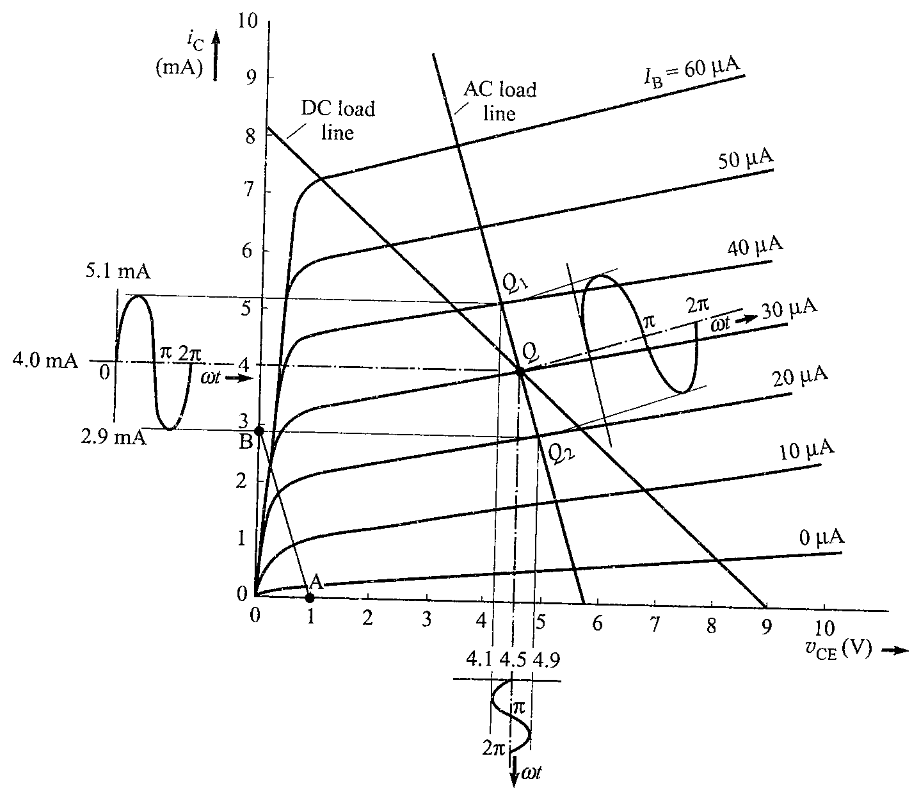


Fig. 8.4 *Calculation of gain by graphical method*

Figure 8.5 explains how a line with the given slope $-1/R_{ac}$ is drawn.

$$\text{Slope} = -\frac{1}{R_{ac}} = \tan\theta = \tan(180° - \alpha) = -\tan\alpha$$

$$\therefore \quad \tan\alpha = \frac{1}{R_{ac}} = \frac{OB}{OA}$$

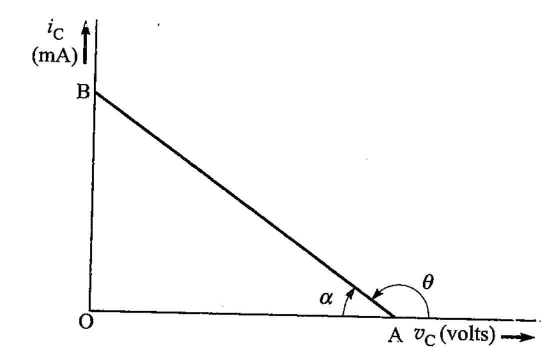


Fig. 8.5 *To draw a line with a slope of $-1/R_{ac}$*

Let us take OA = 1 V. Then

$$OB = \frac{OA}{R_{ac}} = \frac{1}{320} = 0.0031\ A = 3.1\ mA$$

After locating the points B and A in Fig. 8.4, line AB whose slope = $1/R_{ac}$ can be drawn.

Suppose an ac voltage of 5 mV is applied at the input. This corresponds to a peak-to-peak variation of $5 \times \sqrt{2} \times 2 = 14.14$ mV. Assume that the input characteristics of the transistor are such as to produce a 20 μA peak-to-peak variation in the base current corresponding to this input ac voltage. When the base current varies within these limits (from 20 μA to 40 μA), the instantaneous operating point moves along the ac load line between points Q_1 and Q_2. The corresponding variations in collector current and collector-to-emitter voltage are shown in Fig. 8.4. The collector current varies between the limits 2.9 mA to 5.1 mA. The collector-to-emitter voltage varies between the limits 4.1 V to 4.9 V. The current gain and the voltage gain of the amplifier are given as:

$$\text{Current gain} = \frac{I_{C(max)} - I_{C(min)}}{I_{B(max)} - I_{B(min)}} = \frac{(5.1 - 2.9)\ \text{mA}}{(40 - 20)\ \mu\text{A}} = \mathbf{110}$$

$$\text{Voltage gain} = \frac{V_{CE(max)} - V_{CE(min)}}{V_{i(max)} - V_{i(min)}} = \frac{(4.9 - 4.1)\ \text{V}}{14.14\ \text{mV}} = \mathbf{56.58}$$

8.3.3 Are Input and Output in Same Phase?

In Fig. 8.4 observe the waveforms of the base current, the collector current and the collector-to-emitter voltage. When the input voltage increases, the base current also increases. The instantaneous Q point moves towards Q_1; as a result, the current I_C increases, but the voltage V_{CE} reduces. For clear understanding of amplifier operation, the variations of input voltage, base current, collector current, collector-to-emitter (output) are again drawn in Fig. 8.6. From these diagrams, it is clear that *the input voltage and output voltage are out of phase by 180°.*

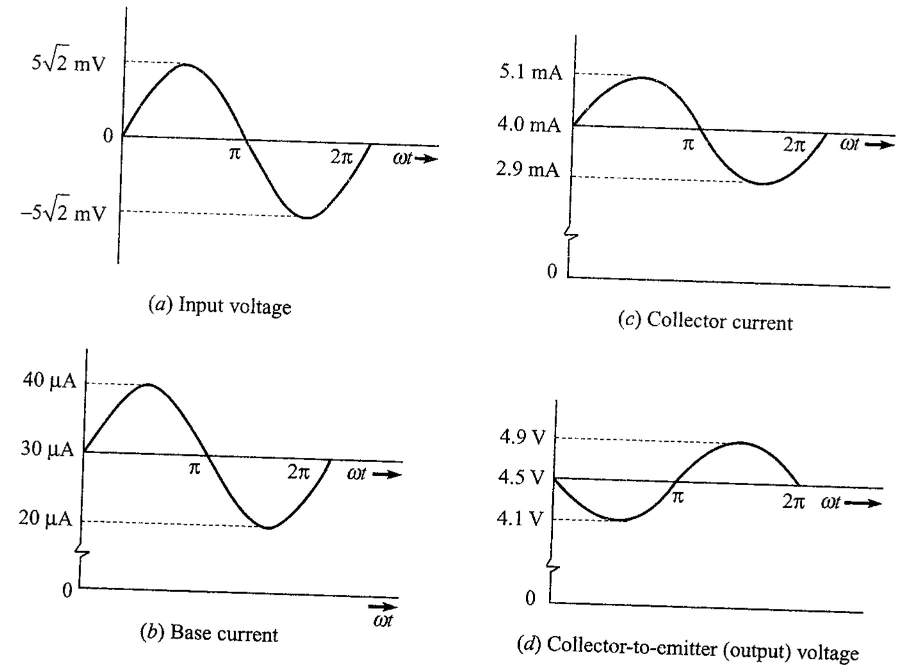


Fig. 8.6 *Phase relationships between input and output*

8.4 EQUIVALENT CIRCUIT METHOD

Our main concern in analysing an amplifier circuit is to determine its ac behaviour. We are interested in calculating the ac current gain, voltage gain, input impedance and output impedance. For this purpose, the given amplifier circuit is converted into its equivalent circuit from the ac point of view. All the capacitors and the dc supplies are replaced by short circuits. The CE amplifier circuit of Fig. 8.1*b* then reduces to the form of Fig. 8.2*b*. In the equivalent circuit method of analysis of the amplifier, the transistor is also replaced by its ac equivalent.

8.4.1 Development of Transistor AC Equivalent Circuit

Figure 8.7 shows the typical output characteristics of a transistor in the CE mode. The curves are almost horizontal. For a given value of base current, the collector current hardly depends upon the value of the collector-to-emitter voltage. Keeping I_B constant, the change in I_C corresponding to certain change in V_{CE} is very small. It means that the output section of the transistor offers very high dynamic resistance. The transistor, therefore, can be replaced by a current source between its output terminals. This is shown in Fig. 8.8.

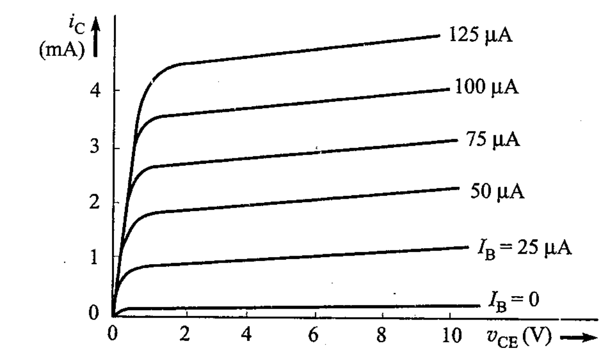


Fig. 8.7 *Typical output characteristics of a transistor in CE mode*

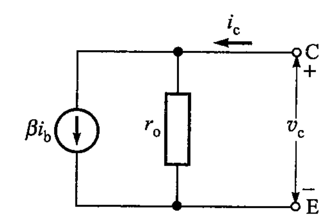


Fig. 8.8 *Transistor equivalent circuit between collector and emitter*

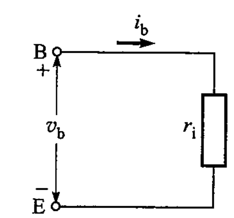


Fig. 8.9 *Transistor equivalent circuit between base and emitter*

The current source βi_b depends, as it should, on the input ac current i_b and the current amplification factor β. The resistance r_o represents the dynamic output resistance of the transistor and its value is quite high (typically 40 kΩ).

In the input section, the emitter-base junction of the transistor is forward-biased. The input characteristic of the transistor is similar to that of a forward-biased diode. The junction, therefore, can be replaced by a resistance r_i. The value of this resistance is low (typically 800 Ω). The input section of the transistor therefore simply becomes the one shown in Fig. 8.9

The complete ac equivalent circuit of the transistor is obtained by combining the input and output section. This is shown in Fig. 8.10.

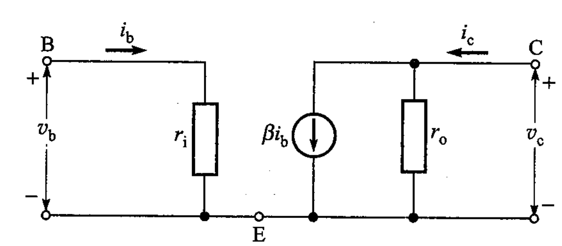


Fig. 8.10 *Transistor equivalent circuit*

8.4.2 *h*-parameter Equivalent Circuit

Quite often, the manufacturers specify the characteristics of a transistor in terms of its h parameters (the letter h stands for hybrid). The word *hybrid* is used with these parameters because they are a mixture of constants having different units. The hybrid parameters have become popular because they can be measured easily.

A transistor is a three-terminal device. If one of the terminals is common between the input and the output, there are two ports (pairs of terminals). See Fig. 8.11. For our purpose, the pair of terminals at the left represents the input terminals and the pair of terminals at the right, the output terminals. For each pair of terminals, there are two variables (current and voltage). There are a number of ways in which these four variables can be related. One of the ways, which is most frequently employed in transistor circuit analysis is as follows:

$$v_1 = h_{11}i_1 + h_{12}v_2 \quad (8.4)$$

$$i_2 = h_{21}i_1 + h_{22}v_2 \quad (8.5)$$

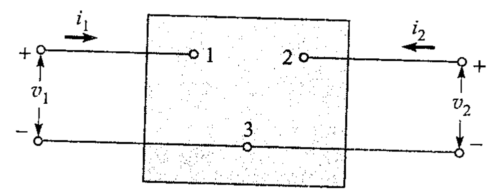


Fig. 8.11 *A transistor as a two-port network*

The parameter h_{11}, h_{12}, h_{21} and h_{22} which relate the four variables of the two-port system by the Eqs. (8.4) and (8.5) are called hybrid parameters. These parameters can be defined from the above equations by first putting $v_2 = 0$ (i.e., short-circuiting the output terminals) and then putting $i_1 = 0$ (i.e., opening the input terminals)

$$h_{11} = \left.\frac{v_1}{i_1}\right|_{v_2=0} = \text{Input impedance (with output shorted)} = h_i$$

$$h_{21} = \left.\frac{i_2}{i_1}\right|_{v_2=0} = \text{Forward current ratio (with output shorted)} = h_f$$

$$h_{12} = \left.\frac{v_1}{v_2}\right|_{i_1=0} = \text{Reverse voltage ratio (with input open)} = h_r$$

$$h_{22} = \left.\frac{i_2}{v_2}\right|_{i_1=0} = \text{Output admittance (with input open)} = h_o$$

It may be noted that h_i, being the ratio of voltage to current, has units of Ω. Similarly, h_o being the ratio of current to voltage has units of siemens (earlier known as mhos). However, h_f and h_r being the ratio of similar quantities are pure numbers having no units. Thus, the parameters are hybrid in nature.

An additional suffix e is added to the symbols of the h parameters to indicate that the transistor is used in the CE mode. In this mode, the terminal 1 is the base terminal and terminal 2 is the collector. Therefore, v_1 and i_1 become v_b and i_b, respectively, and at the output port v_2 and i_2 become v_c and i_c, respectively. With this understanding, Eqs. (8.4) and (8.5) can be written as

$$v_b = h_{ie}i_b + h_{re}v_c \quad (8.6)$$

$$i_c = h_{fe}i_b + h_{oe}v_c \quad (8.7)$$

Since each term of Eq. (8.6) has the units of volts, we can use Kirchhoff's voltage law to find a circuit that 'fits' this equation. The result is shown in Fig. 8.12. Similarly, we observe that each term of Eq. (8.7) has the units of current. Using Kirchhoff's current law, we get the circuit shown in Fig. 8.13 to fit this equation. Combining both of these figures, we get Fig. 8.14. This circuit satisfies both the Eqs. (8.6) and (8.7); and therefore this is the complete ac equivalent circuit of the transistor using h parameters.

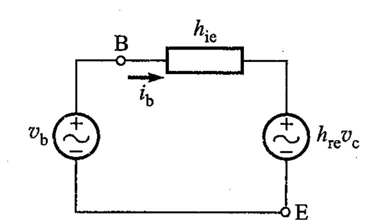


Fig. 8.12 *Hybrid input equivalent circuit*

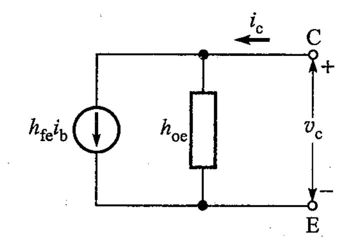


Fig. 8.13 *Hybrid output equivalent circuit*

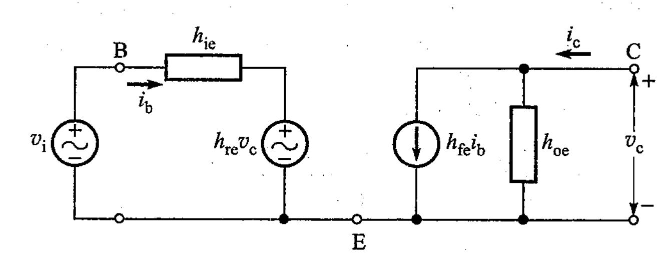


Fig. 8.14 *Complete hybrid equivalent circuit of a transistor*

Let us now compare the hybrid equivalent circuit of Fig. 8.14 with the one developed in Fig. 8.10. We find that

$$h_{ie} = r_i, \text{ the dynamic input resistance}$$

$$h_{fe} = \beta, \text{ the current amplification factor}$$

$$1/h_{oe} = r_o, \text{ the dynamic output resistance.}$$

The only difference in the two circuits is the presence of a voltage source $h_{re}v_c$ in the input of the hybrid model. The magnitude of this voltage source depends upon the output voltage v_c. The parameter h_{re}, therefore, represents a "feedback" of the output voltage to the input circuit. In the normal operation of the transistor, this effect is very small. It will make practically no difference if we neglect the term $h_{re}v_c$ from the hybrid equivalent circuit. The typical values of h parameters are

$$h_{ie} = 1\ \text{k}\Omega$$
$$h_{re} = 2.5 \times 10^{-4}$$
$$h_{fe} = 50$$
$$h_{oe} = 25\ \mu\text{S} \quad (\text{or, } 1/h_{oe} = 40\ \text{k}\Omega)$$

The h parameters at a given operating point can be determined from the static characteristics of the transistor, as illustrated in Example 8.1

Example 8.1 Determine the hybrid parameters from the given transistor characteristics (Figs. 8.15 and 8.16) at an operating point, $I_C = 2.15$ mA, and $V_{CE} = 8.5$ V.

Solution: On the collector characteristics of Fig. 8.15*a*, draw a vertical line corresponding to $V_{CE} = 8.5$ V. Draw a horizontal line corresponding to $I_C = 2$ mA. The intersection of these two lines fixes the operating point. This is marked as Q. Note that this Q point lies in the middle of the two characteristic curves corresponding to base currents $i_{B1} = 10\ \mu$A and $i_{B2} = 20\ \mu$A. This indicates that the base current at the operating point is 15 μA. An additional characteristic curve for $I_B = 15\ \mu$A is drawn.

Refer to Fig. 8.15*a*. At constant V_{CE} of 8.5 V, if i_B changes, say, by a small amount around the Q point from 10 μA to 20 μA, the collector current changes from 1.7 mA to 2.7 mA. Therefore,

$$h_{fe} = \left.\frac{\Delta i_C}{\Delta i_B}\right|_{V_{CE}=\text{const.}} = \left.\frac{(2.7-1.7)\times10^{-3}}{(20-10)\times10^{-6}}\right|_{V_{CE}=8.5\ \text{V}}$$

$$= \frac{10^{-3}}{10\times10^{-6}} = \mathbf{100}$$

Refer to Fig. 8.15*b*. At constant I_B of 15 μA, suppose the voltage v_{CE} changes around the Q point from 7 to 10 V. The corresponding change in collector current is from 2.1 mA to 2.2 mA. Therefore,

$$h_{oe} = \left.\frac{\Delta i_C}{\Delta v_{CE}}\right|_{I_B=\text{const.}} = \left.\frac{(2.2-2.1)\times10^{-3}}{10-7}\right|_{I_B=15\ \mu\text{A}}$$

$$= \frac{0.1\times10^{-3}}{3} = 33\ \mu\text{A/V} = \mathbf{33\ \mu S}$$

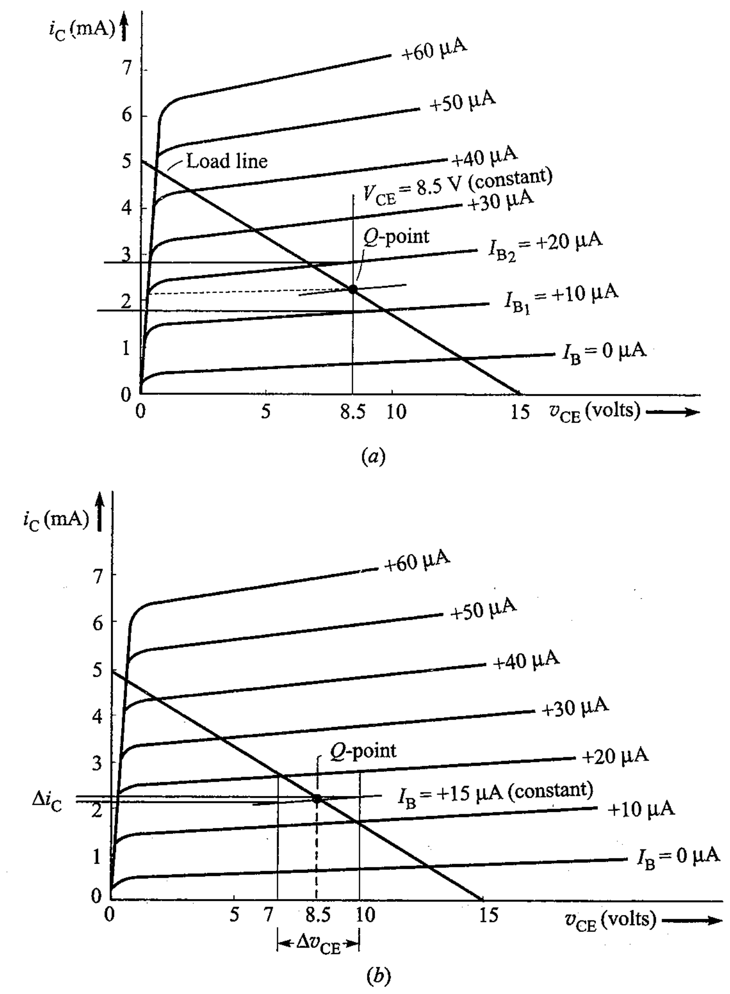


Fig. 8.15 *Collector characteristics of a transistor for the calculation of h_{fe} and h_{oe}*

To determine the parameters h_{ie} and h_{re}, we first fix the Q point on the input characteristics, as shown in Fig. 8.16. As shown in Fig. 8.16*a*, an additional curve corresponding to $V_{CE} = 8.5$ V is drawn. A small change in v_{BE} is then chosen, resulting in a corresponding change in i_B. We may then calculate h_{ie} as follows:

$$h_{ie} = \left.\frac{\Delta v_{BE}}{\Delta i_B}\right|_{V_{CE}=\text{const.}} = \left.\frac{0.730-0.715}{(20-10)\ \mu\text{A}}\right|_{V_{CE}=8.5\ \text{V}}$$

$$= \frac{0.015}{10\times10^{-6}} = 1.5\ \text{k}\Omega$$

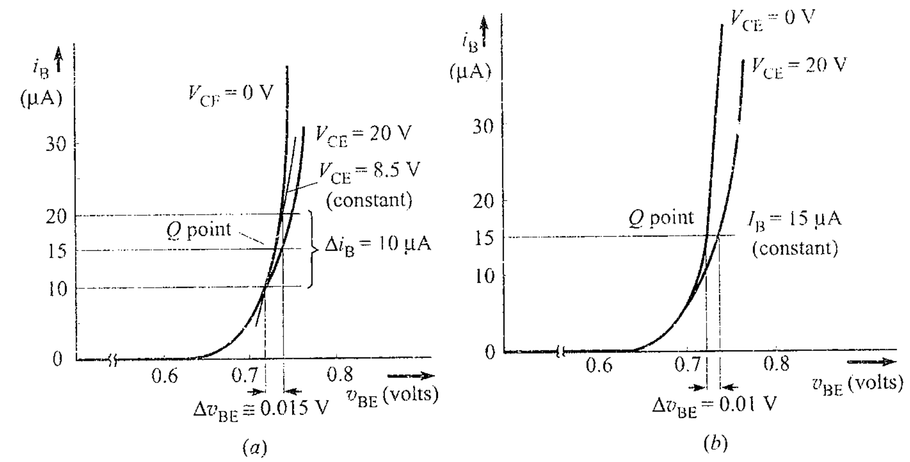


Fig. 8.16 *Input characteristics of a transistor for the calculation of h_{ie} and h_{re}.*

The last parameter h_{re} can be found by first drawing a horizontal line through the Q point of $I_B = 15$ μA. As shown in Fig. 8.16*b*, when v_{CE} changes from 0 V to 20 V, the corresponding change in v_{BE} is from 0.72 V to 0.73 V. Therefore,

$$h_{re} = \left.\frac{\Delta v_{BE}}{\Delta v_{CE}}\right|_{i_B = \text{const.}} = \left.\frac{(0.73 - 0.72)\text{ V}}{(20 - 0)\text{ V}}\right|_{I_B = 15\,\mu\text{A}}$$

$$= \frac{0.01}{20} = 5 \times 10^{-4}$$

The value of the parameter h_{re} is very small. Change in v_{BE} corresponding to a large change in v_{CE} is quite small. Such a small change in the voltage v_{BE} is difficult to determine from the graph. The value of h_{re} obtained from the graphical method may not be very accurate.

8.4.3 Amplifier Analysis

The CE transistor amplifier circuit of Fig. 8.1*b* was redrawn in Fig. 8.2*b* from the point of view of its ac behaviour. In equivalent circuit method of analysing the amplifier, the transistor is also replaced by its ac equivalent. Once the complete ac equivalent circuit is available to us, we can determine the current gain, voltage gain, input impedance and output impedance of the amplifier.

Figure 8.17 shows the complete ac equivalent circuit of the transistor amplifier of Fig. 8.1*b*. On the output side, the two resistors R_C and R_o can be replaced by a single resistor R_{ac} such that

$$R_{ac} = R_C \parallel R_o = \frac{R_C R_o}{R_C + R_o}$$

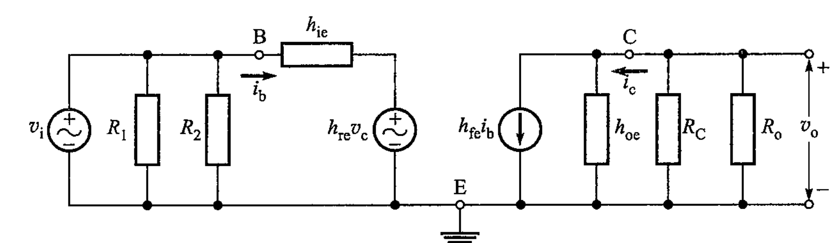


Fig. 8.17 *Complete ac equivalent circuit of the transistor amplifier*

On the input side, if the input voltage source v_i is assumed to be ideal (with zero internal resistance), the presence or the absence of the resistors R_1 and R_2 is immaterial. Whatever may be the values of R_1 and R_2 the current i_b remains the same. We can therefore ignore these resistors altogether. The result is the circuit of Fig. 8.18.

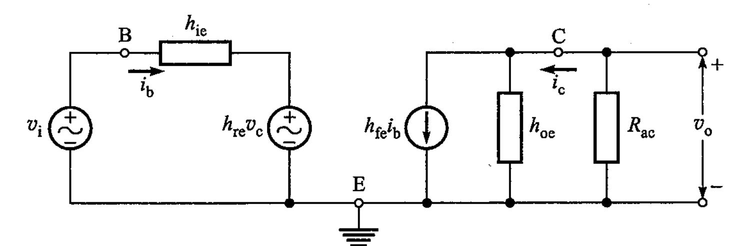


Fig. 8.18 *AC equivalent circuit where R_C and R_o are replaced by R_{ac} and the biasing resistors R_1 and R_2 are omitted*

We can now carry out the analysis of the amplifier using this ac equivalent circuit. Such an exact analysis is found to be very lengthy and tedious. Much effort may be saved if certain approximations are made. The results obtained from such an *approximate analysis* will not be much different from those obtained by the exact analysis.

For a typical transistor amplifier circuit R_{ac} is of the order of 1 kΩ. Whereas, $h_{oe} = 25$ μS, so that $1/h_{oe} = 40$ kΩ. As $1/h_{oe}$ is in parallel with R_{ac}, the equivalent resistance is $(1/h_{oe}) \parallel R_{ac} \simeq R_{ac}$, because $(1/h_{oe})$ is about 40 times greater than R_{ac}. Therefore, for the approximate analysis, we may omit h_{oe} from the equivalent circuit. The value of h_{re} is typically 1×10^{-4}. This means that the feedback voltage $h_{re}v_c$ is very small and therefore can be omitted from the equivalent circuit. With these approximations, the ac equivalent circuit of the amplifier becomes the one shown in Fig. 8.19.

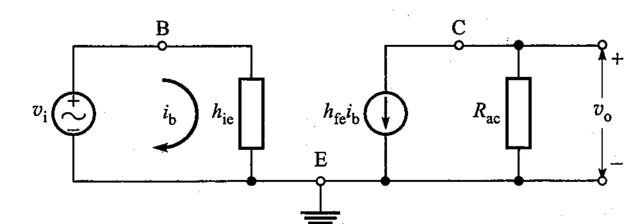


Fig. 8.19 *AC equivalent circuit for approximate analysis of the amplifier*

Current gain From Fig. 8.19, the output current i_c is seen to be the same as $h_{fe}i_b$. The current in the input is i_b. Therefore,

$$\text{Current gain, } A_i = \frac{\text{Output current}}{\text{Input current}} = \frac{i_c}{i_b} = \frac{h_{fe}i_b}{i_b} = h_{fe} \tag{8.8}$$

or

$$A_i = \beta$$

Voltage gain The output voltage $v_o = -h_{fe}i_bR_{ac}$. Note the negative sign. The flow of the output current is such that it makes the collector negative with respect to the ground. The input voltage $v_i = i_bh_{ie}$. Therefore,

$$\text{Voltage gain, } A_V = \frac{\text{Output voltage}}{\text{Input voltage}} = \frac{-h_{fe}i_bR_{ac}}{i_bh_{ie}} = \frac{-h_{fe}R_{ac}}{h_{ie}} \tag{8.9}$$

or simply,

$$A_V = \frac{\beta R_{ac}}{r_i}\angle 180^\circ$$

The angle 180° indicates that output and input voltages have a phase difference of 180°.

Power gain The power gain of the amplifier is simply the product of current gain and voltage gain. Thus, power gain,

$$A_p = A_iA_V \tag{8.10}$$

Input impedance The closed loop equation for input side is

$$v_i = i_bh_{ie}$$

Therefore,

$$\text{Input impedance, } Z_{in} = \frac{\text{Input voltage}}{\text{Input current}} = \frac{v_i}{i_b} = \frac{i_bh_{ie}}{i_b} = h_{ie} \tag{8.11}$$

or simply,

$$Z_{in} = r_i$$

In case, the biasing resistors R_1 and R_2 are to be considered, the input section of the equivalent circuit is as shown in Fig. 8.20. Then, the input impedance becomes

$$Z'_{in} = R_1 \,||\, R_2 \,||\, h_{ie} \simeq h_{ie} \quad (\text{Since } h_{ie} \text{ is much smaller than } R_1 \text{ or } R_2)$$

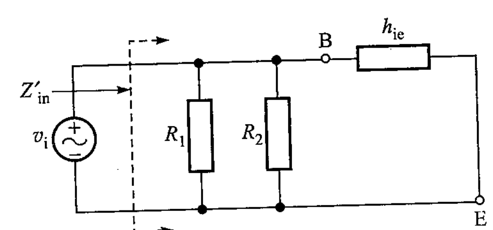


Fig. 8.20 *Input impedance when biasing resistors are considered*

Output impedance The output impedance of an amplifier is defined as the ratio of the output voltage to the output current with the input v_i set at zero. When we set

$v_i = 0$, the current i_b is also zero, and consequently the output current i_c also becomes zero even if we connect an external voltage source at the output terminals. In other words, the output impedance Z_o will be infinite. If we take into account R_{ac}, the output impedance (Fig. 8.21) Z'_o is simply R_{ac}. Had we not neglected h_{oe}, the output impedance would have been $Z_o = 1/h_{oe}$ and the impedance Z'_o would have been

$$Z'_o = (1/h_{oe}) \,||\, R_{ac} \simeq R_{ac} \tag{8.12}$$

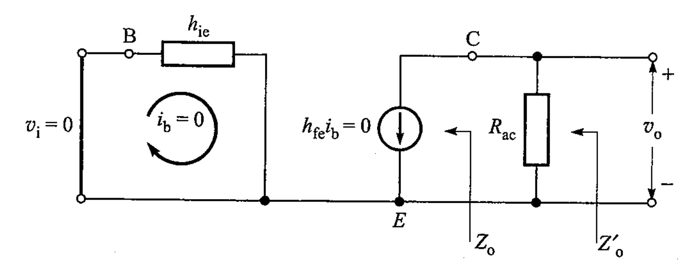


Fig. 8.21 *Output impedance of an amplifier*

Example 8.2 Figure 8.22 shows a common-emitter amplifier using fixed bias. Draw its ac equivalent circuit. Calculate its (*a*) input impedance; (*b*) voltage gain and (*c*) current gain. Assume $h_{ie} = r_i = 2$ kΩ, and $h_{fe} = \beta = 100$.

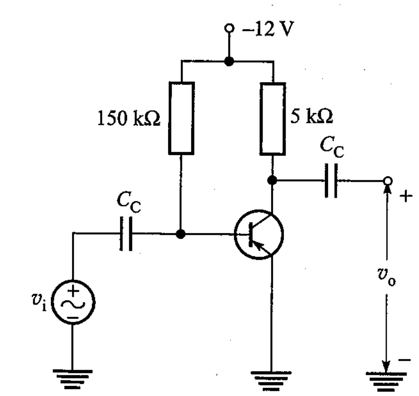


Fig. 8.22 *Common-emitter amplifier*

Solution: The ac equivalent circuit of the given amplifier circuit is shown in Fig. 8.23. It is assumed that the capacitors offer a short circuit at the signal frequency.

(*a*) As is clear in the equivalent circuit, the base resistor (150 kΩ) is much greater than the input resistance of the transistor (2 kΩ). Under this situation, the input impedance of the circuit may be taken as **2 kΩ**. (If exact calculations are made, this impedance is equal to 150 kΩ || 2 kΩ = 1.973 kΩ.)

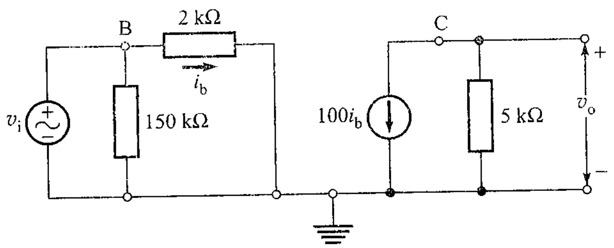


Fig. 8.23

(*b*) For calculation of voltage gain we make use of the formula,

$$A_v = \frac{\beta R_{ac}}{r_{in}} \angle 180°$$

Here β is 100, r_{in} is 2 kΩ and R_{ac} is 5 kΩ. Substituting these values, we get

$$A_v = \frac{100 \times 5 \times 10^3}{2 \times 10^3} \angle 180° = \mathbf{250 \angle 180°}$$

(*c*) Output current is equal to $100 i_b$, whereas input current is i_b. Therefore, current gain is **100**.

Example 8.3 In the single stage amplifier circuit shown in Fig. 8.24, an NPN transistor is used. The parameters of this transistor are $\beta_{ac} = 150$ and $r_{in} = 2\,\text{k}\Omega$. Calculate its (*a*) voltage gain; (*b*) input impedance and (*c*) *Q* point. (neglect V_{BE})

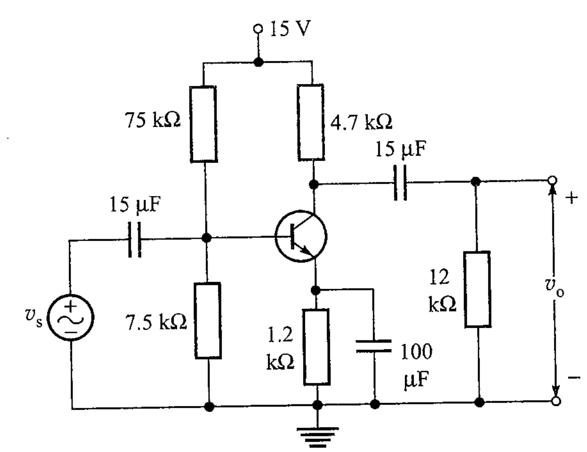


Fig. 8.24

Solution:

(*a*) In the amplifier circuit, the ac load resistance is

$$R_{ac} = 4.7\ \text{k}\Omega \parallel 12\ \text{k}\Omega = \frac{4.7 \times 10^3 \times 12 \times 10^3}{(4.7 + 12) \times 10^3}$$
$$= 3.38\ \text{k}\Omega$$

The voltage gain is given as

$$A_v = \frac{\beta_{ac} R_{ac}}{r_{in}} \angle 180°$$

Here, $\beta_{ac} = 150$; $r_{in} = 2$ kΩ and $R_{ac} = 3.38$ kΩ

$$\therefore \qquad A_v = \frac{150 \times 3.38 \times 10^3}{2 \times 10^3} = \mathbf{253.5 \angle 180°}$$

(*b*) As far as ac operation of the amplifier circuit is concerned, the resistors 75 kΩ and 7.5 kΩ are both connected between the base and ground. Hence, the input impedance of the amplifier is

$$Z_{in} = 75\ \text{k}\Omega \parallel 7.5\ \text{k}\Omega \parallel r_{in}$$
$$= 75\ \text{k}\Omega \parallel 7.5\ \text{k}\Omega \parallel 2\ \text{k}\Omega \simeq 7.5\ \text{k}\Omega \parallel 2\ \text{k}\Omega$$
$$= \mathbf{1.5\ k\Omega}$$

(*c*) The resistors 75 kΩ and 7.5 kΩ make a potential divider. The voltage at the base is given as

$$V_B = V_{CC} \frac{R_2}{R_1 + R_2} = \frac{15 \times 7.5 \times 10^3}{75 \times 10^3 + 7.5 \times 10^3} = \frac{15}{11} = \mathbf{1.36\ V}$$

Since, $V_{BE} = 0$, the voltage at the emitter is same as that at the base, i.e.,

$$V_E = V_B = 1.36\ \text{V}$$

Therefore, emitter current is

$$I_E = \frac{V_E}{R_E} = \frac{1.36}{1.2 \times 10^3} = \mathbf{1.13\ mA}$$

The collector-to-emitter voltage is

$$V_{CE} = V_{CC} - (R_C + R_E) I_E \qquad [\text{since } I_C \simeq I_E]$$
$$= 15 - (4.7 + 1.2) \times 10^3 \times 1.13 \times 10^{-3}$$
$$= \mathbf{8.33\ V}$$

8.5 FET SMALL-SIGNAL AMPLIFIER

Figure 8.25 shows a small-signal amplifier that uses an N-channel JFET. It uses self-bias arrangement* provided by R_S-C_S combination.

* Note that self-bias arrangement cannot be used for an EN MOSFET, as it needs positive dc bias voltage at its gate (for N-channel), instead of negative voltage. For an EN MOSFET, we use potential divider biasing circuit.

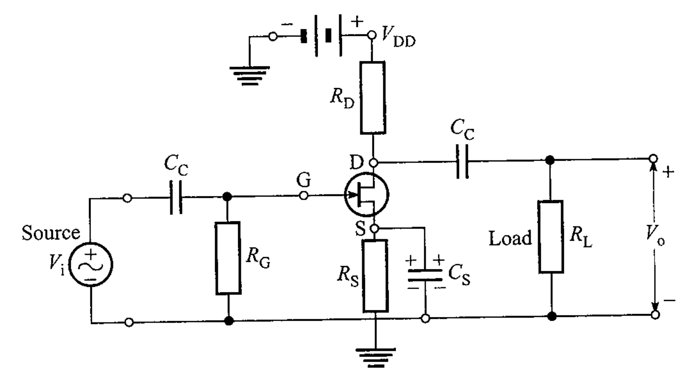


Fig. 8.25 *An FET small-signal amplifier*

8.5.1 AC Analysis of FET Amplifier

Our main concern in analysing an amplifier is to determine its response to an ac signal input. We are interested in finding the ac voltage gain, the current gain, the input resistance and the output resistance. Hence, we convert the given amplifier circuit into its ***ac equivalent circuit.***

The ac input signal varies so fast (i.e., its frequency is so high) that the capacitors effectively behave as short circuits. The voltage between the positive and

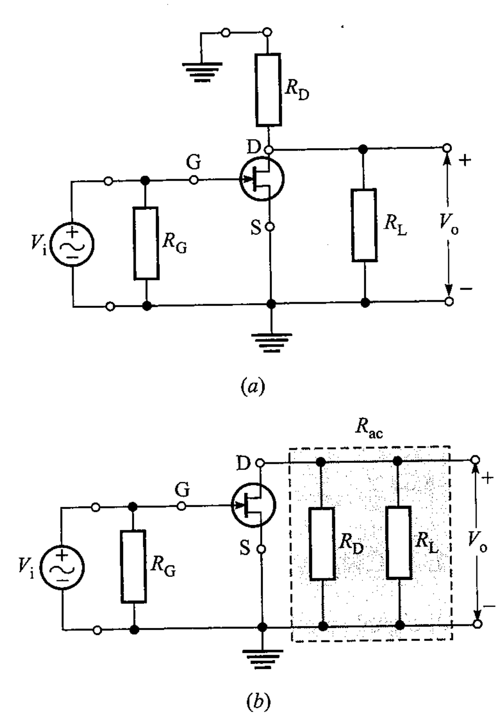


Fig. 8.26 *AC equivalent circuit of the FET amplifier of Fig. 8.25*

negative terminals of the dc supply V_{DD} does not vary (with time) even though the drain current does. Hence, from an ac point of view, these terminals are as good as short circuit. Thus, the ac equivalent circuit of the amplifier becomes that given in Fig. 8.26*a*. The same circuit has been redrawn differently in Fig. 8.26*b*. From the ac point of view, the resistor R_D is in parallel with the load resistor R_L. The ***ac load*** of the amplifier thus becomes

$$R_{ac} = R_D \parallel R_L = \frac{R_D R_L}{R_D + R_L} \tag{8.13}$$

The circuit of Fig. 8.26*b* would be easy to analyse, if we could replace the transistor by its circuit model.

8.5.2 AC Equivalent Circuit Model of an FET

Figure 8.27 shows the ac equivalent circuit model of an FET. Since there is almost no gate current, no resistor has been shown between the gate and the source terminals. At the input, the FET behaves just like an open circuit. At the output side, the FET has been replaced by a ***controlled current source***, $g_m V_{gs}$, in parallel with the resistor r_d. The value of the current source is controlled by the gate-to-source voltage, V_{gs}.

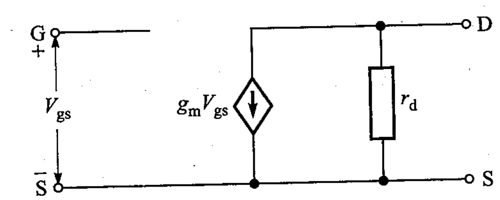


Fig. 8.27 *AC equivalent circuit model of an FET*

Note that the circuit of Fig. 8.27 is not valid for large-signal amplifiers. The parameters g_m and r_d can be assumed constant only if the signal is small.

8.5.3 AC Equivalent of FET Amplifier Circuit

By replacing the FET in the circuit of Fig. 8.26*b* with its equivalent circuit model, we get the ac equivalent of the complete circuit as shown in Fig. 8.28.

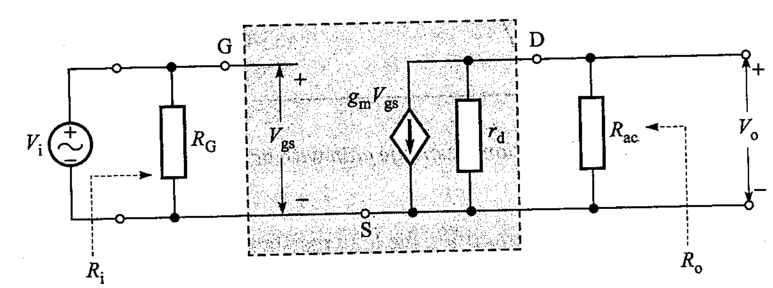


Fig. 8.28 *AC equivalent of the FET amplifier circuit*

It is obvious that $V_{gs} = V_i$. The output voltage V_o is determined by the current source and the effective resistance R_{eff} between the drain and the source,

$$V_o = -(g_m V_{gs})R_{eff} = -g_m V_i R_{eff}$$

where $$R_{eff} = r_d \parallel R_{ac} = r_d \parallel (R_D \parallel R_L)$$

The minus sign comes because the current $g_m V_{gs}$ flows through R_{eff} from bottom to top, making the top point negative. The ***voltage gain*** of the amplifier is

$$A_V = \frac{V_o}{V_i} = -g_m (r_d \parallel R_D \parallel R_L) \tag{8.14}$$

The minus sign reminds us that the output voltage is ***inverted*** with respect to the input. That is, *the output and input are 180° out of phase.*

The ***input resistance*** of the amplifier is

$$R_i = R_G \tag{8.15}$$

The ***output resistance*** of the amplifier is

$$R_o = R_L \parallel R_D \parallel r_d \tag{8.16}$$

Look at the expressions given in Eqs. 8.14 and 8.16. Normally r_d is of the order of 400 kΩ whereas R_D and R_L are only a few kilohms. We can therefore ignore r_d.

Maximum voltage gain The gain of a given FET amplifier varies with the ac load R_{ac}, as

$$A_v = -g_m (r_d \parallel R_{ac}) = -g_m \frac{r_d R_{ac}}{r_d + R_{ac}} = -g_m \frac{r_d}{r_d/R_{ac} + 1}$$

It increases with R_{ac}, as shown in Fig. 8.29. The maximum voltage gain that can be obtained from an FET is when $R_{ac} \to \infty$, and is given as

$$A_{vo} = -g_m r_d = -\mu \tag{8.17}$$

The gain A_{vo} is called open-circuit gain of the FET.

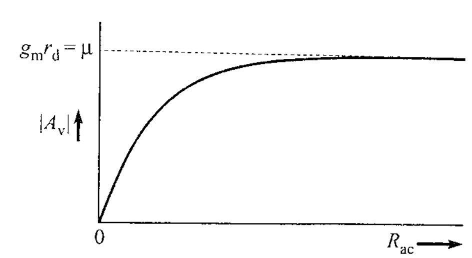


Fig. 8.29 *Variation of voltage gain with ac load resistance*

Example 8.4 In an FET amplifier, the load resistance $R_L = 12$ kΩ; $R_G = 1$ MΩ, $R_S = 1$ kΩ, $C_S = 25$ uF. The FET used has $\mu = 20$ and $r_d = 100$ kΩ. If the input signal voltage is 0.1 V at a frequency of 1 kHz, find the output signal voltage of the amplifier.

Solution: The reactance of the bypass capacitor C_S is

$$X_{CS} = \frac{1}{2\pi f C_S} = \frac{1}{2 \times 3.141 \times 1 \times 10^3 \times 25 \times 10^{-6}}$$
$$= 6.3\ \Omega$$

This is much smaller than $R_S = 1$ kΩ. We can assume R_S to be completely bypassed. The magnitude of the voltage gain of the amplifier is

$$A = \frac{\mu R_L}{R_L + r_d} = \frac{20 \times 12 \times 10^3}{12 \times 10^3 + 100 \times 10^3} = 2.14$$

Therefore the output signal voltage is

$$v_o = A v_i = 2.14 \times 0.1 = \mathbf{0.214\ V}$$

Review Questions

1. Draw the circuit diagram of a single-stage transistor amplifier. State the function(s) of each component used in this circuit.
2. Explain how amplified voltage becomes available at the output points of a single-stage amplifier.
3. Explain how phase reversal of the signal takes place when it is amplified by a single-stage voltage amplifier.
4. Draw an ac equivalent circuit of a common-emitter transistor amplifier. Derive the following expression using this equivalent circuit:
$$A_v = \frac{\beta R_{ac}}{r_{in}} \angle 180°$$
Explain with the help of equivalent circuit, the phase reversal of the signal.
5. State the name of the four *h* parameters for a transistor in CE configuration. Define them. Write down the typical values of these parameters.
6. A step-up transformer can increase the voltage level of an ac signal. This can also be achieved by a transistor voltage amplifier. Explain the difference in the two processes.
7. Explain the following terms in brief (say, within 5 lines), in connection with a transistor voltage amplifier:
   (*a*) Input impedance (Z_i)
   (*b*) Output impedance (Z_o)
   (*c*) Voltage gain
   (*d*) Current gain
   (*e*) Power gain
8. State what will happen to the voltage gain of an amplifier if the bypass capacitor (C_E) is open circuited.

9. Explain the difference between dc load line and ac load line. Why is it necessary to draw ac load line for calculating the voltage gain of an amplifier?
10. Why have you to draw dc load line while you calculate the gain from an ac load line?
11. Using the transistor characteristics and the load line, explain the phase reversal of the signal.

Objective-Type Questions

I. A number of statements are given below. Choose the correct statement.

1. The voltage gain of a transistor amplifier is a constant quantity and is independent of load resistance.
2. The voltage gain of a transistor amplifier increases as ac load resistance increases.
3. The voltage gain of a transistor amplifier in CE mode is always less than unity.
4. The input impedance of a transistor amplifier in CE configuration is very high (say, 5 MΩ).
5. The input impedance of a transistor amplifier in CE mode is low (say, 1.5 kΩ).
6. The input impedance of CE amplifier is extremely low (say, 10 Ω).
7. The output impedance of a transistor amplifier is independent of the transistor configuration.
8. The voltage gain of a transistor amplifier decreases when the emitter-bypass capacitor C_E is present in the circuit.
9. The voltage gain of a transistor amplifier using potential-divider biasing arrangement with emitter-bypass capacitor depends upon the value of R_E.
10. The phase reversal between output and input takes place only for voltage waves and not for current waves, in a transistor amplifier in CE configuration.

II. Below are some incomplete statements. Four alternatives are provided for each. Choose the alternative that completes the statement correctly.

1. In an amplifier, the coupling capacitors are used
   (*a*) to control the output
   (*b*) to limit the bandwidth
   (*c*) to match the impedances
   (*d*) to prevent dc mixing with input or output.
2. If the power gain of an amplifier is X and its voltage gain is Y, then its current gain will be
   (*a*) X/Y (*b*) Y/X
   (*c*) $X \cdot Y$ (*d*) $X+Y$

3. An amplifier circuit of voltage gain 100, gives 2 V output. The value of input voltage is
   (*a*) 200 V (*b*) 50 V
   (*c*) 20 mV (*d*) 2 mV
4. The input signal to an amplifier having a gain of 200 is given as 0.5 cos (313*t*). The output signal may be represented by
   (*a*) 100 cos (313*t* + 90°)
   (*b*) 10 cos (403*t*)
   (*c*) 100 cos (313*t* + 180°)
   (*d*) 200 cos (493*t*)

III. Fill in the blanks in the following sentences using the most appropriate alternative from those given in bracket.

1. For a good voltage amplifier, its input impedance should be ________ compared to the resistance of the source. (high/low/inductive/capacitive)
2. The coupling capacitors mainly affect ________ cut-off frequency of an amplifier. (lower/upper/single/double)
3. In a PNP transistor, the emitter resistor (R_E) keeps the emitter at a ________ voltage compared to its ground potential. (positive/negative/zero)
4. The output current waveform in CE amplifier is ________ with input current wave. (in phase/out of phase by 180°/out of phase by 90°)
5. The output voltage waveform of CE amplifier is ________ with its input voltage wave. (in phase/out of phase by 180°/leading by 90°/lagging by 90°)

Answers

I. 2, 5, 10
II. 1. (*d*) 2. (*a*) 3. (*c*) 4. (*c*)
III. 1. high 2. lower 3. negative 4. in phase
5. out of phase by 180°

Tutorial Sheet 8.1

1. The amplifier circuit given in Fig. T. 8.1.1 uses a transistor AC125. The collector characteristics of this transistor are given in Fig. T. 8.1.2. Locate the Q point if the quiescent base current is 15 μA. Draw the ac load line on the collector characteristics to determine the output voltage, if input base current swing is 5 μA (peak) sine wave. If the dynamic input resistance of the transistor is 800 Ω, calculate the voltage gain.
[**Ans.** $V_o = 0.95$ V (p-p); $A_v = 119$]

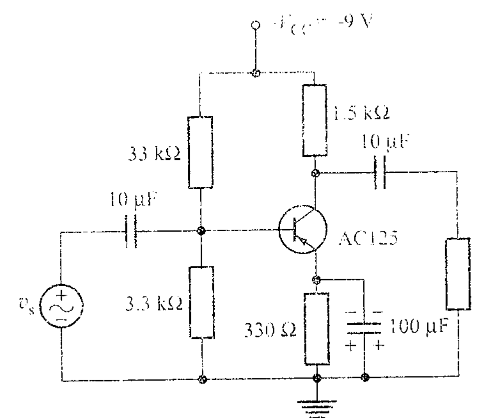


Fig. T. 8.1.1

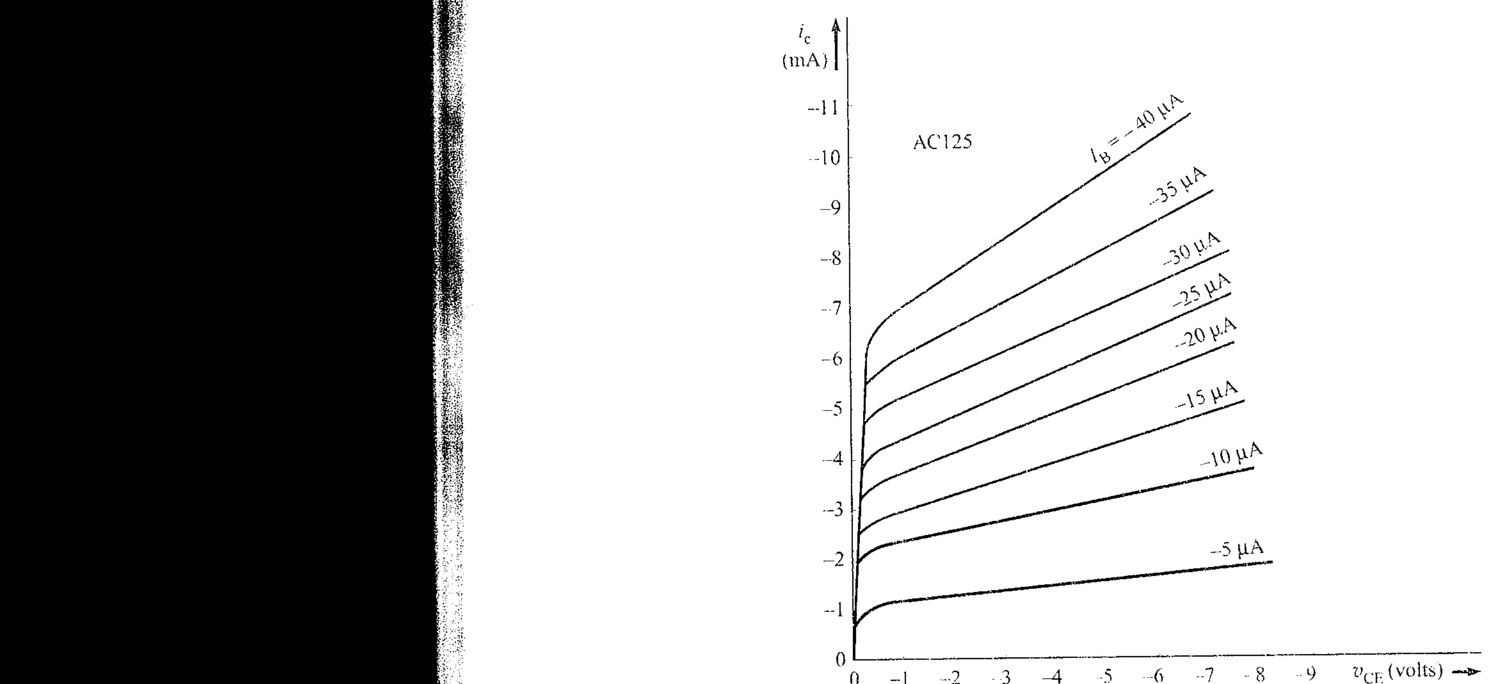


Fig. T. 8.1.2

2. Figure T. 8.1.3 shows the collector characteristics of a transistor used in an amplifier circuit. If the emitter resistor R_E is 200 Ω, determine the value of the collector resistance with the help of dc load line, drawn on the given characteristics. If the biasing resistors are such that the Q point base current is 40 µA, determine the collector current and collector-to-emitter voltage.

[**Ans.** R_C = 2.5 kΩ; I_{CQ} = 5.5 mA; V_{CE} = 10 V]

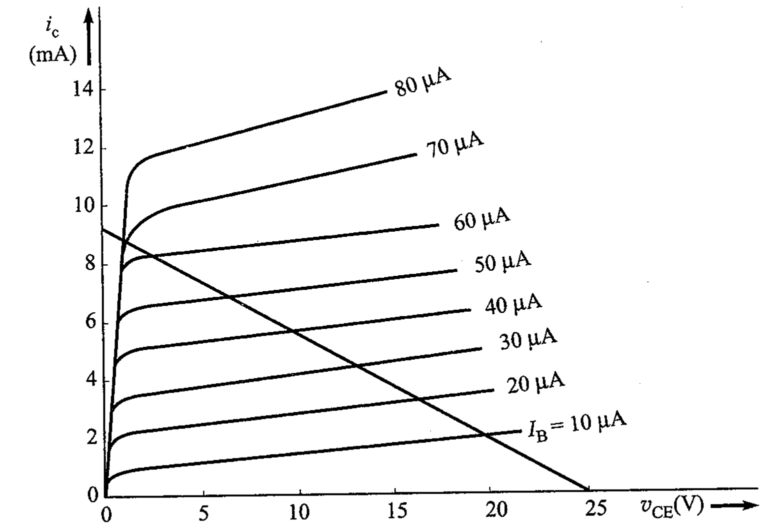


Fig. T. 8.1.3

3. If we connect a resistor across the output of the amplifier in Q. 2 such that the effective ac load resistance becomes 1 kΩ, draw the ac load line and calculate the voltage gain of the amplifier circuit assuming r_{in} to be 1 kΩ. Also calculate the value of the load resistor connected at the output points.

[**Ans.** A_V = 150; R_L = 1.67 kΩ]

• Tutorial Sheet 8.2 •

1. It is desired that the coupling capacitor C_C of Fig. T. 8.2.1 should couple all frequencies from 500 Hz to 1 MHz of the source to the output. Calculate the value of the coupling capacitor such that its impedance is not more than 10 % of the load impedance, at any frequency. [**Ans.** 0.32 µF]

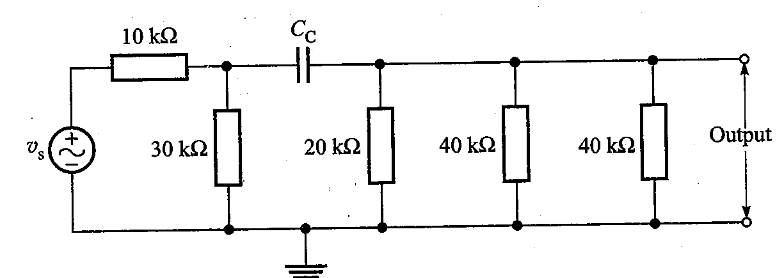


Fig. T. 8.2.1

2. The point A in Fig. T. 8.2.2 is desired to be effectively at the ground potential for the frequency range 40 Hz to 8 kHz. Calculate suitable value of bypass capacitor C_E. Assume that for effective bypassing, the impedance of the capacitor C_E should not be more than 10 % of the resistance R_E.

[**Ans.** 10 μF]

3. Work out the following quantities for the circuit given in Fig. T. 8.2.3: (*a*) ac emitter current; (*b*) ac voltages at emitter, base and collector; (*c*) voltage gain. Assume h_{ie} or r_{in} = 250 Ω.

[**Ans.** (*a*) i_e(peak) = 1.02 mA ≃ 1 mA.
(*b*) v_e = 0 V, v_b = 5 mV (peak);
v_c = 1 V (peak); (*c*) A_v = 200]

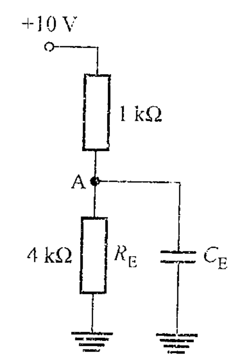


Fig. T. 8.2.2

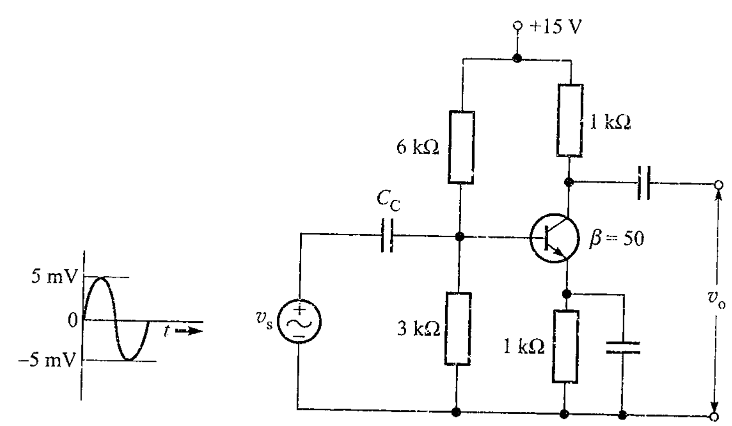


Fig. T. 8.2.3

4. In the single-stage amplifier circuit in Fig. T. 8.2.4, transistor AC126 is used. Draw its ac equivalent circuit and calculate the voltage gain (v_o/v_s) with and without R_L. Assume the following transistor parameters:

h_{fe} or β_{ac} = 150; h_{ie} or r_{in} = 1.5 kΩ

[**Ans.** 50, 100]

5. In the amplifier circuit of Fig. T. 8.2.4, calculate
(*a*) v_o/v_s, if the source resistance is 600 Ω.
(*b*) i_o/i_i, where i_o is the current through the output resistance R_L and i_i is the input current as shown.
(*c*) i_o/i_s, where i_s is the ac current supplied by the input source.

[**Ans.** (*a*) 29.85; (*b*) 75; (*c*) 44.44]

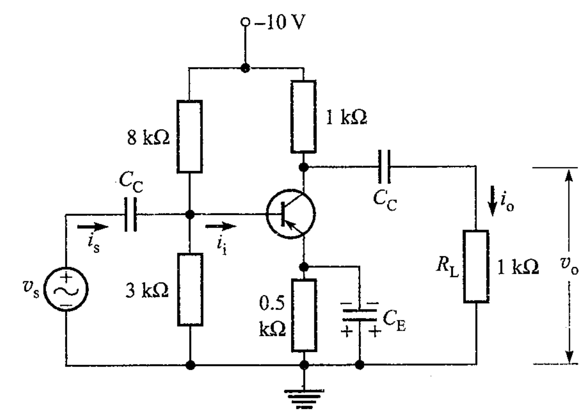


Fig. T. 8.2.4

Tutorial Sheet 8.3

1. A single-stage amplifier uses a JFET whose g_m = 2 mS and r_d = 50 kΩ. Find the gain of the amplifier, if the load resistance is 25 kΩ. [**Ans.** 33.3]
2. In an FET amplifier circuit, R_D = 10 kΩ, R_G = 1 MΩ, R_S = 1 kΩ and C_S = 25 μF. The FET used has μ = 20 and r_d = 100 kΩ. If the input signal voltage is 0.2 V at a frequency of 1 kHz, determine the output signal voltage.

[**Ans.** 0.364 V]

Experimental Exercise 8.1

Title Single-stage transistor amplifier.

Objectives To

1. trace the circuit diagram of single-stage transistor amplifier;
2. measure the *Q* point collector current and collector-to-emitter voltage;
3. measure the maximum signal which can be amplified by the amplifier without having clipped output;
4. measure the voltage gain of the amplifier at 1 kHz;
5. measure the voltage gain of the amplifier for different values of load resistance.

Apparatus Required Amplifier circuit, electronic multimeter, ac millivoltmeter, CRO.

Circuit Diagram As given in Fig. E. 8.1.1 Typical values of the components are also given.

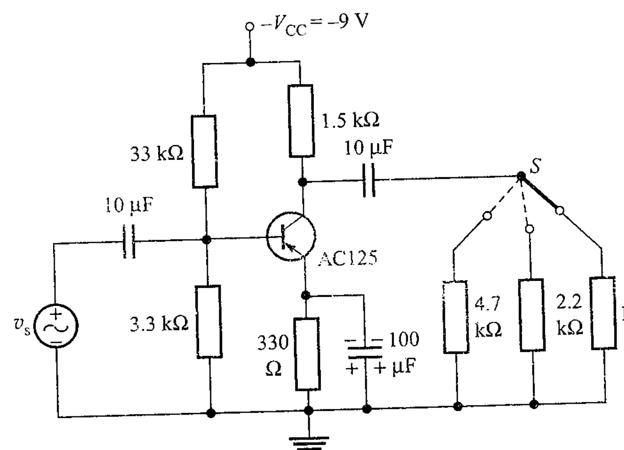


Fig. E. 8.1.1

Brief Theory In the amplifier circuit shown in the figure, the resistors R_1, R_2 and R_E fix a certain Q point. The resistor R_E stabilises it against temperature variations. The capacitor C_E bypass the resistor R_E for the ac signal. As it offers very low impedance path for ac, the emitter terminal is almost at ground potential. When the ac signal is applied to the base, the base-emitter voltage changes, because of which the base-current changes. Since collector current depends upon the base current, the collector current also changes. When this changing collector current passes through the load resistance R_C, an ac voltage is produced at the output. As the output voltage is much more than the input voltage, the circuit works as an amplifier circuit. The voltage gain of this amplifier is given by the formula

$$A_v = \frac{\beta R_{ac}}{r_{in}} \angle 180°$$

where r_{in} is the dynamic input resistance, β is the current amplification factor, and R_{ac} is the ac load resistance in the circuit.

Procedure

1. Look at the circuit and draw it accordingly in your notebook. With the help of the colour code, find the values of every resistor. Note the values of capacitors also.
2. Connect the dc supply V_{CC} (either from the regulated transistorised power supply or from IC power supply). Measure the dc voltage supplied.
3. For the measurement of quiescent collector current, measure the voltage of collector terminal with reference to ground (V_C). Calculate collector current form the formula

$$I_C = \frac{V_{CC} - V_C}{R_C}$$

Also measure V_{CE}, i.e., dc voltage between the collector and the emitter.
4. Make sure that the transistor is operating in the active region by noting that V_{CE} is about half of V_{CC}. Feed ac signal at 1 kHz at the input of the amplifier. Observe the amplified output on the CRO. Increase the input signal till the output waveshape starts getting distorted. Measure this input signal. This is the maximum signal that the amplifier can amplify without giving distorted output.
5. Now feed an ac signal that is less than the maximum signal handling capacity of the amplifier. Fix the frequency of the input signal at 1 kHz. Note the input and output voltages and calculate the voltage gain.
6. Connect different load resistors and find the voltage gain of the amplifier for each.

Observations

1. *Q point of the amplifier*:

V_{CC}	V_C	$V_{CC} - V_C$	$I_C = \frac{V_{CC} - V_C}{R_C}$	V_{CE}

2. Maximum signal that can be handled by the amplifier without introducing distortion = __________ mV. Frequency of the input signal = 1 kHz.
3. *Voltage gain of the amplifier*:

S. No.	*Load resistor*	*Input voltage*	*Output voltage*	*Gain* = $\frac{v_o}{v_i}$

Result

1. Q point of the transistor is

$$I_C = \_\_\_\_ \text{ mA}, V_{CE} = \_\_\_\_ \text{ V}$$

Since $V_{CE} \simeq \frac{1}{2} V_{CC}$, the transistor is biased in the middle of active region.

2. Maximum signal handing capacity of the amplifier (at 1 kHz) = ________ mV.
3. The voltage gain reduces as the load resistance decreases.

9

UNIT

MULTI-STAGE AMPLIFIERS

"Well, the big products in electronics in the '50s were radio and television. The first big computers were just beginning to come in and represented the most logical market for us to work in."

Jack Kilby (1923–2005)
American Engineer

After completing this unit, students will be able to:

- explain the need of multi-stage amplifiers in electronic systems
- calculate the overall gain (as a ratio and also in dB) of a multi-stage amplifier, if the gain of each stage is known
- explain the working of different types of multi-stage amplifiers (resistance-capacitance coupled, transformer-coupled and direct-coupled) using BJTs, and FETs
- state applications of *RC*-coupled, transformer-coupled and direct-coupled multi-stage amplifiers
- explain the frequency response curve of an *RC*-coupled amplifier
- compute the mid-frequency gain of a given two-stage *RC*-coupled amplifier
- compute the lower and upper cutoff frequencies of a single-stage *RC*-coupled amplifier using an FET
- state the effect of cascading a number of stages on the bandwidth of an amplifier
- explain different types of distortion that occur in the signal when it is amplified by an amplifier
- state the classification of amplifiers on the basis of frequency, coupling, purpose, and operating point

9.1 DO WE REQUIRE MORE THAN ONE STAGE?

An amplifier is the basic building block of most electronic systems. Just as one brick does not make a house, a single-stage amplifier is not sufficient to build a practical electronic system. In the last chapter, we had discussed the performance of a single-stage amplifier. Although the gain of an amplifier does depend on the device parameters and circuit components, there exists an upper theoretical limit for the gain obtainable from one stage. The gain of single stage is not sufficient for practical applications. The voltage level of a signal can be raised to the desired level if we use more than one stage. When a number of amplifier stages are used in succession (one after the other) it is called a *multi-stage amplifier* or a *cascaded amplifier*. Much higher gains can be obtained from the multi-stage amplifiers.

9.2 GAIN OF A MULTI-STAGE AMPLIFIER

A multi-stage amplifier (n-stages) can be represented by the block diagram shown in Fig. 9.1. You may note that the output of the first stage makes the input of the second stage; the output of the second stage makes the input of the third stage, ..., and so on. The signal voltage v_s is applied to the input of the first stage. The final output v_o is then available at the output terminals of last stage. The output of the first (or the input to the second stage) is

$$v_1 = A_1 v_s$$

where A_1 is the voltage gain of the first stage. Then the output of the second stage (or the input to the third stage) is

$$v_2 = A_2 v_1$$

Similarly, the final output v_o is given as

$$v_o = v_n = A_n v_{n-1}$$

where A_n is obviously the voltage gain of the last (nth) stage.

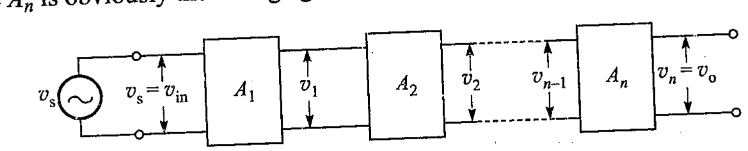


Fig. 9.1 *Block diagram of a multi-stage amplifier having n stages*

We may look upon this multi-stage amplifier as a single amplifier, whose input is v_s and output is v_o. The overall gain A of the amplifier is then given as

$$A = \frac{v_o}{v_s} = \frac{v_1}{v_s} \times \frac{v_2}{v_1} \times \ldots \times \frac{v_{n-1}}{v_{n-2}} \times \frac{v_o}{v_{n-1}}$$

or

$$A = A_1 \times A_2 \times \ldots \times A_{n-1} \times A_n \tag{9.1}$$

The gain of an amplifier can also be expressed in another unit called *decibel*.

9.2.1 Decibel

In many problems it is found very convenient to compare two powers on a logarithmic scale rather than on a linear scale. The telephone industry proposed a logarithmic unit, named *bel* after Alexander Graham Bell. The number of bels by which a power P_2 exceeds a power P_1 is defined as

$$\text{Numbers of bels} = \log_{10}\frac{P_2}{P_1}$$

For practical purposes it has been found that the unit bel is quite large. Another unit, one-tenth as large, is more convenient. This smaller unit is called the *decibel* (abbreviated as dB), and since one decibel is one-tenth of a bel, we have

$$\text{Number of dB} = 10 \times \text{Number of bels} = 10\log_{10}\frac{P_2}{P_1} \tag{9.2}$$

Note that the unit dB denotes a power ratio. Therefore, the specification of a certain power in dB is meaningiess unless a standard reference level is either implied or is stated explicitly. In communication applications, usually 6 mW or 1 mW is taken as standard reference level. When 1 mW is taken as reference, the unit dB is often referred to as dBm. A negative value of number of dB in Eq. (9.2) means that the power P_2 is less than the reference power P_1.

For an amplifier, P_1 may represent the input power and P_2 the output power. If V_1 and V_2 are the input and output voltages of the amplifier, then

$$P_1 = \frac{V_1^2}{R_i}$$

and

$$P_2 = \frac{V_2^2}{R_o}$$

where R_i and R_o are the input and output impedances of the amplifier. Then, Eq. (9.2) can be written as

$$\text{Number of dB} = 10\log_{10}\frac{V_2^2/R_o}{V_1^2/R_i} \tag{9.3}$$

In case the input and output impedances of the amplifier are equal, i.e., $R_i = R_o = R$, Eq. (9.3) simplifies to

$$\text{Number of dB} = 10\log_{10}\frac{V_2^2}{V_1^2} = 10\log_{10}\left(\frac{V_2}{V_1}\right)^2$$

$$= 10 \times 2\log_{10}\frac{V_2}{V_1} = 20\log_{10}\frac{V_2}{V_1} \tag{9.4}$$

However, in general, the input and output impedances are not always equal. But the expression of Eq. (9.4) is adopted as a convenient definition of the decibel voltage gain of an amplifier, regardless of the magnitudes of the input and output impedances. Of course, this usage is technically improper.

As an example; if the voltage gain of an amplifier is 10, it can be denoted on the dB scale as

$$\text{Gain in dB} = 20\log_{10}\frac{V_2}{V_1} = 20\log_{10}10$$

$$= 20 \times 1 = 20\text{ dB}$$

9.2.2 Gain of Multi-Stage Amplifier in dB

The gain of a multi-stage amplifier can be easily computed if the gains of the individual stages are known in dB. If we take logarithm (to the base 10) of Eq. (9.1) and then multiply each term by 20, we get

$$20\log_{10}A = 20\log_{10}A_1 + 20\log_{10}A_2 + \ldots + 20\log_{10}A_n$$

In the above equation, the term on the left is the overall gain of the multi-stage amplifier expressed in dB. The terms on the right denote the gains of the individual stages expressed in dB. Thus, the overall voltage gain in dB of a multi-stage amplifier is the sum of the decibel voltage gains of the individual stages. That is,

$$A_{dB} = A_{dB1} + A_{dB2} + \ldots A_{dB|n|} \tag{9.5}$$

9.2.3 Why is dB Used?

You may wonder why we use a logarithmic scale to denote voltage or power gains, instead of using simpler linear scale. The reasons for the popularity of dB scale are as follows:

1. It permits gains to be directly added when a number of stages are cascaded. (Use of logarithms changes multiplication into an addition).
2. It permits us to denote, both very small as well as very large, quantities of linear scale by conveniently small figures. Thus a voltage gain of 0.000 001 (in fact, it represents a loss instead of gain) may be represented as a voltage gain of –120 dB, or a voltage loss of 120 dB. Similarly, a power gain of 456 000 is simply 56.59 ($\simeq$ 56.6) dB on the logarithmic scale.
3. The output of many amplifiers is ultimately converted into sound and this sound is received by human ear. Experiments show that the ear responds to the sound intensities on a proportional of logarithmic scale rather than the linear scale. If the audio power increases from 4 W to 64 W, the hearing level does not increase by a factor of 64/4 = 16. The response of the ear will increase by a factor of only 3, since $(4)^3 = 64$. Thus, the use of the dB unit is justified on a psychological basis too.

Example 9.1 A multi-stage amplifier consists of three stages. The voltage gains of the stages are 30, 50 and 80. Calculate the overall voltage gain in dB.

Solution: We know that the overall voltage gain in dB of the three-stage amplifier is given as

$$A_{dB} = A_{dB1} + A_{dB2} + A_{dB3}$$

But, we are given the voltage gains of the individual stages as ratios. So, we should first find the gains of the individual stages in decibels. Thus,

$$A_{dB1} = 20 \log_{10} 30 = 29.54 \text{ dB}$$
$$A_{dB2} = 20 \log_{10} 50 = 33.98 \text{ dB}$$
$$A_{dB3} = 20 \log_{10} 80 = 38.06 \text{ dB}$$

Therefore,

$$A_{dB} = 29.54 + 33.98 + 38.06 = \mathbf{101.58 \text{ dB}}$$

Alternatively, we could have determined A_{dB} as follows: The overall voltage gain is

$$A = A_1 \times A_2 \times A_3$$
$$= 30 \times 50 \times 80 = 120000$$

Therefore, the overall voltage gain in dB is

$$A_{dB} = 20 \log_{10} 120000 = \mathbf{101.58 \text{ dB}}$$

9.3 HOW TO COUPLE TWO STAGES?

In a multi-stage amplifier, the output of one stage makes the input of the next stage (see Fig. 9.1). Can we connect the output terminals of one amplifier to the input terminals of the next amplifier directly? This may not always be possible due to practical difficulties. We must use a suitable coupling network between two stages so that a minimum loss of voltage occurs when the signal passes through this network to the next stage. Also, the dc voltage at the output of one stage should not be permitted to go to the input of the next. If it does, the biasing conditions of the next stage are disturbed.

The coupling network not only couples two stages; it also forms a part of the load impedance of the preceding stage. Thus, the performance of the amplifier will also depend upon the type of coupling network used. Three generally used coupling schemes are:

1. Resistance-capacitance coupling
2. Transformer coupling
3. Direct coupling

9.3.1 Resistance-Capacitance Coupling

Figure 9.2 shows how to couple two stages of amplifiers using resistance-capacitance (*RC*) coupling scheme. This is the most widely used method. In this scheme, the signal developed across the collector resistor R_C of the first stage is coupled to the

base of the second stage through the capacitor C_C. The coupling capacitor C_C blocks the dc voltage of the first stage from reaching the base of the second stage. In this way, the dc biasing of the next stage is not interfered with. For this reason, the capacitor C_C is also called a *blocking capacitor*.

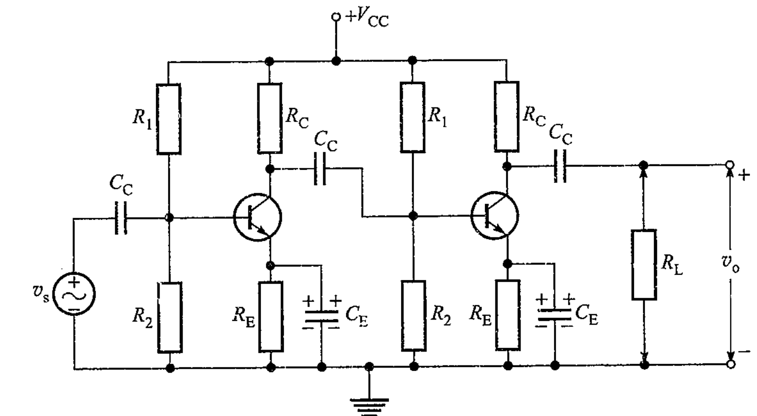


Fig. 9.2 *Two-stage RC-coupled amplifier using BJTs*

Some loss of the signal voltage always occurs due to the drop across the coupling capacitor. This loss is more pronounced when the frequency of the input signal is low. (This point is discussed in more detail in Section 9.4.1.) This is the main drawback of this coupling scheme. However, if we are interested in amplifying ac signals of frequencies greater than about 10 Hz, this coupling is the best solution. It is the most convenient and least expensive way to build a multi-stage amplifier.

RC coupling scheme finds applications in almost all audio small-signal amplifiers used in record players, CD players, public address systems, radio receivers, television receivers, etc. Figure 9.3 illustrates the use of *RC* coupling in the case of two stages of FET amplifiers.

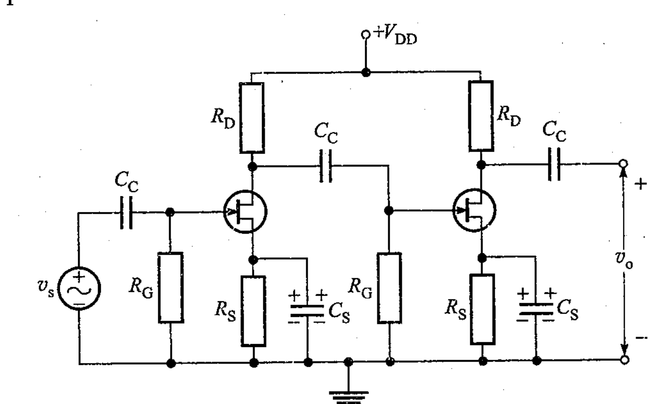


Fig. 9.3 *Two-stage RC-coupled amplifier using FETs*

9.3.2 Transformer Coupling

In this type of coupling, a transformer is used to transfer the ac output voltage of the first stage to the input of the second stage. The resistor R_C (see Fig. 9.2) is replaced by the primary winding of a transformer. The secondary winding of the transformer replaces the wire between the voltage divider (of the biasing network) and the base of the second stage. Figure 9.4 illustrates the transformer coupling between the two stages of amplifiers, using BJTs.

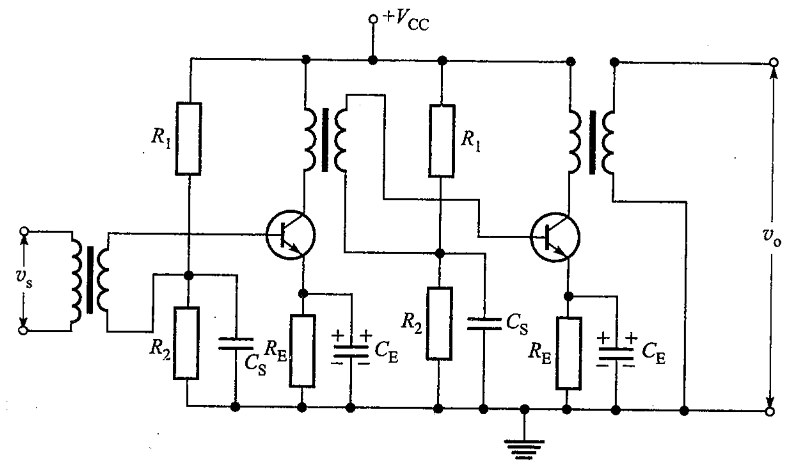


Fig. 9.4 *Two stages, using BJTs, are coupled by a transformer*

Note that in this circuit there is no coupling capacitor. The dc isolation between the two stages is provided by the transformer itself. There exists no dc path between the primary and the secondary windings of a transformer. However, the ac voltage across the primary winding is transferred (with a multiplication factor depending upon the turns-ratio of the transformer) to the secondary winding.

The main advantage of the transformer coupling over *RC* coupling is that all the dc voltage supplied by V_{CC} is available at the collector. There is no voltage drop across the collector resistor R_C (of *RC*-coupled scheme). The dc resistance of the primary winding is very low (only a few ohms). The ac impedance across the primary depends upon the turns-ratio of the transformer and the input impedance of the second amplifier; and it can be made sufficiently high. The absence of resistor R_C in the collector circuit also eliminates the unnecessary power loss in this resistor. These considerations of power are important when the amplifier is to work as a power amplifier (see Unit 10).

The transformer coupling scheme has some disadvantages also. The most obvious disadvantage is the increased size of the system. The transformer is very bulky as compared to a resistor or a capacitor. It is also relatively costlier. Another disadvantage of this scheme arises from the fact that the transformer used differs in its working from an ideal one. In the transformer, there is some leakage inductance

and interwinding capacitances. Because of these stray elements, the transformer-coupled amplifier does not amplify the signals of different frequencies equally well. The interwinding capacitance may give rise to a phenomenon of resonance at some frequency. This may make the gain of the amplifier very high at this frequency. At the same time, the gain may be quite low at other frequencies.

Because of the above drawbacks, the transformer-coupling scheme is not used for amplifying low frequency (audio) signals. However, they are widely used for amplification of radio-frequency signals. Radio frequency means anything above 20 kHz. In radio receivers, the rf ranges from 550 kHz to 1600 kHz for the medium-wave band; and from 3 MHz to 30 MHz for the short wave band. In TV receivers, the rf signals have frequencies ranging from 54 MHz to 216 MHz. By putting suitable shunting capacitors across each winding of the transformer, we can get resonance at any desired rf frequency. Such amplifiers are called *tuned-voltage amplifiers*. These provide high gain at the desired rf frequency. For this reason, the transformer-coupled amplifiers are used in radio and TV receivers for amplifying rf signals. (Such amplifiers are discussed in Unit 11 of this book.)

The use of a transformer for coupling not only saves power loss in the collector resistor R_C, but also helps in proper impedance matching. By suitably selecting the turns ratio of the transformer, we can match any load with the output impedance of the amplifier. This helps in transferring maximum power from the amplifier to the load. This is discussed in more details in Unit 10 on power amplifiers.

A tuned transformer-coupled amplifier using FETs is shown in Fig. 9.5.

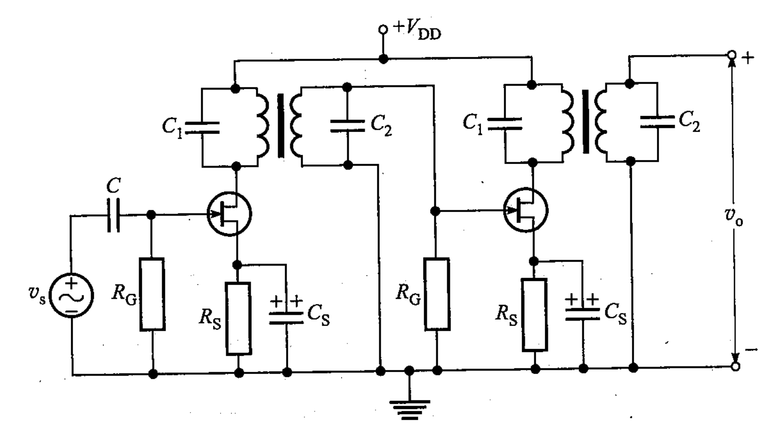


Fig. 9.5 *Transformer-coupled, tuned voltage amplifier using FETs*

9.3.3 Direct Coupling

In certain applications, the signal voltages are of very low frequency. For example, thermocouples are used for the measurement of temperature in furnaces. The voltage induced in the thermocouple is very small in magnitude (of the order of μV). This voltage needs to be amplified to a suitable level before it can be used to deflect the needle of a meter. The temperature of the furnace may change very slowly. The indicating meter should respond to such slowly varying changes. The amplifier used for the amplification of such slowly varying signals makes use of direct coupling. In this type of coupling scheme, the output of one stage of the amplifier is connected to the input of the next stage by means of a *simple connecting wire*.

For applications where the signal frequency is below 10 Hz, coupling capacitors and bypass capacitors cannot be used. At low frequencies, these capacitors can no longer be treated as short circuits, since they offer sufficiently high impedance. On the other hand, if coupling and bypass capacitors are to serve their purpose, their values have to be extremely large. Such capacitors are not only very expensive, but also are inconveniently large in size. For example, to bypass a 100 Ω emitter resistor at a frequency of 10 Hz, we need a capacitor of about 1000 μF. The lower the frequency, the worse the problem becomes. To avoid this problem, direct coupling is used. Figure 9.6 shows a two-stage direct-coupled amplifier, using BJTs. Note that no coupling and bypass capacitors are used. Therefore, both dc as well as ac are coupled to the next stage. The dc voltage at the collector of the first stage reaches the base of the second stage. This should be taken into account while designing the biasing circuit of the second stage (see Example 9.2).

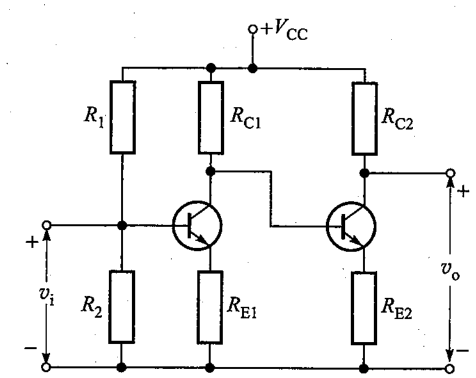


Fig. 9.6 *Two-stage, direct-coupled amplifier using BJTs*

The direct coupling scheme has a serious drawback. The transistor parameters like V_{BE} and β vary with temperature. This causes the collector current and voltage to change. Because of the direct coupling, this voltage change appears at the final output. Such an unwanted change in output voltage which has no relationship with input voltage is called *drift*. The drift in direct-coupled amplifiers is a serious problem. It can be wrongly interpreted as a genuine output produced by the input signal. There are some specially designed direct-coupled amplifier circuits in which the problem of drift is minimised to a considerable extent.

Figure 9.7 shows how FETs can be used in making a two-stage direct-coupled amplifier.

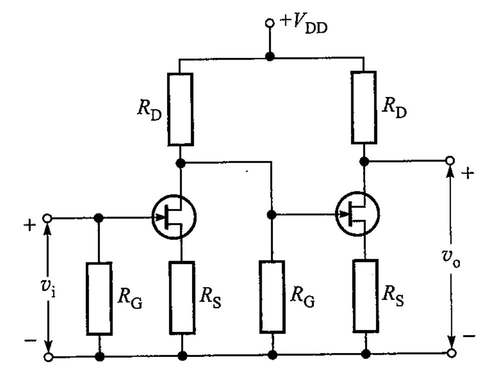


Fig. 9.7 *Two-stage direct-coupled amplifier using FETs*

Example 9.2 A two-stage direct-coupled amplifier is shown in Fig. 9.8. The transistors used in the circuit have $V_{BE} = 0.7$ V and $\beta = 300$. If the voltage at the input is +1.4 V, calculate the voltage at the output terminal.

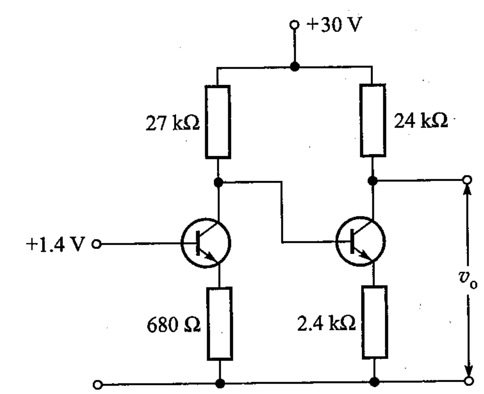


Fig. 9.8 *A direct-coupled amplifier*

Solution: The input voltage is +1.4 V. The voltage V_{BE} being 0.7 V, the voltage at the emitter terminal will be

$$V_E = 1.4 - 0.7 = 0.7 \text{ V}$$

Therefore, the emitter current (first stage),

$$I_{E1} = \frac{0.7}{680} \simeq 1 \text{ mA}$$

Since $I_{C_1} \simeq I_{E_1}$, the collector voltage is

$$V_{C1} = V_{CC} - I_{C_1} \times 27 \times 10^3 = 30 - 1 \times 10^{-3} \times 27 \times 10^3$$
$$= 3 \text{ V}$$

For this circuit, the base voltage of the second stage is the collector voltage of the first stage. Thus,

$$V_{B2} = V_{C1} = 3\ \text{V}$$

Since, $V_{BE} = 0.7$ V, the emitter voltage of the second stage is

$$V_{B2} = 3 - 0.7 = 2.3\ \text{V}$$

The emitter current,

$$I_{E2} = \frac{2.3}{2.4 \times 10^3} \simeq 1\ \text{mA}$$

Since, $I_{C2} \simeq I_{E2}$, The voltage at the collector is

$$V_{C2} = V_{CC} - I_{C2} \times 24 \times 10^3 = 30 - 1 \times 10^{-3} \times 24 \times 10^3$$
$$= 6\ \text{V}$$

Thus, the voltage at the output terminal is

$$V_o = \mathbf{6\ V}$$

9.4 FREQUENCY RESPONSE CURVE OF AN *RC*-COUPLED AMPLIFIER

A practical amplifier circuit is meant to raise the voltage level of the input signal. This signal may be obtained from the piezoelectric crystal of a record player, the sound head of a tape recorder, the microphone in case of a PA system, or from a detector circuit of a radio or TV receiver. Such a signal is not of a single frequency. But it consists of a band of frequencies. For example, the electrical signal produced by the voice of human being or by a musical orchestra may contain frequencies as low as 30 Hz and as high as 15 kHz. Such a signal is called audio signal. If the loudspeakers are to reproduce the original sound faithfully, the amplifier used must amplify all the frequency components of the signal equally well. If it does not do so, the output of the loudspeaker will not be an exact replica of the original sound. When this happens, we say that *distortion* has been introduced by the amplifier.

The performance of an amplifier is judged by observing whether all frequency components of the signal are amplified equally well. This information is provided by its frequency response curve. This curve illustrates how the magnitude of the voltage gain (of amplifier) varies with the frequency of the input signal (sinusoidal). It can be plotted by measuring the voltage gain of the amplifier for different frequencies of the sinusoidal voltage fed to its input (see Fig. 9.9).

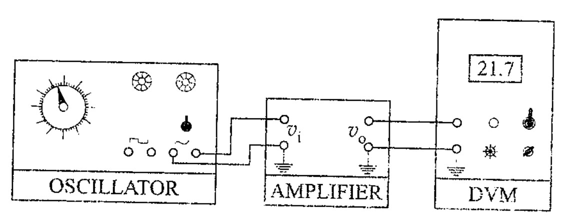


Fig. 9.9 *Measurement of voltage gain for plotting frequency-response curve*

Figure 9.10 shows a frequency response curve of a typical *RC*-coupled amplifier. This curve is usually plotted on a semilog graph paper with frequency on logarithmic scale so as to accommodate large frequency range. Note that the gain is constant only for a limited band of frequencies. This range of frequencies is called the mid-frequency range and the gain is called mid-band gain, A_{vm}. On both sides of the mid-frequency range, the gain decreases. For very low and for very high frequencies, the gain of the amplifier reduces to almost zero.

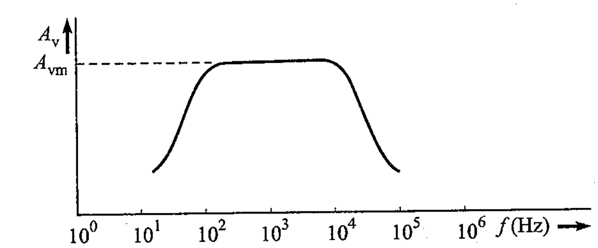


Fig. 9.10 *Frequency response curve of an RC-coupled amplifier*

9.4.1 Fall of Gain in Low-frequency Range

In the last section, we analysed an amplifier circuit to determine its voltage gain. This was the mid-frequency gain. In mid-frequency range, the coupling and bypass capacitors are as good as short circuits. But, when the frequency is low, these capacitors can no longer be replaced by the short-circuit approximation. The lower the frequency, the greater is the value of reactance of these capacitors, since

$$X_C = \frac{1}{2\pi fC}$$

Let us first examine how the coupling capacitor C_C affects the voltage gain of the amplifier at low frequencies. The output section of the first stage of the two-stage *RC*-coupled amplifier of Fig. 9.2 is redrawn in Fig. 9.11*a*. The output voltage v_o of this stage is the input to the second stage. The resistors R_1 and R_2 are the biasing resistors for the second stage. From the ac point of view, this circuit is equivalent to the one drawn in Fig. 9.11*b*. Assume for the time being, that the capacitor C_E is replaced by a short circuit. The resistors R_1, R_2 and input impedance h_{ie} of the next stage are in parallel and are equivalent to a resistor R. This resistance forms a part of the load resistance of the previous (first) stage. It really does not matter whether the output voltage v_o is taken at the left side or at the right side of the resistor R.

The capacitor C_C is in series with the resistor R, and this series combination is in parallel with the collector resistor R_C. The whole of this impedance forms the ac load for the preceding stage. But the effective output of the stage is the ac voltage developed across the resistor R (see Fig. 9.11*c*). At mid-frequencies (and also at high frequencies), the reactance of the capacitor C_C is sufficiently small compared to R. We can treat it as a short circuit so that the resistor R comes in parallel with the resistor R_C. In such a case, the voltage v_i across resistor R_C will be the same as the voltage v_o

across R. However, at low frequencies, the reactance of C_C [$= 1/(2\pi f C_C)$] becomes sufficiently large. This causes a significant voltage drop across C_C. The result is that the effective output voltage v_o decreases. The lower the frequency of this signal, the higher will be the reactance of the capacitor C_C and the more will be the reduction in output voltage v_o. At zero frequency (dc signals), the reactance of capacitor C_C is infinitely large (an open circuit). The effective output voltage v_o then reduces to zero. Thus we see that the output voltage v_o (and hence the voltage gain) decreases as the frequency of the signal decreases below the mid-frequency range.

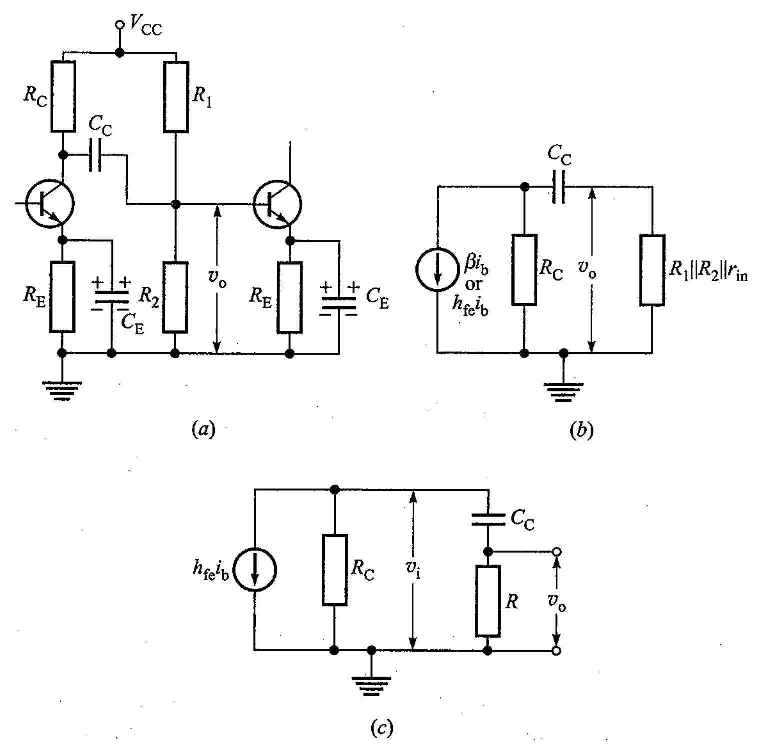


Fig. 9.11 *(a) Output section of an RC-coupled amplifier (b) Its ac equivalent; (c) The same equivalent circuit redrawn in another way*

The other component, due to which the gain decreases at low frequencies, is the bypass capacitor C_E. Figure 9. 12 shows the input section of the amplifier. The capacitor C_E is connected across the emitter resistor R_E. This capacitor is meant to bypass the ac current to ground. The impedance of this capacitor is quite low (as good as a short circuit) in the mid-frequency range as well as in high-frequency range. Therefore, at these frequencies, the emitter is effectively grounded for ac current. However, as the frequency decreases,

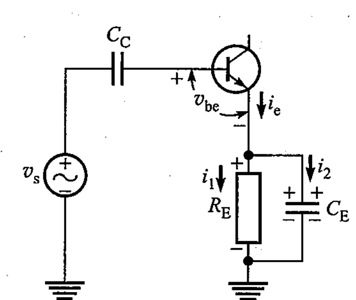


Fig. 9.12 *Input section of an RC-coupled amplifier*

the reactance of the capacitor C_E becomes comparable to resistance R_E. The bypassing action of the capacitor is no longer as good as at mid- and high-frequencies. The emitter is not at ground potential for ac. The emitter current i_e divides into two parts, i_1 and i_2. A part of current i_1 passes through the resistor R_E. The rest of the current i_2 ($= i_e - i_1$) passes through the capacitor C_E. Due to current i_1 in R_E, an ac voltage $i_1 \times R_E$ is developed.

When the polarity of the input signal voltage is as shown in figure, the current i_1 flows from the emitter to ground. The polarity of the voltage $i_1 R_E$ is also marked in the figure. Then, the effective input voltage to the amplifier (that is the voltage between the base and emitter of the transistor) becomes

$$v_{be} = v_s - i_1 R_E \tag{9.6}$$

The effective input voltage is thus reduced. The output voltage v_o of the amplifier will now naturally be reduced. In other words, the gain of the amplifier ($= v_o/v_i$) reduces. This reduction in gain occurs due to the inability of the capacitor C_E to bypass ac current. The lower the frequency, the higher is the impedance of the capacitor C_E, and the greater is the reduction in gain.

Note that the resistor R_E is not only a part of the input section, but also is a part of the output section. The voltage $i_1 R_E$ developed across the resistor R_E depends upon the output ac current. In this way, the effective input to the amplifier depends on the output current. The reduction in gain due to such a process is technically described as *negative current feedback effect*.

In Fig. 9.12, there is also a coupling capacitor C_C in the input section of the amplifier. Due to this capacitor, the effective input voltage is reduced at low frequencies in much the same way as the effective output voltage v_o is reduced due to the coupling capacitor in the output section. Thus, the coupling capacitor in the input side is also responsible for the decrease of gain at low frequencies.

In practical circuits, the value of the bypass capacitor C_E is very large ($\simeq 100\ \mu F$). Therefore, it is the coupling capacitor that has the more pronounced effect in reducing the gain at low frequencies.

9.4.2 Does Gain Fall at High Frequencies?

As the frequency of the input signal increases, the gain of the amplifier reduces. Several factors are responsible for this reduction in gain. Firstly, the beta (β) of the transistor is frequency dependent. Its value decreases at high frequencies (see Fig. 9.13). Because of this, the voltage gain of the amplifier reduces as the frequency increases.

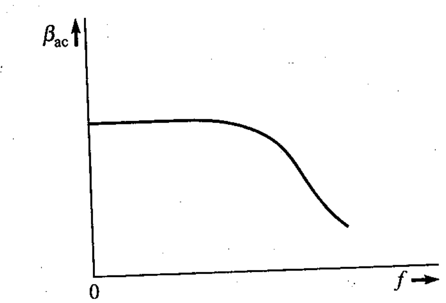


Fig. 9.13 *Variation of short-circuit current gain β_{ac} with frequency*

Another important factor responsible for the reduction in gain of the amplifier at high frequencies is the presence of the device. In case of a transistor, there exists some capacitance due to the formation of a depletion layer at the junctions. These inter-electrode capacitances are shown in Fig. 9.14. Note that the connection for these capacitances are shown by dotted lines. This has been done to indicate that these are not physically present in the circuit, but are inherently present with the device (whether we like it or not).

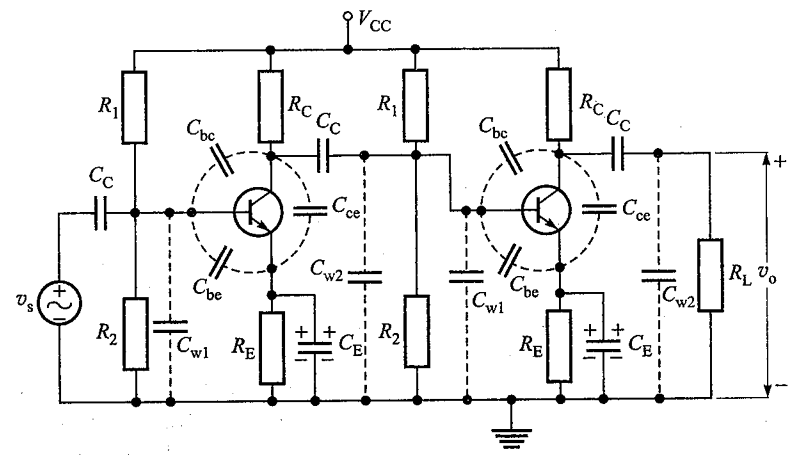


Fig. 9.14 *RC-coupled amplifier. Capacitances that affect high-frequency response are shown by dotted connections*

The capacitance C_{bc} between the base and collector connects the output with the input. Because of this, negative feedback takes place in the circuit and the gain decreases. This feedback effect is more, when the capacitance C_{bc} provides a better conducting path for the ac current. Such is the case at high frequencies. As the frequency increases, the reactive impedance of the capacitor becomes smaller.

The capacitance C_{be} offers a low-impedance path at high frequencies in the input side. This reduces the input impedance of the device, and the effective input side signal is reduced. So the gain falls. Similarly, the capacitance C_{ce} produces a shunting effect at high frequencies in the output side.

Besides the junction capacitances, there are wiring capacitances C_{w1} and C_{w2}, as shown in Fig. 9.14. The connecting wires of the circuit are separated by air which serves as a dielectric. This gives rise to some capacitance between the wires, though the capacitance value may be very small. But at high frequencies, even these small capacitances C_{ce}, C_{w2}, and the input capacitance C_i of the next stage can be represented by a single shunt capacitance

$$C_s = C_{ce} + C_{w2} + C_i \tag{9.7}$$

The output section of the amplifier is shown in Fig. 9.15 from the ac point of view, for high-frequency considerations. The capacitance C_s is the equivalent shunt capacitance as given by Eq. (9.7). Note that the coupling and bypass capacitors do not appear in the figure, because they effectively represent short circuits at these frequencies.

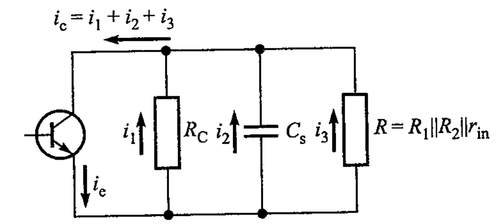


Fig. 9.15 *Output section of an RC-coupled amplifier at high frequencies*

As can be seen from Fig. 9.15, the collector current i_c is made up of three currents i_1, i_2 and i_3. As the frequency of the input signal increases, the impedance of the shunt capacitance C_s decreases, since

$$X_{C_s} = \frac{1}{2\pi f C_s}$$

As a result, the current i_2 through this capacitance increases. This reduces both the currents i_1 and i_3, since the total current i_c $(= i_1 + i_2 + i_3)$ is almost constant. It means that the output voltage v_o $(= i_3 R)$ decreases. The higher the frequency, the lower is the impedance offered by C_s and the lower will be the output voltage v_o.

9.4.3 Bandwidth of an Amplifier

Frequency response curve of an *RC*-coupled amplifier is shown in Fig. 9.16. The gain remains constant only for a limited band of frequencies. On both the low-frequency side as well as on the high-frequency side, the gain falls. Now an important question arises—where exactly should we fix the frequency limits (of input signal) within which the amplifier may be called a good amplifier? The limit is set at those frequencies at which the voltage gain reduces to 70.7 % of the maximum gain A_{vm}. These frequencies are known as the *cut-off frequencies* of the amplifier. These frequencies are marked in Fig. 9.16. The frequency f_1 is the *lower cut-off frequency* and the frequency f_2 is the *upper cut-off frequency*. The difference of the two frequencies, that is $f_2 - f_1$, is called the *bandwidth* (*BW*) of the amplifier. The mid-frequency range of the amplifier is from f_1 to f_2. Usually, the lower cut-off frequency f_1 is much lower than upper cut-off frequency f_2, so that we have

$$BW = f_2 - f_1 \simeq f_2 \tag{9.8}$$

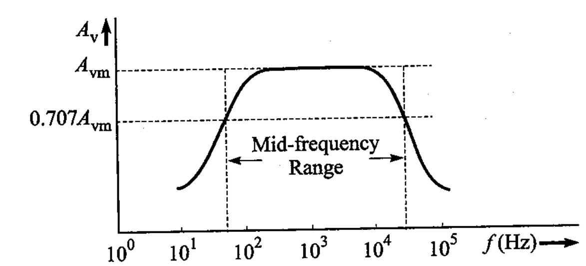


Fig. 9.16 *Bandwidth of an RC-coupled amplifier*

Why the limit is set at 70.7 % of maximum gain At the cut-off frequencies the voltage gain is

$$0.707A_{vm}\ [(=1/\sqrt{2})A_{vm}]$$

where A_{vm} is the maximum gain or the mid-frequency gain of the amplifier. It means at these frequencies, the output voltage is $1/\sqrt{2}$ times the maximum voltage. Since the power is proportional to the square of the voltage, the output power at these cut-off frequencies becomes one-half of the power at mid-frequencies. On the dB scale this is equal to a reduction in power by 3 dB. For this reason, these frequencies are also called *half-power frequencies*, or 3 *dB frequencies*.

We have taken a difference of 3 dB in power to define the cut-off frequencies, because this represents an audio-power difference that can just be detected by the human ear. For the frequencies below f_1 and above f_2, the output power will reduce by more than 3 dB.

Example 9.3 An *RC*-coupled amplifier has a voltage gain of 100 in the frequency range of 400 Hz to 25 kHz. On either side of these frequencies, the gain falls so that it is reduced by 3 dB at 80 Hz and 40 kHz. Calculate gain in dB at cut-off frequencies and also construct a plot of frequency response curve.

Solution: The gain in dB is

$$A_{dB} = 20 \log_{10}A = 20 \log_{10}100 = 40 \text{ dB}$$

This is the mid-band gain. The gain at cut-off frequencies is 3 dB less than the mid-band gain, i.e.,

$$(A_{dB}) \text{ (at cut-off frequencies)} = 40 - 3 = 37 \text{ dB}$$

The plot of the frequency response curve is given in Fig. 9.17.

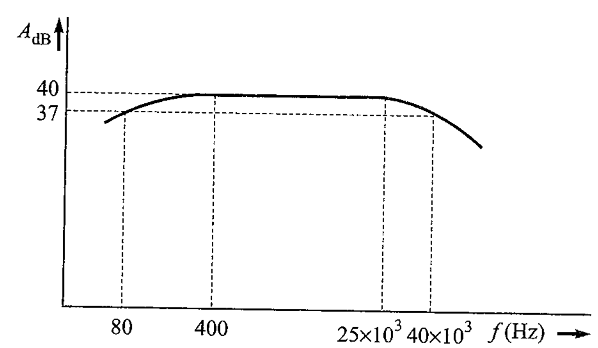


Fig. 9.17

Effect on bandwith when stages are cascaded A number of stages are cascaded to obtain higher values of voltage gain. But then the bandwidth of the amplifier does not remain the same. It decreases. The upper cut-off frequency decreases and the lower cut-off frequency increases. It happens because greater number of stages means a greater number of capacitors in the circuit. And each capacitor adversely affects the frequency response.

If n identical stages are cascaded, the overall mid-band voltage gain becomes

$$A'_v = (A_{vm})^n \tag{9.9}$$

where A_{vm} is the mid-band voltage gain of an individual stage. If f_1 and f_2 are the lower and upper cut-off frequencies of an individual stage, the overall all cut-off frequencies are given by

$$f_1' = \frac{1}{\sqrt{(2^{1/n}-1)}} f_1 \tag{9.10}$$

and

$$f_2' = \sqrt{(2^{1/n}-1)}\, f_2 \tag{9.11}$$

9.5 ANALYSIS OF TWO-STAGE RC-COUPLED AMPLIFIER

We have complete information about an amplifier if following parameters are known:

1. Mid-band voltage gain A_{vm}
2. Bandwidth, i.e., $f_2 - f_1$
3. Input impedance Z_i
4. Output impedance Z_o

The analysis of a two-stage amplifier will depend upon what active device is used. We shall analyse a two-stage amplifier circuit using BJTs as well as JFETs.

9.5.1 Two-Stage BJT *RC*-Coupled Amplifier

A two-stage *RC*-coupled amplifier circuit using identical BJTs is shown in Fig. 9.18. For its analysis, we first draw the circuit from the ac point of view. This is shown in Fig. 9.19. Since we wish to determine the gain in the mid-frequency range, all the coupling and bypass capacitors are replaced by short circuits. The dc power supply is also replaced by a short circuit. Next, we replace the transistors by their h-parameter approximate models. The result is the Fig. 9.20. In the approximate model of the transistor, small parameters h_{re} and h_{oe} are neglected.

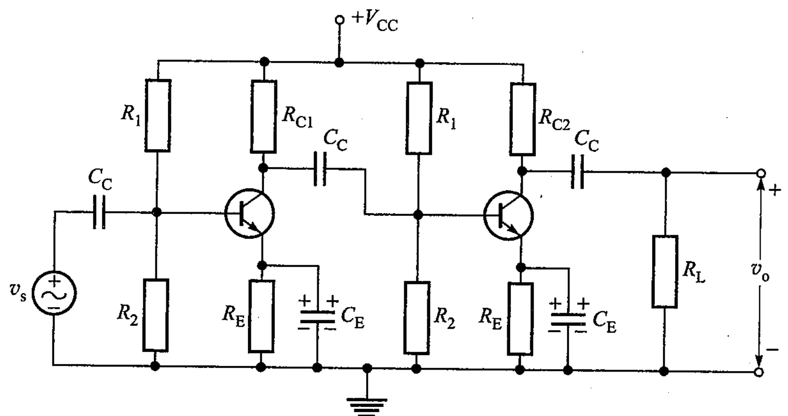


Fig. 9.18 *Two-stage BJT RC-coupled amplifier*

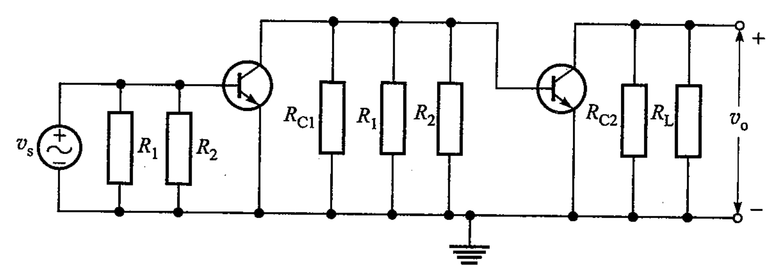


Fig. 9.19 *Circuit of Fig. 9.18 from ac point of view*

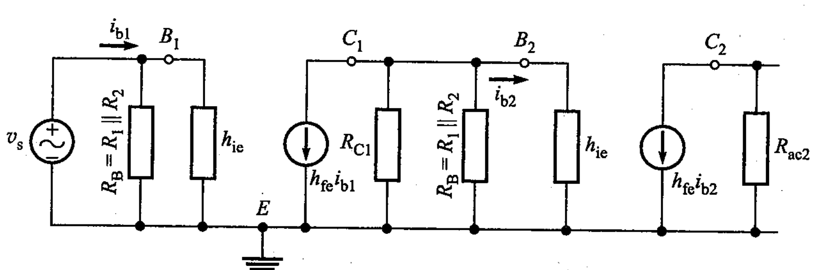


Fig. 9.20 *BJTs of Fig. 9.19 replaced by their approximate h-parameter model*

The parallel combination of resistors R_1 and R_2 is replaced by a single resistor R_B, i..e.,

$$R_B = R_1 \| R_2 = \frac{R_1 R_2}{R_1 + R_2} \tag{9.12}$$

For finding the overall gain of the two-stage amplifier, we should know the gains of the individual stages. Let us first find the gain A_2 of the second stage. We can use the formula derived in Unit 8 (see Eq. (8.9)) and write

$$A_2 = -\frac{h_{fe} R_{ac2}}{h_{ie}} \tag{9.13}$$

where R_{ac2} is the ac load resistance for the second stage, given by

$$R_{ac2} = R_{C2} \| R_L = \frac{R_{C2} R_L}{R_{C2} + R_L} \tag{9.14}$$

For determining the gain of the first stage, let us have a closer look into what constitutes its ac load resistance, R_{ac1}. From Fig. 9.20, the resistance R_{ac1} is the parallel combination of R_{C1}, R_B and h_{ie} (the input impedance of the second stage). That is,

$$R_{ac1} = R_{C1} \| R_B \| h_{ie}$$

The voltage gain of the first stage is then given as

$$A_1 = -\frac{h_{fe} R_{ac1}}{h_{ie}} \tag{9.15}$$

Using Eqs. (9.13) and (9.15) the overall gain A_{vm} can be easily determined, since

$$A_{vm} = A_1 \times A_2 \tag{9.16}$$

It should be noted that the voltage gain of the first stage A_1 is always less than the voltage gain A_2 of the second stage. This is because R_{ac1} is very much reduced, as the input impedance h_{ie} is in parallel with R_{C1} and R_B. This effect is called the *loading effect in multi-stage amplifiers.*

The overall input impedance is simply the input impedance of the first stage. If the biasing resistors R_1 and R_2 are large compared to h_{ie}, the overall impedance Z_i is simply h_{ie}.

The output impedance of the amplifier is R_{ac2}.

9.5.2 Two-Stage JFET *RC*-Coupled Amplifier

A two-stage JFET amplifier appears in Fig. 9.21. Unlike BJTs, the input impedance of a JFET is very high. Also the gate resistor R_G connected in the input of a stage is much greater (of the order of 1 MΩ) than the load resistance R_L (of the order of 10 kΩ) of the preceding stage. Therefore, a stage does not load its preceding stage. Each of the stages can be considered quite independently. For determining the overall gain of the amplifier, we shall consider the first stage only (shown within the dotted box in Fig. 9.21).

In Fig. 9.22*a* the ac equivalent of the first stage of the *RC*-coupled amplifier of Fig. 9.21 is shown. The JFET is replaced by its voltage-source equivalent. From the ac point of view, the dc supply V_{DD} offers a short circuit. The parallel combination $R_S C_S$ does not appear in the equivalent circuit. The capacitor C_S is assumed to be large enough to put the source at ground potential, at the operating frequencies. The capacitor C_{sh} represents the total shunt capacitance. This includes inter-electrode capacitance C_{ds}, wiring capacitance C_w and the effective input capacitance C_i of the next stage. Typical value of the shunt capacitor C_{sh} is 200 pF.

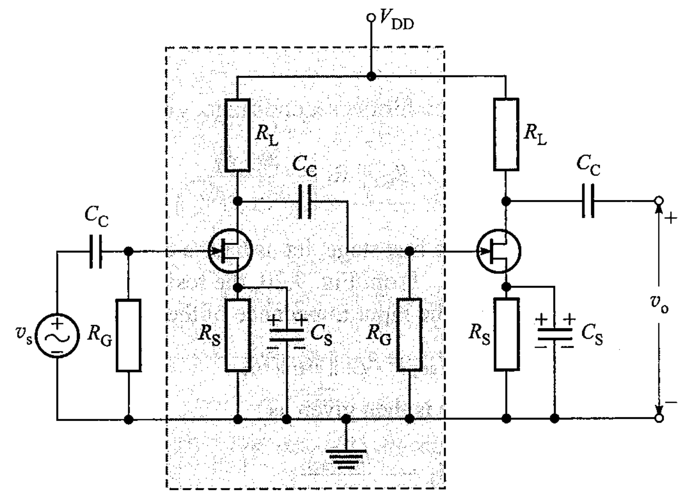


Fig. 9.21 *Two-stage RC-coupled amplifier using JFETs*

In mid-frequency range, the frequency is high enough to make the reactance of capacitor C_C small, compared to R_G. Therefore, it may be replaced by a short-circuit. For example, if C_C = 0.05 µF, at a frequency of 1 kHz, the reactance $X_C = 1/2\pi f C_C = 3.18$ kΩ. This is much smaller than the value of R_G (usually 1 MΩ). In this frequency range, the shunt capacitor C_{sh} offers very high impedance in parallel with the resistors R_L and R_G.

Hence it is neglected. The ac equivalent circuit for the mid-frequency range does not contain any capacitor, as shown in Fig. 9.22*b*. If the parallel combination of R_L and R_G is replaced by a single resistor R, so that

$$R = R_L \parallel R_G = \frac{R_L R_G}{R_L + R_G}, \text{ or } \frac{1}{R} = \frac{1}{R_L} + \frac{1}{R_G}$$

the mid-band voltage gain is given as

$$A_{vm} = -\frac{\mu R}{r_d + R} \tag{9.17}$$

Since $\mu = r_d g_m$, the above equation becomes

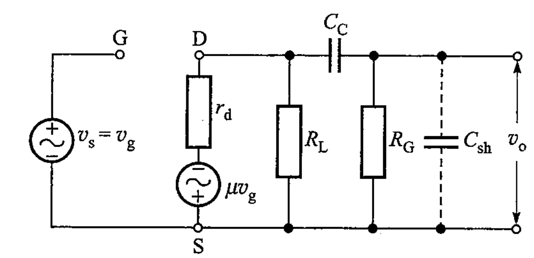
$$A_{vm} = -\frac{r_d g_m R}{r_d + R}$$

We now divide both numerator and denominator by $r_d R$ so as to get

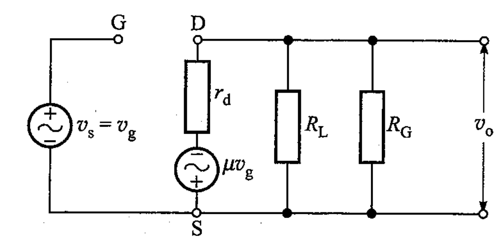
$$A_{vm} = -\frac{g_m}{\frac{1}{R} + \frac{1}{r_d}} = -\frac{g_m}{\frac{1}{R_L} + \frac{1}{R_G} + \frac{1}{r_d}} = -\frac{g_m}{\frac{1}{R_{eq}}}$$
$$= -g_m R_{eq} \tag{9.18}$$

Obviously, R_{eq} is the equivalent parallel combination of r_d, R_L and R_G, i.e.,

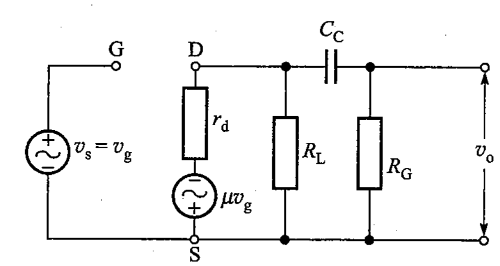
$$\frac{1}{R_{eq}} = \frac{1}{r_d} + \frac{1}{R_L} + \frac{1}{R_G} \tag{9.19}$$



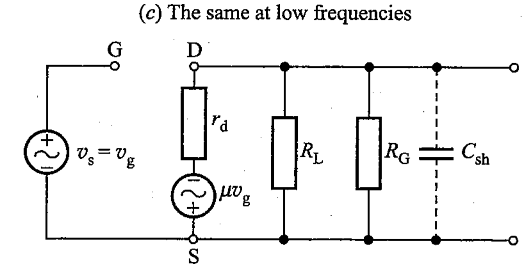
(*a*) The ac equivalent of one stage of *RC*-coupled BJT amplifier



(*b*) The same at mid-frequency



(*c*) The same at low frequencies



(*d*) The same at high frequencies

Fig. 9.22

Figure 9.22*c* shows the ac equivalent circuit in the low-frequency range. For low frequencies, the reactance of C_C becomes comparable to R_G. Hence, it cannot be neglected. The capacitor C_{sh} need not be considered, as its reactance becomes much higher than what it was in mid-frequency range. The frequency response of the amplifier in low-frequency depends upon C_C, and the cut-off frequency f_1 is given by

$$f_1 = \frac{1}{2\pi C_C R'} \tag{9.20}$$

where,

$$R' = (r_d \parallel R_L) + R_G \tag{9.21}$$

In the high-frequency range, the ac equivalent circuit becomes as shown in Fig. 9.22*d*. Now, capacitor C_C does not appear, but the capacitor C_{sh} will have a shunting effect at the output. The fall in gain at high frequencies is due to the shunting effect of this capacitor C_{sh}. The upper cut-off frequency is given by

$$f_2 = \frac{1}{2\pi C_{sh} R_{eq}} \tag{9.22}$$

where R_{eq} is given by Eq. (9.19).

Equations (9.17), (9.20) and (9.22) give the performance of a single-stage *RC*-coupled amplifier. Using these values the overall performance of the two-stage *RC*-coupled amplifier can be determined with the help of Eqs. (9.9), (9.10) and (9.11).

Example 9.4 A two-stage *RC*-coupled BJT amplifier is shown in Fig. 9.23. Calculate (*a*) input impedance Z_i, (*b*) output impedance Z_o, and (*c*) voltage gain A_{vm}. For both the transistors, h_{fe} or $\beta = 120$ and r_{in} or $h_{ie} = 1.1$ kΩ.

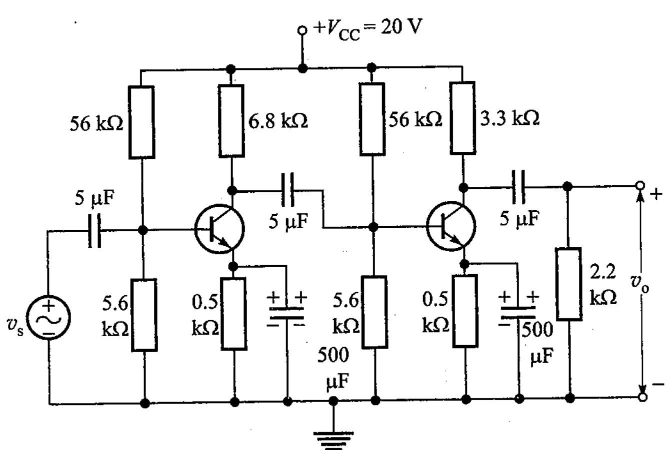


Fig. 9.23 *A two-stage RC-coupled amplifier*

Solution:

(*a*) *Input impedance*: From the knowledge of single-stage analysis, it is clear that for ac response, both 5.6-kΩ and 56-kΩ resistors will appear in parallel if

the network is redrawn from an ac point of view. They are also in parallel with the input impedance of the first transistor, which is approximately $h_{ie} = 1.1$ kΩ, since the emitter resistor is bypassed by C_E.

$$\therefore \quad Z_i = 5.6\text{ k}\Omega \parallel 56\text{ k}\Omega \parallel 1.1\text{ k}\Omega = \mathbf{0.905\text{ k}\Omega}$$

(*b*) *Output impedance*: Recall that the approximate collector-to-emitter equivalent circuit of a transistor is simply a current source $h_{fe}i_b$(or βi_b). The dynamic output resistance r_o(which is $1/h_{oe}$), being very large, was neglected in the approximate analysis. Therefore, R_{eq2}, which is a parallel combination of 3.3 kΩ and 2.2 kΩ, is the effective output impedance.

$$\therefore \quad Z_o = 3.3\text{ k}\Omega \parallel 2.2\text{ k}\Omega = \mathbf{1.32\text{ k}\Omega}$$

(*c*) *Voltage gain*: For calculating the voltage gain, the ac equivalent circuit of the given two-stage amplifier is drawn (Fig. 9.24).

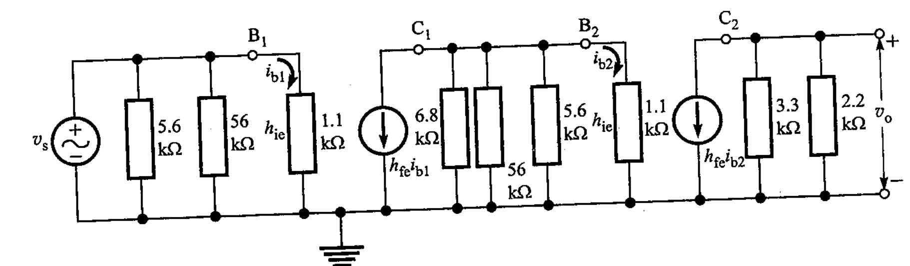


Fig. 9.24 *AC equivalent circuit of a two-stage BJT RC-coupled amplifier using approximate h-parameter model*

The voltage gain of the second stage is given by

$$A_2 = \frac{-h_{fe}R_{ac2}}{h_{ie}}$$

Here $h_{fe} = 120$; $h_{ie} = 1.1$ kΩ; and

$$R_{ac} = 3.3\text{ k}\Omega \parallel 2.2\text{ k}\Omega = \frac{3.3\times10^3 \times 2.2\times10^3}{(3.3+2.2)\times10^3}$$

$$= 1.32\text{ k}\Omega$$

$$\therefore \quad A_2 = \frac{-120\times1.32\times10^3}{1.1\times10^3} = -144$$

The voltage gain of the first stage is given by

$$A_1 = \frac{-h_{fe}R_{ac1}}{h_{ie}}$$

Here, $R_{ac1} = 6.8\text{ k}\Omega \parallel 56\text{ k}\Omega \parallel 5.6\text{ k}\Omega \parallel 1.1\text{ k}\Omega = 0.798\text{ k}\Omega$

$$\therefore \quad A_1 = \frac{-120\times0.798\times10^3}{1.1\times10^3} = -87.05$$

$\therefore$ Overall gain $A = A_1 \times A_2 = -87.05 \times (-144)$

$= \mathbf{12\ 535}$

Example 9.5 For a two-stage JFET amplifier shown in Fig. 9.25, calculate the maximum voltage gain and the bandwidth. Assume the following:

$R_L = 10\ \text{k}\Omega$; $R_G = 470\ \text{k}\Omega$; $C_{sh} = 100$ pF; $\mu = 25$; $r_d = 8\ \text{k}\Omega$

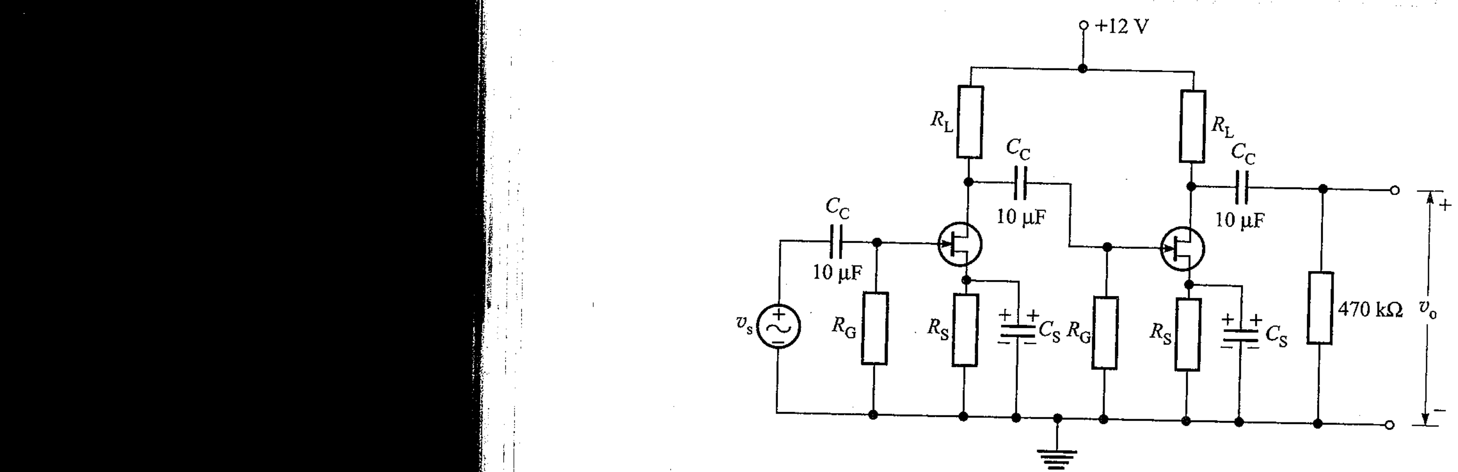


Fig. 9.25 *A two-stage RC-coupled amplifier using JFETs*

Solution: In order to calculate the maximum voltage gain of the given two-stage amplifier, the voltage gains of the individual stages are calculated. Since the two stages are identical, and cascading of the stages does not affect the performance of the preceding stage, the voltage gain of both the stages is same. The voltage gain of a JFET amplifier is

$$A_{vm} = g_m R_{eq} \angle 180°$$

Here, $$g_m = \frac{\mu}{r_d} = \frac{25}{8 \times 10^3}\ \text{S}$$

and $$R_{eq} = r_d \parallel R_L \parallel R_G = 8\ \text{k}\Omega \parallel 10\ \text{k}\Omega \parallel 470\ \text{k}\Omega$$

$$\simeq 8\ \text{k}\Omega \parallel 10\ \text{k}\Omega = 4.44\ \text{k}\Omega$$

$$\therefore \quad A_{vm} = \frac{25 \times 4.4 \times 10^3}{8 \times 10^3} = 13.75$$

The voltage gain of the two-stages, A'_{vm}, is given by

$$A'_{vm} = (A_{vm})^2 = (13.75)^2 = \mathbf{189.06}$$

The lower cut-off frequency of the amplifier is given by

$$f_1 = \frac{1}{2\pi C_C R'}, \text{ where } R' = (r_d \parallel R_L) + R_G$$

Here, $C_C = 10\ \mu$F and $R' = \left(\frac{10 \times 8}{18}\text{k}\Omega + 470\ \text{k}\Omega\right) = 474.44\ \text{k}\Omega$

$$\therefore \quad f_1 = \frac{1}{2\pi \times 10 \times 10^{-6} \times 474.44 \times 10^3} = \mathbf{0.0335\ Hz}$$

The lower cut-off frequency of the two stages is given by

$$f_1' = \frac{f_1}{\sqrt{(2^{1/2} - 1)}} = \frac{0.0335}{\sqrt{1.414 - 1}} = \mathbf{0.052\ Hz}$$

The upper cut-off frequency of the amplifier is given by

$$f_2 = \frac{1}{2\pi C_{sh} R_{eq}}, \text{ where } R_{eq} = r_d \parallel R_L \parallel R_G$$

Here, $C_{sh} = 100$ pF and $R_{eq} = 4.44\ \text{k}\Omega$. Therefore,

$$f_2 = \frac{1}{2\pi \times 100 \times 10^{-12} \times 4.44 \times 10^3} = \mathbf{358.5\ kHz}$$

The upper cut-off frequency for a two-stage amplifier is given by

$$f_2' = \sqrt{(2^{1/2} - 1)} \times f_2 = 0.6436 \times 358.5\ \text{kHz} = 230.72\ \text{kHz}$$

$$\therefore \quad \text{Bandwidth} = f_2' - f_1' = 230.72\ \text{kHz} - 52.05\ \text{Hz} \cong \mathbf{230\ kHz}$$

9.6 DISTORTION IN AMPLIFIERS

The purpose of an amplifier is to boost up the voltage or power level of a signal. During this process. the waveshape of the signal should not change. If the waveshape of the output is not an *exact replica* of the waveshape of the input, we say that *distortion* has been introduced by the amplifier. An *ideal amplifier* will amplify a signal without changing its waveshape at all. Such an amplifier *faithfully* amplifies the signal, and we say it has a good *fidelity*. Such an amplifier is called Hi-Fi (high fidelity) amplifier.

A number of factors may be responsible for causing distortion. It may be caused either due to the reactive components of the circuit, or due to imperfect (non-linear) characteristics of the transistor. There are three types of distortion. These may exist either separately or simultaneously in an amplifier.

1. Frequency distortion
2. Phase or time-delay distortion
3. Harmonic, amplitude, or non-linear distortion

9.6.1 Frequency Distortion

In practical situations, the signal is not a simple sinusoidal voltage. It has a complex waveshape. Such a signal is equivalent to a signal obtained by adding a number of sinusoidal voltages of different frequencies. These sinusoidal voltages are called the frequency components of the signal. If all the frequency components of the signal are not amplified equally well by the amplifier, *frequency distortion* is said to occur. The cause for this distortion is non-constant gain for different frequencies. This occurs due to the inter-electrode capacitances of the active devices and other reactive components of the circuit. For example, an RC-coupled amplifier can amplify signals whose frequency lie within its bandwidth (Fig. 9.20). Let us arbitrarily state that the input signal to this amplifier contains many equal-amplitude frequency components spread over a large band (even beyond f_2). The higher frequency components will not be amplified to the same extent as the lower frequency components (see Fig. 9.26). Due to such a distortion, the speech or music (produced by the amplified electrical signal) appears to be quite different from the original one.

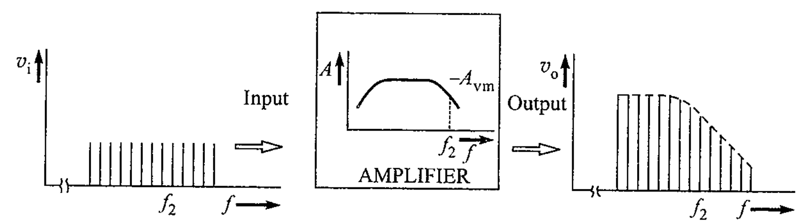


Fig. 9.26 *Illustration of frequency distortion*

9.6.2 Phase Distortion

Phase distortion is said to occur if the phase relationship between the various frequency components making up the signal waveform is not the same in the output as in the input. It means that the time of transmission or the delay introduced by the amplifier is different for various frequencies. The reactive components of the circuit are responsible for causing this type of distortion.

This distortion is not important in audio amplifiers. Our ears are not capable of distinguishing the relative phases of different frequency components. But this distortion is objectionable in video amplifiers used in television.

9.6.3 Harmonic Distortion

This type of distortion is said to occur when the output contains new frequency components that are not present in the input signal. These new frequencies are the

harmonics of the frequencies present in the input. For example, the input signal may consist of two frequency components, say 400 Hz and 500 Hz. If the amplifier gives rise to harmonic distortion, the output will contain

400 Hz (f_1, the fundamental)
500 Hz (f_2, the fundamental)
800 Hz (second harmonic of $f_1 = 2 \times 400$ Hz)
1000 Hz (second harmonic of $f_2 = 2 \times 500$ Hz)
1200 Hz (third harmonic of $f_1 = 3 \times 400$ Hz)
1500 Hz (third harmonic of $f_2 = 3 \times 500$ Hz)
1600 Hz (fourth harmonic of $f_1 = 4 \times 400$ Hz)
... ...
... ...

This is illustrated in Fig. 9.27.

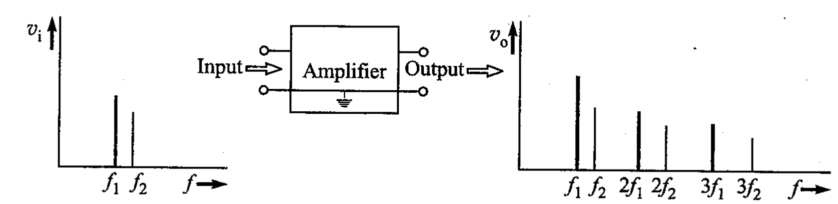


Fig. 9.27 *Illustration of harmonic distortion*

Harmonic distortion in an amplifier occurs because of the *nonlinearity* in the dynamic transfer characteristic curve. Hence this distortion is also called *nonlinear distortion.* In small-signal amplifiers (voltage amplifiers), the amplified signal is small. Only a small part of the transfer characteristic curve is used. Because of this, the operation takes place over an almost linear part of the curve. It is called linear because changes in the input voltage produce proportionate changes in the output current. It means that the shape of the amplified waveform is the same as the shape of the input waveform. Thus, in case of voltage amplifiers, where small signals are handled, no harmonic distortion occurs.

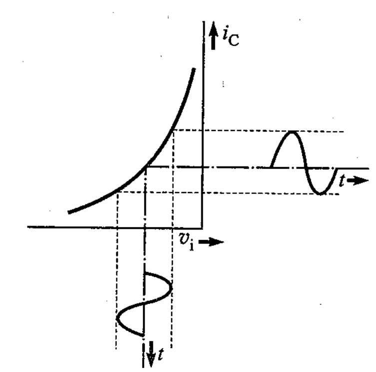


Fig. 9.28 *Harmonic distortion occurs due to nonlinearity of the transfer characteristics*

In power amplifiers, the input signal is large. The change in the output current is no longer proportional to the changes in input voltage (*see* Fig. 9.28). If the input is a sinusoidal voltage, the output is no longer a pure sine wave. This type of distortion is also sometimes called *amplitude distortion.*

9.7 CLASSIFICATION OF AMPLIFIERS

An amplifier is a circuit meant to amplify a signal with a minimum of distortion, so as to make it more useful. The classification of amplifiers is somewhat involved. A complete classification must include information about the following:

1. Active device used
2. Frequency range of operation
3. Coupling scheme used
4. Ultimate purpose of the circuit
5. Condition of dc bias and magnitude of signal

An amplifier may use either a BJT or an FET.

Based on frequency range of operation, the amplifiers may be classified as follows:

1. DC amplifiers (from zero to about 10 Hz)
2. Audio amplifiers (30 Hz to about 15 kHz)
3. Video or wide-band amplifiers (up to a few MHz)
4. RF amplifiers (a few kHz to hundreds of MHz)

Usually, in an amplifier system, a number of stages are used. These stages may be cascaded by either direct coupling, *RC* coupling, or transformer coupling. Sometimes, *LC* (inductance capacitance) coupling is also used. Accordingly, The amplifiers are classified as:

1. Direct-coupled amplifiers
2. *RC*-coupled amplifiers
3. Transformer-coupled amplifiers
4. *LC*-coupled amplifiers

Depending upon the ultimate purpose of an amplifier, it may be broadly classified as either voltage (small-signal), or power (large-signal) amplifier. Till now, we had considered the voltage amplifiers. In the next unit, we shall discuss the power amplifiers.

The amplifiers may also be classified according to where the quiescent point is fixed and how much the magnitude of the input signal is. Accordingly, four classes of operation for amplifiers are defined as follows:

Class A: In class A operation, the transistor stays in the active region throughout the ac cycle. The Q point and the input signal are such as to make the output current flow for 360°, as shown in Fig. 9.29*a*.

Class B: In class B operation, the transistor stays in the active region only for half the cycle. The Q point is fixed at the cut-off point of the characteristic. The power drawn from the dc power supply by the circuit, under quiescent conditions is small. The output current flows only for 180° (see Fig. 9.29*b*).

Class AB: This operation is between class A and B. The transistor is in the active region for more than half the cycle, but less than the whole cycle. The output current flows for more than 180° but less than 360° (see Fig. 9.29*c*).

Class C: In a class C amplifier, the Q point is fixed beyond the extreme end of the characteristic. The transistor is in the active region for less than half cycle. The output current remains zero for more than half cycle, as shown in Fig. 9.29*d*. The dc current drawn from the power supply is very small.

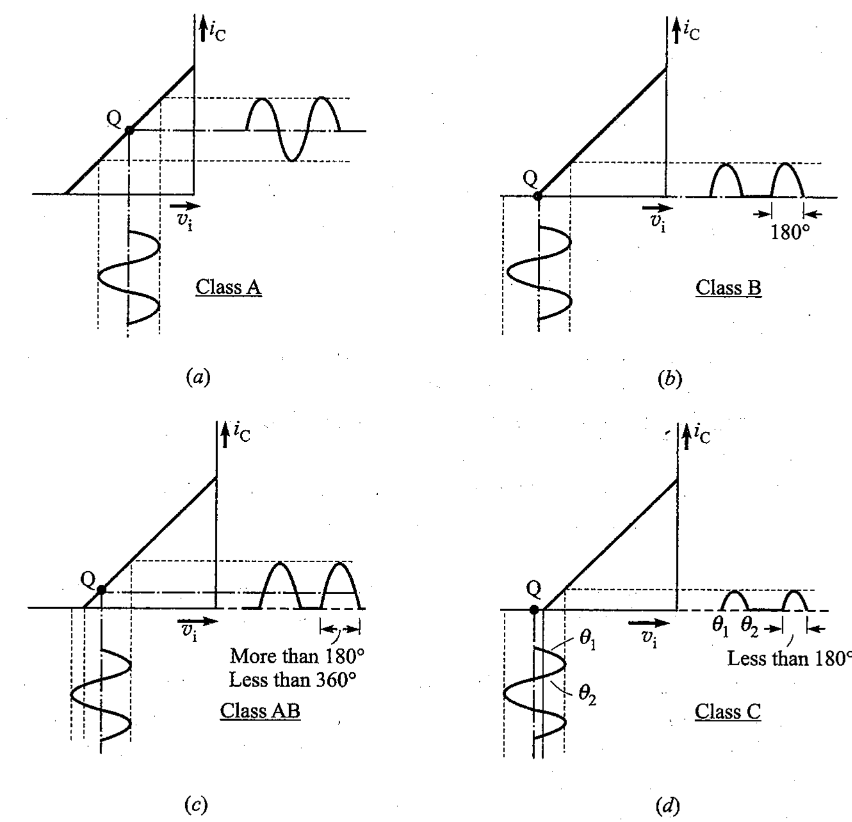


Fig. 9.29 *Classification of amplifiers based on the biasing condition*

In this unit, as well as in the previous unit, we have considered small-signal amplifiers under class A operation. In the units to follow, we will see different applications of other classes of operation.

Review Questions

1. Why do you need more than one stage of amplifiers in practical circuits?
2. State the reasons why we prefer expressing the gain of an amplifier on a logarithmic scale rather than on a linear scale.

3. Define the unit *decibel* for expressing (*a*) voltage; (*b*) current; and (*c*) power.
4. What are the various coupling schemes of two stages of amplifiers?
5. (*a*) Draw the circuit diagram of a two-stage *RC*-coupled amplifier using BJTs. Give the typical values of the components used; (*b*) Explain why *RC*-coupled amplifier circuits cannot be used to amplify slowly varying dc signals.
6. Draw the circuit diagram of a two-stage *RC*-coupled amplifier using JFETs. Give the typical values of the components used.
7. State the applications where you would prefer using transformer-coupling scheme.
8. State the advantages and disadvantages of a transformer-coupled amplifier.
9. What type of coupling scheme would you use for amplifying a signal obtained from a thermocouple meant to measure the temperature of a furnace? Give reasons.
10. Draw the frequency response curve of a typical *RC*-coupled amplifier. Mark the gain-axis in dB. Why do you prefer using logarithmic scale for the frequency axis ? How will you find the 3-dB frequencies from this curve?
11. Why does the gain of an *RC*-coupled amplifier fall in (*a*) low-frequency range, (*b*) high-frequency range?
12. How do you define the cut-off frequencies of an amplifier and what do you understand by the bandwidth of an amplifier?
13. While defining the cut-off frequencies of an amplifier, why do we take 70.7 % of the mid-band gain?
14. When more stages are cascaded to obtain high gain, does the bandwidth of the multi-stage amplifier remain the same as that of the individual stages? If not, why?
15. What do you understand by the loading effect in a multi-stage BJT amplifier? Why does such a loading effect not occur in the case of a multi-stage FET amplifier?
16. Draw the ac equivalent circuit of one stage of a multi-stage *RC*-coupled amplifier using JFETs for (*a*) mid-frequencies, (*b*) low frequencies, and (*c*) high frequencies. Give the expressions of lower and upper cut-off frequencies in terms of circuit components.
17. What do you understand by Hi-Fi amplifier system?
18. What are the different types of distortions that can occur while a signal is amplified by an amplifier? Give the reasons for each type of distortion.
19. It is said that phase distortion does not have any importance in the case of audio amplifiers. Why?
20. Harmonic distortion is also called nonlinear distortion. Why?
21. State at least one typical application of each type of coupling.
22. "In a multi-stage amplifier, the input impedance of an amplifier stage should be very high, and output impedance must be very low." Justify this statement.

23. In a two-stage amplifier, each stage uses identical BJTs and components; yet the voltage gain of the first stage is much less than that of the second stage. Explain why this happens. If FETs are used in this circuit, then it is not so. Why?

Objective-Type Questions

I. Below are some incomplete statements. Four alternatives are provided for each. Choose the alternative that completes the statement correctly.

1. Two identical stages of BJT amplifiers are cascaded by *RC*-coupling. If 10 is the mid-band voltage gain of each stage, the overall gain of the cascaded amplifier will be
   (*a*) 40 dB (*b*) 100 dB
   (*c*) 20 dB (*d*) ($20 \log_{10} 20$) dB
2. For amplifying a signal containing frequency components from 450 kHz to 460 kHz, the most appropriate amplifier is
   (*a*) *RC*-coupled amplifier using MOSFETs
   (*b*) *RC*-coupled amplifier using JFETs
   (*c*) direct-coupled amplifier using BJTs
   (*d*) transformer-coupled tuned amplifier using transistors
3. The coupling capacitor C_C in an *RC*-coupled FET amplifier is usually a
   (*a*) 5 μF; mica capacitor
   (*b*) 0.05 μF; paper capacitor
   (*c*) 0.1 μF; electrolytic capacitor
   (*d*) 50 μF; electrolytic capacitor
4. The main component responsible for the fall of gain of an *RC*-coupled amplifier in low-frequency range is
   (*a*) the active device itself (BJT or FET)
   (*b*) stray shunt capacitance C_{sh}
   (*c*) coupling capacitor C_C
   (*d*) the gate resistor R_G
5. The overall gain of a two-stage *RC*-coupled amplifier is 100. A signal voltage of 10 V, 1 kHz is applied across the output terminals of this amplifier. Then, the voltage obtained across the input terminals will be
   (*a*) 0.1 V, 1 kHz (*b*) 0 V
   (*c*) 100 V, 1 kHz (*d*) 10 V, 1 kHz
6. Harmonic distortion of the signal is produced in an *RC*-coupled transistor amplifier. The probable component responsible for this distortion is
   (*a*) the transistor itself
   (*b*) the power supply V_{CC}
   (*c*) the coupling capacitor C_C
   (*d*) the biasing resistors. R_1 and R_2.

II. Some statements are given below. Choose the correct statement. Rewrite those statements which are false after making necessary corrections.

1. In a multi-stage amplifier there are two or more stages.
2. The overall voltage gain of a multi-stage amplifier is obtained by adding the voltage gains of each stage when expressed as a voltage ratio.
3. When you connect an identical second-stage BJT amplifier to the first stage, the voltage gain of the first stage increases.
4. The lower cut-off frequency of a two-stage *RC*-coupled amplifier is higher than its value for the single-stage amplifier.
5. By cascading the second stage of an identical transistor amplifier, the upper cut-off frequency increases.
6. In a multi-stage amplifier, transformer coupling is used whenever we want to amplify very low frequency or dc signals.
7. *RC*-coupling is the best coupling scheme when frequency of the signal is in the range of 60 Hz to 20 kHz.
8. In a multi-stage voltage amplifier transformer coupling is usually used to amplify audio signals.
9. We always use *RC* coupling for amplifying a small band of rf (1400 kHz to 1410 kHz) signal.
10. While amplifying weak audio signals by a multi-stage amplifier, we should make sure that the coupling network does not disturb the biasing of the next stage.
11. From the point of view of "loading effect" an FET is a better active device than a bipolar junction transistor.
12. The coupling capacitors and bypass capacitors are responsible for the decrease of voltage gain at high frequencies in a multi-stage amplifier.
13. In a multi-stage amplifier, the voltage gain decreases at low frequencies mainly because of junction capacitances of the active device used.

Answers

I. 1. (*a*) 2. (*d*) 3. (*a*) 4. (*c*) 5. (*b*) 6. (*a*)

II. 1. T.

2. The overall voltage gain of an amplifier is obtained by *multiplying* the voltage gains of each stage when expressed as voltage ratio.

   OR

   The overall voltage gain of an amplifier is obtained by adding the voltage gains of each stage when expressed *in terms of dB*.
3. When you connect an identical second-stage BJT amplifier to the first stage, the voltage gain of the first stage *decreases*.
4. T.
5. By cascading the second stage of an identical transistor amplifier, the upper cut-off frequency *decreases*.

6. In a multi-stage amplifier, *direct coupling* is used whenever we want to amplify low frequency or dc signals.
7. T.
8. In a multi-stage voltage amplifier, *RC coupling* is usually used to amplify audio signals.
9. We always use *transformer* coupling for amplifying a small band (1400 kHz to 1410 kHz) of rf signal
10. T.
11. T.
12. The *shunt capacitance* is responsible for the decrease of voltage gain at high frequencies in a multi-stage amplifier.

    OR

    The coupling capacitors and bypass capacitors are responsible for the decrease of voltage gain at *low* frequencies in a multi-stage amplifier.
13. In a multi-stage amplifier, the voltage gain decreases at *high* frequencies mainly because of junction capacitances of the active device used.

Tutorial Sheet 9.1

1. A transistor multi-stage amplifier contains two stages. The voltage gain of the first stage is 50 dB and that of the second stage is 100. Calculate the overall gain of the multi-stage amplifier in dB. **[Ans.** 90 dB]
2. The overall voltage gain of a two-stage *RC*-coupled amplifier is 80 dB. If the voltage gain of the second stage is 150, calculate the voltage gain of the first stage in dB. **[Ans.** 36.47 dB]
3. The voltage gain of a multi-stage amplifier is 65 dB. If the input voltage to the first stage is 5 mV, calculate the output voltage of the multi-stage amplifier. **[Ans.** 8.89 V]
4. An audio signal contains 40 Hz as the lowest frequency and 10 kHz as the highest frequency. This signal is amplified by an amplifier whose maximum gain at 1 kHz is 20 dB. Draw the frequency response curve of this amplifier, indicating lower and upper cut-off frequencies.

Experimental Exercise 9.1

Title Two-stage BJT *RC*-coupled amplifier.

Objectives To

1. trace the given circuit and note down the value of each component;
2. measure the operating-point collector current and collector-to-emitter voltage for both the amplifier stages;

3. measure the voltage gain of the first stage with and without connecting second stage;
4. explain the loading effect of the second stage on the first stage;
5. measure the maximum signal which can be fed to the input of two-stage *RC*-coupled amplifier without causing distortion in the output;
6. measure the overall gain of the two-stage *RC*-coupled amplifier.

Apparatus Required Experimental board of two-stage amplifier, signal generator, CRO, ac millivoltmeter and electronic multimeter.

Circuit Diagram As given in Fig. E. 9.1.1 (typical values of the components are also given).

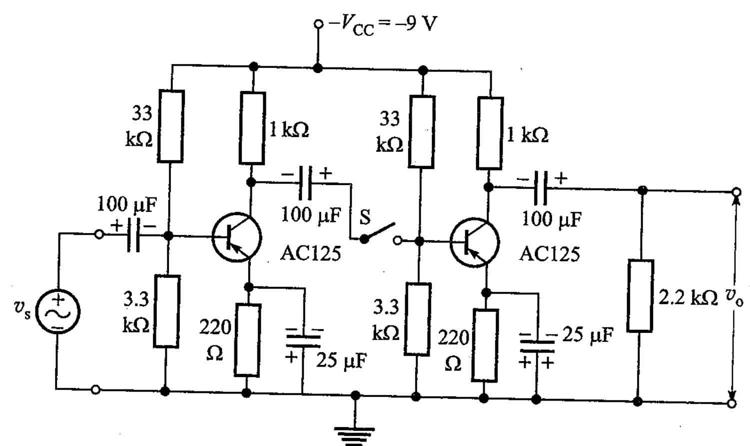


Fig. E. 9.1.1

Brief Theory When the voltage gain provided by a single stage is not sufficient, we have more than one stage in the amplifier. The overall gain of the two stages is given as

$$A = A_1 \times A_2$$

where A_1 is the voltage gain of the first stage and A_2 is the voltage gain of the second stage.

In Fig. E. 9.1.2, two stages are shown connected through a switch S. It is observed that the voltage gain of the first stage depends upon whether the switch S is closed

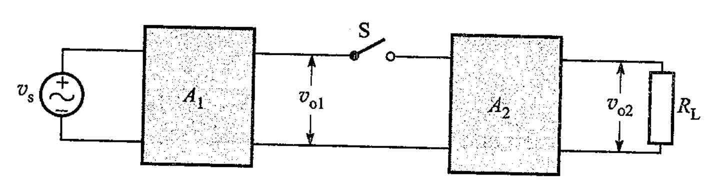


Fig. E. 9.1.2

or open. When the switch S is open, the voltage gain is high. If the switch S is closed the output voltage v_{o1} is very much reduced. This effect is known as *loading of first stage.* When the switch S is closed, the input impedance of the second stage comes in parallel with the load resistance of the first stage. Because of this, the effective load resistance of the first stage is reduced and hence the gain (or output voltage) also decreases.

Procedure

1. Take the experimental board and trace the given circuit. Find out whether the transistor is PNP or NPN. Note down the values of all the resistors and capacitors.
2. Connect the dc voltage from the regulated power supply. Select a voltage, say 9V. You may use 6V or 15V in case single-voltage IC power supply is available.
3. Note the Q point for both the transistors, i.e., find I_C and V_{CE} for both the transistors.
4. Feed ac signal from an audio oscillator to the input of the first stage. Adjust the frequency at 1 kHz. See the output waveshape on the CRO. Go on increasing the input ac voltage till distortion starts appearing in the output waveshape. Note the value of this input signal. This is the maximum signal handling capacity. Repeat the same experiment with a single stage (open the switch S).
5. Measure the signal voltage at:
   (*i*) the output of the first stage
   (*ii*) the output of the second stage
   Calculate from these readings, the voltage gain of first stage, second stage and also overall voltage gain of the two stages.
6. Disconnect the second stage and then measure the output voltage of the first stage. Calculate the voltage gain of the first stage under this condition and compare it with the voltage gain obtained when the second stage was connected.

Observations

1. *Q point of the transistor*:

   V_{CC} = ________ V

   For first stage:

   V_{C1} = ________ V; $I_{C1} = \dfrac{V_{CC} - V_{C1}}{R_{C1}}$ = ________ mA

   V_{CE1} = ________ V

   For second stage:

   V_{C2} = ________ V; $I_{C2} = \dfrac{V_{CC} - V_{C2}}{R_{C2}}$ = ________ mA

   V_{CE2} = ________ V

2. (*a*) Maximum signal handling capacity of the two-stage amplifier =___ mV
   (*b*) Maximum signal handling capacity of single-stage amplifier = ____ mV.
3. *Voltage gain*:

S. No.	*Input voltage*	*Output of first stage*	*Output of second stage*	A_1	A_2	$A = A_1 \times A_2$
1.						
2.						

4. *Voltage gain with second stage disconnected*:

S. No.	*Input voltage*	*Output voltage*	*Gain A_1*
1.			
2.			

Result

1. Both the transistors are operating in active region, as shown by the Q-point readings.
2. The two-stage amplifier can handle a signal of _____ mV only. This is so, because the overall gain of the two stages is very high. The single-stage amplifier can handle a signal of ____ mV.
3. Loaded gain of the first stage (   ) is much less than its unloaded gain (   ).
4. The overall gain A = ____

   = 20 $\log_{10}$ (   )

   = ______ dB

Experimental Exercise 9.2

Title Frequency response curve of two-stage RC-coupled amplifier.

Objectives To

1. identify the values of all the resistors and capacitors used in the given circuit;
2. make sure that the transistors are working in active region by measuring the Q point of the amplifier;
3. make sure that excessive signal is not fed to the input, by seeing undistorted output waveshape on the CRO;
4. plot the frequency response curve of the single-stage amplifier;
5. determine the values of upper and lower cut-off frequencies of single-stage amplifier;
6. plot the frequency response curve of two-stage RC-coupled amplifier;
7. determine the values of upper and lower cut-off frequencies for a two-stage RC-coupled amplifier;

8. compare the bandwidth of a two-stage amplifier with that of single-stage amplifier.

Apparatus Required Experimental board, signal generator, electronic multimeter, ac millivoltmeter, CRO, and transistorised (or IC) power supply.

Circuit Diagram This is same as Fig. E. 9.1.1.

Brief Theory The voltage gain of an RC-coupled amplifier is maximum around 1 kHz. As frequency decreases, the gain starts falling. This decrease in gain at low frequencies is mainly because of coupling capacitors C_C. At low frequencies, coupling capacitors offer sufficiently high impedance. There occurs a voltage drop across these capacitors and hence the output voltage decreases. Also, the bypass capacitors at very low frequencies are no longer effective short circuits. Because of this, the ac current passes through the resistor R_E. This gives rise to negative feedback, and the voltage gain reduces. The voltage gain also decreases at high frequencies because of (*i*) the shunt capacitances made up of junction capacitances and wiring capacitances, and (*ii*) the decrease in β at such frequencies.

The frequencies where voltage gain falls by 3 dB or becomes 70.7 % of the maximum value are called cut-off frequencies. These frequencies can be determined from the frequency response curve.

Procedure

1. You are already familiar with the experimental board. After making sure that transistors are biased in the active region, feed the input signal such that the output is undistorted at 1 kHz. Find the overall gain of the two-stage amplifier.
2. For plotting the frequency response curve of the first stage, disconnect the second stage. Now find the voltage gain of the amplifier at 1 kHz. Change the signal generator frequency on the lower frequency side. You will observe that the voltage gain is decreasing. Find such a frequency (f_1) where gain becomes 70.7 % of maximum gain. In a similar manner, find the signal frequency on the high frequency side (f_2) where voltage gain is reduced to 70.7 % of the maximum gain. Calculate $f_2 - f_1$. This gives the value of bandwidth of the single-stage amplifier.
3. Now connect the second stage with the first stage. By feeding a 1-kHz signal, find the maximum gain. Now decrease the frequency such that the gain is reduced to 70.7 % of the maximum gain. This gives the value of lower cut-off frequency (f'_1) of the two-stage amplifier. By changing the frequency of the signal generator on the higher frequency side, determine the value of the upper cut-off frequency (f'_2). The bandwidth of the two-stages amplifier ($f'_2 - f'_1$) can now be calculated.
4. Most of the times, the signal generator does not supply constant output when the frequency is changed. It is possible that when measurements at f_1, f_2, f'_1 and f'_2 are made, the input would have changed. Now measure the voltage gain of a single stage as well as the two-stage amplifier near these frequencies. Take less