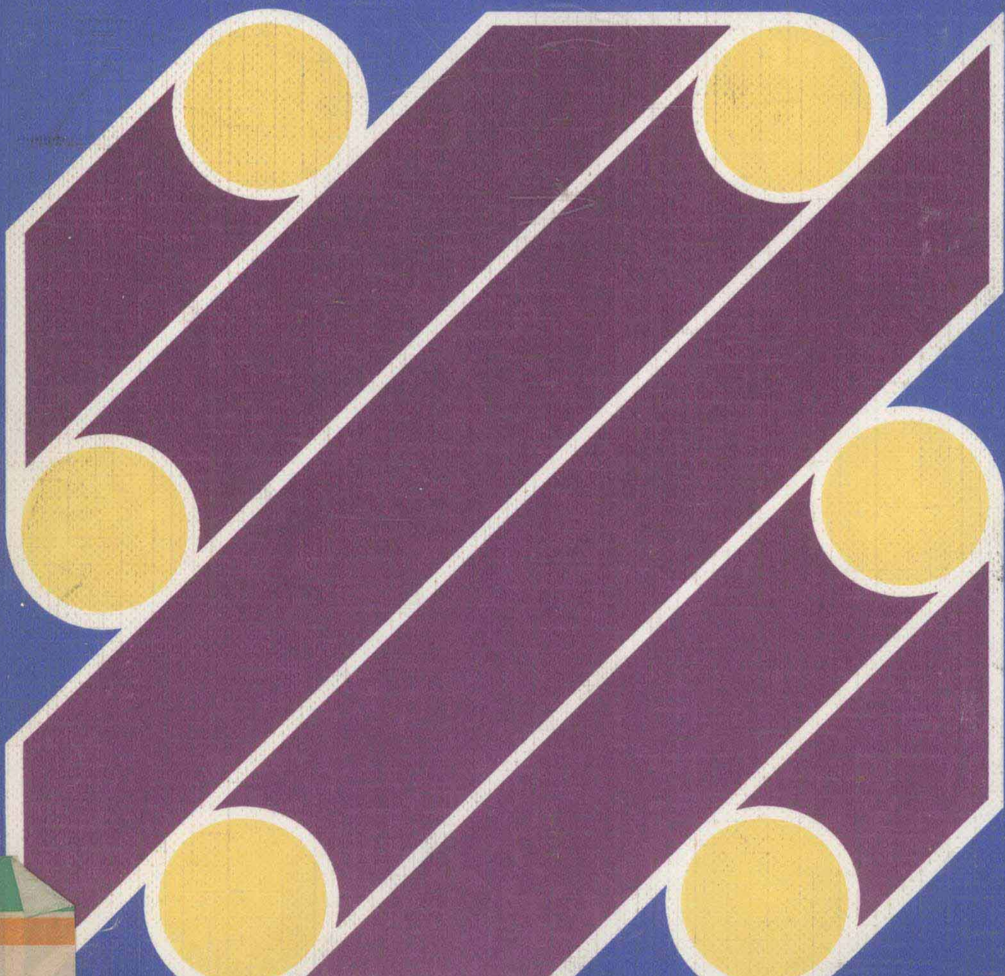
DIGITAL ELECTRONICS FOR WORKS ELECTRICIANS

Noel M Morris



Digital electronics for works electricians

Noel M. Morris

McGraw-Hill Book Company (UK) Limited

London · New York · St Louis · San Francisco · Auckland · Bogotá
Guatemala · Hamburg · Johannesburg · Lisbon · Madrid · Mexico
Montreal · New Delhi · Panama · Paris · San Juan · São Paulo
Singapore · Sydney · Tokyo · Toronto

Published by McGraw-Hill Book Company (UK) Limited
MAIDENHEAD · BERKSHIRE · ENGLAND

British Library Cataloguing in Publication Data
Morris, Noel Malcolm
Digital electronics for works electricians
1. Digital electronics
I. Title
621.3815 TK7868.D5 78–41105

ISBN 0–07–084523–9

12345 WC&S 81079

PRINTED AND BOUND IN GREAT BRITAIN

Preface

The remarkable advances in electronics are typified by the explosive growth of digital electronics. Digital systems have developed to a point where industry is looking to the computer to carry out many functions previously performed by electro-mechanical devices. This book has been written to satisfy the need for information in the field of digital electronics. In writing it I have endeavoured to strike a balance between the complementary functions of education and training, for both are vital to the works electrician of today.

The book should suit many industrial training schemes and will enhance the digital electronics knowledge of electricians and others who have to install and service digital equipment. The synopsis of the book owes much to discussions between Mr R. Wander, Head of Production Training at the Production Engineering Research Association, and myself.

The book may be broadly divided into the chapters dealing with semiconductor devices and integrated circuits (chapters 1 and 10), logic gates (chapters 2–4, inclusive), bistable elements and memory systems (chapters 5 and 9), binary numbers, counters and shift registers (chapters 6–8, inclusive), computers and microprocessors (chapter 11), and test equipment (chapter 12).

I would like to express my thanks to the individuals and organizations who have materially contributed to the book. Particular thanks are due to Mr R. Wander (PERA), to Mr S. Rakowski, MSc, Senior Lecturer at the North Staffordshire Polytechnic, and to other reviewers of the manuscript. The manufacturers who have contributed to the book include A V R Electronic Services, GenRad Ltd., G R Electronics Ltd., Hewlett Packard, Instem Ltd., Intel Corporation, National Semiconductors (UK) Ltd., Quarndon Electronics, R S Components Ltd., and Teknis Ltd. I would also like to thank the electronics industry at large for supplying other information used in the book. I also acknowledge the assistance given by the editorial and production staff of the McGraw-Hill Book Company.

I am also greatly indebted to my wife for her not inconsiderable efforts in the preparation of the manuscript, and for her patience with my preoccupation while it was being written.

I hope that readers will enjoy this book, which is dedicated to the continuing development of electronics.

Noel M. Morris

Contents

1

Semiconductor devices and integrated circuits

1.1 Electrons and holes

Current flow is explained in terms of the movement of electrical *charge carriers* between points in a circuit. To explain current flow, it is necessary to know something about the atomic structure of materials used in electronic circuits.

From an engineering standpoint, an atom consists of two types of charged particles, namely ***electrons*** and ***protons***. The mass of a proton is 1840 times greater than that of an electron, and having the greater mass the protons are concentrated in the centre or *nucleus* of the atom, as shown in Fig. 1.1. A proton has a positive electrical charge, while an electron has an equal and opposite charge. Each atom has as many electrons as it has protons, so that the net electrical charge on it is zero; should the atom lose an electron, it is left with a net positive charge. Should the atom gain an electron, it has a net negative charge.

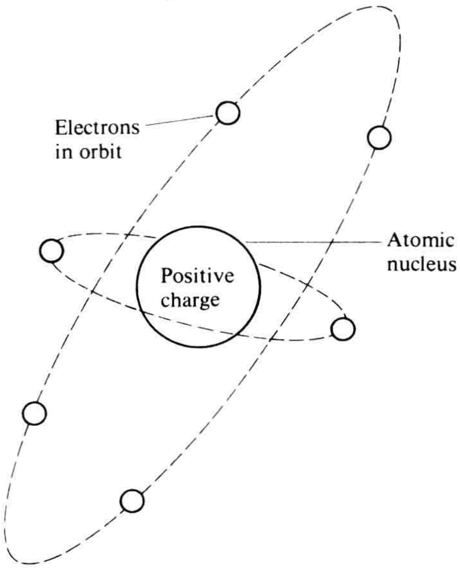


Fig. 1.1 Electrons in orbit around the nucleus

Electrons orbit around the nucleus (Fig. 1.1) in what are known as ***layers***, ***bands*** or ***shells***. A simple analogy of energy layers in an atom are parking levels in a multistorey car park. Ground level can be regarded as the nucleus of the atom, while higher parking levels in the car park represent energy levels within the atom at which electrons can be expected to be found. In nature the lower energy levels in the atom are 'filled' with electrons before electrons appear in higher levels (also, as is usually the case with a multistorey car park, the lower levels are filled first). When an energy level is 'full', no movement of electrons can take place on that level. However, at the highest energy level at which electrons are found it is unlikely that the level is completely full. The outermost shell in the atom which contains electrons is known as the ***valence shell*** or ***valence energy band***; the electrons which ultimately take part in the process of conduction orbit in the valence shell.

For an electron to take part in conduction, it must acquire more energy than that required to keep it orbiting in the valence band. The energy may come from one of many sources including heat, light and atomic radiation. The application of an electrical voltage to a semiconductor results in the electrons in the conduction band being subjected to an electrical force, which propels them towards the positive pole of the supply. The electrons arriving at the positive pole constitute current flow in the circuit; at the same time an equal number of electrons are injected into the circuit from the negative pole of the battery to maintain electronic charge balance in the semiconductor.

When an electron escapes from an atom, it becomes a *mobile* negative charge; since the atom has lost an electron (a negative charge), it is left with a net positive charge of one electronic unit. As a result of this, it will try to attract to itself a mobile electron which is in the vicinity of the atom. On this basis it is possible to describe the positive charge on the atom as an electronic ***hole***; a hole can be regarded as a positive charge carrier much as an electron is a negative charge carrier. *A hole is simply the absence of an electron at a point where one would normally be found.*

1.2 Semiconductors

A semiconductor is a material whose resistivity lies between that of a good conductor and a good insulator. Commonly used semiconductor materials in diodes and transistors are ***silicon*** and ***germanium***, whilst other materials such as gallium arsenide and gallium phosphide are used in other devices such as light-emitting diodes (LED). Silicon is the most widely used semiconductor material and is an element found in many forms of rocks and stones (sand is silicon dioxide).

Current flow in a pure semiconductor occurs as a result of the movement of electrons and holes which are generated mainly by thermal effects in the manner outlined in section 1.1. As the ambient temperature increases, the number of electrons which are able to take part in conduction increases. Hence at a given value of applied voltage, the current flowing through the semiconductor material increases with temperature, i.e., its resistance decreases with increasing temperature. In other words, semiconductors have a ***negative resistance-temperature coefficient*** (n.t.c.). The current flow from this effect is, in the case of diodes and transistors, a nuisance since it gives rise to a leakage current which is dependent on temperature. Devices made from germanium have a higher value of leakage current due to this cause than those made from silicon. Conductivity from the above cause is known as ***intrinsic conductivity***; for this reason, pure semiconductors are sometimes described as ***i-type*** semiconductors.

In order to introduce conductivity which can be controlled by the user of semiconductor devices, the manufacturer introduces a controlled amount of impurity during manufacture. The amount introduced is typically one part per ten million parts of pure material. Depending on the type of impurity added, the semiconductor is described as either a p-type material or an n-type material.

In a ***p-type*** material, the added impurity causes the semiconductor to have an excess of 'free' holes. The name p-type implies mobile ***P***ositive charge carriers. Gallium and indium are materials often used to dope silicon to give a p-type semiconductor. The reader is reminded that in an isolated piece of p-type material there are equal numbers of electrons and holes, so that electrical neutrality exists within the slice as a whole; however, not all the charge carriers are 'free' to act as current carriers since the majority are tied by atomic binding forces. Thus current flow in a p-type material is largely due to the movement of mobile positive charge carriers towards the negative pole of the supply. Hence in p-type semiconductors, holes are the ***majority charge carriers***. At the same time a very small number of mobile electrons are generated by thermal effects (see above), and these make a small contribution to current flow. In p-type material, electrons are described as ***minority charge carriers***.

In an ***n-type*** material, the dopant used (typically arsenic or antimony) causes it to have mobile electrons (n-type implies *N*egative charge carriers). In n-type material, the majority of current flow is due to the movement of electrons through the material, and *in this case electrons are the majority charge carriers*. Movement of holes (*which are minority charge carriers in n-type material*), which are generated by thermal effects also make a small contribution to the total current flow.

Both p-type and n-type semiconductors are used in the manufacture of semiconductor devices. A few specialized devices use i-type semiconductors.

1.3 Diode characteristics

A diode is a two-terminal electronic device, one electrode being known as the ***anode*** and the other as the ***cathode*** (see Fig. 1.2(a)). The diode presents a low resistance to the flow of current when the anode potential is positive with respect to the cathode, when it is said to be ***forward biased***. When the anode is negative with respect to the cathode the diode presents a very high resistance to the flow of current, and is said to be ***reverse biased***. Thus the diode may be regarded as a ***voltage sensitive switch*** which is ON when it is forward biased, and is OFF when reverse biased.

On a practical semiconductor diode (see also section 1.4), the cathode

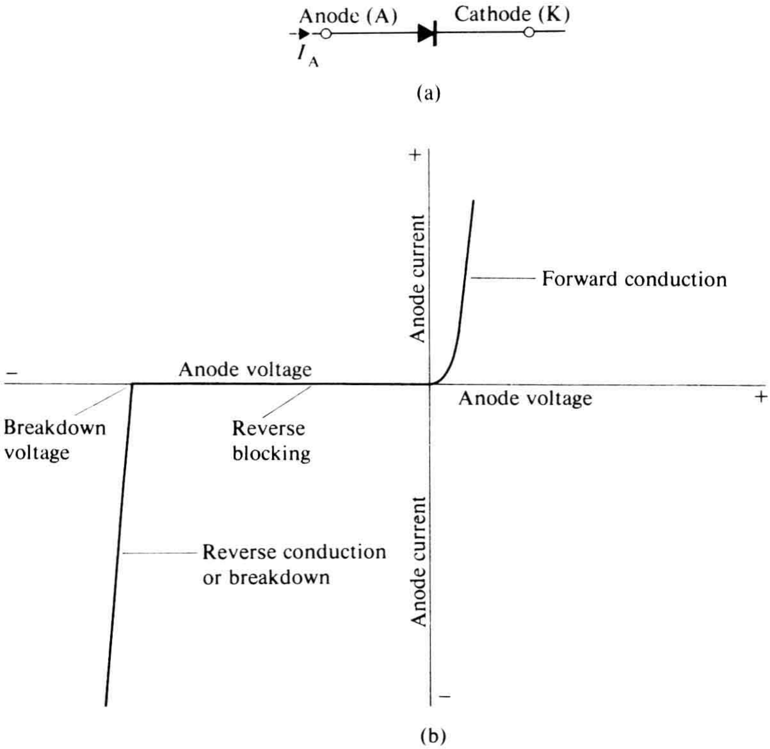


Fig. 1.2 (a) Circuit symbol for diode, and (b) its characteristic

electrode is usually clearly marked, sometimes by means of a coloured band or dot. Diodes can be tested using a multimeter switched to its OHMS range; the reader should note that the '+' terminal of an analogue instrument is in fact *negative* with respect to the '−' terminal due to the connections of the internal battery (for details see *Electronics for Works Electricians*, by Noel M. Morris, (McGraw-Hill, Maidenhead)). When the anode of a healthy diode is connected to the '−' terminal of the meter and the cathode to the '+' terminal, the resistance indicated by the meter should be low; when the connections are reversed the instrument should indicate infinity ohms.

When forward biased the ***forward voltage drop*** across the diode is typically 0·2 to 0·6 V for a germanium diode, and 0·5 to 0·8 V for a silicon diode (see Fig. 1.2(b)).

When reverse biased at a voltage below the breakdown voltage, the diode is said to be in its ***reverse blocking mode***; in this operating state only ***leakage current*** flows through the diode. In diodes normally associated with logic circuits, the value of the leakage current is typically a few nanoamperes (1 nA = one thousandth of one millionth of one ampere); for the reasons given above, the leakage current increases slightly with increase in temperature.

There is a value of reverse bias voltage known as the ***breakdown voltage*** (see

Fig. 1.2(b)) beyond which the reverse current increases rapidly; at voltages greater than the breakdown voltage the diode is said to work in its ***reverse breakdown mode***. In the case of signal diodes used in logic circuits and also in rectifier diodes, this operating state causes the power dissipation in the diode to rise to a dangerously high level and may result in the diode being damaged. Special diodes known as Zener diodes (see section 1.5) are designed to work in the reverse breakdown mode.

1.4 Semiconductor p-n junction diodes

The p-n junction diode consists of a single crystal of semiconductor material containing a p-type anode region and an n-type cathode region (see Fig. 1.3).

When the diode is connected as shown in Fig. 1.3(a), with the p-type anode positive with respect to the n-type cathode (i.e., forward biased), it is found that a large current flows in the circuit. The reason is as follows. In p-type material the

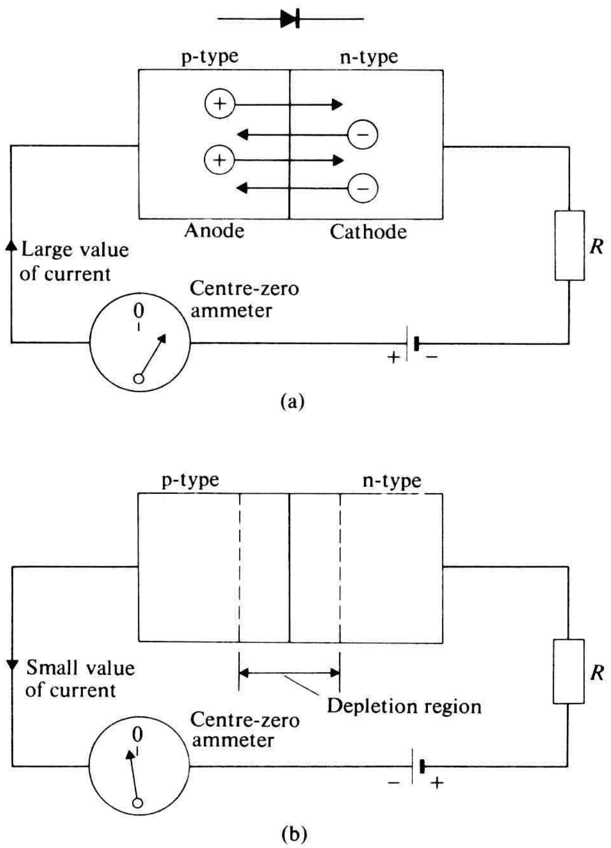


Fig. 1.3 (a) Forward biased p-n junction diode, and (b) reverse biased diode

majority charge carriers are holes, and in n-type material are electrons. When the p-region is connected to the positive pole of the battery, holes are injected into the anode from the positive supply; these mobile holes are attracted into the n-region by the negative polarity of the supply. At the same time, electrons are injected into the n-region from the negative pole of the battery and these mobile charges are attracted to the p-region by the positive polarity of the power supply. The flow of these charge carriers constitutes current flow in the circuit. Thus ***the diode is forward biased when the p-type anode is positive with respect to the n-type cathode.***

If the polarity of the supply is reversed as in Fig. 1.3(b), the circuit current becomes practically zero (the leakage current). In this mode (reverse biased), the negative potential applied to the p-type anode attracts mobile holes (positive charge carriers) away from the p-n junction. At the same time the positive potential applied to the n-type cathode attracts mobile electrons away from the p-n junction. The net result is that there remain no mobile charge carriers in the region of the p-n junction; that is the region around the junction is depleted of charge carriers, and is known as a ***depletion region***. Since the depletion region does not contain any mobile charge carriers, it is effectively an insulating region which prevents current flow through the diode (other than leakage current). In the case of rectifier diodes, reverse breakdown of the diode occurs when the reverse bias voltage is sufficiently high to rupture the depletion region. In Zener diodes (section 1.5), reverse breakdown occurs at a fairly low voltage.

1.5 Zener diodes

A Zener diode is a p-n junction diode whose impurity doping is much greater than that of a normal diode, resulting in reverse breakdown occurring at fairly low voltage (Zener diodes with breakdown voltages in the range between a few volts and a few hundred volts are obtainable). In the case of a particular device, the breakdown voltage has a stable value which is reasonably constant over a fairly wide range of reverse current values.

Zener diodes are almost invariably used in the reverse breakdown mode, with the anode being connected to the negative pole of the supply. Providing that the rating of the diode is not exceeded, then any value of reverse current up to the rated value does not damage the device; for example, a 1 W, 4 V Zener diode can carry a reverse current of

$$I = \frac{\text{power dissipation}}{\text{breakdown voltage}} = \frac{1}{4} = 0{\cdot}25\ \text{A}$$

A popular application of these diodes in electronic circuits is as a voltage reference source, whose function it is to provide a stable value of voltage. For this type of application, standard cells have been largely superceded by Zener diode circuits.

A simple voltage reference source is shown in Fig. 1.4. Here an unstabilized supply voltage V_S is supplied from a bridge rectifier and smoothing circuit. The value of V_S is typically twice the value of the breakdown voltage (V_Z) of the Zener

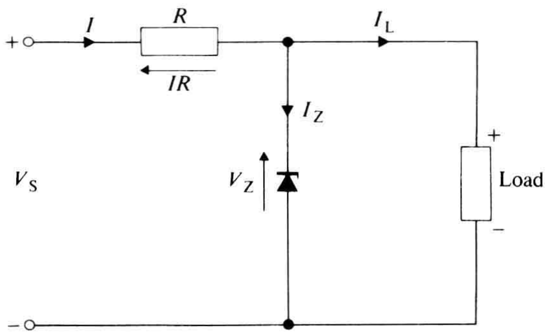


Fig. 1.4 Zener diode voltage reference source

diode. The Zener diode symbol is characterized by the 'cranked' cathode electrode (symbolizing the breakdown effect). Resistor R fulfils the dual functions of being a current limiting device and that of a voltage 'dropper'; the difference in potential between V_S and V_Z appears across the resistor. Since the Zener diode breakdown voltage is constant, the voltage across the load also remains constant despite variations in the supply voltage; variations in V_S appear across R and not across the Zener diode.

1.6 Storage time in diodes

In order to turn OFF a p-n junction diode which is conducting, it is first necessary to remove all the mobile charge carriers in the region of the p-n junction. Only when this has been done can the insulating depletion region be built up at the p-n junction. The time taken for this to take place is known as the *storage time* of the device. The storage time in a diode used in a logic circuit may be as small as a fraction of a nanosecond (1 ns $= 10^{-9}$ s or one thousandth of one millionth of one second), and up to one millisecond in a high current diode (1 ms = one thousandth of one second). While the periods of time mentioned are small, the storage time imposes a limit on the maximum speed at which p-n junction devices can be switched. When a diode carries a large current, it takes longer to remove the charge carriers and the storage time is increased above the value for a low value of current.

Although the description above refers to junction diodes, charge carrier storage effects are present in all p-n junction devices (such as transistors, thyristors, triacs, etc.) when they are switched from the ON state to the OFF state.

1.7 Schottky barrier diodes

A Schottky diode is a rectifying metal-to-semiconductor junction device. The semiconductor material is usually n-type, but p-type may also be used. Due to the different mechanism of operation when compared with p-n junction devices, Schottky diodes do not exhibit charge carrier storage effects and their storage time is zero.

In order to reduce the overall switching time of some types of logic gates, Schottky diodes are incorporated in their circuits.

1.8 Bipolar junction transistors

The name transistor is a contraction of **TRAN**sfer re**SISTOR**, which the inventors coined to describe their operation. The 'bipolar' part of the name comes from the fact that both polarities of charge carrier (bi-polarity), i.e., holes and electrons, are involved in current flow through the device.

The bipolar transistor is a three-layer semiconductor device manufactured in a single piece or 'chip' of semiconductor material such as silicon or germanium. The three regions are known as the ***emitter***, the ***base*** and the ***collector*** regions, respectively, illustrated in Fig. 1.5.

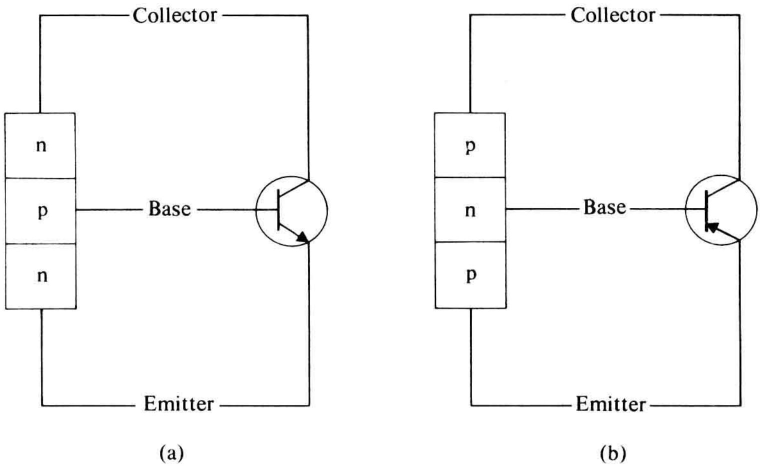


Fig. 1.5 Transistor symbols: (a) n-p-n, and (b) p-n-p

The transistor symbol is designed to give information about the transistor as follows. An arrow is always drawn on the emitter electrode, and points in the direction of 'conventional' current flow (i.e., hole flow). Thus if the arrow points away from the base region, then the base region is p-type; if the arrow points towards the base region, then the base region is n-type. The emitter region gets its name from the fact that charge carriers are emitted from it; the collector region is so named because the majority of the charge carriers leaving the emitter are collected there. The thin base region provides an electrical means of controlling the current flow through the transistor; it gets the name from the fact that in early transistors it was the strong physical base which supported the transistor.

Two principal forms of junction transistor are in use, namely n-p-n types (Fig. 1.5(a)) and p-n-p types (Fig. 1.5(b)). The actual size of a transistor is very small, one type having an area of $7 \times 30\ \mu m$ ($1\ \mu m$ = one millionth of a metre). Both forms of transistor are in use and each is constructed in a single chip of semiconductor material. One method of production is described later in the chapter.

In use, two voltage sources are used to supply the bipolar transistor, one supplying the input circuit and the other the output circuit and both having a common electrical connection. In a practical circuit the two voltages mentioned above are obtained from a single supply. It is the transistor electrode which is connected to this common connection which is used to describe the circuit; thus we have the ***common-emitter*** circuit, the ***common-base*** circuit and the ***common-collector*** circuit. The common-emitter circuit is the most popular, and a simplified version is shown in Fig. 1.6.

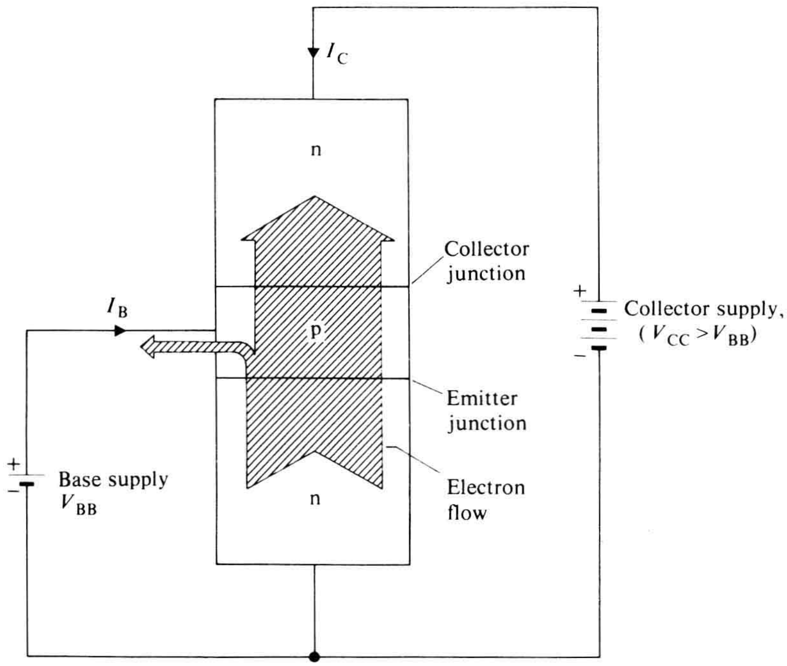


Fig. 1.6 The common-emitter connection

When used as a logic switching element, the base current I_B either has a large value (large, that is, in electronic terms) or it is zero. Let us consider first of all the condition where I_B has a large value, that is when the value of the base supply voltage V_{BB} forward biases the p-n emitter junction. To look further into the operation of the transistor it is necessary to know a little more about the doping levels in the transistor regions. When transistors are first manufactured it is always arranged that the emitter region is very heavily doped so that it has an abundance of mobile charge carriers; the emitter in Fig. 1.6 is n-type, so that it has an abundance of mobile electrons. Also, the base region is very lightly doped and has only a few mobile charge carriers (these are holes in the p-type base region); in fact the base region is so lightly doped that it is almost an i-type