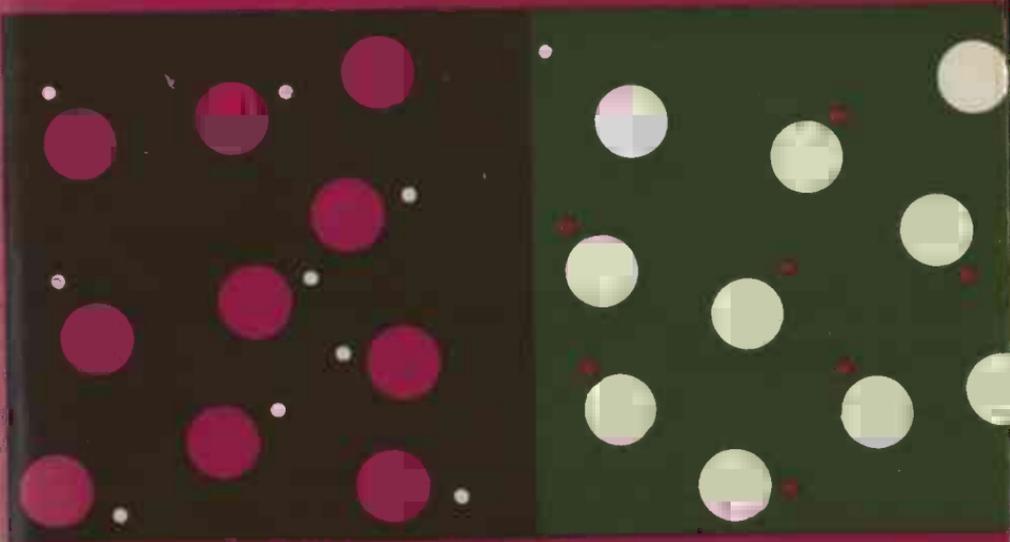


BEGINNER'S GUIDE TO TRANSISTORS



J. A. Reddihough

BEGINNER'S GUIDE TO TRANSISTORS

by

J. A. Reddihough

Transistors, today, dominate the field of electronics. This addition to Newnes' series of Beginner's Guides provides a readable introduction to transistors and their applications for the younger reader who intends to make a career in electronics as well as for the layman of any age who takes an interest in technical matters and who requires a simple but comprehensive account of this modern device.

The book describes what transistors are, how they work, the many types available and their many applications. In doing this, it serves also to introduce the reader to many basic techniques used in electronics engineering at the present time.

The treatment is non-mathematical, but a few simple formulae are given to indicate important relationships. A chapter is included on integrated circuits.

15/-
NET

Newnes books of allied interest

- Beginner's Guide to Electronics
by Terence L. Squires, A.M.I.E.R.E.
- Beginner's Guide to Electricity
by Clement Brown
- Beginner's Guide to Television
by Gordon J. King
- Questions and Answers on Transistors
by Clement Brown
- Questions and Answers on Electronics
by Clement Brown
- Questions and Answers on Radio and Television
by H. W. Hellyer
- Questions and Answers on Audio
by Clement Brown
- Rapid Servicing of Transistor Equipment
by Gordon J. King
- Transistor Pocket Book
by R. G. Hibberd, B.Sc., M.I.E.E., Sen.M.I.E.E.E.
- Electronics Pocket Book
edited by J. P. Hawker and J. A. Reddihough
- Television Engineers' Pocket Book
edited by J. P. Hawker and J. A. Reddihough
- Radio Servicing Pocket Book
edited by J. P. Hawker
- Outline of Radio and Television
by J. P. Hawker

BEGINNER'S GUIDE TO TRANSISTORS

by
J. A. Reddihough

NEWNES BOOKS

© The Hamlyn Publishing Group Ltd. 1968

First published 1968

Published for
NEWNES BOOKS
by The Hamlyn Publishing Group Ltd.,
42 The Centre, Feltham, Middlesex.

MADE AND PRINTED IN GREAT BRITAIN BY
RICHARD CLAY (THE CHAUCER PRESS), LTD., BUNGAY, SUFFOLK

CONTENTS

<i>Preface</i>	vii
1 <i>Introduction to Semiconductor Devices</i>	9
2 <i>Types of Transistor and Associated Devices</i>	34
3 <i>Basic Transistor Circuits and Characteristics</i>	49
4 <i>A.F. Techniques</i>	67
5 <i>R.F. Techniques</i>	89
6 <i>Electronic Circuits</i>	118
7 <i>Power Supplies</i>	138
8 <i>Integrated Circuits</i>	144
9 <i>Servicing Transistorised Equipment</i>	149
<i>Index</i>	159

PREFACE

THIS addition to Newnes' series of Beginner's Guides takes the subject of transistors, describing what they are, how they work, the main types and the many applications in which they are used. The transistor today dominates the field of electronics, being used in vast quantities in radio receivers and other domestic equipment, computers, data-processing equipment, instruments, industrial automatic control systems, air and sea navigational and communications equipment, telecommunications links and so on, so that this book goes quite a way towards introducing the reader to the basic techniques used in electronics at the present time.

There is much to be said for an introductory book written when the technology it covers has become well established. For one thing the book can be based on what has become established practice, which cannot in the nature of things be foreseen in the early days; and for another the next generation of devices will have begun to emerge so that some indication can be given of the way in which the technology is developing. In the case of the present book, we already know that the future lies with the integrated circuit, and a chapter is accordingly included covering this subject.

In common with the other Beginner's Guides, the book assumes that the reader starts with negligible knowledge of the subject, commencing with a brief account of the nature of electric currents before going on to the ways in which semiconductor devices respond to these. The treatment throughout is non-mathematical, though one or two simple mathematical formulae and examples are given in places to indicate the relationships between certain quantities and the magnitudes of voltages, amplification and so on involved.

The aim has been to cover thoroughly the circuits in which

transistors are used, concentrating initially on domestic equipment such as transistor radios and record reproducers, and subsequently going on to the other main applications. The final chapter provides a practical guide to what to look for when confronted with faulty equipment and how to go about fault location.

J. A. R.

INTRODUCTION TO SEMICONDUCTOR DEVICES

AN electric current consists of the organised movement of electrons, so that to understand electrical and electronic devices, such as transistors, semiconductor diodes and similar devices, it is first necessary to know a little about the structure of the atom. The atom consists of a central nucleus around which rotate in orbit one or more electrons. The simplest atom, that of hydrogen, consists of a nucleus with a single electron in orbit. The larger number of electrons in more complex atoms are arranged in a series of shells, and the number of electrons in each shell or orbit obeys a definite law. Thus there are never more than two electrons in the innermost shell, never more than eight in the next, never more than eighteen in the next, and so on. As regards germanium and silicon, the two most commonly used semiconductor materials, there are in the former 32 electrons arranged in four shells and in the latter 14 electrons arranged in three shells. The silicon atom is shown diagrammatically in Fig. 1.1.

A force of attraction exists between the nucleus of an atom and its electrons, this force being the basis of all electrical phenomena. The electron is said to carry a negative electrical charge and the nucleus a positive electrical charge, and an atom that has its complete complement of electrons—it is possible, as we shall see in a moment, for an atom to gain or lose electrons—is electrically neutral, as the positive charge carried by the nucleus is equal to the total of the negative charges carried by its electrons.

The electrons in the orbit closest to the nucleus of an atom are

tightly bound to the atom. Conversely, those in the outermost orbit—called *valence electrons* as they form valence bonds with atoms of different material in chemical reactions—are more loosely bound to the nucleus and in fact in many materials are so loosely bound that many of them escape from their parent atoms and drift around in the substance of which they form a part. An

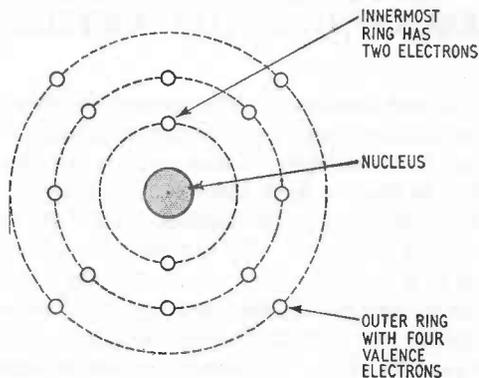


Fig. 1.1. Diagrammatic representation of the silicon atom, which has fourteen electrons orbiting in three rings around its central nucleus. The outer, valence, ring has four electrons in orbit. The representation here is two-dimensional.

atom that has lost an electron is no longer electrically neutral: it carries a net positive charge, and is called a *positive ion*. An atom to which an extra electron has attached itself carries a net negative charge, and is called a *negative ion*.

At absolute zero temperature (-273°C) the atomic structure of matter would be complete and intact, each atom existing stably and at rest, with its correct complement of electrons. Consequently this condition would provide electrical insulation, since there are no free electrons present to act as current carriers. In-

ulators—for example, air, wood, mica and most plastics—differ from conductors of electricity in that this condition continues to exist at normal temperatures. In the case of *conductors*—most metals, for example—at normal temperatures enough energy is given to the material for some of the electrons to break free from their parent atoms and move around freely. In a good conductor—for example, silver, copper or aluminium—there is a very large number of these free electrons which, on application of an electric voltage, will move and act as current carriers to provide a flow of current.

Semiconductors differ from other electrically conductive materials in that other current movement mechanisms exist within them, giving rise to one of the characteristic features of semiconductors—their conductivity increases with rise in temperature. These mechanisms result from the crystalline nature of semiconductor material and the effect on this of the presence of impurities. In fact the different electrical characteristics of different types of diodes, transistors and other semiconductor devices are largely the result of the material first being purified and then dosed—or *doped* as it is called—with a controlled quantity of some specific impurity. Thus to understand the action of transistors and similar devices we must first know something of their physical structure.

Crystal structure

Both germanium and silicon—and indeed all other semiconductor materials—are crystalline; that is, their atomic structure conforms to a regular pattern. Germanium is a metallic crystal substance, silicon a non-metallic crystal substance. We have seen that the germanium atom has 32 electrons and the silicon atom 14, and applying the law previously mentioned concerning the number of electrons in each shell around the nucleus we see that in the case of both germanium and silicon the outer shell consists of four electrons. In the crystalline structure of these materials these valence electrons form covalent bonds, as shown

in Fig. 1.2, with the valence electrons of adjacent atoms, each atom being equidistant from its four adjoining atoms and each valence electron forming a pair—a covalent bond—with one from

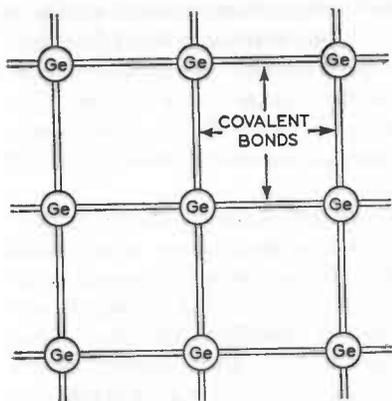
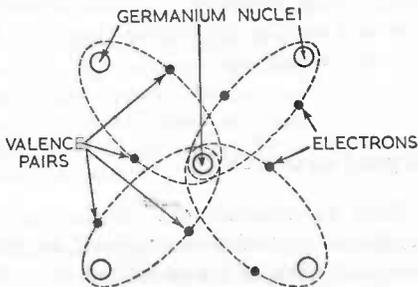


Fig. 1.2. Diagrammatic representation of the germanium crystal lattice structure at absolute zero temperature, showing the two-electron covalent bonds that exist between adjacent atoms.

an adjacent atom. What actually happens is that each pair takes up an orbit around the nuclei of two adjacent atoms. This is shown in Fig. 1.3.

Fig. 1.3. The covalent bonds between adjacent atoms are formed by valence pairs of electrons which orbit the nuclei of adjacent pairs of atoms as shown here.



Effect of temperature: the creation of holes

The condition just described would exist at absolute zero temperature. At this temperature semiconductor materials are electrical insulators. However, at normal temperatures imperfections arise in the crystal lattice structure. The atoms are agitated

by the heat and in the process some of the covalent electrons break free from their bonds. This gives rise to two electrical effects: the electron that breaks free carries a negative electrical charge which, being free, represents a minute movement of current; and on the other hand a 'hole' is created, as shown in Fig. 1.4, in the crystal structure. The idea of a hole is extremely

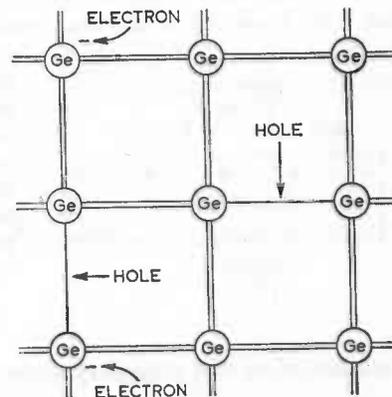


Fig. 1.4. Germanium crystal lattice structure at room temperature, showing the free electrons and holes in the crystal structure that arise at normal temperatures due to the effect of heat.

important and must be clearly grasped since it is fundamental to the operation of most types of transistor. The hole created when an electron breaks free from a covalent bond represents a positive charge equal to the negative charge carried by the electron. It will therefore attract a free electron. The process whereby an electron 'fills' a hole is called *recombination*, and it will be appreciated that at temperatures above absolute zero the freeing of electrons, creation of holes and subsequent recombination is a process that goes on continuously. And just as electrons can be made to move in a given direction to provide current-flow by applying an electrical potential—from a battery, say—to the material, so can holes, for as electrons move from hole to hole through recombination so the holes appear to move in the opposite direction, a process depicted in Fig. 1.5.

The thermal generation of holes and free electrons in this way is the reason for the increase in conductivity (or, put the other way, the decrease in electrical resistance) of semiconductor material with increase in temperature, which we noted earlier was one of the characteristics of semiconductors. It is generally, however, more of nuisance value than anything else. Holes and free electrons are also created in semiconductor material through the addition of certain impurities, that is, doping, and it is this process that is the basis of practical semiconductor devices.

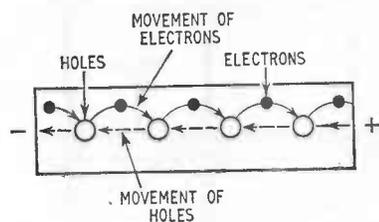


Fig. 1.5. If an electrical potential is applied across a piece of semiconductor material, the free electrons present will be attracted towards the positive side of the potential while the holes will move towards the negative side as shown here.

Creation of *n*- and *p*-type semiconductor material

In pure semiconductor material at temperatures above absolute zero there will clearly be equal numbers of holes and free electrons, since the holes have been created by valence electrons becoming free. By introducing into the crystal lattice structure a material that has a different number of valence electrons to the semiconductor material, however, a preponderance of free electrons or holes is established. The two possible conditions are illustrated in Figs. 1.6 and 1.7.

In the former (Fig. 1.6) an atom of antimony (Sb) is shown incorporated in the crystal lattice. The antimony atom has five valence electrons. When introduced as shown into a germanium crystal lattice four of these valence electrons will form covalent bonds with the valence electrons of the four adjacent atoms of germanium, but the fifth will be free of any such bond and will provide a free negative current carrier. Other atoms with five valence electrons that may be used in this way to create an excess

of free electrons in semiconductor material are arsenic and phosphorous. A substance used in this way is called an *impurity* and an impurity that provides free electrons is called a *donor impurity*—it donates electrons to act as current carriers. Such an impurity is said to be *n* type (*n* stands for negative—the free electrons donated carry negative electric charges), and germanium or silicon doped with donor atoms is therefore known as *n*-type germanium or silicon.

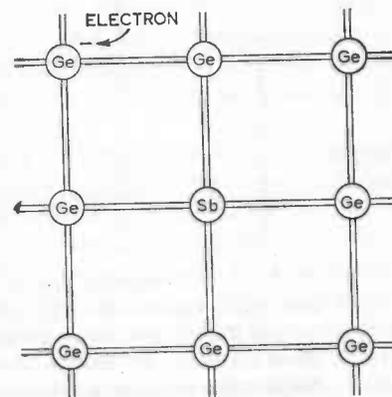


Fig. 1.6. Germanium crystal lattice incorporating a pentavalent antimony atom (Sb), showing the free electron donated to the material by the donor impurity atom when it has formed covalent bonds with the adjacent germanium atoms. (Elements having five valence electrons are termed 'pentavalent'.)

A preponderance of holes is achieved by introducing into the semiconductor crystal lattice structure a substance whose atoms have three valence electrons. Suitable materials include indium, boron and aluminium. In Fig. 1.7 the effect of incorporating an atom of indium into a germanium crystal lattice is illustrated. As the indium atom has only three valence electrons only three covalent bonds will initially be formed with the adjacent germanium atoms. To complete the crystal symmetry, however, a fourth covalent bond will be created by the indium atom capturing an electron from a nearby atom. In this way holes in the semiconductor crystal lattice are created. Trivalent impurity atoms are called *acceptors*: they accept an electron from a nearby atom to create a hole. Such an impurity is said to be a *p*-type impurity

(p for positive, since the hole represents a positive electric charge), and in this case the doped germanium or silicon is called p -type germanium or silicon.

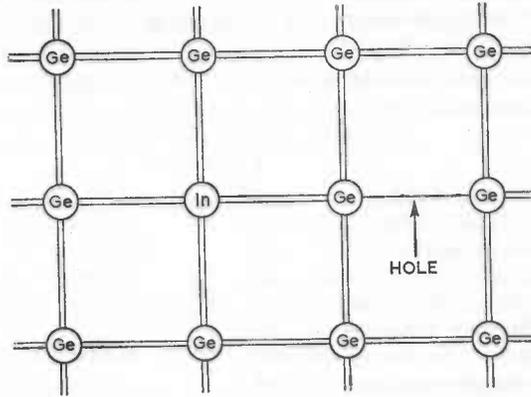


Fig. 1.7. Germanium crystal lattice incorporating a trivalent indium atom (*In*), showing the hole in the crystal lattice structure created by the impurity atom accepting an electron to complete its covalent bonds with adjacent germanium atoms. (Elements having three valence electrons are termed 'trivalent'.)

Photoconduction

In addition to the creation of holes and free electrons in semiconductor material through the introduction of controlled quantities of impurity materials and through the effect of heat, one other cause of free electron and hole generation is of practical importance. The influence of light on many semiconductor materials results in ionisation—the creation of holes or setting free of electrons. Light, in other words, has a similar effect to heat. For this reason most semiconductor devices are provided with a light-proof coating or case. Alternatively the effect can be put to use in devices intended to react to changes in the intensity of illumination, such as light-sensitive cells and phototransistors.

Extrinsic and intrinsic semiconductors, minority and majority carriers

At this point some other terms commonly used in connection with semiconductors can conveniently be introduced. Semiconductor material that has not been doped is sometimes referred to as intrinsic semiconductor material. Current flow in semiconductor material, whether intrinsic or not, due to the effect of heat or light is called intrinsic conduction. Doped semiconductor material, on the other hand, is sometimes referred to as extrinsic semiconductor material, and current flow due to the effect of doping is called extrinsic conduction or, alternatively, impurity conduction.

Because of the effect of heat, at normal temperatures holes and free electrons will both be present in n - and p -type semiconductor material. In n -type semiconductor material there will be a greater number of free electrons than holes, while in p -type semiconductor material there will be more holes than free electrons. The terms *majority* and *minority carriers* are used to refer to the type of current carriers existing in greater and smaller numbers respectively in semiconductor material. Thus electrons are the majority carriers in n -type semiconductor material and holes the majority carriers in p -type semiconductor material; conversely, holes are minority carriers in n -type semiconductor material and in p -type material electrons are the minority carriers. These points are summarised in Table 1.

The action of semiconductor devices is largely based on the

Table 1. Current carriers in semiconductor material

Type of semiconductor	Impurity added	Majority carrier	Minority carrier
n	Donor	Electrons (n , -ve charge)	Holes (p , +ve charge)
p	Acceptor	Holes (p , +ve)	Electrons (n , -ve)

injection of majority carriers from one type of semiconductor into the opposite type, i.e. from an n to a p region or from a p to an n region, in order to establish a current flow through the device. Majority carriers on moving across the junction of course add to the number of *minority* carriers on the side of the junction to which they have moved.

Preparation of semiconductor material

To complete our picture of semiconductor material a word should be said about the preparation of the material for use in semiconductor device fabrication. The first operation is chemical refinement of the raw material—generally germanium dioxide, which can be derived from the flue dust produced by burning certain types of coal, or from copper or zinc ores, in the case of germanium; and sand, which is mainly silicon dioxide, or various silicate compounds, in the case of silicon. This, however, does not provide material of the degree of purity required for transistor fabrication. By further techniques the impurity level is reduced to the order of one part in 10^{10} . These techniques are mainly based on the fact that impurities concentrate most readily in molten material. A molten zone is passed progressively through the material (which is usually in the form of a rod) so that the impurities are carried to one end which, after solidification, can be removed and discarded.

A controlled amount of acceptor or donor impurity must then be added to produce the required electrical characteristics. The amount required for transistor use is about one part in 10^7 , and this has to be levelled, that is, uniformly distributed throughout the semiconductor crystal structure. This process is generally combined with a process of recrystallising the pure material as a large, single crystal—a single crystal structure is necessary in semiconductor devices. Recrystallisation can be achieved by lowering a seed crystal into molten material and then slowly withdrawing it: the material grows on to the seed following the same crystal structure as the seed.

The pn junction

The operation of semiconductor devices depends on the effects that occur at the junction between regions of p - and n -type semiconductor material. The simplest semiconductor devices, small signal semiconductor diodes, consist of a single pn junction. Most types of transistor (exceptions are the unijunction and some forms of field-effect transistor, about which more will be said in Chapter 2) consist of two such junctions in some sort of 'sandwich' form to give pnp or npn arrangements. Other devices, such as the thyristor or silicon-controlled rectifier much used in power circuits, consist of four regions in $pnpn$ form giving three pn junctions. A pn junction is formed basically by introducing impurity of the opposite type into a wafer of p - or n -type semiconductor material prepared along the lines just mentioned (the large single crystal having been sliced into a number of thin wafers). In this way part of the original wafer is converted from n to p type, or vice versa, giving a junction between p and n regions *within the crystal structure*. Note that the junction is a transition from p - to n -type semiconductor material within a continuous crystal structure: merely to join physically p - and n -type material will not result in a structure having the electrical characteristics of a pn junction.

There are several different techniques of junction fabrication, including grown, alloyed, diffused and epitaxial ones—the reader will probably have seen these terms used in describing different types of transistor. More will be said of them in the following chapter.

Electrical characteristics of the pn junction

The electrical characteristics of a pn junction depend on what happens when the junction is first formed, on the degree of doping used in initially preparing the material and then forming the junction(s), and on the potentials applied to the junction(s) in use. When a pn junction is first formed some of the electrons in the n region near the junction will be attracted across the junction by

the holes in the p region; and in doing so they will give rise to the appearance of holes in the n region close to the junction. After this initial movement a state of equilibrium is achieved in which a net *positive* charge is established on the n side of the junction and a net *negative* charge on the p side of the junction. The effect is shown in Fig. 1.8 (a): between A and B on each side of the junction

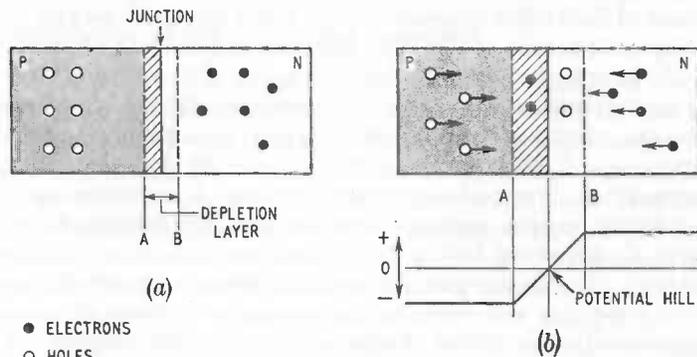


Fig. 1.8. Properties of a pn junction. (a) A depletion layer comparatively free of charge carriers exists on either side of the junction. (b) The migration of charge carriers across the junction when it is first formed, holes from the p side being attracted to the n side and electrons from the n side moving to the p side, sets up a potential hill at the junction, the p side being given a negative charge with respect to the n side, which prevents further movement of charge carriers across the junction.

tion exists a *depletion layer*, so called because the concentration of holes and free electrons is less in this area on each side of the junction than throughout the rest of the block. The combined effect of the negative and positive charges on each side of the junction gives rise to a *potential barrier*, or *potential hill*—see Fig. 1.8 (b)—at the junction. This barrier then opposes further migration of holes and electrons across the junction.

An external d.c. supply may be connected to the pn junction (providing *bias*, as it is called) in either of the two ways shown in

Fig. 1.9 (a) and (b). In case (a) the bias supply reinforces the potential barrier—adds to it in effect—acting further to retard any movement of charge carriers (holes or electrons) across the junction. What happens is that holes in the p region and electrons in the n region are attracted towards the bias supply terminals, thus increasing the width, as shown, of the depletion region. This is

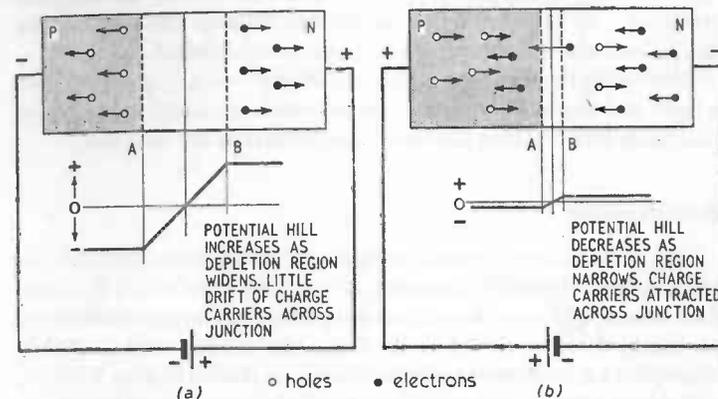


Fig. 1.9. By biasing the junction, the potential hill is either increased, as shown at (a), the depletion layer widening, or decreased if the polarity of the applied bias potential is reversed as shown at (b), the depletion layer then narrowing or being completely cancelled if the bias is sufficient. Applying bias as shown at (a) is termed *reverse biasing* the junction; applying the bias as shown at (b) is called *forward biasing*.

called *reverse biasing* the junction. In case (b) the opposite happens, the bias supply reducing the effect of the barrier so that the flow of charge carriers is increased. In this case the bias supply repels electrons in the n region and holes in the p region so that they move towards the junction where they decrease the depletion region and its associated potential barrier. Applying bias to the junction in this way is termed *forward biasing*. If the forward bias is increased sufficiently the resultant positive potential at the p side of the junction will attract electrons across

the junction from the n region, while simultaneously a greater number of holes will appear on the n side. The minority carrier holes in the n region will move towards the negative supply terminal, where they will draw electrons to fill them from the supply. At the same time minority carrier electrons in the p region will move towards the positive supply terminal and out to the battery to replenish the electrons drawn from its negative terminal. In this way a flow of current through the pn junction device and around the external circuit is established.

By doping the n and p regions in different ways, e.g. giving one a light and the other a heavy concentration of charge carriers, pn junctions with various electrical characteristics are obtained.

Rectification

One of the most common operations throughout electronics is rectification—basically changing an a.c. waveform into a d.c. one. And one of the most important properties of the semiconductor pn junction is its ability to do this. As readers will probably know, the a.c. voltage waveform follows, as shown in Fig. 1.10 (a), a sinewave pattern, varying above and below zero (earth potential) voltage. If, instead of biasing the pn junction with a d.c. supply as in Fig. 1.9, we apply an a.c. supply as shown in Fig. 1.10 (b), the positive half-cycles of the supply will forward bias the junction while the negative half-cycles will reverse bias the junction. As the forward bias will allow current to flow across the junction while the reverse bias will prevent this, current will flow round the circuit only during positive half-cycles of the applied a.c. voltage. A rectified output, consisting of a series of pulses as shown at (c), will thus appear across the load resistor R . Note that if the pn device is reversed it is the negative half-cycles of the a.c. input waveform that will be passed to the output, i.e. current will flow round the circuit during the negative half-cycles of the a.c. input waveform. The symbol used in circuit diagrams to represent a semiconductor diode (single junction, two 'layer' device) is shown in Fig. 1.10 (d). The bar section represents the n -type portion of the device while the arrow section represents the p -type portion.

Note, however, that the symbol is also used to represent other types of diode in circuit diagrams—mainly the metal rectifier, a device that operates on similar principles to semiconductor diode rectifiers—so that it must not be automatically assumed that the device represented by this symbol is in fact a semiconductor junction device.

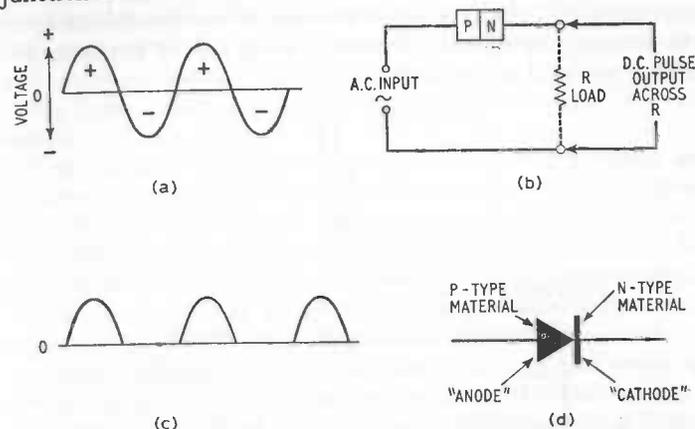


Fig. 1.10. (a) The a.c. voltage waveform of the mains supply is sinusoidal in shape as shown here, with alternate positive and negative excursions above and below earth potential (zero volts). (b) If the a.c. waveform is used to bias a pn junction, connected with the polarity shown here, the positive excursions will forward bias the junction so that current will flow across the junction while the negative excursions will reverse bias it so that current will then not flow. The result is that the pn junction rectifies the a.c. input, providing an output as shown at (c) consisting of a series of positive pulses (or negative pulses if the pn junction is connected the other way round into the circuit). (d) The symbol used in circuit diagrams to denote a semiconductor diode. Note that the arrowhead represents the p -type region and the bar the n -type region.

The rectifying properties of a pn junction can be summed up by saying that the junction has a low resistance to current flow in one direction and a high resistance to current flow in the other direction: it is in this respect a *unidirectional device*.

pn junction characteristics

The electrical characteristics of a typical *pn* junction are shown in greater detail in Fig. 1.11. With forward bias (i.e. forward voltage) applied a current, called forward current, flows increasing with increase in the applied forward voltage in a linear manner after the initial rise (the initial increase is non-linear because of the necessity, as we have previously seen, first to overcome the barrier potential at the junction).

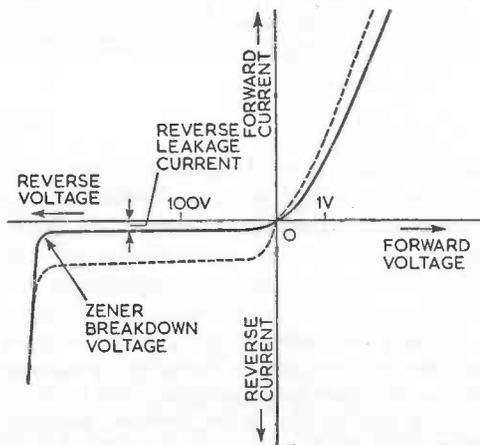


Fig. 1.11. Characteristics of a *pn* junction. The broken-line curves show the effect of increase in temperature on the characteristics.

With reverse bias (reverse voltage) applied to the junction we would not expect to find a flow of current (reverse current) since the *pn* junction has just been described as a unidirectional device allowing current flow in one direction only. There is, however, a small flow of reverse current with reverse bias applied due to the fact that there is a small continuous generation of holes and free electrons on each side of the junction through the effect of heat. Electrons will break away from their parent atoms in the *p*-type

material, while holes will for the same reason appear in the *n*-type material, and these minority carriers will be attracted across the junction by the reverse bias, forming a reverse current generally called the *reverse leakage current* or *reverse saturation current*. This, as shown, remains steady up to a voltage known by a number of terms including *zener voltage*, *breakdown voltage*, *reverse breakdown voltage*, *avalanche breakdown voltage*, etc. As the reverse leakage current is due to the effect of heat, increase in temperature will result in a marked increase in it as shown by the dotted curve in Fig. 1.11. Forward current will also increase slightly for this reason.

The ease with which minority carrier holes and electrons are generated through the effect of heat depends on the *energy gap* of the material. The energy gap is expressed in electron volts (abbreviation eV) and is a measure of the energy that has to be given to an electron before it will break free from its parent atom. As the energy gap is higher for silicon than for germanium, 1.08 eV as opposed to 0.72 eV, in silicon devices the reverse leakage current is less at a given temperature, and the effect of temperature on the operation of silicon semiconductor devices is less than with germanium, silicon devices operating satisfactorily at higher temperatures (silicon transistors operate satisfactorily up to about 200° C, germanium ones up to about 85° C). The disadvantage of the higher energy gap is that higher forward voltages are required to overcome the potential barrier of the junction.

Reverse breakdown voltage

The reason for the zener breakdown voltage that occurs at a certain reverse voltage is that at this—usually quite high—voltage the electrons forming the reverse leakage current are accelerated to such a speed that they begin to knock more electrons out of their covalent bonds, rapidly increasing the number of free electrons and holes available to act as current carriers. The effect is cumulative—it is called an *avalanche effect*—as the greater the number of current carriers generated in this way the greater the number of collisions, etc. This avalanche increases rapidly, as

the reverse characteristic in Fig. 1.11 shows: a substantial reverse current starts to flow and this may destroy the junction.

Some diodes, called zener diodes, are specially designed to operate in the zener region. They are so made that the zener breakdown voltage occurs at a low reverse voltage—typically about -6 V (for most semiconductor devices the reverse breakdown voltage is comparatively high).

pn junction capacitance

Before going on to the transistor, which is a two-junction device (with one or two exceptions), one other characteristic of the basic *pn* junction should be noted. This is that as a reverse biased *pn* junction consists of two regions of conductor material—the main *p* and *n* regions—separated by an area of comparative insulation (the depletion layer on each side of the junction where there are very few current carriers), a *pn* junction so biased forms a capacitor. This gives rise to effects that need to be taken into account in some applications. It also forms the basis of a useful device, the varactor diode, in which the variation in the capacitance of the junction resulting from variation of the biasing voltage is made use of as a compensating device, since the variation in capacitance is inversely proportional to the variation in reverse bias. There are also other specialised applications of the variable capacitance diode.

The transistor

Except for one or two at present rather specialised types that we shall consider in the next chapter, transistors basically consist of two *pn* junctions in a 'three-layer' arrangement giving either a *pnp* or *npn* configuration, as shown in Fig. 1.12. For reasons that will be clear shortly, the three regions are called the *emitter*, *base* and *collector*. It is convenient and the common practice to refer to the emitter-base junction as the *emitter junction*, and the base-collector junction as the *collector junction*, and we shall follow this convention from now on. The symbols used in cir-

cuit diagrams to represent transistors are shown in Fig. 1.13, (a) being used to indicate a *pnp* transistor and (b) an *npn* transistor.

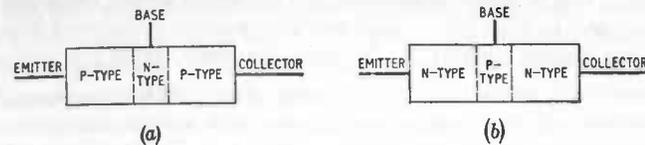
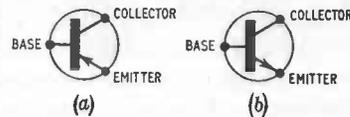


Fig. 1.12. Diagrammatic representation of the transistor, which consists of three regions in *pnp* or *npn* formation giving two *pn* junctions.

The first point to notice is that the two junctions are 'back-to-back': in the case of the *pnp* transistor we have a *pn* junction followed by an *np* junction. The device can thus be considered

Fig. 1.13. Symbols used to represent transistors in circuit diagrams. (a) *pnp* transistor, (b) *npn* transistor. Note different directions of the emitter arrowhead.



as two semiconductor diodes (Fig. 1.14) connected back-to-back, and looked at this way will block current flow in one direction because each diode being unidirectional, one will prevent current flow in one direction

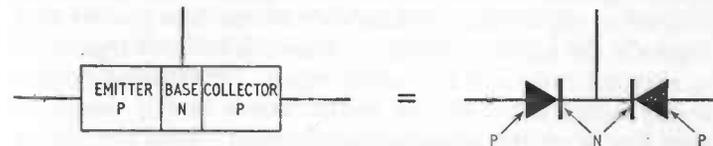


Fig. 1.14. The transistor, looked at from emitter to collector disregarding the base connection, is equivalent to two *pn* junctions connected back-to-back as shown on the right (*pnp* transistor in this example).

and the other will prevent current flow in the opposite direction. Clearly, to establish a flow of current through the device so that we can make practical use of it we must make use of the common centre connection, the base connection, in order separately to bias each junction.

Obtaining flow of current through a transistor

How this is done, to take the *npn* transistor as our example, is illustrated in Fig. 1.15. The emitter junction is forward biased so that charge carriers, in this case electrons, cross the emitter

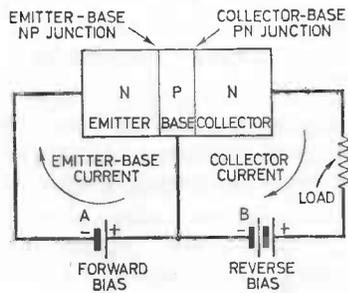


Fig. 1.15. Biasing a transistor (npn type in this example) in order to obtain current flow through it. Battery A forward biases the emitter-base junction so that electrons are injected from the emitter region into the base region. By reverse biasing the base-collector junction by battery B so that the collector region is positive with respect to the base, the electrons injected into the base will then be attracted across the collector-base junction into the collector region and into the load in the external circuit.

junction. We can here see the reason for the term 'emitter': as a result of forward biasing the emitter junction, electrons are injected from the emitter into the base region, attracted across the junction by the fact that the bias has made the base positive with respect to the emitter. Holes, of course, in the base region will be attracted across into the emitter region: but if the base region is only lightly doped and the emitter region heavily doped the main flow of current resulting from forward biasing the emitter junction will be a flow of electrons across it into the base region. Now to obtain current flow through the transistor these electrons, which in the base region are minority carriers, must next be attracted across the collector junction. Remembering that electrons carry a negative electrical charge, this means that we must apply bias to the collector junction so that the collector is positive with respect to the base: the electrons will then flow across the collector junction, into the collector region, and then out through the load to the positive supply battery terminal. The

reason for the term 'collector' is also now apparent: its function is to collect charge carriers from the base region.

We have thus established a flow of collector current. But to do so, as we can see from Fig. 1.15, we have had to reverse bias the collector junction. This is the whole essence of *transistor action*: a flow of current across a low-resistance, because forward biased, junction becomes a flow of current across a high-resistance, reverse biased junction. Hence the term transistor itself, an abbreviation of 'transfer resistor'.

With a *pnp* device, operation is similar except that the biasing arrangements are of the opposite polarity. Thus to forward bias the emitter junction the supply is connected negative to base, positive to emitter, and the minority carriers injected into the base from the emitter are holes carrying a positive charge. To reverse bias the collector junction to attract these holes across into the collector region the collector is made negative with respect to the base. Thus in the *pnp* transistor the main action depends on holes moving across from the emitter region through the base region to the collector. At the collector terminal they represent a positive charge which attracts electrons from the power supply into the collector region: a flow of electrons is thus established through the device through hole movement from emitter to collector, holes moving in one direction while the electrons, moving from hole to hole, move in the opposite direction. The main current carriers in the *npn* transistor on the other hand are, as we have seen, electrons.

Common-base circuit: amplification

The circuit we have so far described is shown in Fig. 1.16 (for a *pnp* transistor), and is called the common-base circuit as the base is common to both the emitter (input) and collector (output) circuits.

If we consider for a moment what happens when the minority carriers flow from the emitter to the collector region it will be apparent that not all of them will do so, for some are bound to recombine with the majority carriers present in the base region.

In a *pnp* transistor some of the holes from the emitter will combine with the free electrons in the *n*-type base region, while in the *npn* transistor some of the electrons from the emitter will combine with the holes in the base region. For this reason, among others, the

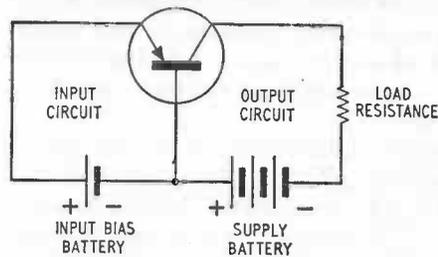


Fig. 1.16. Circuit diagram of the type of circuit, the common-base circuit, shown diagrammatically in Fig. 1.15. In this case, however, a *pnp* transistor is shown, with the polarities of the bias batteries reversed accordingly.

base region is made as narrow as possible. The collector current, therefore, must inevitably be slightly less than the current flowing across the emitter junction. Nevertheless, because of the transistor action amplification will have been achieved. For if we remember Ohm's Law ($V = IR$), the transistor action, in changing a current flow from a low-resistance circuit to a high-resistance circuit, has produced a voltage and a power gain (though not a current gain). Voltage gains of about 250 times are usual with the common-base circuit.

Common-emitter circuit

The common-base circuit is much used in high frequency radio applications—for v.h.f. r.f. amplifiers and frequency changers, in u.h.f. television tuners and so on, where it has certain advantages. Far more common throughout radio and electronics, however, is the common-emitter circuit shown in Fig. 1.17 (a). As can be seen, once again the emitter junction is forward biased (by the small bias battery, A) and the collector junction reverse biased (by the larger supply battery, B). But in this case we are injecting an input current into the base region. Now the amplification provided by an electronic device is the ratio of the change in the input to the device to the variation obtained at the output. Only

a very small current may be fed into the base, while in comparison the collector current may be quite large, so that with this circuit configuration we have not only voltage, and power amplification but current amplification as well. Current gain of about 50 and voltage gain of about 250 times are typical for this arrangement.

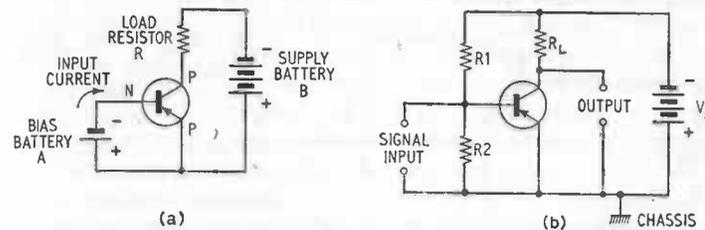


Fig. 1.17. More commonly used than the common-base circuit previously shown is the common-emitter circuit shown here. (a) Basic principle: the bias battery A forward biases the emitter junction so that current from the supply flows through the transistor and the load resistor, the amount of current flowing through the transistor and its load depending on the amount of forward bias applied by the bias battery to the emitter junction. (b) Practical circuit in which a single battery is used to provide the supply and bias voltages. The forward emitter junction bias is determined by the potential divider resistor chain R_1 , R_2 . The signal input current is fed to the base in parallel with the bias current thereby varying the d.c. base bias.

The use of two batteries is not essential (and would not in practice be used). Consider the arrangement shown in Fig. 1.17 (b). Here a small current flows from the supply battery through resistor R_1 into the base, and by selection of the correct values for the voltage divider resistance chain R_1 , R_2 the voltage at the base will be of suitable value to correctly bias the base-emitter junction. This is, in fact, basically the arrangement most commonly used, with the signal current it is required to amplify applied to the base along with the bias current flowing in through R_1 . (R_2 is not strictly necessary, but helps stabilise the bias potential, as outlined in Chapter 3.)

Common-collector circuit

The third possible way of making use of a transistor as an amplifier is to feed the input signal to the base and take the output from across a load resistor connected in the emitter lead; see Fig. 1.18. With this circuit there will again be a current gain, since

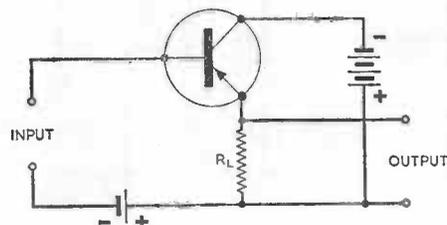


Fig. 1.18. The third possible way of connecting a transistor into circuit to obtain amplification, the common-collector circuit, with the input fed to the base and the output taken from the emitter.

the emitter current is greater than the base current, and in fact the current gain will be slightly higher than in the case of the common-emitter circuit, as the emitter current is slightly greater than the collector current. There will also be power gain. There will not, however, be voltage gain since we are taking the output from the low-resistance (because forward biased) base-emitter circuit. The voltage 'gain' will, in fact, be about unity.

Linear amplification

What we require of an electronic amplifier is that it provides an output that is directly proportional to its input, so that its output is an amplified replica of the input. Let us see how the transistor does this. To take the circuit shown in Fig. 1.17 (b), with no signal applied to the signal input terminals a certain current flows through the transistor from emitter to collector: this current is governed by three factors, (1) the current available from the supply battery and its voltage; (2) the value of the load resistor; and (3) the bias current fed to the base via R_1 to forward bias the emitter junction.

In a normal amplifier, such as that shown in Fig. 1.17 (b), the first two factors are fixed. Amplification is obtained by using the input signal to vary item (3), the base-emitter bias. As the for-

ward bias required at the base is negative with respect to the emitter in the case of a *pnp* transistor and positive with respect to the emitter with an *npn* transistor, for a *pnp* transistor a *negative base drive* is required, while for an *npn* transistor a *positive base drive* is required. Thus, to take the case of the *pnp* transistor as shown in Fig. 1.17 (b), using the input signal to increase the negative potential at the base (relative to the emitter) further decreases the barrier potential of the emitter junction, thus increasing the current flow through the transistor; if, on the other hand, the signal is used to reduce the negative potential at the base—relative to the emitter—the barrier potential of the emitter junction will increase so that the current through the transistor is reduced. In this way the current applied to the base controls the much larger emitter-collector current, and in fact a current of a few microamperes fed to the base to alter the base-emitter potential will control a current of several milliamperes through the transistor. Also this control is linear: the variation in output is proportional to the variation of the input current.

For voltage amplification the output is generally taken as shown from between the collector and chassis, i.e. the 'earthy' side of the supply. Thus one side of the output is tied to the supply, while at the other—collector—side the voltage varies in accordance with the variation in collector current. Considering the load resistor R_L and the transistor as a potential divider network connected across the supply, it is clear that when the transistor is conducting heavily, i.e. passing nearly its maximum emitter-collector current, its resistance will be low. Thus with, say, a 9-V supply voltage V_S —a typical figure for transistor equipment—connected across them and the transistor conducting heavily the voltage across R_L will be nearly 9 V and that across the low-resistance transistor very little. On the other hand, with the transistor cut off, i.e. passing no emitter-collector current, its resistance will be very high and most of the voltage V_S will then appear across the transistor instead of across R_L . The output may alternatively be taken from across R_L , i.e. from between the collector of the transistor and the negative side (in the case of a *pnp* transistor as shown) of the supply.

TYPES OF TRANSISTOR AND ASSOCIATED DEVICES

As we have seen, most types of transistor (the exceptions are rather specialised types such as the field effect and unijunction transistors) consist basically of two pn junctions in pnp or npn formation, and almost all types available to date are made from either silicon or germanium in single crystal form. The various main types of transistor—the alloy, alloy-diffused and planar—differ in the way in which the pn junctions are formed within the crystal structure. There are today three main methods of junction fabrication in use: the alloy, diffusion and epitaxy techniques; and the two junctions in a transistor may be fabricated differently, as in the case of the alloy-diffused transistor in which one junction is made through alloying and the other through diffusion, or the planar epitaxial transistor in which the two junctions are made by diffusion but the collector region consists of low- and high-resistance regions combined through the epitaxial process.

Initial Treatment

The single crystal semiconductor material used for transistor manufacture must, as mentioned in the last chapter, first be refined to a high degree of purity and then doped so as to be of either p or n type. The crystal is then sliced into a number of wafers. To form a pn junction, it is necessary to take the initial p - or n -type wafer and then treat it by means of one of the techniques listed above so that a part of it is converted from p to n type or vice versa. There will then be p and n regions within the same crystal structure.

Alloy-junction transistors

The simplest technique, that used for the first mass-produced transistors, is the alloy process. An alloy-junction transistor, with two pn junctions formed by alloying, is shown in Fig. 2.1. With this type of transistor the wafer forms the base region of the transistor and a region on each side of the wafer is converted to the opposite polarity to give emitter and collector regions, the end product being a sandwich in pnp or npn form with the base at the centre. In the example shown the base consists of an n -type germanium wafer and an acceptor impurity (indium) is used to form p -type emitter and collector regions through alloying. The collector region is made several times larger than the emitter region for maximum efficiency of performance.

The alloy process consists of placing on the surface of the original wafer a small quantity of the required type of impurity and raising the assembly to a temperature at which the impurity melts and begins to dissolve part of the wafer. When cooled, recrystallisation takes place, with some of the impurity—indium in the example shown—alloyed into the original wafer to form regions of different polarity semiconductor material. In this way a junction between n and p regions is formed within the original single crystal structure. Clearly for an alloy-junction transistor it is necessary to go through the alloying process twice, once to form the collector-base junction and the second time to form the emitter-base junction. Both germanium and silicon alloy-junction transistors are made in this way. By careful control of the quantities of impurity, the temperature of the process, and the time taken for alloying and recrystallisation, the electrical characteristics of the junction can be established as required (though extreme precision is not possible, hence the 'parameter spreads' in transistor characteristics).

The pnp or npn wafer is then mounted in a suitable case and supplied with lead-out wire connections.

The germanium alloy transistor constructed as just outlined costs relatively little to make, has the advantage of low resistance and hence small voltage across it when fully conducting, and

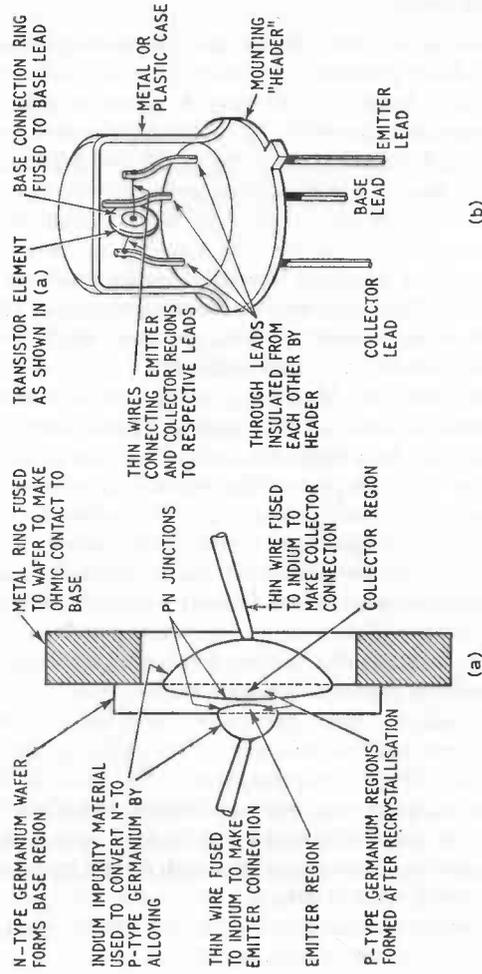


Fig. 2.1. (a) Construction of a germanium pnp alloy junction transistor. The metal ring and thin wires form ohmic (i.e. non-junction) contacts to the base, collector and emitter regions of the transistor. (b) How the active part of the transistor shown in (a) is encapsulated and provided with lead-out wires for connecting into circuit. Note that the active part of the transistor occupies only a small portion of its 'package'.

its emitter can withstand a high reverse voltage, of the same order as the collector. Disadvantages are both its unsuitability for high frequency work (mainly because of the wide base region) and a limit to the collector voltage. Both these disadvantages are caused by the method of fabrication. A similar process of construction is followed in the making of germanium alloy power transistors, capable of operation at several amperes collector current, but with power types the collector is generally bonded to the case—which therefore forms the collector connection—to assist with heat dissipation. Allowable power dissipation is dependent on ambient temperature and exact details of heatsink construction and operating conditions are usually given by the makers. A simplified cross-sectional view of a germanium alloy-junction power transistor is shown in Fig. 2.2.

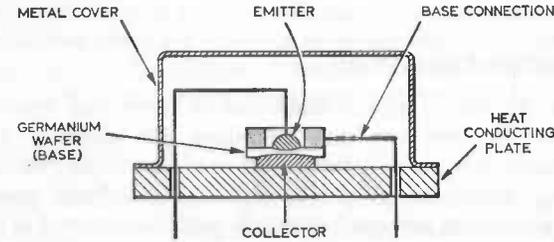


Fig. 2.2. Construction of a typical germanium power transistor. Note that in this case the collector is bonded to the case, which forms the collector connection, to assist in dissipating the considerable heat that arises at the collector junction.

Silicon alloy transistors are generally made by alloying aluminium in thin discs backed by molybdenum electrodes to *n*-type silicon. Main advantages are the low collector leakage current and a high operating temperature (see note on energy gap in the previous chapter). Like the germanium alloy transistor, its relatively simple construction makes it a low-cost device. Principal disadvantage is the low-frequency cut-off.

Diffused junctions

Diffused junctions are made without recourse to the liquid phase of melting acceptor or donor impurities used in the alloy process. Instead, the acceptor or donor atoms are made to move into the initial semiconductor wafer when it is heated to a temperature near its melting point in a gaseous atmosphere containing impurity atoms. The surface then changes from an n - to a p -type layer or vice versa, forming a pn junction just within the wafer. An n -type wafer can be taken, an acceptor diffused in to give a p -type layer, then a donor diffused in again to produce an npn structure. Masking can be used to define the areas where diffusion occurs. More careful control is possible with diffusion than with alloyed junctions, giving improved electrical characteristics.

Alloy-diffused transistors

While the alloy-junction transistor is cheap and robust and acceptable for low frequency operation, for which it is very widely used, it is less satisfactory for operation at r.f., the cut-off frequency being relatively low. The alloy-diffused transistor was introduced to provide better h.f. performance and is widely used, for example, in the frequency changer and i.f. stages of radio receivers.

As the name implies, alloy-diffused transistors are made by a combination of diffusion and alloying techniques. Instead of the original semiconductor wafer forming the base, as with the alloy-junction transistor, the wafer in the alloy-diffused transistor forms the collector. The base layer is formed by diffusion as outlined above, and subsequently an emitter region is established by the alloy process. This procedure gives a narrower base width with much improved high frequency performance. Alloy-diffused transistors have a high current gain and an upper cut-off frequency of around 100 Mc/s. A typical construction is shown in Fig. 2.3.

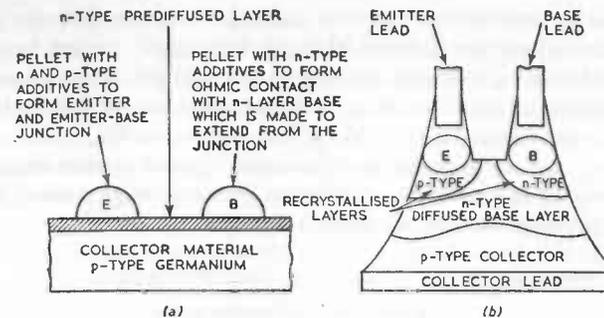


Fig. 2.3. Stages in the manufacture of a germanium alloy-diffused transistor. (a) p -type wafer forms collector region, n -type layer being diffused into it to provide the base region. E and B are pellets which are alloyed on to the n -type diffused layer, the former to provide an emitter region and the latter to provide an ohmic connection to the base region. (b) Active part of the alloy-diffused transistor after alloying, with connecting leads to the base and emitter regions.

Mesa transistors

Among subsequent developments of importance is the mesa transistor, in which a photo-etching process is used after diffusion and alloying—or alternatively double diffusion—to leave the active transistor element as a raised portion above the original wafer. The term mesa signifies this, coming from the Spanish word for table. Germanium mesa transistors are perhaps capable of the highest frequency performance of any types of transistor available today.

Planar transistors

With the application of photo-etching techniques to transistor fabrication we come to the most important type of transistor to arrive on the scene in recent years—the silicon planar transistor. In this both the collector and emitter junctions are formed by diffusion, but the important difference is that a layer of silicon

dioxide is first formed on the surface. In manufacture, the various regions are diffused through 'windows' etched in the oxidised surface, the surface being re-oxidised after each process. As a result of this even more precise definition of the various regions than previously achieved is made possible, while the oxidised surface protects the junctions against contamination. This results in very low and stable collector leakage current and other improvements in the electrical characteristics.

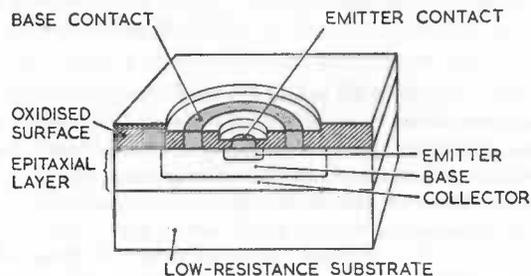


Fig. 2.4. Construction of a small-signal a.f. silicon epitaxial planar transistor. The various regions are formed by diffusion through 'windows' etched in the oxidised protective surface, the surface being re-oxidised after each operation. The collector is a composite region consisting of a high-resistance layer grown epitaxially on a low-resistance substrate, this being done to reduce the voltage drop across the collector region. In r.f. types and power types the same basic processes are used but the geometry of the regions differs.

A drawback of diffused transistors in which the original wafer is used as the collector is the resistance of the collector region. The epitaxial process enables this to be overcome, resulting in the planar epitaxial transistor. In this the collector/wafer is a composite structure consisting of a low-resistance substrate (see Fig. 2.4) with a high-resistance layer nearer the junctions. Epitaxial refers to the way in which the high-resistance collector junction region is formed on the low-resistance substrate, the process con-

sisting of growing a thin film of semiconductor on to a single crystal wafer of the same material, with the crystal orientation of the original wafer maintained into the layer (hence the word epitaxial, meaning in the same axis). The planar epitaxial transistor element shown in Fig. 2.4 is a small-signal, a.f. type. A more complex layout is used for r.f. types.

The latest development in this type of transistor is the incorporation of an integral shield beneath the base connection region to reduce the collector-base feedback capacitance; this type of transistor is particularly suitable for use in the wideband i.f. stages required in television receivers.

Production of transistors

In the U.S. during 1966 some 481 million silicon and 369 million germanium transistors of the types so far described were produced. These figures put in perspective the relative importance of more specialised types of transistor, of which at the present time the field effect transistor is the most important. 2.2 million field effect transistors were produced in the U.S. in 1966, though this was an increase of 267% over 1965, which shows the growing interest in these devices.

Field effect transistors

The field effect transistor differs from the types of transistor so far described in two respects: first, it is a voltage-controlled device, control over the flow of current through it being achieved by applying a bias voltage to the control electrode; and secondly, its operation depends on majority carriers only, hence its alternative name *unipolar* transistor. As a result of being voltage instead of current controlled, it has a high input impedance (*bipolar* transistors, on the other hand, have a low input impedance), which is a great advantage in certain applications.

There are two main types of field effect transistor, the junction-gate field effect transistor and the insulated-gate field effect transistor. Fig. 2.5 (a) shows the construction of a simple

junction-gate field effect transistor based on an n -type silicon substrate (wafer). If an external supply is connected between the source and drain, current (majority carriers) will flow through the transistor from source to drain via the channel. Bias applied between the source and gate, however, will produce an electric field in the channel that will impede this flow of current (hence the name field effect transistor). Current flow through the

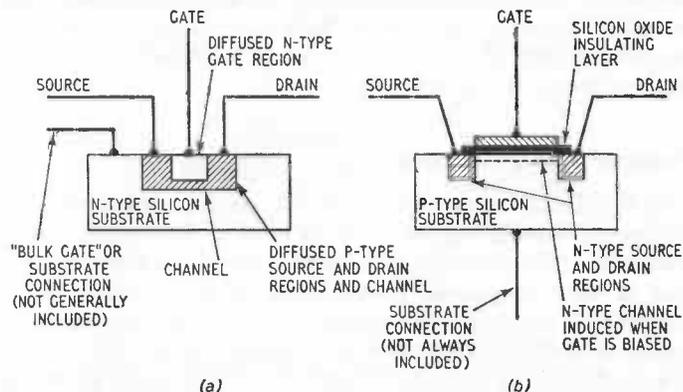


Fig. 2.5. Block outlines of the two main types of field effect transistor. (a) Junction-gate field effect transistor; (b) insulated-gate field effect transistor, which is alternatively known as the metal oxide semiconductor transistor.

device is thus controlled by varying the bias applied to the gate. The type shown is termed a p -channel field effect transistor: an n -channel version can equally well be made by diffusing n -type source, drain and channel regions into a p -type substrate.

If the gate region is separated from the channel by means of an insulating layer, the action of the device operates on principles similar to the capacitor—the capacitance formed by the insulating layer as the dielectric and the gate and channel as the 'plates'. We then have the insulated-gate field effect transistor, with, once again, control of the flow of current along the channel determined by the bias applied to the gate. A more elaborate form

of insulated-gate field effect transistor construction is, however, generally used, as shown in Fig. 2.5 (b). As can be seen, there are here two pn junctions, the drain and source regions being separate. Thus no current flows through the device on connecting a supply between the source and drain because the junctions

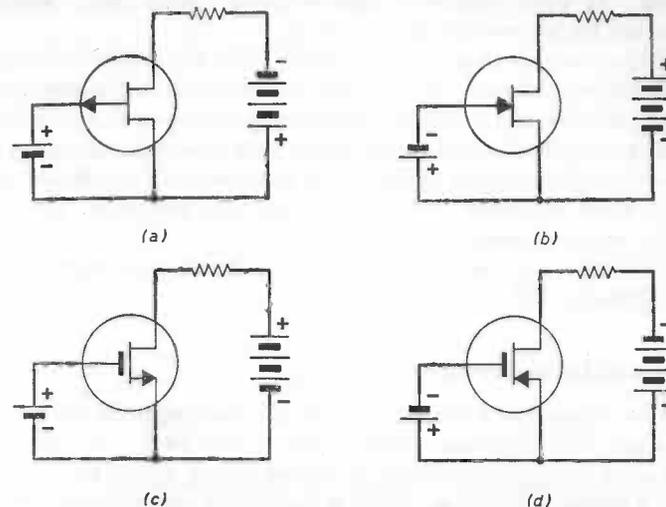


Fig. 2.6. Biasing arrangements for field effect transistors. (a) p -channel junction gate f.e.t.; (b) n -channel junction gate f.e.t.; (c) n -channel insulated gate f.e.t.; (d) p -channel insulated gate f.e.t. Arrangements (a) and (b) constitute depletion-mode operation, where increasing the gate bias reduces the flow of current through the transistor. Arrangements (c) and (d) constitute enhancement-mode operation, where increasing the gate bias increases the flow of current through the transistor.

are back-to-back. Consider, however, what happens when, in the arrangement shown in Fig. 2.5 (b) with a p -type substrate and n -type source and drain regions, positive bias is applied to the gate. Free electrons, since unlike poles attract, will move towards the area just beneath the insulating layer. In this way a negatively charged region appears beneath the layer of insulation,

and this induced n -type channel enables current to flow from source to drain. Varying the gate-source bias alters the flow of current through the device. An n -channel version is shown, but p -channel versions are equally possible. This type of field effect transistor is also known as the metal oxide semiconductor transistor. In some versions a lightly doped 'initial layer' exists between the source and drain regions.

Where current flow through a field effect transistor increases with increase in gate bias (forward gate bias), as in the arrangement just described, this is called *enhancement-mode operation*. Where, on the other hand, increased gate bias reduces current flow through the device (reverse gate bias), as in the junction-gate field effect transistor and simpler insulated-gate type, this is called *depletion-mode operation*.

Field effect transistor circuit symbols and biasing arrangements are shown in Fig. 2.6.

The unijunction transistor

The unijunction transistor consists of a base region in the form of a bar, with two base contacts, one at each end. An emitter region of the opposite polarity is formed on one side of the bar to give a single pn junction. Biasing the emitter results in minority carrier injection into the base region, altering the conductivity of the base. The characteristic exhibits a negative resistance region (i.e. fall in emitter-base voltage with increase in emitter-base current) which makes the unijunction transistor a useful device in certain types of oscillator circuit.

ASSOCIATED DEVICES

Having briefly considered the main types of transistor in use today, it is worth while adding notes on some of the more common semiconductor devices that are used with them, in particular junction and point-contact diodes; two more special types of diode, the zener and tunnel diode; a four-layer ($pnpn$) device, the

silicon-controlled rectifier or thyristor; and one or two special types of resistor, such as the thermistor, which owe their characteristics to the fact that they are made of semiconductor material. Two other devices, the variable capacitance diode and phototransistor, were mentioned in Chapter 1.

Junction and point-contact diodes

The junction diode as its name implies is simply a pn junction which, as we saw in Chapter 1, has rectifying properties so that it can be used to perform such small-signal functions as detection,

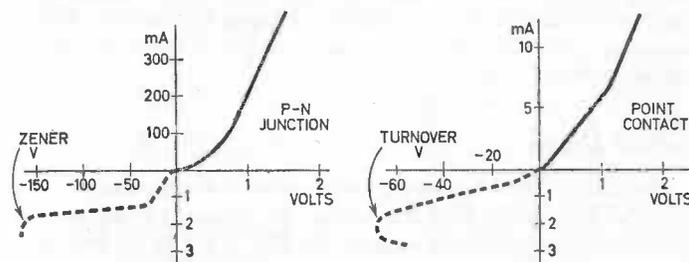


Fig. 2.7. Comparison between pn junction diode (left) characteristics and point-contact diode (right) characteristics.

signal clamping, etc., or, if more generously rated, as a power rectifier. The earlier point-contact diode is still very widely used: it has similar rectifying properties to the pn junction but instead of a pn junction it consists of a piece of semiconductor material—germanium or silicon—on the surface of which a pointed metal wire presses. Typical materials used for the wire are tungsten or platinum. Two connections are made, to the wire and to the semiconductor material. The characteristics of the junction and point-contact diode are compared in Fig. 2.7.

In the case of the commonly used germanium point-contact small-signal diode the wire point contact forms the 'anode' and the semiconductor portion consists of n -type germanium to form the 'cathode', which is often colour-coded red.

Diodes are also commonly given + and - markings to indicate the 'cathode' and 'anode' respectively; thus with bias applied corresponding to these markings, the diode is reverse biased.

Zener diodes

The zener diode differs from other types of silicon-junction diode in that its reverse breakdown voltage (zener voltage) occurs at a fairly low voltage. Units are available with zener voltages at various standard voltages from about 2.5 V up. The slope of the reverse voltage/reverse current characteristic increases very sharply after the zener voltage (see Fig. 7.3). This characteristic means that it can be used to stabilise voltages in relation to current variations (see Chapter 7), or as a stable 'reference voltage' source.

Tunnel diodes

The tunnel diode has very heavily doped *p* and *n* regions. As a result of this as the forward voltage is increased from zero the forward current increases rapidly (see Fig. 2.8). This increase is due to a movement of majority carriers. Beyond a certain point this movement of majority carriers ceases, giving the characteristic a negative resistance region as shown. At a higher forward voltage, forward current due to normal minority carrier movement commences and increases.

Thyristors

The thyristor, or silicon-controlled rectifier, is a four-layer device with three *pn* junctions. Normally, therefore, it would block current in either direction. A gate connection is made, as shown in Fig. 2.9, to one of the centre regions.

The device thus presents a high resistance to current flow in either direction. The resistance, however, falls suddenly to a low value if a forward bias voltage exceeding a certain value is applied across it: then the action is that of a rectifier, with forward current flow only.

The controlled rectifier can also be switched to its low-resistance condition by the application of a small trigger pulse to its gate connection. The interesting fact here is that the device remains 'on', i.e. conducting, after the cessation of the trigger pulse provided that the current flow is not interrupted: if the current ceases, the thyristor returns to its high-resistance condition, i.e. 'off', until triggered once more.

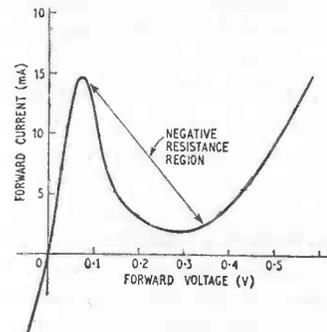


Fig. 2.8. Tunnel diode characteristic.

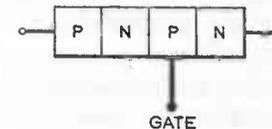


Fig. 2.9. Block schematic representation of a thyristor, which has four regions, two *p*-type and two *n*-type ones, giving three *pn* junctions. A gate connection is made to one of the centre regions.

The device is thus a rectifier that may be controlled either by varying the forward voltage applied to it or the voltage applied to the gate or a combination of both. The main applications are in the switching and regulation of industrial plant, and in motor control. Domestic uses include light dimmer controls, and it has been suggested for use as an output stage in the line timebases of fully transistorised television receivers.

Recent developments

With continuing research it is likely that many new devices will in time be introduced. Particular attention is at present being paid to the development of semiconductor devices for use at microwave frequencies, where thermionic devices have so far generally proved more successful. The step recovery diode, for example, has the ability to switch from reverse conduction to cut

off in a time measured in femtoseconds (10^{-15} seconds). Some such devices, such as the hot carrier diode and metal base transistor, are based on metal to n -type semiconductor junctions. Other devices are based on the use of newer semiconductor materials. The Gunn diode, for example, which with a sufficiently high bias applied produces microwave oscillations, uses n -type gallium arsenide.

Gallium arsenide is one of the 'compound semiconductors' that have been the subject of considerable research in recent years. They are mostly compounds formed from the combination of trivalent and pentavalent elements. Other examples are gallium phosphide and cadmium phosphide.

Semiconductor resistors

Semiconductor materials are the basis of a number of special types of resistor with very useful characteristics. As we have already seen, the resistance of most semiconductor materials decreases with increase of temperature: this negative temperature coefficient, as it is called, is made use of in the *thermistor* (made of a mixture of the oxides of certain metals), a device that can in consequence be used for stabilisation purposes to compensate against the effect of increase in temperature. The *voltage dependent resistor* (made of silicon carbide) has the useful feature that its resistance varies with change of applied voltage. Consequently it is used for voltage stabilisation (see Chapter 7), and its 'rectifying (i.e. non-linear) characteristic' is used, for example, in the e.h.t. stabilisation circuits of modern television receivers.

BASIC TRANSISTOR CIRCUITS AND CHARACTERISTICS

As we saw in Chapter 1 transistors may be of pnp or nnp formation and there are three basic circuit configurations, the common-emitter, common-base and common-collector circuits, as shown in Fig. 3.1. The three basic configurations have different characteristics, for example, different input and output impedances, each having certain advantages for different applications. Typical input and output impedances are given in Fig. 3.1. As can be seen with the common-emitter circuit (a) the input is

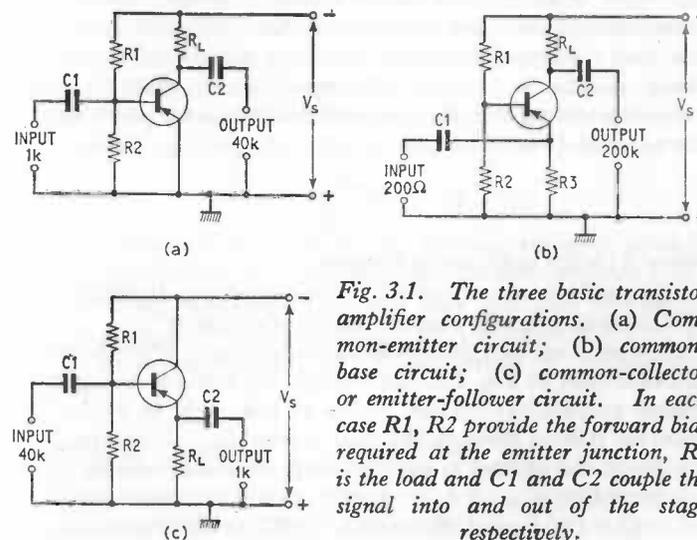


Fig. 3.1. The three basic transistor amplifier configurations. (a) Common-emitter circuit; (b) common-base circuit; (c) common-collector or emitter-follower circuit. In each case R_1 , R_2 provide the forward bias required at the emitter junction, R_L is the load and C_1 and C_2 couple the signal into and out of the stage respectively.

applied between the base and emitter and the output taken from between the collector and emitter. Fig. 3.1 (b) and (c) show the common-base and common-collector circuits respectively. The circuits are shown with *pnp* transistors: with *npn* transistors the polarity of the supply voltage V_S would be reversed. In each case the potential divider network R_1, R_2 provides forward bias for the emitter-base junction, and R_L is the output load resistor. An additional resistor R_3 is required in the common-base circuit to prevent the input to the emitter being short-circuited. It is the general practice to couple the signal into and out of the stage by means of input and output coupling capacitors, C_1 and C_2 respectively in each case. These capacitors prevent the d.c. conditions at one stage affecting the following stage so that only the a.c. signal fluctuations are passed from one stage to the next one.

Gain

A simple common-base arrangement is shown in Fig. 3.2 and will serve as an introduction to transistor characteristics. Because the collector current will have passed through the emitter and base regions, any change in emitter current will result in a change in collector current. The ratio between these changes is called the *current transfer ratio*, and is designated, in the case of the common-base circuit, by the symbol α (alpha). Thus

$$\alpha = \frac{\delta I_C}{\delta I_E}$$

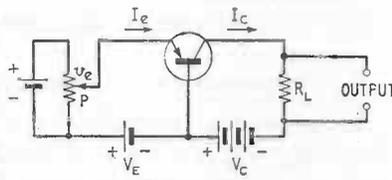
where δ (delta) indicates an increase.

Figures for α are very near but not quite unity, typically about 0.98, because of recombination in the base region.

Now let us see how this works out in practice. With the simple circuit shown in Fig. 3.2, the voltage V_E establishes a suitable emitter current, say 0.5 mA. Most of this—0.98 of it—flows as collector current through the load resistor R_L . A second input voltage to the emitter is derived from a variable resistor across another voltage source v_e . Suppose that this variable resistor is set to give a change at the input of 1 mV, producing a change in

emitter current of 10 μ A. As α is 0.98 the change in collector current will be 9.8 μ A. The emitter junction resistance of a typical small-signal transistor is about 100 ohms, and let us suppose that the load resistor is 10,000 ohms. Thus the change at the input is 10 μ A through a resistance of 100 ohms, but the

Fig. 3.2. Simple common-base amplifier stage to demonstrate how gain is achieved. The arrows here indicate what is known as 'conventional current' flow, electron flow being in the opposite direction.



change at the output is 9.8 μ A through a resistor of 10,000 ohms. The result, from Ohm's Law, is a voltage change across R_L of 98 mV for a change in voltage at the input of 1 mV. This gives us a voltage gain of 98 times in spite of a slight loss of current flowing out through the base contact.

With the more frequently used common-emitter circuit, the current transfer ratio is the ratio of current change at the base to current change at the collector, and is designated β (beta) (or sometimes α'). These parameters, β and α , are mathematically related, the relationships, for those wishing to look further into this, being

$$\beta = \frac{\alpha}{(1 - \alpha)} \text{ and } \alpha = \frac{\beta}{(1 + \beta)}$$

In practice, the abbreviations h_{fe} and h_{FE} are more generally used by manufacturers in quoting common-emitter current gain, h_{fe} indicating small-signal current gain and h_{FE} large-signal current gain, and both implying output short-circuited to a.c. These abbreviations form part of the system of 'hybrid parameters'.

Hybrid parameters

Hybrid parameters are widely used in the semiconductor industry to specify transistor characteristics. They are a set of

resistance, admittance, and voltage and current ratios for given conditions (hence the term hybrid). In the case of the common-emitter circuit these h parameters are:

h_{ie} : input resistance, common emitter, output short-circuited.

h_{re} : reverse voltage feedback ratio, common emitter, input open circuit.

h_{fe} : forward current gain, common emitter, output short-circuited.

h_{oe} : output admittance, common emitter, input open circuit.

If one group of parameters is known, for one circuit configuration, it is possible to work out the others. In practice, however, such information is more useful to the designer than to the engineer or constructor, who chooses transistors on more empirical lines than a comparison of detailed parameters.

Commonly used symbols and abbreviations

The subscripts E , B and C are commonly used in quoting transistor characteristics. Thus, for example, V_{CE} indicates collector to emitter voltage, $V_{CE(max)}$ maximum collector to emitter voltage, I_C collector current, etc. A list of commonly used symbols and abbreviations is given in Table 3.1. Note that transition frequency f_T is the frequency at which the common-emitter current gain falls to unity, and gives an indication of the frequency limitations of a transistor. Thermal resistance, generally quoted in mW per °C, is the power per °C that a transistor will dissipate. Suppose, for example, that a transistor is quoted as having a thermal resistance θ_{j-amb} of 0.5 mW per °C and a maximum junction temperature $T_{j(max)}$ of 70° C. Then at an ambient temperature of 20° C the maximum safe collector dissipation would be $50 \times 1/0.5 = 100$ mW.

Characteristics of the three basic circuit configurations

Approximate characteristics of the three basic circuit configurations are summarised in Table 3.2. These must be taken as a guide only, not as accurate conditions for a particular transistor.

Table 3.1. Common transistor symbols and abbreviations

BV_{CBO}	Collector-base breakdown voltage	$I_{CM(max)}$	Maximum peak collector current
CB	Common-base circuit	I_E	Emitter current
C_{eb}	Collector-base capacitance	I_{EBO}	Emitter-base leakage current, collector open-circuit
CC	Common-collector circuit	I_F	Forward current
CE	Common-emitter circuit	I_R	Reverse current
c_{ob}	Maximum common-base output capacitance	$P_{c(max)}$	Maximum collector dissipation
c_{te}	Collector depletion capacitance	$P_{tot(max)}$	Maximum total dissipation
f_T	Transition frequency	T_{amb}	Ambient temperature
h_{fe}	Small signal common-emitter signal current gain with output short-circuited to a.c.	T_c	Case temperature
h_{FE}	Large signal common-emitter signal current gain with output short-circuited to a.c.	T_j	Junction temperature
h_{FEL}	Large signal current amplification factor	$T_{j(max)}$	Maximum junction temperature
I_B	Base current	V_{BE}	Base-emitter voltage
I_C	Collector current	V_{CB}	Collector-base voltage
$I_{C(AV)max}$	Maximum mean collector current	$V_{CBM(max)}$	Maximum peak collector-base voltage
I_{CBO}	Common-base collector-base current with emitter open-circuit (leakage current)	$V_{CB(max)}$	Maximum collector-base voltage
$I_{CBO(max)}$	Maximum collector-base cut-off current with emitter open-circuit	$V_{CE, etc.}$	Collector-emitter voltage, etc., as for collector-base voltages
		$V_{CE(sat)}$	Collector-emitter voltage for saturated (fully conducting) operation
		V_F	Forward voltage
		V_R	Reverse voltage
		θ_{j-amb}	Thermal resistance in free air
		θ_{j-case}	Junction to case thermal resistance

The table indicates the orders of magnitude of the main parameters with the three configurations: actual figures will of course vary considerably with different types of transistor.

Table 3.2. Characteristics of circuit configurations for comparison

Characteristic	Common base	Common emitter	Common collector
Input to Output from Current gain	Emitter Collector Less than unity	Base Collector About 50	Base Emitter About 50
Voltage gain	High (about 250)	High (about 250)	Low (about 1)
Input impedance	Low (200 Ω)	Medium (1,000 Ω)	High (100 k Ω)
Output impedance	High (200 k Ω)	Medium (40 k Ω)	Low (1,000 Ω)
Power gain	Medium (30 dB)	High (40 dB)	Low (16 dB)
H.F. response (power 3dB down)	High	Low	Dependant on source and load resistances
Phase shift	0°. Output and input in phase	180° (inverse)	0° (in phase)

Phase shift

In Table 3.2 a 180° phase shift is shown for the common-emitter circuit. Let us see how this arises, taking the common-emitter stage with *pnp* transistor shown in Fig. 3.1 (a) as our example. We know that to increase the flow of current through the transistor the base must be made more negative with respect to the emitter, i.e. an increased negative voltage is required at the input, which is across R₂. A greater current then flows through the transistor and its load resistor R_L. Now the transistor and its load R_L form a potential divider across the supply voltage. This increase in current flow means that the resistance of tran-

sistor falls, so that there is an increase in voltage across R_L and a corresponding fall in voltage across the transistor. The output voltage is that across the transistor, between its collector and emitter, and this will be positive going (i.e. going less negative). Thus input and output voltages are in opposite phase.

In the case of the common-base circuit the input is between the emitter and base, i.e. across R₂ and R₃, Fig. 3.1 (b). Now the junction of R₂ and R₃ is tied to the positive side of the supply. This means that to make the base more negative with respect to the emitter to increase the current flowing through the transistor, the emitter must be made more positive with respect to the junction R₂, R₃. Thus a positive-going voltage across R₃ increases the current flowing through the transistor. Once again the voltage across R_L increases and that across the transistor falls, i.e. the output voltage taken from the points shown is positive going. In this case, then, the input and output voltages are in phase with each other.

With the common-collector circuit, Fig. 3.1 (c), the input is as at (a) with an increasingly negative-going signal across R₂ required to increase the current through the transistor. Again the voltage across R_L increases and that across the transistor falls. This means that with the positive side of the supply at earth potential we have an increased negative output across the load resistor R_L. Input and output are thus in phase—and because of this the circuit is often called the emitter-follower circuit.

Input characteristics

The input characteristics of a transistor depend on whether it is operated in the common-base, common-emitter or common-collector mode. Since the common-emitter configuration is that most often used, this will be used as the basis of the following notes on biasing.

As we have seen, with the normal type of bipolar *pnp* or *npn* transistor in the common-emitter mode a current must be fed into the base for the transistor to be brought into operation. In Fig. 3.2 this current was derived from batteries V_E and

v_e . The input characteristic of the common-emitter circuit is a plot of the base current variation against base voltage variation. A typical characteristic is shown in Fig. 3.3. This corresponds to the forward voltage/forward current characteristic of a pn junction—compare with Fig. 1.11. If distortion of the signal to be amplified is to be avoided, a standing d.c. bias must

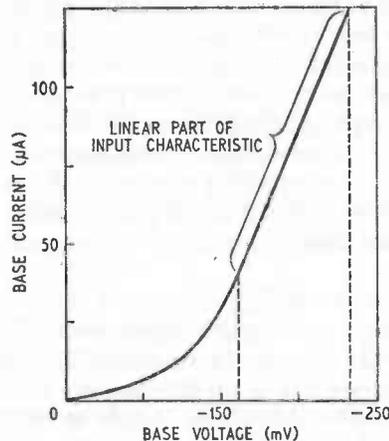


Fig. 3.3. Typical input characteristic for a transistor operating in the common-emitter mode. Note how the characteristic has the same form as the basic forward pn junction characteristic shown in the top right quadrant in Fig. 1.11.

clearly be applied to the base of the transistor so that the a.c. signal variations take place on the linear part of the input characteristic. Put another way, the emitter junction must be forward biased so that current flows across it before the signal it is desired to amplify is applied to the base. Thus in Fig. 3.2, for example, the current from V_E must be such that signal variations—variations in v_e in this case—are centred on the straight part of the input characteristic—say between -160 and -225 mV, Fig. 3.3. In this way linear—i.e. non-distorted—amplification of the signal is obtained. Germanium transistors need a d.c. base-emitter bias voltage of about 0.1 V, silicon transistors needing a base-emitter bias voltage of about 0.6 V.

D.C. bias

The simplest method of biasing a common-emitter stage using a single battery—instead of the various batteries in Fig. 3.2.—is shown in Fig. 3.4 (a). Here a small current from the supply is applied via the bias resistor R_{BIAS} to the base. This will make the base negative with respect to the emitter, the situation we require, with the npn configuration shown, to forward bias the emitter junction.

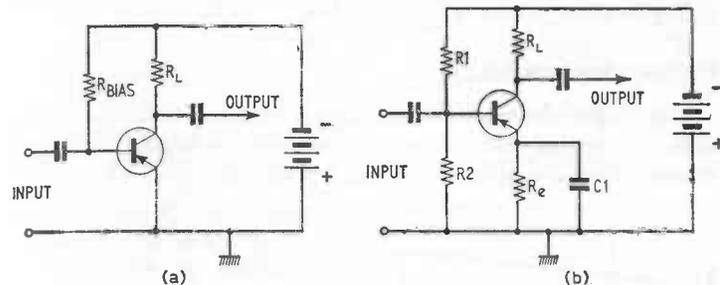


Fig. 3.4. Biasing arrangements for a common-emitter transistor amplifier stage. (a) Simplest possible case in which a bias current to forward bias the emitter junction is obtained from the supply line. (b) More usual arrangement in which the base is biased by means of a potential divider network R_1 , R_2 and an emitter resistor R_e is included to provide bias stabilisation. C_1 decouples R_e at signal frequency.

In practice, more elaborate biasing arrangements are generally used in order to overcome the increase in current flow with rise in temperature that occurs in semiconductor material. As a transistor heats up in operation, so the current flowing through it will increase; the greater current flow means further increase in heat and so on, so that a condition known as *thermal runaway* can occur if precautions—called bias stabilisation—are not taken.

The usual steps taken to overcome this are shown in Fig. 3.4 (b). First, a potential divider (R_1 , R_2) is used to provide a stable base bias voltage: the potential divider stabilises the base voltage against variations that would otherwise occur with changes

a point on the collector voltage (V_C) axis corresponding to the supply voltage (V_S)—say -9 V, a typical value for transistor equipment—and a point on the collector current (I_C) axis corresponding to V_S/R_C , i.e. a collector current determined, in accordance with Ohm's Law, by dividing the supply voltage by the value of the load resistor R_C . Suppose, in the example given, the load resistor is 1,000 ohms: then $9/1,000 = 9$ mA, the load line terminating on the I_C axis at this point.

Clearly to obtain linear amplification at the output the collector load resistor must be chosen in conjunction with the supply voltage, and the variations of base input current must be such that operation of the transistor will be centred upon linear portions of the output characteristics. In the example shown in Fig. 3.6 based, as we have seen, on a 9-V supply voltage and 1-k collector load resistor, swings in collector current and voltage over the portion A, B, C of the load line will provide linear amplification so far as the output is concerned. Base current (I_b) variations between point C, about $18 \mu\text{A}$, and point A, about $55 \mu\text{A}$, will thus result in collector voltage variations between -3 and -7 V and collector current variations between 2 and 6 mA, all over linear parts of the output characteristics. To obtain these conditions the d.c. base bias current needs, as can be seen, to be about $35 \mu\text{A}$ (I_b at point B) to centre the a.c. signal swings at this point.

Two-stage amplifier circuit

A simple two-stage amplifier using *pn*p transistors is shown in Fig. 3.7. The output signal of the first stage is developed across its load resistor R_2 and fed to the base of the second transistor via capacitor C , which passes on the a.c. signal variations but prevents the d.c. supply at the collector of the first transistor appearing at the base of the second transistor. Because of the relatively low input and output impedances of bipolar transistors, large value coupling capacitors are generally needed with them and electrolytics are, as shown, often used. It is important to ensure that they are connected into circuit with the correct polarity.

This circuit also illustrates an alternative form of bias stabilisation to that shown in Fig. 3.4 (b). The technique is to connect the bias resistors (R_1 and R_3) direct to the transistor collectors instead of to the supply. The principle is similar to that previously described: as collector current increases, because of

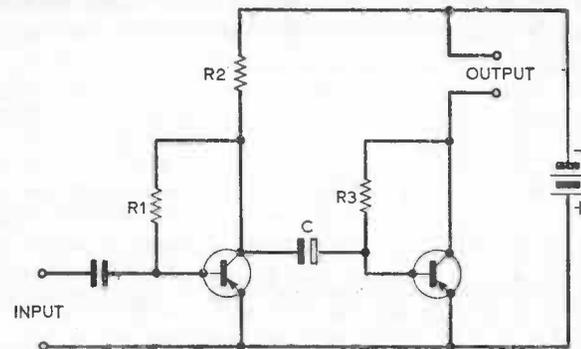


Fig. 3.7. Simple two-stage transistor amplifier. Note the polarity of the electrolytic coupling capacitor: with *pn*p transistors, the base of the following stage is generally positive with respect to the collector of the previous stage. The output driven by the second transistor could be an earpiece, forming its load, to give a hearing-aid output stage. Stabilisation is achieved in this circuit by connecting the base bias resistors R_1 and R_3 direct to the collectors of the transistors.

temperature increase, so the collector voltage will fall. This reduces the base bias, so that there is decreased base, and hence collector, current. Though less effective, this simpler method is adequate for some applications.

D.C. coupled stages

With the low voltages used in transistor circuits, it is common to dispense with coupling capacitors and use direct coupling between stages instead. This saving also has the advantage that the

signal phase shifts that can result from passing the signal through a capacitor are avoided. The problem, however, arises that the full collector voltage of one stage appears at the base of the next stage, and this will be too large for correct biasing of the following stage. The situation can be remedied by using a larger value emitter resistor in the second stage. Consider the circuit shown in Fig. 3.8, using two direct-coupled germanium *pnp* transistors

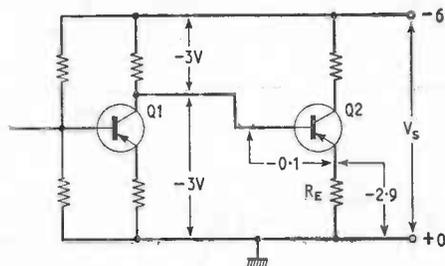


Fig. 3.8. Directly coupled transistor amplifier stages. The value of R_E is chosen to establish the correct base bias for the second stage when it is fed direct from the collector of the first stage.

Q1 and Q2. Transistor Q1 is so biased that the voltage across its load resistor, with a 6-V supply as shown and no signal applied to its base, is -3 V and that between its collector and the positive side of the supply is -3 V. Thus -3 V also appears between Q2 base and the positive side of the supply. If now a largish value resistor R_E is connected in Q2 emitter lead of such value that the voltage across it is -2.9 V, then Q2 base-emitter voltage will be 0.1 V, about right to forward bias the emitter junction of a small-signal germanium transistor. Note that with this arrangement the need for a separate biasing network for the second transistor is removed.

Negative feedback

The effect of a large value emitter bias resistor, however, is that substantial signal variations will, unless steps are taken to avoid this, occur across this resistor; i.e. variation in the current through the transistor will produce variation in the voltage across the emitter resistor. These variations across the emitter resistor will be in phase with the input voltage and thus, in a common-emitter

stage, will be in opposite phase to the output signal. The effect, looking at the circuit from the input side, is that when the signal increases the forward bias of the emitter junction the effect of the voltage change across the emitter resistor will be to reduce this bias, thus reducing stage gain. This is termed negative feedback, and may be introduced deliberately where maximum gain is not required in order to reduce distortion (since harmonics of the input signal arising in the stage will be cancelled by the effect of the negative feedback). Negative feedback introduced in this way also increases the input resistance of the stage, an advantage in some applications.

Bypass capacitors

When the input signal is small and it is desired to avoid negative feedback, a 'bypass' or 'decoupling' capacitor is added across the emitter bias resistor. This smooths out the signal frequency variations at the emitter, its value being determined by the frequency of the signals being handled. The emitter resistor bypass capacitor is C1 in Fig. 3.4 (b).

R.F. amplification

Amplifier stages intended to give amplification at radio frequencies (r.f.) or intermediate frequencies (i.f., see Chapter 5) generally have a frequency selective load. An example is shown in Fig. 3.9 (a). Here instead of a simple resistor the collector load consists of an inductor L and capacitor C in parallel. This is termed a parallel tuned circuit. The values of L and C are chosen to give maximum impedance at the required frequency, i.e. the frequency of the signal to be amplified. At this frequency, the parallel-tuned circuit is the equivalent of a suitable load resistor, but at other frequencies its impedance is low: thus an output signal is only obtained at the required frequency. This frequency is termed the resonant frequency of the circuit, which is said to be 'tuned' to this frequency. The values of L and C are given by the formula $f_0 = 1/(2\pi\sqrt{LC})$, where f_0 is the frequency at which an output is required.

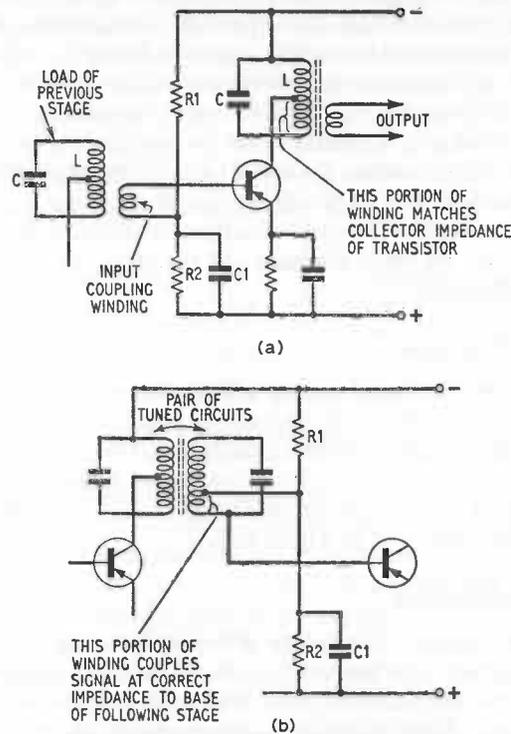


Fig. 3.9. R.F. amplifier stages usually have a frequency selective load, which may be provided by the parallel LC-tuned circuits shown here in the collector leads. In (a) the coupling to the base is by means of a small coupling winding. R1 and R2 provide d.c. base bias, R2 being decoupled so that the other end of the base coupling winding is at earth potential signalwise. In (b) interstage coupling is by means of a pair of tuned circuits, to give improved selectivity. Note how the input to the base is coupled to the second tuned circuit.

The sharpness of the frequency selectivity of a tuned circuit is determined by its Q value, which in practice may be taken as the ratio of the reactance to the resistance of the inductor L : clearly any resistance in the load will allow signals at other frequencies to develop.

Since $L \times C$ is a constant for a given resonant frequency, the value of L can be chosen to match the output impedance of the transistor so that there is maximum signal transfer from the transistor to the tuned circuit, and the value of C then chosen so as to combine with L to be at resonance at the required frequency.

The calculated values of L , however, following the above formula, tend to be rather small with transistors, making it difficult to construct an inductor having a good Q figure. A technique commonly used to overcome this problem is to connect the collector to a tapping on the inductor, as shown (Fig. 3.9 (a)). When this is done only the lower portion of the inductor winding needs to match the transistor's collector impedance, and a larger inductor can therefore be used with the value of C reduced appropriately to restore resonance.

The base impedance of a transistor is much lower than its collector impedance. The most common technique used to match the tuned circuit load impedance of one stage to the base impedance of the following stage is, as shown in Fig. 3.9 (a), to use a small winding to couple the signal from the tuned circuit to the base of the next stage.

Improved frequency selectivity is obtained by using a pair of tuned circuits to couple r.f. amplifier stages together as shown in Fig. 3.9 (b). Here again the collector of the first stage is generally taken to a tapping on the coil in the first tuned circuit, and the signal can be fed to the base of the second stage at the correct impedance by connecting the base to a lower tapping point, as shown, on the coil in the second tuned circuit.

As with other amplifier stages, emitter junction forward bias is provided by a potential divider network R1, R2, and bias stabilisation is effected by means of an emitter resistor with bypass capacitor. With this type of circuit the input coupling winding is connected between the base of the transistor and the junction of

the base bias potential divider network. Loss of signal will occur across R2 unless the junction of R1, R2 is at earth potential as far as the signal is concerned, and for this reason R2 must be bypassed by a capacitor (C1) of suitable value.

R.F. oscillators

If a little of the output of a tuned amplifier is fed back to its input we have an oscillator, a stage providing a sustained signal, sinusoidal in form as shown (Fig. 3.10), at a given frequency. In

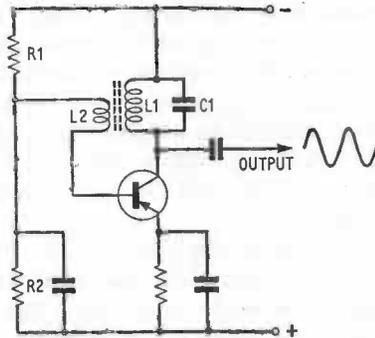


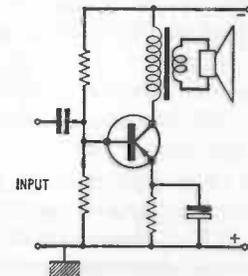
Fig. 3.10. If the LC-tuned circuit load is coupled back to the base by means of a small coupling winding L2, as shown here, a simple r.f. oscillator is obtained providing a sinewave output at a frequency determined by the values of L1 and C1.

the simple r.f. oscillator circuit shown in Fig. 3.10 the emitter-base junction is forward biased by the potential divider network R1, R2 so that current flows through the transistor. Part of the output developed across the tuned circuit L1, C1 is fed back to the base via the small coupling winding L2, resulting in sustained oscillation at a frequency dependent on the values of L1 and C1, provided that the feedback winding is arranged with the correct phase relationship to ensure that the feedback is positive. The oscillatory nature of the output is the result of the action of the inductor L1 and C1, the transistor providing gain to maintain oscillation in the tuned circuit. As shown, the oscillator output is taken from the collector via a coupling capacitor.

A.F. TECHNIQUES

In considering transistor a.f. amplifier techniques it is convenient to start at the output end since the output required from an amplifier is the main feature determining its design. In the last chapter the principle of biasing a transistor amplifier stage to overcome non-linearities in its input and output characteristics was outlined. This method of biasing is termed Class A amplifier operation. Working along these lines a simple Class A transistor output stage feeding a loudspeaker may be produced as shown in Fig. 4.1. The only difference between this and previously

Fig. 4.1. Simple Class A transistor a.f. output stage. A transformer is used to match the loudspeaker to the output transistor. In some designs an auto-transformer or choke is used instead of the double-wound transformer shown here. In practice, one side of the transformer secondary winding is generally connected to earth (i.e. chassis).



illustrated transistor a.f. amplifier stages is the use of a transformer in the collector circuit to match the transistor output impedance to the impedance of its load—the loudspeaker—maximum power being fed to the loudspeaker when these two impedances are the same. To obtain maximum power gain in the output stage, the common-emitter configuration is used, and a step-down transformer, as shown, is needed to match the low impedance of most types of loudspeaker to the relatively high

impedance of the common-emitter collector circuit. Output stages of this type have been used in a number of car radios, and in other equipment also. Using a power transistor such as the AD140, outputs of 2-3 W are usual. In some models the output transformer is of the autotransformer variety.

Push-pull stages

To reduce distortion and obtain reasonable outputs using less-expensive lower-power transistors push-pull output stages are, however, generally used for transistor audio amplifiers. They also have the advantage in battery equipment that if operated in

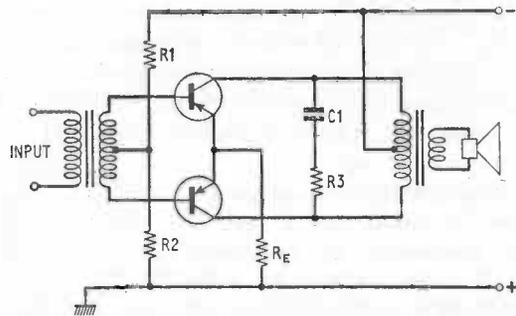


Fig. 4.2. Basic push-pull output stage. The input is applied via a driver transformer with a centre tapped secondary winding to provide opposite phase signals at either end of the secondary for the two transistors. The output transformer has a centre-tapped primary winding, and again in practice it is usual for one side of the secondary winding to be connected to chassis.

what is termed the Class B condition the current drain from the battery is much reduced. A simple example is shown in Fig. 4.2. Under Class B bias conditions when there is no input signal both transistors are biased 'off' so that they draw no current from the supply. As shown the input is applied via a transformer (or

phase-splitter stage, see later) with a centre-tapped secondary winding. This means that oppositely phased signals will appear at the bases of the two transistors. As both, being *pnp* types in this example, require a negative-going drive signal, the result is that when one transistor conducts the other remains cut-off and vice versa—and of course with no signal neither transistor conducts. Thus one transistor will conduct on the positive-going excursions of the signal applied to the primary of the input transformer, while the other transistor conducts on the negative-going excursions of the signal applied to the primary of the input transformer. The output signals are developed across an output transformer with centre-tapped primary winding and applied via the secondary winding to the loudspeaker.

Because of the non-linearity in the input characteristic when a transistor begins to conduct, strict Class B operation is not generally adopted. Instead, a small standing bias is applied to the bases of the two transistors. In Fig. 4.2 this bias is provided by the potential divider R1, R2. Distortion produced if this bias is not present, or is incorrect, is termed *cross-over distortion*, i.e. non-linearity in the operation of the stage in the region of the combined characteristics of the two transistors where one transistor is beginning to conduct and the other is ceasing to conduct. It has a very marked effect on the quality of sound reproduction, hence the need to take steps to overcome it. To avoid unwanted negative feedback with this arrangement R2 must be kept small in value. Resistor R_E provides bias stabilisation and is usually of the order of 1-10 ohms in value. An RC network (C1, R3) is often included across the output to reduce the output at the higher frequencies. Since many transistor radios use small loudspeakers not capable of handling the lower frequencies, this attenuation of the output at the higher frequencies is necessary to avoid an output that over-emphasises the higher frequencies.

An alternative input arrangement commonly used is shown in Fig. 4.3. Here each transistor is fed from a separate winding on the input transformer and has its own base bias potentiometer network. Thermistors, which have a negative temperature coefficient (see Chapter 2), are often used as shown across the lower

arm of each potential divider network to provide stabilisation against changes in the bias conditions with changes in temperature. They may also, of course, be used for the same purpose across the lower arm of the base bias network in the type of circuit shown in Fig. 4.2, i.e. across R2.

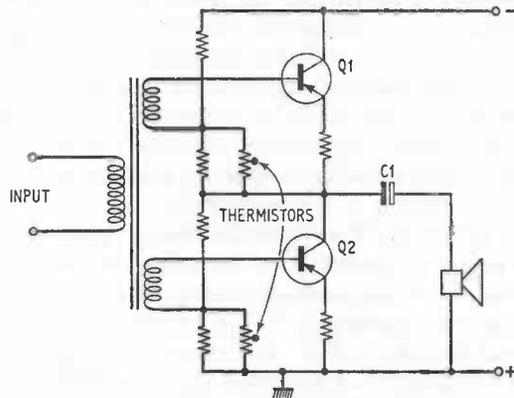


Fig. 4.3. In this push-pull output stage the two transistors are fed from separate windings on the driver transformer, thermistors are included in the base bias networks to provide stabilisation against the effects of temperature variation, and the loudspeaker is coupled by means of a capacitor (C1) to the centre-point of the output stage—an arrangement sometimes referred to as 'single-ended output'.

Transformerless output

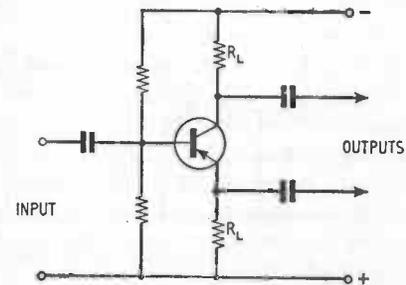
Fig. 4.3 also illustrates an output arrangement which dispenses with the need for an output transformer. The two transistors are connected in series across the supply, but in parallel with the load—the loudspeaker. Since the output impedance of this parallel arrangement is much less than that of a push-pull circuit with the output taken from between the collectors of the two output transistors, this circuit enables the output to be coupled to the

load without the need for a matching transformer. The output is, in this arrangement, coupled to the loudspeaker by the high-value (electrolytic) capacitor C1. This capacitor charges when Q1 conducts and discharges when Q2 conducts. Thus, charging and discharging at signal frequency, the capacitor provides the required a.c. coupling to the loudspeaker.

Phase-splitter stage

In some circuits the oppositely phased input signals required at the bases of the two output transistors in a push-pull output stage are derived from a transistor phase-splitter stage instead of from a centre-tapped input transformer. One possibility is shown in Fig. 4.4. As can be seen, load resistors are connected in

Fig. 4.4. Transistor phase-splitter stage. This may be used instead of a driver transformer with centre-tapped secondary or two secondary windings. Opposite phase output signals are obtained from the collector and emitter by connecting equal value load resistors in the collector and emitter leads. In practice the values might be varied to compensate for the different output impedances of the collector and emitter.



both the collector and emitter circuits. Since there is 180° phase shift between the base and collector but no phase shift between the base and emitter, outputs taken from across the collector and emitter load resistors will be oppositely phased and can be used to drive a push-pull output stage. There are, however, snags with this arrangement. For one thing the current gain of the collector and emitter circuits differs slightly, and for another the impedances of the two outputs are widely different.

Complementary-symmetry circuits

The need for a split-phase input can be completely avoided by using a complementary-symmetry circuit in which the output pair consists of an *npn* and a *pnp* transistor with matched characteristics, the configuration being as shown in Fig. 4.5. The principle

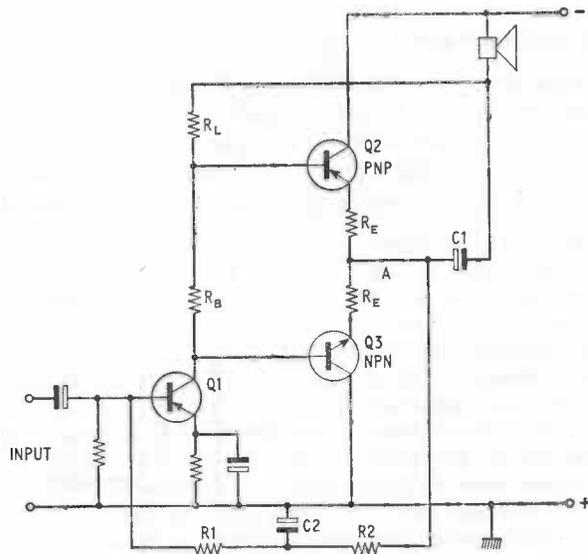


Fig. 4.5. Complementary symmetry output stage, using a *pnp* and an *npn* transistor, with single-ended output connection. The load R_L of the driver transistor $Q1$ is a.c. coupled to the emitters of the output transistors via $C1$ so that the output transistors operate in the common-emitter mode. The d.c. feedback network $R1, C2, R2$ is often used with this circuit.

here is that a *pnp* transistor requires a negative-going drive signal, while an *npn* transistor requires a positive-going drive signal. Consequently, with the input from the driver stage $Q1$ applied simultaneously to both bases the *pnp* transistor ($Q2$) will conduct

on negative-going excursions of the input signal waveform and the *npn* transistor ($Q3$) will conduct on positive-going excursions of the input. This is nowadays a very widely used arrangement.

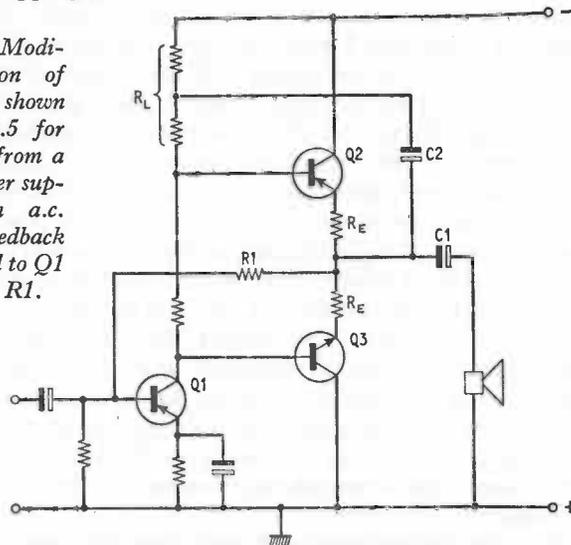
Three other points arise with this circuit configuration. First, as the polarity of the supplies required by the two output transistors is opposite, the collector of one (the *pnp* transistor) can be taken direct to the negative side of the supply, while the collector of the *npn* transistor is taken direct to the positive side of the supply. The output is again the voltage swing at point A, with $Q2$ and $Q3$ conducting alternately to provide the signal swings at this point which are coupled by $C1$ to the load. But as the output is taken from between the two emitters its impedance is further reduced (in comparison with the configuration shown in Fig. 4.3) permitting matching to any ordinary loudspeaker.

Secondly, this may at first appear to be a common-collector configuration. For maximum power gain, however, the common-emitter circuit is required, and in actual fact the two output transistors in this circuit do operate in the common-emitter mode. This is because, as shown, the load (R_L) of the driver stage $Q1$ is taken to the 'live' terminal of the loudspeaker, and is consequently a.c. coupled to the emitters of the output transistors. That is, the output from $Q1$ is direct coupled to the bases of the output transistors and RC coupled to their emitters. In this way the input to the output stage is applied between the bases and emitters of the output transistors so that in effect they are acting in the common-emitter mode, the output being taken from between the emitters and collectors of the output transistors.

The third important point to note about this type of circuit is the d.c. biasing of the output transistors. This is provided by resistor R_B , which is much smaller in value than the driver stage load resistor R_L . The effect of this bias resistor is that the base of the *pnp* transistor $Q2$ will be slightly negative (about 0.1 V for a germanium transistor) with respect to its emitter, while the base of the *npn* transistor $Q3$ will be slightly positive with respect to its emitter. These are the conditions we require to bias the emitter junctions of the two output transistors for linear operation.

The very small value resistors R_E in the emitter leads provide stabilisation as in other types of amplifier circuits. It is a common practice with this circuit to provide a d.c. feedback network (R_1 , C_1 , R_2) to further assist in stabilising the operating conditions: this is described later under the heading 'negative feedback'. Another factor in this type of circuit is the effect on the driver of coupling its load to the output of the amplifier (via C_1); the result is an increase in the drive applied to the output stage, increasing the efficiency of the circuit (a technique known as 'bootstrapping').

Fig. 4.6. Modified version of the circuit shown in Fig. 4.5 for operation from a mains power supply. Both a.c. and d.c. feedback are applied to Q1 base via R_1 .



Mains operation

One or two minor variations are generally adopted, as shown in Fig. 4.6, in the case of equipment intended for operation from the a.c. mains supply. It is found that connecting the loudspeaker to chassis instead of to the negative side of the supply reduces the effect of residual 50 c/s ripple (or 100 c/s ripple with full-wave rectification, as is generally used) at the loudspeaker. This

involves modifying the way in which the signals from the driver stage Q1 are applied to the output transistors. As shown, the driver load resistor R_L is tapped (two separate resistors being used as shown) and the centre point a.c. coupled via C_2 to the emitters of the output transistors. D.C. feedback from the output stage via R_1 to the base of the driver stage assists in providing stabilisation of the operating conditions of the circuit.

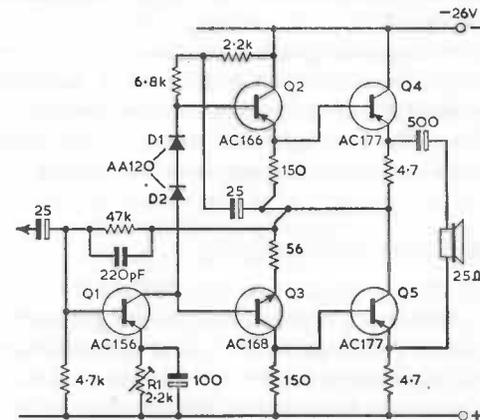


Fig. 4.7. A commonly used circuit where higher output power is required. A complementary-symmetry driver stage (Q2, Q3) feeds a pair of pnp power output transistors. Stabilisation against the effects of heat and power-supply variations on the bias voltages is provided by the bias stabiliser diodes D1 and D2 and the various d.c. feedback paths.

Complementary-symmetry driver stage

To provide increased power output it is common practice to use a complementary-symmetry driver stage feeding a Class B output stage as shown in Fig. 4.7. In this case, the complementary-symmetry transistors Q2 and Q3 conduct on alternate half-cycles

to switch the *pnp* power output transistors Q4 and Q5 on and off alternately. Direct coupling is used between the complementary-symmetry driver stage and the output transistors, and cross-over distortion is removed by arranging for a small standing current to flow in the base circuits of the output stage. Note that, as both output transistors are *pnp* types, both require a negative-going drive. For this reason, Q4 is fed from Q2 emitter while Q5 is fed from Q3 collector. The loudspeaker is again capacitively coupled.

Since with this type of circuit Q1 is also generally required to provide a certain amount of power there are in this arrangement two driver stages, Q1 and the complementary-symmetry pair (it is conventional to refer to stages that provide power amplification but precede the actual output stage as driver stages).

Heatsinks and bias stabilisation

To assist in dissipating the heat that arises at the collector of a power output transistor, such transistors are generally provided with a heatsink. This consists of a metal device clamped to the transistor (the collectors of power transistors are generally bonded to the case) and may, for example, be finned so as to present a large surface to the surrounding air to increase the transfer of heat away from the transistor. A coating of silicon grease is also frequently applied to power transistors to increase thermal conductivity. It is important to maintain these arrangements in the event of a replacement being made.

Since the conductivity of semiconductors increases with increase in their temperature (and conversely falls with decrease in temperature) temperature variations will interfere with the biasing arrangements of the type of circuit we have been considering unless measures are taken to counteract this. When we remember that the bias voltages required in transistor equipment are quite small, e.g. 0.1 V to forward bias a germanium transistor emitter junction correctly, it will be appreciated that this is a serious problem. In Fig. 4.5, for example, an increase in Q1 collector current due to rise in temperature will alter the base biasing of the

output transistors Q2 and Q3 by altering the voltage across R_B , and this will result in cross-over distortion. Such a situation in the circuit shown in Fig. 4.7 will disturb the base biasing of both the complementary-symmetry driver stage and the following output stage.

The use of a thermistor in parallel with R_B , Fig. 4.5, will compensate for this drift in the biasing arrangements. Since the resistance of a thermistor falls with increased temperature, the voltage across it will fall, thereby compensating for the increased voltage across R_B due to the increased collector current of Q1. And conversely at low temperatures the voltage across the thermistor will increase, moving the output transistors away from the cut-off point on their input characteristics. Thermistors have been used in this way in a number of designs, often with the parallel resistor made variable to enable the biasing arrangement to be preset for optimum performance. This provision of a preset adjustment is desirable so that the circuit can be individually adjusted to take into account the slight differences in characteristics that exist between transistors of the same type and the tolerances in the ratings of other components. It is also worth pointing out here that in view of the small bias voltages used it is necessary when replacing components to keep to the same values and tolerance ratings, and to follow any instructions for presetting that may accompany a particular equipment.

A similar and now widely used approach is to use one or more bias stabilisation diodes or transistors. In the circuit shown in Fig. 4.7 a pair of bias stabilisation diodes, D1 and D2, is used. Such diodes must of course be connected so that they are forward biased. Again, with increased current flowing through them as a result of increase in the driver transistor Q1 collector current, so the voltage across them will drop, thus pulling back the d.c. base bias of Q2 and Q3. Bias stabilisation diodes have the advantage that they also provide stabilisation against changes in the supply voltage, a point that is particularly important in battery-operated equipment where the supply voltage falls towards the end of the battery's life. A fall in supply voltage will decrease the forward bias applied to the diode, reducing the current through it and

increasing its resistance. The resultant increase in the voltage across it increases the base bias of the output stage and avoids the early onset of cross-over distortion as the battery voltage falls. Once again, to compensate for variations in transistor characteristics and component tolerances, a preset adjustment is generally provided at some point to enable the circuit to be individually set up for optimum performance. The preset control is usually in the base or emitter circuit of the single transistor driver stage. In Fig. 4.7, R1 in Q1 emitter provides this function. A bias stabilisation diode or transistor may similarly be used across R2 in the type of circuit shown in Fig. 4.2, and R1 made adjustable to provide presetting.

Bias stabilisation diodes are frequently fitted to the heatsink used for the power output transistors.

Negative feedback can also be used to increase bias stabilisation, as we shall see.

Negative feedback

We saw in Chapter 3 how negative feedback may be introduced in an individual amplifier stage by using an un-bypassed emitter resistor. Alternatively, since the collector and base signal voltage variations of a common-emitter stage are in opposite phase, negative feedback can be applied in a common-emitter stage by feeding back to the base a portion of the output signal appearing at the collector. A d.c. blocking capacitor may be needed in the feedback circuit to maintain correct base bias conditions. Negative feedback is widely used to improve reproduction by cancelling spurious harmonics of the signals being amplified. In multi-stage amplifiers intended for high quality performance the use of negative feedback over several stages is commonly found. For example, in the circuit shown in Fig. 4.7 negative feedback from the collector of one of the output transistors (Q5) is fed back to the base of the driver transistor Q1 via the parallel-connected 47-k resistor and 220-pF capacitor. The proportion of the signal fed back is set by the feedback resistor (47 k in this case); a capacitor is generally found in parallel with it to provide phase correction.

Note also that in this circuit the feedback is applied direct to the emitters of the complementary-symmetry driver stage Q2, Q3.

Depending on the circuitry, it may be necessary to include a capacitor in the feedback path to block d.c. so that only signal variations are fed back. On the other hand, it is also common practice to use d.c. feedback in transistor amplifiers to stabilise the d.c. conditions. A change in the d.c. conditions at one point in the circuit will result in a change of opposite polarity at a later point. This later change can be fed back so as to oppose the original one, thus providing d.c. bias stabilisation by means of feedback. Such feedback is via R1 in Fig. 4.6. In this type of circuit it may be desired to remove signal variations in the feedback path by including in it a bypass capacitor: an example is C2 in Fig. 4.5.

'Overall', i.e. multi-stage, feedback is generally taken from the output stage and applied, as in Fig. 4.7, to the base of the first driver stage. It may, alternatively, be applied to a preceding voltage amplifier stage, and the injection point may be an emitter circuit. Including frequency-conscious components in the feedback path enables the frequency characteristics of the amplifier output to be tailored to suit particular needs, and may also be employed over a single stage for tone control purposes or to correct the characteristics of the input signal (e.g. record pickup and recording characteristics).

Small-signal stages

The number of 'small-signal' voltage amplifying stages required preceding the driver stage depends on the input signal available—which may be from a gramophone pickup, microphone, tape head or, in 'industrial' applications, some other form of transducer—and the input needs of the driver stage. Small-signal stages are, of course, biased for Class A operation. Generally, for gramophone amplifiers one to three stages are needed. A third stage is required only where elaborate tone control circuits are incorporated in the amplifier since these introduce some loss of gain which must be made up. Tone control circuits may be of

the 'passive' variety, in which, for example, to increase bass output the treble output is reduced, or of the feedback variety in which, again taking bass boost as our example, the output of a stage at the higher frequencies is reduced by increasing the amount of negative feedback at the higher frequencies. Often a mixture of these two approaches is found to be convenient. In the example shown in Fig. 4.8, a feedback circuit via P1 is used to

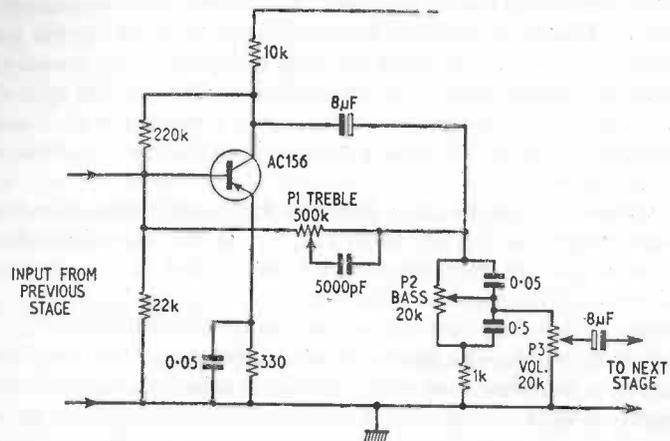


Fig. 4.8. Voltage amplifier stage incorporating tone control networks. Note polarity of electrolytic coupling capacitors. The treble tone control takes the form of a frequency-conscious variable feedback network, the bass control being in a 'passive' circuit.

provide treble adjustment, while a passive network incorporates the bass control P2. Overall output is adjusted by means of the volume control P3. In Fig. 4.9 both the treble and bass controls are incorporated in a comprehensive feedback tone control network.

It is usual, as shown, to couple the signal into and out of tone control networks and volume controls by means of coupling capacitors. This is done in order to avoid interference to the base bias conditions of the following stage. It will be appreciated that

a volume control could seriously disturb the base biasing of the following stage, so that a d.c. blocking capacitor is usually present in series with the slider.

In order to match the high output impedance of the commonly used ceramic type of gramophone pickup, a number of models use a common-collector input stage. Since the output from a cera-

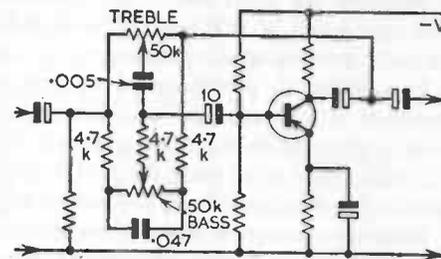


Fig. 4.9. In this circuit both the treble and bass controls are incorporated in a feedback circuit between the collector and base of the transistor.

mic or crystal pickup is fairly high, this may be the only stage used prior to the driver and output stage in low-power record reproducers.

As in all electronic equipment, the conditions of the input stage are most important, since distortion and noise introduced in this stage will be amplified in all succeeding stages. Techniques used to obtain maximum performance in this respect include operating the input stage with low emitter current and the use of a silicon planar transistor for the input stage.

TAPE-RECORDER CIRCUITS

For recording purposes a tape-recorder amplifier is required to accept a signal from a microphone or other signal source and use this to drive the tape head which establishes on the tape a permanent (though erasable) magnetic pattern proportional to the electrical signal applied to the head. For playback purposes the tape-recorder amplifier is required to accept the signal generated in the tape head when the magnetised tape is drawn across it and amplify this to a suitable level to drive a loudspeaker. Because

the magnetic characteristics of the tape are non-linear, the record signal is applied to the tape along with a sinusoidal h.f. bias signal at about 50–75 kc/s which acts to linearise the recording characteristic. In addition, improvement in signal-to-noise ratio is obtained by boosting the higher frequencies when recording, so that treble boost is used in recording and compensating bass boost is required on playback. The compensation required varies with tape speed, so that separate, switchable networks are needed for each speed at which the recorder operates.

In addition to providing record and playback amplification, therefore, a tape-recorder amplifier must include an oscillator stage to generate the h.f. bias signal required. The oscillator output is also used as an erase signal, fed to a separate erase head, to remove previous recordings or residual magnetism from the tape.

Most tape-recorders use a common record/playback amplifier, the circuit changes required for the two different functions being introduced by record/playback switching. Some more expensive machines, however, use separate record and playback amplifiers. Since power amplification is not required for recording, the playback output stage is generally switched to act as bias oscillator in the record position. Apart from the features noted above there is little in tape-recorder amplifier circuits that differs from a.f. amplifier techniques as already outlined in this and the previous chapter. The output on record and the input on playback must, of course, be matched to the impedance of the record/playback head.

Playback amplifier

Fig. 4.10 (a) shows a section of a playback amplifier incorporating a simple bass boost negative feedback network (R_1 , C_1) to provide the required playback compensation. The negative feedback is applied via R_1 , C_1 from collector to base of TR2. Since the reactance of C_1 decreases with increase in frequency, the feedback will be greater at the higher frequencies reducing the gain at these frequencies and thereby providing bass boost. The time constant ($C \times R =$ time constant in seconds, with C in

farads and R in ohms) of the network is chosen to match the recording characteristic, which is an international standard, one for each standard tape speed. For $7\frac{1}{2}$ in./sec, the time constant is $100 \mu\text{s}$. Since low noise is the most important need in the input stage, TR1 is a low-noise silicon transistor operated at low

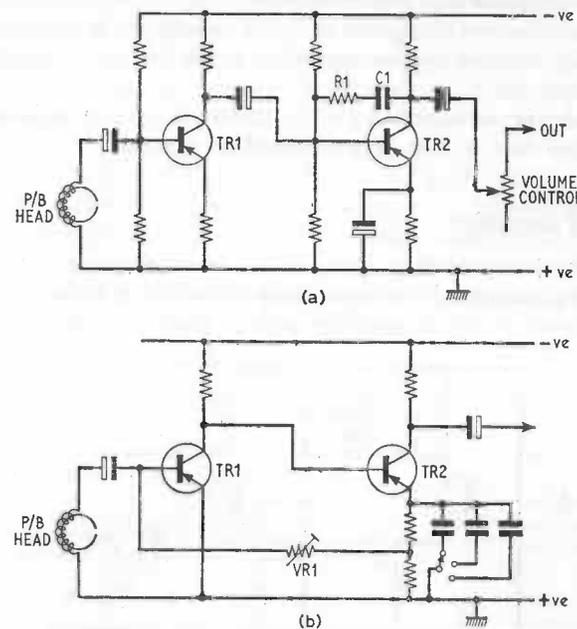


Fig. 4.10. Tape recorder playback-amplifier input circuits.

emitter current and with its emitter resistor un-bypassed to provide negative feedback and to increase the input impedance of the stage to match the tape head impedance. An input stage with too low input impedance can result in mismatch producing a fall off in high frequency response. Treble boost is sometimes used to compensate for this.

Fig. 4.10 (b) shows an alternative playback preamplifier circuit. TR1 and TR2 are a directly coupled pair with a feedback path

from TR2 emitter to TR1 base via potentiometer VR1 which is used to preset the overall gain for optimum noise performance. TR1 d.c. bias is stabilised by the d.c. feedback from TR2. Playback compensation in this circuit depends on the reactance of the winding of the tape head: as frequency increases, the head reactance increases thus neutralising the rising output. The necessary time constant change for different tape speeds is provided by switching different bypass capacitors across the resistors in TR2 emitter circuit.

The driver and output stages of a tape-recorder playback amplifier follow normal transistor a.f. amplifier practice.

Record amplifier

Again, for recording, a low-noise input stage is required, especially as microphone signals may be as little as $1-2 \mu\text{V}$. Fig. 4.11 shows a record amplifier output stage incorporating the

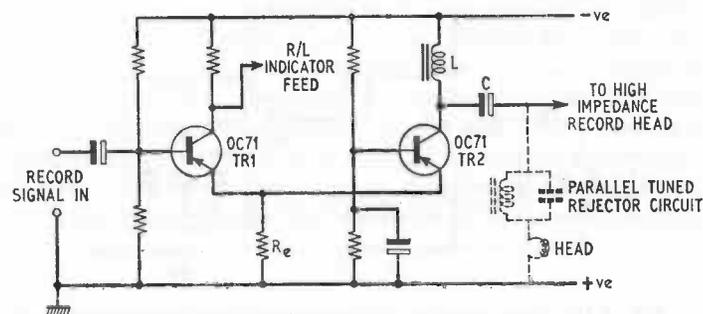


Fig. 4.11. Record output stage using a long-tailed pair (TR1, TR2) in which the coupling between the transistors is via the shared emitter resistor R_e . With this arrangement the output transistor TR2 operates in the common-base mode.

features required for recording. Treble boost is needed to compensate for losses incurred in the recording process and to obtain an improved signal-to-noise ratio, and for best results a reasonably constant current must be applied to the head. The

circuit consists of a 'long-tailed pair', that is two transistors, TR1 and TR2, coupled by means of a common emitter resistor R_e across which the input signal for the second transistor is developed. This means that the input is applied to the emitter of the second stage so that it operates in the common-base mode, thus providing a high-impedance output to match to a high-impedance record head. The load inductor L is used to achieve constant current recording. As its reactance at low frequencies is lower than at high frequencies, the signal voltage rises with frequency.

To obtain constant current recording with a low-impedance record head, the circuit shown dotted in Fig. 4.11 may be used, with the collector load again an inductor, but with a larger value capacitor in position C and with a parallel-tuned circuit in series with the feed to the head. TR2 must, of course, be operated in the common-emitter mode when used to drive a low-impedance head.

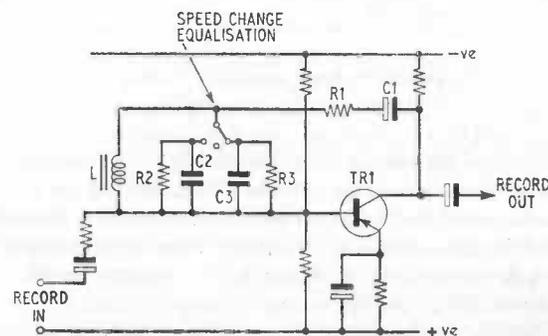


Fig. 4.12. Equalisation by means of a frequency-conscious feedback loop.

Frequency selective circuits are necessary in a record amplifier to compensate for high frequency losses, different compensation being required at different tape speeds. The circuit shown in Fig. 4.12 provides both treble and bass correction by means of a feedback circuit incorporating a parallel-tuned circuit. The

feedback is applied from TR1 collector to its base via C1, R1 and L. C2 and C3 form the parallel-tuned circuit with L, either one being selected by means of the speed-change switch. The impedance of a parallel-tuned circuit being maximum at its resonant frequency, the feedback will be least and the gain greatest at this frequency. R2 and R3, again selected by the speed-change

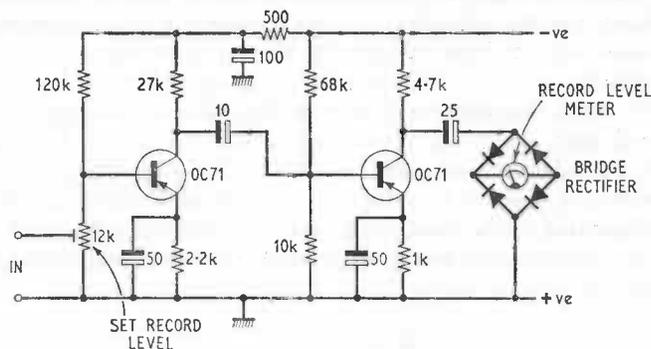


Fig. 4.13. Two-stage voltage amplifier feeding a record level meter via a bridge rectifier.

switch, are used to reduce the Q of the feedback tuned circuit, giving it a 'flat', bandpass characteristic instead of a sharply selective characteristic. In this way the response of the amplifier is adjusted by means of the frequency selective characteristic of the feedback network: the feedback is minimum within the required bandwidth so that the gain of the stage increases, but outside this bandwidth the feedback increases and the gain is thus reduced.

Record level indication is obtained in the circuit shown in Fig. 4.11 by tapping a portion of the signal appearing at the collector of TR1 and feeding this to a record level indicator arrangement. A small meter is often used for this purpose, a suitable record level meter indicator circuit being shown in Fig. 4.13. The signal tapped off from TR1 (Fig. 4.11) collector is amplified by the two OC71 transistors and rectified by the bridge rectifier circuit

shown, giving a d.c. meter reading proportional to the a.c. signal current (the action of a bridge rectifier is described in Chapter 7).

Most forms of sinewave oscillator are suitable for providing the h.f. bias needed for recording. In the example shown in Fig. 4.14 two complementary transistors are used in a push-pull

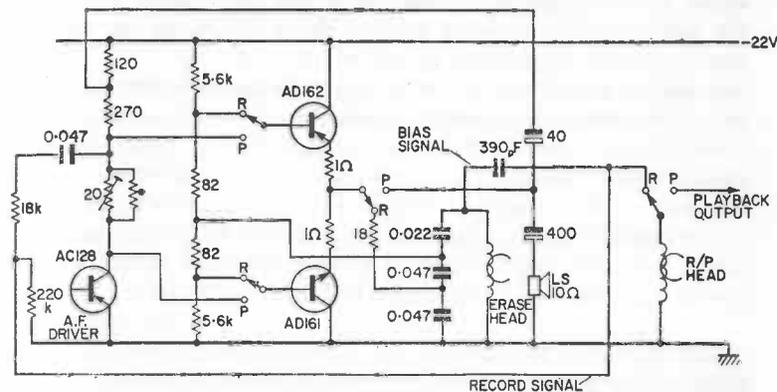


Fig. 4.14. A complementary-symmetry stage (AD161 and AD162) which acts as push-pull output stage on playback and bias oscillator on record. On both playback and record, input from the voltage amplifier stages is to the base of the AC128 audio-driver stage. On record the resistor chain in the base circuit of the complementary-symmetry stage provides the forward bias required at the bases of these transistors to forward bias them into conduction to provide sustained oscillation. Record output is from the collector of the driver transistor via the 18-k resistor which provides constant current drive to the record/playback head. The bias oscillator tuned circuit comprises the erase head winding and associated parallel capacitors.

arrangement and a saving in components is achieved by using the erase head winding as the inductive part of the oscillator tuned circuit. The tuned circuit is connected to the emitters of the transistors, with feedback to the base circuit. The advantage of this type of circuit is that by including as shown switching it can also be used on playback as the output stage. The bias for the

record head is tapped off via a 390 pF capacitor. To avoid feedback of the bias signal to the audio circuits a filter tuned to the bias frequency may be incorporated in the record head signal feed, a filter of this type being shown dotted in Fig. 4.11.

In the arrangement shown in Fig. 4.14 the output signal on record is taken from the collector circuit of the AC 128 driver stage: the relatively high-value 18k resistor in the signal feed to the head provides constant current drive which, as we have previously seen, is required by the record head.

A further use of transistors in tape-recorders is in automatic tape-drive motor speed control systems.

5

R.F. TECHNIQUES

TRANSISTOR radio receivers have been produced and sold in vast quantities all over the world, and considerable variation exists in their design. There is, however, a simple 'basic' type intended for reception of a limited number of stations that follows a fairly standard pattern. In this chapter we shall consider first this simple, standard type of transistor radio, then consider some of the circuits that have been used where a more elaborate specification is called for, and finally look briefly at television receiver applications.

Simple receivers intended for reception of amplitude modulated transmissions on the medium- and long-wavebands generally consist of a ferrite rod aerial to pick up the transmissions, a mixer stage which converts the various stations the set will receive to a standard intermediate frequency (i.f.) of, generally, 470 kc/s, one or two stages of i.f. amplification, a demodulator (or detector as it is also called) to recover the original audio signals, and an audio section which uses the types of technique described in the last chapter.

Input and mixer stage

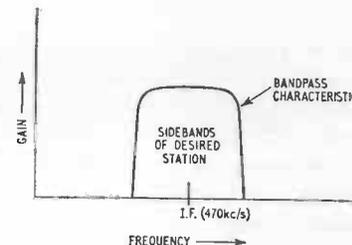
A simple receiver aerial input and mixer stage is shown in Fig. 5.1. The transistor's d.c. conditions are established along the lines previously outlined with a potential divider network (R1, R2) in the base circuit and an emitter resistor R3 with by pass capacitor C3. The input tuned circuit L1, VC1 (with trimmer T1 providing preset adjustment) is coupled to the ferrite rod aerial. The reactance of the input tuned circuit is varied by means of VC1 to enable different stations to be selected. The

six-transistor plus diode-detector circuit. In this two input circuits are included to increase the available tuning range. L1 and its associated components cover the medium waveband; L2 and its associated components (extra 82 pF capacitor) the long waveband. The switching shown enables either waveband to be selected. The oscillator circuit is modified for long waveband reception by switching the 230-pF tuning capacitor and trimmer T3 into circuit. A further modification is the provision of an aerial socket to enable a car aerial (see later) to be connected to the set. Note that in this design the signal coupling windings to TR1 base are in series with the base bias network (56 k and 10 k resistors): the lower arm of the bias network is therefore bypassed.

The mixer output is developed across its LC load circuit, which is tuned to the i.f. This is called the first i.f. transformer (IFT1). Its output is applied to the base of TR2, the first i.f. amplifier. The output of this is similarly developed across the tuned load circuit IFT2 and then applied to a second i.f. amplifier stage TR3 with further tuned load circuit IFT3 tuned to the i.f. Note that, as mentioned in Chapter 3, the collectors are connected to taps on the coils of the tuned circuits to simplify the design of the tuned circuits, the bases being fed from small windings on the tuned circuits to provide impedance matching at the inputs to the transistors. The output from the second i.f. stage is fed to the detector diode X1. Thus it will be seen that the main pre-detection amplification takes place in the i.f. amplifier section of the receiver. The selectivity of the receiver, i.e. its ability to select the wanted signal and reject those at other frequencies, is also mainly determined by the i.f. amplifier section. Because of this, the i.f. transformers are designed to have a 'bandpass characteristic' of the form shown in Fig. 5.4, with maximum gain at the i.f. and close to it and sharp rejection, i.e. steep fall in gain, to each side so that unwanted signals are rejected. The 'band' in this case refers to the fact that the wanted audio information is carried in the region just around the actual i.f. Quite such sharp rejection as that shown in Fig. 5.4 is not, in practice, possible, but the designer approximates as closely as he can to it. The i.f. transformers in this example (Fig. 5.3) consist of a single tuned

circuit (collector winding plus 250 pF tuning capacitor in each case) with a small coupling winding to couple the signal to the base of the following stage. While this is a very common arrangement, in many receivers the i.f. transformers consist of two tuned circuits coupled together, as shown in the design in Fig. 5.6

Fig. 5.4. Ideal bandpass characteristic giving good selectivity, i.e. amplification of wanted signals and rejection of unwanted ones.



(C4/L7 and C7/L8). A double-tuned transformer gives improved selectivity, though the technique used in the design in Fig. 5.3 is adequate for most purposes. As with the mixer output circuit, the other i.f. transformers are also provided with ferrite tuning cores for alignment purposes.

It will be noticed in Fig. 5.3 that an RC feedback network is connected between the output and the base of each i.f. amplifier stage (1.2 k resistor and 56 pF capacitor in the case of TR2, 3.9 k resistor and 18 pF capacitor in the case of TR3). This is termed a unilateralisation network and is used to neutralise feedback via the internal collector-base capacitance of the transistor. The principle is that feedback in opposite phase to the internal feedback in the transistor results in cancellation of the unwanted internal capacitive feedback. In some designs neutralisation is achieved by means of a feedback capacitor instead of an RC unilateralisation network. These techniques were commonly employed when alloy-junction transistors were used in the i.f. stages, since internal capacitance was a problem with them. With the use of alloy-diffused and planar transistors in i.f. stages, however, the internal feedback capacitance is negligible and neither neutralisation nor unilateralisation is required.

In all other respects the i.f. stages follow normal amplifier

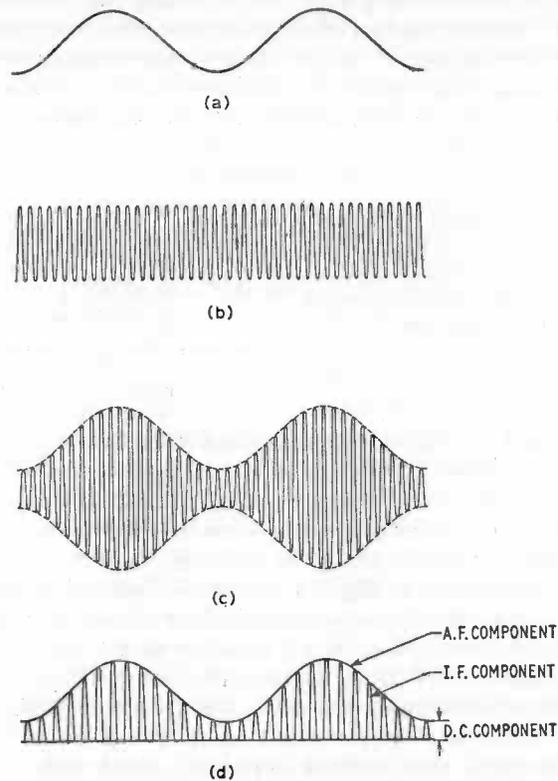


Fig. 5.5. Modulation and demodulation (detection). (a) A.F. waveform. (b) R.F. carrier wave. (c) R.F. carrier wave modulated by a.f. waveform. (d) The output from the demodulator, which rectifies the input applied to it, consists of a d.c. component which is blocked from the audio stages by means of a capacitor, an i.f. component which is filtered out, and the original a.f. waveform.

practice as outlined elsewhere in this book, except that automatic gain control is applied to TR2 base. This is a form of feedback. Before going into this, however, the operation of the detector stage must be described.

Detection

A word is first necessary on the process of modulation at the transmitter. Suppose, see Fig. 5.5, that we have an audio signal (a), which might, for example, be derived from a microphone. To transmit this it is necessary to impress it on a radio frequency 'carrier' wave. Suppose that such a carrier is as shown at (b). Mixing the audio signal and the r.f. carrier will result in a modulated carrier wave as shown at (c). This is the type of signal received by the receiver. To recover the audio signal, two processes are necessary, first to remove half of the incoming signal so that we get an output as shown at (d), and secondly, to remove the r.f. component of this signal. The detector diode (X1) rectifies the signal to give the type of output shown at (d)—the process of rectification was described in Chapter 1—and a simple filter circuit is used to remove the r.f., or rather i.f. as it will be at the detector, component of the signal. In the circuit shown in Fig. 5.3 the two 0.04- μ F capacitors and 470-ohm resistor at the detector output comprise the i.f. filter. The detector output signal is developed across its 5-k load resistor, which also serves as the set's volume control, and will also have, as shown in Fig. 5.5 (d), a d.c. component. The output at the slider of the volume control is fed via the 8- μ F capacitor, which blocks the d.c. component of the signal, passing on only the a.f. signal fluctuations, to the base of the audio amplifier TR4. Note that the detector diode is often mounted inside the screening can of the final i.f. transformer, and its presence is not always therefore evident.

Audio stages

The audio stages are conventional, with a driver transformer providing the input required by the Class B push-pull output

stage. Negative feedback from TR6 collector is applied to TR6 base (via the 15-k resistor) and TR4 base (via the 1-M resistor).

A.G.C.

Automatic gain control (a.g.c.) is used to overcome fluctuations in the strength of the received signal. The mean level of the signal at the output of the detector is proportional to the strength of the input signal, and can therefore be used in a feedback loop to provide a.g.c. Referring to Fig. 5.3, the signal at the detector output after i.f. filtering is fed back as a bias voltage to TR2 base via a network (8.2-k resistor and 8- μ F by-pass capacitor) which smooths out the a.f. signal variations. This feedback bias reduces the gain of the first i.f. stage on increase in input signal strength, but allows it to rise to maximum when the input signal is weak. In this way the effect of signal strength variations on the receiver output is cancelled.

This type of a.g.c. is called *reverse* a.g.c. An alternative technique, *forward* a.g.c., is used in transistor television i.f. stages (and a few radio receivers), and will be described later.

Power supply

A 9-V battery is used to provide the d.c. supply required to power the receiver (Fig. 5.3), and the positive side of the supply is 'earthed', i.e. forms the 'common' chassis side of the supply. To prevent feedback via the supply line, the supply line is provided with decoupling components. The 100- μ F electrolytic capacitor on the right, next to the on/off switch, decouples the battery. Since the battery has some internal resistance, signal fluctuations will exist across it following the variations in current drawn by the output transistors: these are smoothed out or decoupled at the negative, collector side of the supply by this 100- μ F electrolytic capacitor. The 680-ohm resistor and second 100- μ F electrolytic capacitor likewise prevent signal feedback via the supply line from the output stage to earlier parts of the receiver.

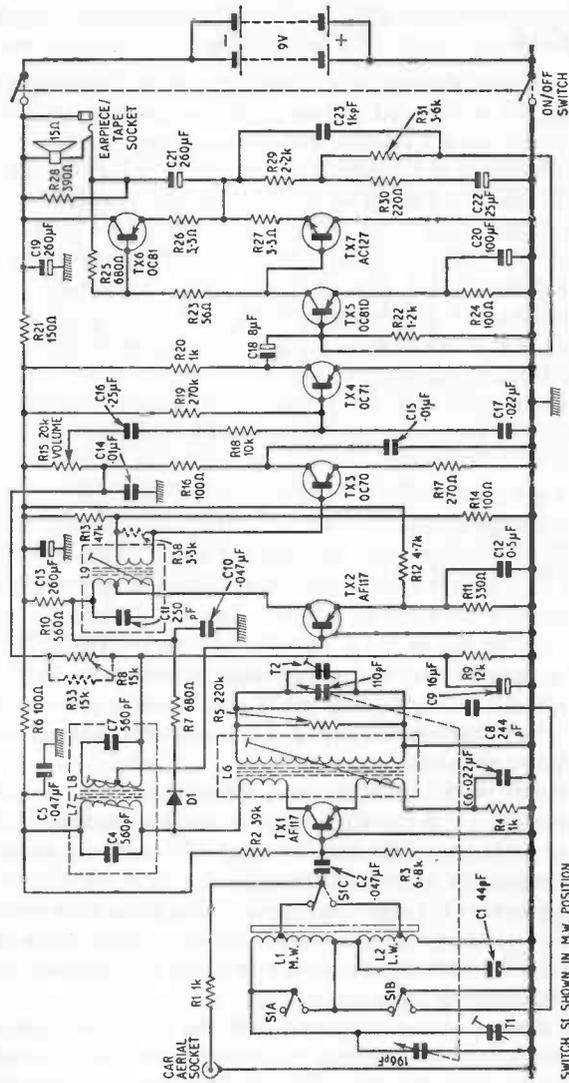
Circuit Fig. 5.6

To illustrate some alternative receiver techniques, the complete circuit in Fig. 5.6 is included. Once again the receiver is a two-waveband model using a self-oscillating mixer stage TX1. The first i.f. transformer is of the double-tuned variety, with two tuned circuits C4/L7 and C7/L8. The output from this is applied to a single i.f. amplifier stage TX2 using an alloy-diffused transistor, neutralisation not being required. Detection is performed by a transistor 'collector-bend' detector TX3. The output from this is fed to a stage of a.f. amplification TX4, then to a driver stage TX5 which feeds a complementary-symmetry output stage, all these stages following techniques described in Chapter 4. Feedback from the output stage is applied to the base of the driver stage, and a.g.c. is applied to the base of the i.f. amplifier stage via R8.

Because ferrite rod aerials are directional—they need to be aligned with the geographical location of the transmitter to give maximum signal pickup—and because the metal body of a car tends to act as a screen against radio transmissions, provision is made to connect a car aerial to the set. In most designs the car aerial input is coupled in via a separate coil on the ferrite rod. Alternatively separate tuning arrangements may be used so that the ferrite rod and the windings on it can be switched out of circuit. In the design shown in Fig. 5.3 the car aerial is capacitively coupled to the tuned input circuit.

The operation of the collector-bend detector is fairly simple. The transistor is biased (by the potential divider network R13, R14) so as to make use of the non-linearity of the reverse-biased collector characteristic to rectify the signal fed to it, i.e. the collector junction acts as the signal rectifier. This form of detector has the advantage of providing a certain amount of amplification. C14, R16 and C15 form the i.f. filter, and R8 and C9 remove the a.f. variations from the a.g.c. line.

A crystal diode and series resistor, D1 and R7 in the circuit shown, are often used in transistor receivers to increase the range of the automatic gain control. The diode is normally reverse



SWITCH S1 SHOWN IN N.W. POSITION

Fig. 5.6. Seven-transistor, two-waveband radio receiver circuit using alloy-diffused transistors in the mixer (TX1) and i.f. amplifier (TX2) stages. TX3 collector bend detector, TX4 a.f. amplifier, TX5 audio driver, TX6 and TX7 complementary-symmetry audio output stage.

biased so that it is cut off. A very strong signal, however, will result in it being forward biased into conduction. The result is that the low value resistor R7 is introduced across the collector output circuit, and this will reduce the gain at the first i.f. transformer.

Modules

Many transistor receivers today are built around small 'modules', self-contained assemblies on a small printed board. Both a.f. and i.f. modules are employed. An example of an i.f. module is shown in Fig. 5.7. This is a pre-aligned unit containing three

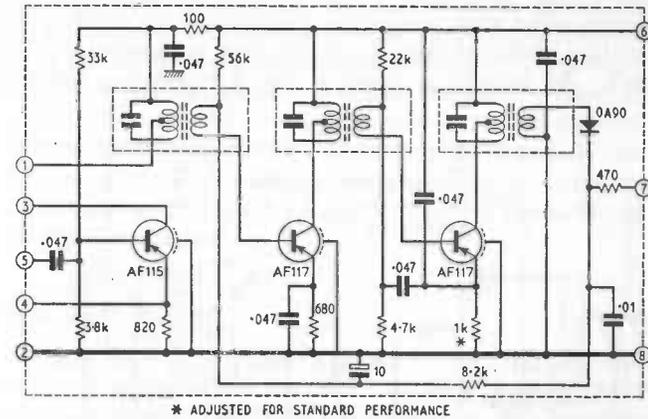


Fig. 5.7. Pre-assembled modules are used in many transistor receivers today. In this example the AF115 transistor is intended to operate as a self-oscillating mixer followed by two AF117 i.f. amplifier stages with a.g.c. applied to the first OA90 detector.

alloy-diffused transistors and a diode for detection and a.g.c., with all external connections to the module made via eight tags. The AF117 transistors act as i.f. amplifiers, and the AF115 as mixer. Tags 6, 8 and 2 are used for d.c. supply purposes, and the a.f. output is available at tag 7. Tags 1, 3, 4 and 5 are used for whatever

tuning requirements are decided upon at the input end, with the input signal applied to tag 5, tag 4 providing emitter coupling to the oscillator tuned circuit and tags 1 and 3 providing collector coupling to the oscillator tuned circuit, to form a self-oscillating mixer.

Separate receiver oscillator

While the vast majority of transistor radio receivers use the self-oscillating mixer arrangement, in a few, intended mainly where short-wave reception is to be included, a separate oscillator stage is employed. Since the various tuned circuits in a self-oscillating mixer stage tend to affect each other, the effect of using a separate oscillator stage is to improve performance in this respect. An example is shown in Fig. 5.8: this is a simplified

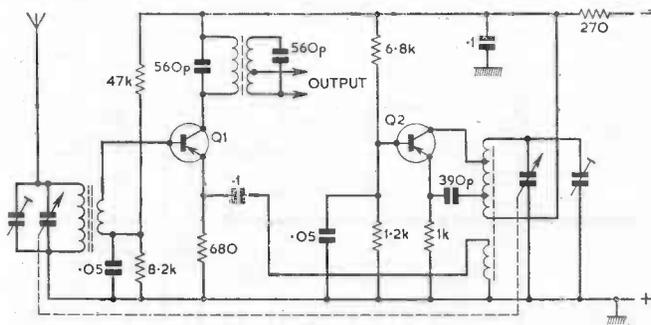


Fig. 5.8. Simplified circuit of a mixer stage (Q1) coupled to a separate local oscillator stage (Q2).

circuit, showing the tuning arrangements for one waveband only. Q2 is the oscillator, with the tuned circuit in its collector lead and feedback to its emitter via the 390-pF capacitor (thus the oscillator is operating in the common-base mode, as does the oscillator part of a self-oscillating mixer). The oscillator signal is coupled to

the emitter of the mixer stage via the small winding and 0.1- μ F coupling capacitor. A number of different types of separate oscillator circuits has been used.

To give further improvement in sets intended for long-distance reception, a stage of r.f. amplification preceding the mixer stage is included in some models.

Bandspreading

Because of the crowded conditions of the broadcast wavebands, in many receivers bandspread tuning is incorporated, that is provision is made for additional fine tuning after the main tuning control has been set as closely as possible to the correct station setting. In many sets this bandspread feature is restricted to the part of the medium waveband around Radio Luxembourg. In others the feature is available over the entire tuning range of the receiver. A number of techniques has been used to provide this improved tuning capability. Most rely on provision in the oscillator circuit of an additional variable capacitor which can be used to vary the tuning a few kc/s either side of the setting of the main tuning capacitor.

Negative chassis

In both the receiver circuits shown in Figs. 5.3 and 5.6, the positive or emitter side of the power supply is 'earthed', i.e. connected to chassis. This is the usual practice. In some designs, however, it is the negative or collector (assuming *pn*p transistors) side of the supply that is connected to chassis, and in this case the supply line decoupling components may be connected in the positive supply line and the various decoupling or bypass capacitors connected to the 'earthed' collector negative supply line instead of to the positive or emitter side of the supply. Circuits following this practice are sometimes so drawn that they appear to be 'upside down', with the emitters at the top of the diagram and the collectors at the bottom, or with the thick chassis line at the top and the bypass capacitors taken to this. Fig. 5.9

illustrates these points: in each case C_1 , R_1 and C_2 are supply decoupling components and C_B is an emitter-resistor bypass capacitor. Other arrangements have been used.

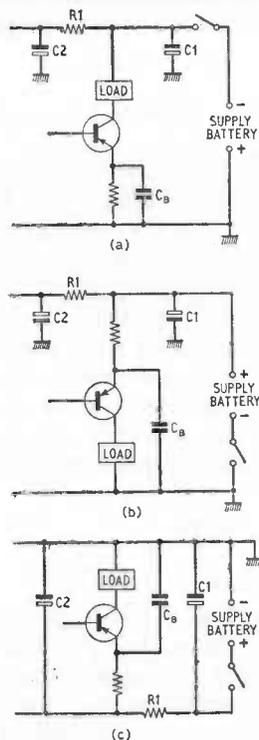


Fig. 5.9. Either the positive or negative side of the supply may be connected to chassis, and circuit drawings will vary accordingly. (a) Usual arrangement with pnp transistors and the positive side of the supply taken to chassis. (b) and (c) Ways of drawing the circuit when the negative side of the supply is taken to chassis. In each case C_1 , R_1 , and C_2 are the supply decoupling components and C_B is an emitter decoupling capacitor. When servicing, the polarities of the supply components should be carefully noted to ensure that electrolytics are connected into circuit correctly. When, as shown in these examples, a single-pole on/off switch is used, this may be found in either the positive or negative side of the supply.

V.H.F./F.M. RECEIVERS

In addition to a.m. reception many receivers incorporate further circuitry for the reception of the frequency modulated (f.m.) broadcast transmissions at v.h.f. in the band 87.5–100 Mc/s. The additional requirements to receive these are: (a) a tuner unit incorporating input, amplification and mixer circuitry

capable of v.h.f. reception and tuning over the broadcast v.h.f. waveband; the v.h.f./f.m. mixer stage generally provides an output at 10.7 Mc/s, the standard i.f. for this purpose; (b) additional tuned circuits in the i.f. section of the receiver to provide amplification at 10.7 Mc/s; and (c) a type of detector capable of recovering the audio signals from the f.m. signal appearing at the output of the i.f. amplifier.

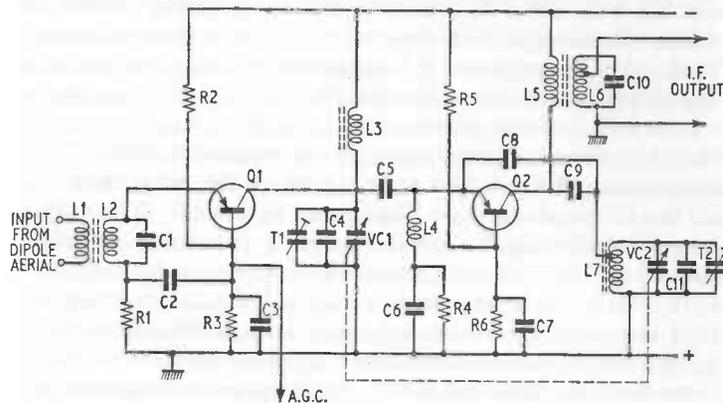


Fig. 5.10. Typical Band II tuner unit circuit for v.h.f./f.m. reception. Q_1 r.f. amplifier, Q_2 self-oscillating mixer. Both transistors operate in the common-base mode.

V.H.F. tuner unit

The circuit of a typical transistor tuner unit for v.h.f./f.m. reception is shown in Fig. 5.10. While this is the basic arrangement generally used in such units, in practice a few minor additions are usually made in order to provide improved stabilisation.

In view of the low signal strengths at v.h.f., a stage of r.f. amplification is necessary prior to the mixer stage, and as optimum results at v.h.f. are obtained by operating transistors in the common-base configuration, both the r.f. amplifier (Q_1) and the mixer (Q_2) transistors are operated in this mode. The input from the aerial is coupled to the emitter of the r.f. amplifier by a wideband

transformer L1, L2, the tuned circuit L2, C1 in this transformer being designed for maximum gain over the complete bandwidth of the tuner. R1 provides emitter bias and is bypassed by C2. Base bias is provided by the potential divider network R2, R3 with C3 to bypass R3. In some tuner units a.g.c. is applied to the r.f. amplifier stage, and this is generally injected in the base circuit as shown. Q1 output is developed across the tuned circuit L3, VC1, C4, T1, with VC1 the manual tuning control for station selection ganged to the oscillator variable tuning capacitor VC2. The output from Q1 is capacitively matched to the input impedance of the mixer emitter by C5. L4, C6 comprise a resonant filter circuit tuned to the i.f. of 10.7 Mc/s and are included to prevent signals at the i.f. breaking through to the following circuits. R4 provides emitter bias for the mixer, with R5, R6 and C7 providing base biasing and bypassing. The output from Q2 is developed across L5 to which is coupled the tuned circuit L6, C10. L5 and L6 form the first i.f. transformer, tuned to 10.7 Mc/s. It is the general practice to feed the output from the v.h.f. mixer to the base of the a.m. mixer stage, which is switched to act as an additional i.f. amplifier for v.h.f./f.m. reception, all that is necessary being to switch out of circuit its oscillator circuits and include a 10.7-Mc/s i.f. transformer in its collector load circuit in series with the 470-kc/s a.m. i.f. transformer.

The oscillator-tuned circuit in the v.h.f. tuner is L7, VC2, C11, T2, with VC2 ganged to VC1 as we have seen for station selection. The practice in the rather different circumstances of v.h.f. reception is for the oscillator frequency to be *below* the input signal. Thus with an input signal of 98 Mc/s the oscillator frequency would be 87.3 Mc/s to give an i.f. output at 10.7 Mc/s. Q2 collector is connected to the oscillator-tuned circuit by C9, and feedback to its emitter is via C8. Since the feedback should be positive with no phase shift, but a phase difference of about 90° between the emitter and collector signal currents exists at the frequencies the transistor is handling (the phase relationship between input and output signals at these frequencies does not exactly follow the relationship given in Table 3.2); C8 is chosen so as to provide the required phase correction.

Dual-channel i.f. stages

As already mentioned, the a.m. mixer stage in an a.m./f.m. receiver forms the first i.f. stage when the receiver is switched to v.h.f./f.m. reception. It is a fairly simple matter to arrange for the whole i.f. section of the receiver to operate at the two intermediate frequencies, 10.7 Mc/s and 470 kc/s, without switching. Two sets of i.f. transformers are required, of course, to couple the i.f. amplifiers together, and these are connected in series in the i.f. amplifier collector circuits. Such an arrangement is generally termed a dual-channel i.f. amplifier. An example is shown in Fig. 5.11. The load circuits tuned to the higher frequency are in each case connected nearest to the collector, the technique being based on the fact that the reactance of the 10.7-Mc/s tuned circuit is negligible at 470 kc/s, the a.m. i.f., while the reactance of the 470-kc/s tuned circuit is negligible at 10.7 Mc/s. Quite complex impedance matching arrangements may be needed with the use of pairs of series-connected tuned circuits. It is common practice to connect small value—about 220 ohms—resistors, marked R, in the collector leads to prevent unwanted oscillations developing, and in some receivers capacitors of about 0.002 μ F, marked C, are included to bypass the 470 kc/s transformers for 10.7 Mc/s operation. The a.m. detector follows conventional practice as previously described, with a.g.c. taken from the point shown.

F.M. detector

Of a number of possible types of f.m. detector, that used in the overwhelming majority of a.m./f.m. receivers is the ratio detector employing a pair of germanium diodes as shown in Fig. 5.11. Ratio-detector circuits may be balanced or unbalanced. That shown in Fig. 5.11 is of the unbalanced variety. The balanced circuit differs simply in that the ratio-detector load resistor R3 and i.f. bypass capacitor C3 are centre-tapped with the centre-tapping point connected to earth.

Points to note about the ratio-detector circuit are that the i.f. transformer uses a pair of loosely coupled tuned circuits (L1, C1

and L2, C2) with the secondary centre-tapped and connected to a third winding coupled to the primary, and that the two diodes are connected in series with the secondary and the load resistor R3. The frequency modulated signal varies about a centre frequency. At this centre frequency, to which the ratio-detector i.f. trans-

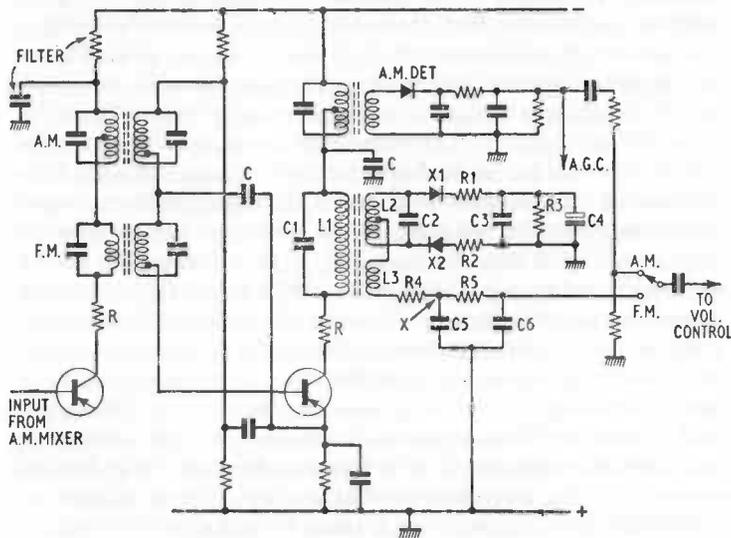


Fig. 5.11. Typical dual-channel i.f. stages of an a.m./f.m. receiver. Standard i.f.s are 10.7 Mc/s for f.m. and 470 kc/s for a.m. The higher frequency i.f. transforms are connected nearest to the collectors of the transistors. The f.m. demodulator comprises a ratio-detector circuit X1, X2.

former is tuned, equal and opposite voltages appear at each end of the secondary (because it is centre-tapped), both diodes conduct equally (a matched pair is used) and a steady current flows through the load resistor R3, this being stabilised by the large-value electrolytic capacitor C4. As the frequency of the input signal varies in accordance with the modulation, however, difference voltages arise between the centre-tapping and the ends of the

secondary winding—because of the phase shifts produced by the frequency variations—and these result in unequal conduction of the two detector diodes. These difference voltages result in an amplitude varying a.f. signal that is the ratio between them—hence the name ratio detector—appearing across the third winding L3.

One of the main reasons for the use of f.m. for broadcasting is its freedom from most types of interference, which are of similar form to a.m. signals, and the reason for the popularity of the ratio detector is its good a.m. rejection. Since, however, parts of the circuit are sensitive to a.m. it is important to design such stages for maximum a.m. rejection. Steps that may be taken include the incorporation of balancing resistors—R1 and R2 in the circuit shown—and the inclusion of an RC time-constant amplitude limiter filter—R4, C5 in this example. In some circuits one of the balancing resistors is made variable to enable the circuit to be individually adjusted for maximum a.m. rejection.

R5 and C6 form a 'de-emphasis' filter. This is included to compensate for the high frequency boost that is given to the signal at the transmitter to improve the signal-to-noise ratio.

In some current a.m./f.m. receivers the transistors in the i.f. amplifier stages are operated in the common-base mode on f.m. and in the common-emitter mode on a.m.

TELEVISION RECEIVERS

At present the tendency is still for the use of transistors in television receivers to be limited to the small-signal stages—the r.f. and i.f. sections of the receivers, that is. The first large-scale use of the transistor in television receivers was for u.h.f. tuner units since at these frequencies recent types of transistor provide appreciably better performance than do valves.

U.H.F. tuners

As in the case of v.h.f. tuners for f.m. sound broadcast reception, the general arrangement used is to have a stage of r.f.

amplification followed by a self-oscillating mixer stage, with the transistors operating in the common-base mode. Physically, however, there are considerable differences in the associated circuitry. This is because the very small values that would be required for the tuning components—coils and capacitors—at u.h.f. make it necessary to use alternative circuit techniques. A conventional coil, for example, to tune to u.h.f. channels could be smaller than the associated lengths of wiring if conventional techniques were employed. Use is made instead of the fact that a short length of transmission line—a resonant line—acts as a tuned circuit at u.h.f., the frequency at which it resonates being determined by its length. Resonant lines are simulated in u.h.f. tuners by short lengths of stout wire, the chassis forming the other section of the 'transmission line', and variable tuning is achieved by means of ganged capacitors connected at the ends of the lines.

Further, because of the need to screen the various circuits from each other, u.h.f. tuners are constructed using a rigid steel box with separate compartments for the various tuned circuits.

An example is shown in Fig. 5.12, in which VT1 is the r.f. amplifier and VT2 the mixer. There are four variable-tuned circuits tuned by the ganged capacitor C4a-d. The signal from the aerial is coupled to the input resonant line X2, which is tuned by C4a, by the small coupling wire loop X1, while X3 couples the input tuned circuit to VT1 emitter. VT1 tuned load circuit consists of X4 tuned by C4b, and this is coupled to the mixer input tuned circuit X5/C4c by two coupling slots—small apertures drilled in the screening partition between the two circuits. X4/C4b and X5/C4c thus form a bandpass-tuned transformer. X6 couples the input to VT2 emitter. The oscillator-tuned circuit consists of X7 tuned by C4d, and is coupled to VT2 collector by C17 with feedback to the emitter via C12. The i.f. output, at 39 Mc/s vision and 33.5 Mc/s sound, the standard i.f.s for u.h.f. television reception, is developed across the i.f. coil L5 and fed to the v.h.f. tuner unit where, on u.h.f., the v.h.f. mixer is generally used as an additional i.f. stage.

Quite a variety in detail exists between different designs. In some tuners a wideband input circuit is used for the r.f. amplifier

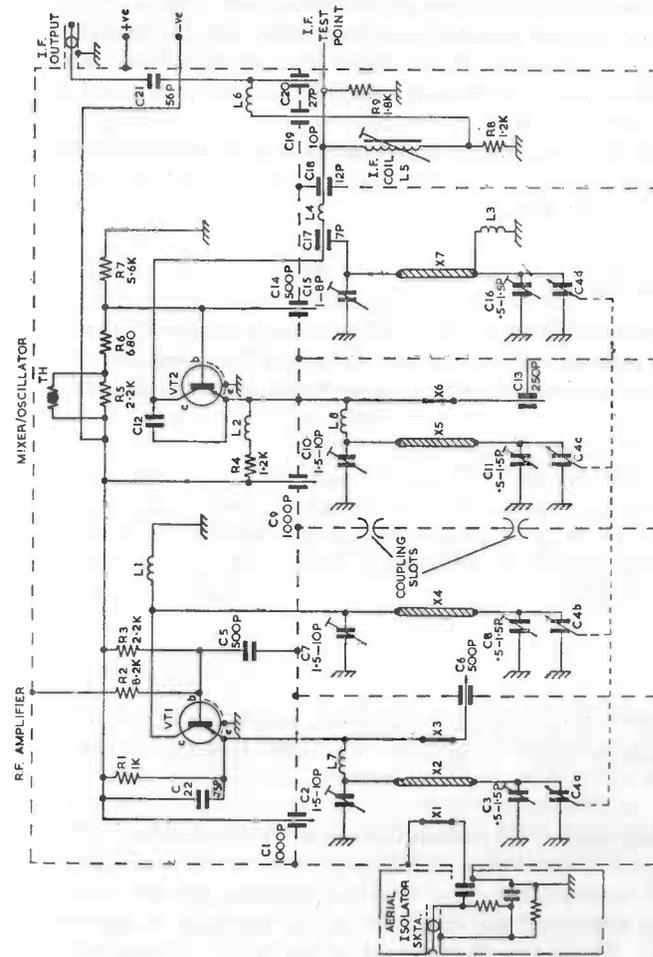


Fig. 5.12. Representative television receiver u.h.f. tuner unit circuit. VT1 r.f. amplifier and VT2 self-oscillating mixer operate in common-base mode. At u.h.f. inductive parts of the tuned circuits of loops or lengths of stout wire (X2, X4, X5, X7, with signal coupling loops X1, X3, X6). Coupling between the bandpass tuned circuit between VT1 and VT2 is through small coupling slots—apertures drilled in the screening plates. Tuned circuits comprise X2, C4a input; X4, C4b and X5, C4c bandpass coupling transformer; and X7, C4d oscillator tuned circuit.

stage, so that only three variable-tuned circuits are involved. The transistors, unlike the arrangement shown in Fig. 5.12, where they are mounted in a separate compartment, are often mounted in the tuned circuit screened compartments, the r.f. amplifier transistor being mounted in the aerial input tuned circuit compartment and the mixer transistor in the oscillator tuned circuit screened compartment.

Note that the positions of the loops and wiring (which must be kept as short as possible) are critical at these frequencies, and must not, therefore, be disturbed.

V.H.F. tuners

Combined transistor u.h.f./v.h.f. tuner units are used in some television receivers—Fig. 5.13 gives a typical block outline—but it is still more common to find separate u.h.f. and v.h.f. tuner

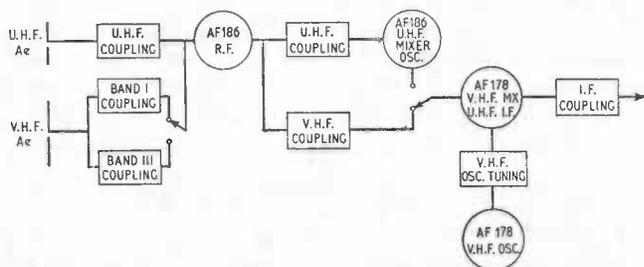


Fig. 5.13. Integrated u.h.f./v.h.f. tuner units are used in some television receivers. This block schematic shows a typical arrangement.

units being used. The main difference between Band I/III television v.h.f. tuner units and Band II f.m. sound v.h.f. tuners is that in television tuners a separate oscillator stage is used so that three transistors are required—as r.f. amplifier, mixer and oscillator. The reason for this is that the self-oscillating mixer circuit does not give satisfactory results where the signal and intermediate frequencies are in close proximity as they are in the

case of Channel 1, Band I, which extends down to 41 Mc/s, which is close to the standard intermediate frequencies for Bands I/III of 38.15 Mc/s sound, 34.65 Mc/s vision.

While the mixer and oscillator stages are operated in the common-base mode, in some transistor v.h.f. tuners the r.f. amplifier stage is operated in the common-emitter mode. Continuous tuning over the bands, with switching between them, is usually provided by making the permeability of the tuning coils variable.

I.F. stages

The requirement in this part of the receiver, as with other types of receiver, is to provide a large measure of the overall gain and selectivity: the transistors provide the gain and the tuned circuits the selectivity. Three vision i.f. stages are generally needed, with a further two sound-only i.f. stages, the Band I/III a.m. sound being amplified in addition in one or two of the vision i.f. stages while the Band IV/V f.m. sound is passed through the entire vision i.f. amplifier chain. Because of dual-standard 405/625 line operation and the need for adjacent channel rejector circuits rather complex tuned circuits are found in the i.f. sections of the receiver, with considerable differences between different models.

Forward a.g.c.

In the type of a.g.c. previously described feedback is used to reduce the current flowing through the controlled stage and consequently the gain it provides. While satisfactory for radio reception, this technique has the drawback of biasing the transistor back into a less linear portion of its input characteristic, and for this reason is not a suitable technique for television receiver a.g.c. Instead, in television receivers a technique known as forward a.g.c. is used. In this the feedback is used to increase the current flowing through the controlled transistor. A resistor is included in the output load circuit so that as the current through the transistor increases the collector voltage falls, pulling back the

gain of the stage. Transistors intended for use with this form of a.g.c. are designed to have appropriate output characteristics. To increase the range of automatic gain control, a stage of d.c. amplification is generally included in the a.g.c. circuit. Where a.g.c. is applied to r.f. stages as well it is usual to clamp the a.g.c. to the r.f. stage so that it does not come into operation until the voltage on the a.g.c. line reaches a certain level, remaining inoperative when weak signals are being received.

Other parts of the receiver

There is little advantage at present to be gained from the use of transistors in other parts of the receiver. An exception is the decoder circuitry in colour television receivers. Here the large number of stages required to process the colour signals means that transistors, on the grounds of size and low power consumption with consequent less generation of heat, are being used from the start. Before leaving receiver circuits, however, it is worth briefly describing the two types of oscillator, the multivibrator and blocking oscillator, generally used in television receiver timebases in their transistor forms. While the multivibrator is the most widely used timebase oscillator in valved receivers, in fully transistorised receivers the blocking oscillator has been found to perform better as a timebase oscillator.

The basic requirement of a timebase waveform generator is that it should provide a sawtooth waveform—unlike the sinusoidal waveform generated by the types of oscillator so far considered—of the type shown at the output in Fig. 5.14. The reason for this, to take the line timebase as our example, is that the waveform is needed to deflect the beam linearly across the face of the picture tube with a subsequent much faster retrace period—blanked out—at the end of each line when the beam returns to the other side of the screen ready to scan out the next line. The usual way of obtaining this sawtooth type of waveform is through the slow charge and rapid discharge of a capacitor, since capacitors charge through a series resistor in an exponential fashion that approximates sufficiently closely the required sawtooth waveform. The

timebase oscillator is actually used as an electronic switch to control the charge and discharge periods of the charging capacitor.

Blocking oscillator

A simple transistor-blocking oscillator circuit is shown in Fig. 5.14, C2 being the charging capacitor across which the output waveform is generated. With the transistor Q1 cut off (non-

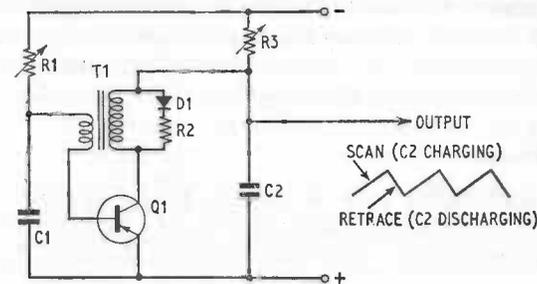


Fig. 5.14. The blocking oscillator, a common timebase waveform generator. The waveform is developed across C2, which charges when Q1 is held non-conducting and discharges when Q1 briefly conducts. The repetition frequency of the circuit is determined by the time constant of C1, R1.

conducting) because of the bias at its base represented by the charge on C1, the charging capacitor C2 charges via R3 to provide the forward scanning portion of the output waveform. The charge on C1 gradually leaks away via R1 and the point is reached where Q1 conducts. This effectively short-circuits C2 so that it discharges rapidly via Q1 giving the retrace or flyback portion of the output waveform. When Q1 conducts, feedback via the base winding of the tightly coupled transformer T1 rapidly drives Q1 into saturation, i.e. maximum collector current. Since there is no further change in the current through T1 the field in it collapses, and the resulting reverse voltage across the base winding charges C1 sufficiently to cut Q1 off. This charge holds Q1 cut off or blocked until the charge on it has leaked away again sufficiently for the whole sequence of operations to be repeated. The

time constant of the network R_1, C_1 determines the frequency at which the half-cycles of conduction occur, so that making R_1 variable provides a hold control. The value of R_3 determines the amplitude of the charge established across C_2 , and thus the amplitude of the output waveform, and can be made variable to provide a height control (width is generally controlled in the output stage). The oscillator can be synchronised by feeding a train of sync pulses to the base to initiate the conduction periods of Q_1 . The diode across T_1 primary is used to prevent excessive negative excursions of the collector voltage when the field established in the transformer collapses on Q_1 being driven into saturation.

Multivibrator

The multivibrator may be used in much the same way to control the charge and discharge periods of a capacitor to obtain a sawtooth waveform. The basic circuit is shown in Fig. 5.15,

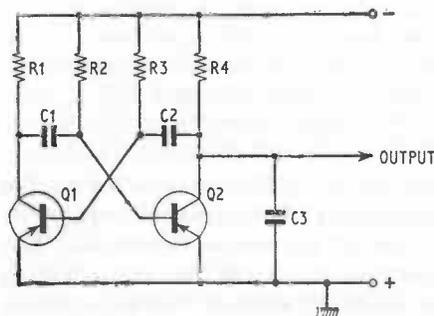


Fig. 5.15. The basic multivibrator circuit, used as a timebase waveform generator. The waveform is again developed as a result of C_3 charging and discharging. The multivibrator consists of two RC coupled amplifier stages cross-coupled—from collector to base—to each other.

with C_3 the charging capacitor across which the sawtooth waveform is generated. The circuit, as can be seen, consists of two resistance-capacitance coupled common-emitter amplifier stages cross-coupled, from collector to base, to each other. The principle is that the transistors conduct alternately. Q_2 controls the charging and discharging of C_3 as in the previous arrangement: when it conducts, C_3 discharges and when it is cut off C_3 charges via R_4 , which again can be made variable to control the amplitude

of the scanning waveform. The time constants of the cross-coupling networks determine the time during which each transistor conducts. To provide the required scan/retrace periods of the output waveform, they are adjusted so that Q_1 conducts about 90% of the time and Q_2 about 10% of the time. Making R_3 variable is a convenient way of providing a hold control to vary the frequency of the output waveform and once again the circuit can be synchronised by applying a train of sync pulses to one of the base or collector circuits to initiate changeover of conduction from one transistor to the other.

The operation of the circuit is as follows. Assume that Q_1 has just started to conduct. Its collector voltage falls, i.e. goes more positive with the *pnp* transistor circuit shown, and a positive potential appears at Q_2 base as a result of the capacitive coupling between Q_1 collector and Q_2 base, cutting Q_2 off. Q_2 collector voltage rises, i.e. goes more negative, and a negative potential appears at Q_1 base. In this way Q_1 is very rapidly driven into saturation, i.e. the condition in which it passes maximum collector current, and Q_2 is held cut off. C_1 then charges negatively via R_2 , in an exponential manner, and at a point determined by the time constant of C_1, R_2 , the base of Q_2 is sufficiently negative for it to begin to conduct. The sequence of events is then reversed, with Q_2 rapidly driven into saturation and Q_1 cut off. The cycles of operation continue, with Q_1 and Q_2 switching on and off alternately at a frequency determined, as we have seen, by the time constants of the cross-coupling networks.

ELECTRONIC CIRCUITS

THE last two chapters have covered most of the circuits used in domestic electronic equipment of the entertainment variety—radio receivers, record-players and so on. There is, however, much wider scope for the use of transistors and it is the purpose of this chapter to describe some of the commonly used transistor circuits in other branches of electronics.

Astable multivibrator

At the end of the last chapter the operation of the basic multivibrator circuit was described in connection with its use in generating a sawtooth waveform. It is, however, basically a square wave generator. If the charging capacitor is omitted (C_3 , Fig. 5.15) the output from either collector, is, as shown in Fig. 6.1, a square waveform or series of pulses. Assuming that the output as before is taken from Q2 collector, and that the supply voltage for the collectors is -6 V , then when Q2 is cut off the voltage at its collector will rise to (almost) -6 V , while when it is conducting it will fall to (almost) 0 V . Because Q2 still has some resistance when conducting its collector voltage will not fall to 0 V , while the presence of R_4 means that when cut off its collector voltage will not rise to the full supply voltage. Also, since the transistors do not switch on and off instantaneously the waveform pulses will not in practice have the perfectly perpendicular sides shown in Fig. 6.1. The output is nevertheless substantially square and suitable for many purposes in electronics.

The circuit continues, as we saw in the last chapter, to switch from one condition, e.g. Q1 on and Q2 off, to the other, Q1 off

and Q2 on, indefinitely so long as the power supply is connected. For this reason it is termed 'astable', i.e. it has no stable condition in which it will remain. The mark-space times are determined by the time constants of the RC cross-coupling networks, and consequently it is possible to obtain a wide variety of pulse widths (mark times) by selecting coupling components of different values. Pulse amplitude is determined by the value of R_4 .

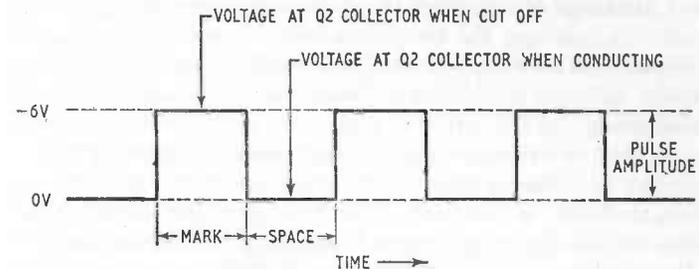


Fig. 6.1. The output from the collector of an astable symmetrical (i.e. $R_1 = R_4$, $R_2 = R_3$, $C_1 = C_2$ in Fig. 5.15) multivibrator is a square wave as shown here when the charging capacitor shown in Fig. 5.15 (C_3) is omitted. This is an ideal waveform to show the principle: in practice, the sides of the waveform would not be perfectly vertical nor the top of the pulses perfectly flat.

The astable, or free-running, multivibrator is one of a whole family of similar circuits all based on cross-coupling between two stages. Three other important ones will be described: the bi-stable and monostable multivibrators and the Schmitt trigger. Before passing on to them, however, it is worth noting that the blocking oscillator circuit described at the end of the previous chapter is also astable and is also used as a pulse generator in electronics, with the omission of the sawtooth charging network R_3 , C_2 (Fig. 5.13). Because of the speed at which the transistor in this circuit switches off after beginning to conduct, however, there are limitations to the range of possible mark-space ratios obtainable from it.

Bistable circuit

The basic bistable multivibrator circuit is shown in Fig. 6.2 (a). As can be seen, it differs from the astable multivibrator in that the cross-couplings are resistive—R1, R2 and R3 on one side, and R4, R5, R6 on the other—and a standing positive bias (assuming *pnp* transistors as shown) is applied to the bases. Thus we have entirely resistive cross-couplings, with no charge and discharge of capacitors to influence the operation of the circuit. In this case the transistors, when conducting or cut off, influence the potential distribution across the potential divider networks, and once the circuit has been placed in one condition, Q1 conducting and Q2 cut off or vice versa, no further change takes place until an external signal—in the form of a triggering pulse—appears to initiate a changeover. For this reason an input circuit, as shown, is provided. The bistable circuit takes its name from the fact that it has two stable states, i.e. it is stable either with Q1 conducting and Q2 cut off or with Q2 conducting and Q1 cut off.

Assume that Q1 is cut off and Q2 conducting. As Q1 is cut off its collector voltage will rise almost to the (negative) collector supply potential. As a result of the coupling from Q1 collector to Q2 base via R2, Q2 base will be sufficiently negative to overcome the positive bias applied via R3. Thus Q2 will be held 'on'. Its collector voltage will fall, i.e. be positive going, and Q1 is thus held cut off by the positive bias applied to its base via R6. This situation remains stable. Suppose now that either (a) a positive-going pulse is applied via C2 to Q2 (which is conducting) base, or (b) a negative-going pulse is applied via C1 to Q1 (which is cut off) base. Provided that the pulses are of sufficient amplitude, in case (a) Q2 will cut off or in case (b) Q1 will be switched on. The resulting alteration to the collector voltages and base biasing will mean that the new situation, with Q1 now on and Q2 cut off, will remain stable until a further input pulse arrives to alter the situation back again.

Let us now look at the output side of the circuit. The output is taken as shown from Q2 collector (an output may also be ob-

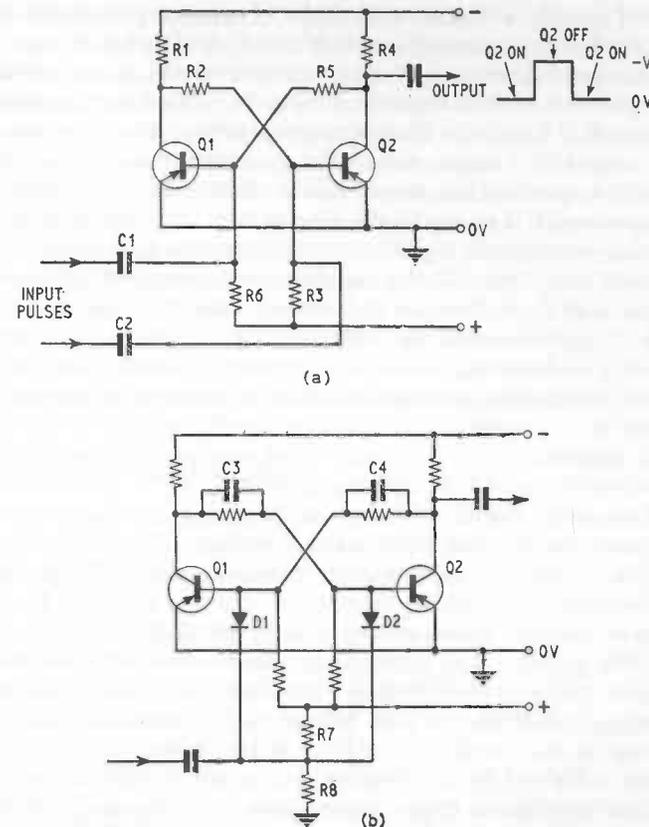


Fig. 6.2. (a) The basic bistable cross-coupled circuit. (b) In practice, the modifications shown here are generally found, with the input pulses applied via steering diodes (D1, D2) and speed-up capacitors (C3, C4) added in the cross-coupling networks. Bias for the steering diodes is often derived from the collectors of the associated transistors instead of, as shown here, a separate bias network R7, R8. In that case, polarity of the diodes is reverse of that shown, and both diodes are reverse biased, one much more heavily than the other. Also, collectors are generally clamped as in Fig. 6.7.

tained from Q1 collector, and will be of opposite polarity to that obtained at Q2 collector) and is of the form shown. As can be seen, when Q2 switches off the output (across Q2) alters, moving from 0 V to $-V$ in this case, and when it switches on the output reverts to 0 V again. These changes represent the generation at the output of a single pulse. But two input pulses have been needed to produce this output pulse. In this lies the usefulness of the circuit: it counts by dividing by two. The circuit is used in vast quantities in digital counting applications, in which connection it is often called a flip-flop since two input pulses are required to flip it from one stable state to the other and then flop it back again in order to obtain one output pulse. The term flip-flop is, however, also commonly used to refer to the monostable circuit, the operation of which it rather more accurately describes; because of this the term flip-flop should be viewed with suspicion until it is quite clear whether the bistable or monostable circuit is actually being referred to.

In practice, two modifications to the circuit just described are generally used to give improved performance. These are shown in Fig. 6.2 (b). The diode input arrangement (D1, D2) is used to improve the switching time when, as would normally be the case in practice, pulses of one polarity are used to trigger the bistable circuit. The arrangement shown is for negative-going trigger pulses applied to *pnp* transistors. The biasing of the diodes depends on the base voltage of the transistors and the voltage at the junction R7, R8: it is here arranged so that the diode connected to the transistor that is 'off' is forward biased, so that the negative trigger pulse passes to its base and switches it 'on', while the diode connected to the transistor that is 'on' is reverse biased, blocking the input pulses. The input pulse is thus steered to the base of the transistor to be triggered next. For positive-going input pulses the polarity of the diodes is reversed and R7 connected to a point at negative potential. The positive trigger pulse will then be steered to the 'on' transistor to switch it 'off'.

The other modification is the inclusion of two small value capacitors C3 and C4 across the cross-coupling resistors. These

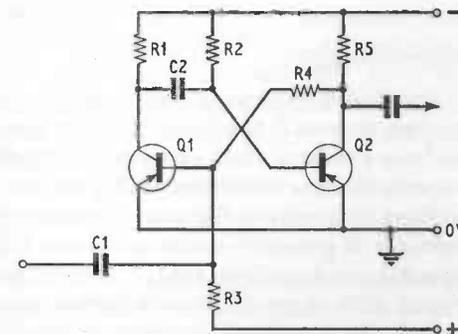
are of insufficient value to switch the transistors on and off by charging and discharging as in the case of the astable circuit: what they achieve is to communicate more rapidly to the base of one transistor the change of collector voltage at the other when it changes from cut off to conduction and vice versa. They are called speed-up capacitors for obvious reasons.

A 'domestic' application of the bistable multivibrator is found in decoders for PAL colour television receivers. These require a phase reversing system on alternate lines. Square wave outputs of opposite polarity are taken from each collector of the bistable multivibrator and used to control the alternate line phase reversing system in the decoder, the bistable multivibrator being switched from one stable state to the other by a series of pulses at line frequency derived from the receiver's line timebase.

Monostable multivibrator

The monostable multivibrator is a combination of the astable and bistable configurations, as shown in Fig. 6.3, and has one stable and one unstable state. A trigger pulse is used to trigger

Fig. 6.3. Basic monostable circuit. As can be seen, one side follows the astable configuration and the other the bistable configuration. After being triggered the circuit reverts to its initial stable state at a time determined by the time constant of R1, C2, R2.



the circuit from the stable to the unstable state, and after a time determined by the time constant of the single RC coupling network the circuit reverts to its stable condition.

The stable state is with Q1 cut off by the positive bias applied to its base via R3 and Q2 conducting because of the forward, i.e. negative, bias applied to its base across R2. A negative-going trigger pulse applied to Q1 base via C1 will result in Q1 conducting. Its collector voltage falls, and a positive voltage is applied to Q2 base as a result of the coupling capacitor C2, cutting Q2 off. Q2 collector voltage rises, i.e. goes more negative, and as a result of the cross-coupling via R4 the voltage at Q1 base is negative, holding Q1 on. This is the unstable condition. In this condition C2 charges negatively via R2 until the point is reached, determined by the time constant C2, R2, when Q2 base is negative again and it conducts once more. Its collector voltage falls, i.e. becomes less negative, and as a result of this the bias at Q1 base is positive again and Q1 is cut off. We are then back at the stable state which continues until the arrival at Q1 base of a further trigger pulse. The result of all this is that a square-wave output is generated at Q2 collector from a short trigger pulse applied to Q1 base, the width of the square-wave pulse being determined by the time constant C2, R2. The circuit can also be used to introduce a time delay equal to the time constant C2, R2 in a system.

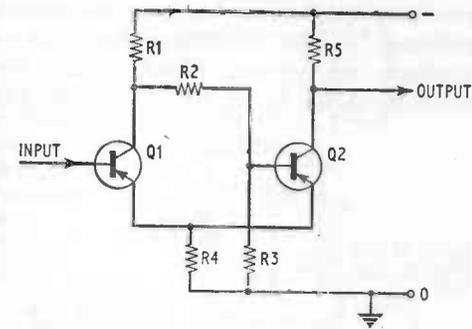
Schmitt trigger

The fourth commonly used member of this family of cross-coupled circuits is the Schmitt trigger shown in Fig. 6.4. Here we have a resistor chain cross-coupling network, R1, R2 and R3, between the two transistors, and a further coupling between the emitters by means of the shared emitter resistor R4 (a speed-up capacitor is generally connected across R2, as with the bistable circuit, but is not shown here). This is basically another bistable circuit, with at any time one transistor conducting and the other cut off. Its function is, however, rather different from that of the bistable multivibrator previously described.

Suppose that we feed a sinewave input to Q1 base. When this is negative-going, Q1 conducts (assuming *pnp* transistors as shown). The voltage across Q1 drops, while that across R1 and

R4 increases. This means that Q1 collector will become more positive with respect to the negative side of the supply line, making Q2 base, via R2, more positive as well, while Q1 emitter is going more negative with respect to the positive side of the supply, carrying Q2 emitter voltage with it because of the common

Fig. 6.4. The basic Schmitt trigger circuit. This cross-coupled circuit has a single resistive cross-coupling from the collector of one transistor (Q1) to the base of the other (Q2), and emitter coupling via the common emitter resistor R4.



emitter resistor R4. The result of this is a substantial reverse bias at Q2 base-emitter junction so that Q2 is quickly switched 'hard off'. The sinewave input to Q1 is, however, a little later positive-going, cutting Q1 off. Q2 biasing arrangement then works in the opposite sense so that it is suddenly provided with a substantial forward bias and rapidly switches fully on. The result of all this is that at Q2 collector a good square wave output is obtained by feeding to Q1 base a sinewave, or, alternatively, a square waveform of poor shape.

Transistor switches

In the circuits we have been discussing the transistors act as switches, changing from cut off to fully on and subsequently reverting to cut off. And, in fact, transistors are widely used as switches, their cut off condition being the equivalent of an open switch, while when fully conducting they are the equivalent of a closed switch. In this way a small signal applied to the base so that the transistor changes from cut off to fully on, or vice versa,

controls a much larger flow of current through the transistor from emitter to collector. The transistor as a 'static' or 'contactless' switch is in this way widely used today in many applications where electro-mechanical relays were previously used, i.e. for sequential control, data routing and such applications.

One difference worth noting between the use of transistors in switching applications and their use as amplifiers is that when switched to fully conducting both the emitter and collector junctions are forward biased. Take, for example, the *pn*p transistor shown in Fig. 6.5. The base is biased negatively with

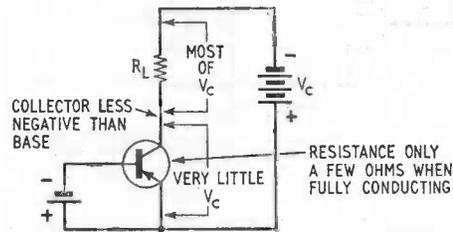


Fig. 6.5. When fully conducting both junctions of a switching transistor are forward biased, as shown here. In this condition the transistor is said to be 'bottomed'.

respect to the emitter to forward bias the emitter junction, and the usual negative collector supply provided. When the transistor is fully conductive its resistance falls to a very small value—a few ohms—so that, again viewing the transistor and its load as a potential divider, almost the entire collector supply voltage appears across the load R_L . This means that the transistor collector is only very slightly negative—less negative in fact than the base. Thus the collector is actually positive with respect to the base and the collector junction is also forward biased.

Relay operation

In addition to replacing relays in some applications, in others transistors are used to increase relay sensitivity or speed of operation. Take, for example, the case where it is desired to use a relay but only a very small control signal, insufficient itself to energise the relay coil, is available. A transistor may be used as

shown in Fig. 6.6 with the relay coil as its load and the control signal applied to its base. The control signal switches the transistor on, and the resultant collector current is sufficient to energise the relay coil. As in the case of the blocking oscillator

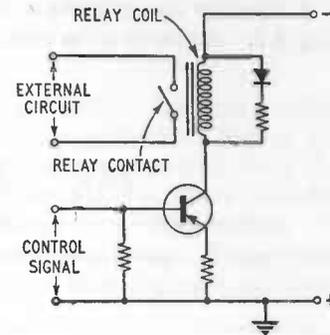


Fig. 6.6. With the relay's energising coil used as the transistor's collector load a simple transistor amplifier can be used as shown here to increase relay sensitivity.

circuit, it is usual to connect, as shown, a diode across the coil to overcome the potentially damaging voltage peak that occurs when the field in the relay coil collapses on Q1 cutting off and the relay de-energising.

Logic circuits

A very extensive application of transistors as switches is in computer logic circuits, where they are widely used in conjunction with bistable flip-flop circuits. A logic circuit is one used to 'recognise' changes in its input conditions and provide an appropriate output. The general scheme is that the circuit has several input terminals and a single output terminal. There are two basic logic functions that such an arrangement can provide, termed the AND and OR functions. A circuit which provides a change at its output only when a signal appears at all its input terminals provides the AND function, while one which produces a change at its output when a signal appears at any one of its input terminals provides the OR function. When signal in-

version occurs in the circuit, as will be the case if it includes a single common-emitter transistor, the circuits are termed NAND and NOR circuits. These logic circuits provide gating, i.e. signal routing, and are used to gate the inputs and outputs of bistable counting arrangements.

A common arrangement is the transistor NOR circuit shown in Fig. 6.7. A change at its output occurs when a signal appears at

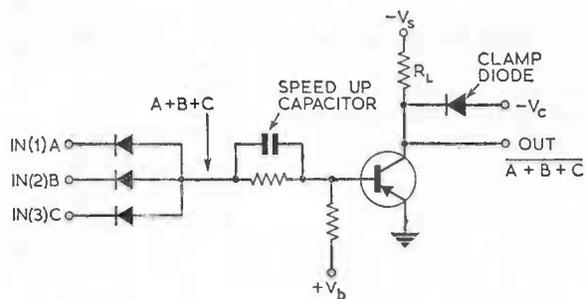


Fig. 6.7. A very common transistor logic circuit, the NOR gate, which is widely used in computers and automatic control/static switching applications. The circuit is shown using a germanium pnp transistor. In practice, for this type of application silicon npn transistors are generally used today, with reverse supply voltage and diode polarity.

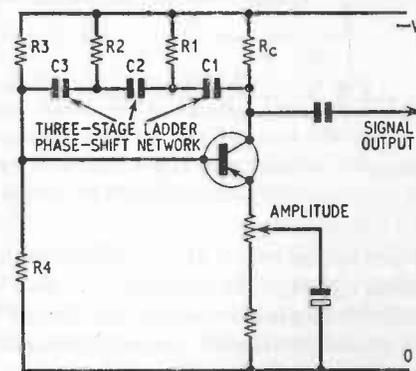
one of its input terminals, and because of the signal inversion it is said to be a NOR gate. The NAND circuit is much the same except for the base biasing network. The inputs are applied via diodes which isolate the inputs from one another, and a speed-up capacitor is used in the same way as with the bistable and other cross-coupled switching circuits previously described. The clamp diode is used to clamp the collector voltage at a certain level (the voltage $-V_c$). At this voltage the diode is forward biased and conducts, preventing further change in collector voltage. This commonly used practice is again intended to increase the speed at which the circuit operates. With diode

input and a transistor providing gain and inversion, we have what is called a diode-transistor logic circuit. Similar results can be obtained using resistors or transistors as the input elements, or by arranging the circuit in other alternative ways, giving rise to such terms as resistor-transistor logic, transistor-transistor logic, emitter-coupled logic and so on.

Sinewave oscillators

We have seen in recent pages the transistor acting as a pulse, square-wave and sawtooth waveform generator. Earlier we saw, in Chapter 3 (see Fig. 3.10 and accompanying text) the transistor operating as a sinewave generator or oscillator, with the sinewave oscillations produced across a tuned collector load circuit and feedback to the base to sustain oscillation. This type of sinewave oscillator—the LC oscillator—is widely used and here

Fig. 6.8. The basic RC sinewave oscillator. Feedback from collector to base is via the three-stage phase-shift ladder network C1 R1, C2 R2, C3 R3. Each RC pair has the same values and contributes a phase shift of 60° so that at one frequency only, depending on the values of C and R, positive feedback is obtained and oscillation occurs.



are many variations of it. One important use of the principle is the self-oscillating mixer which was described in Chapter 5 (Fig. 5.1). There are, however, a number of other approaches to sinewave oscillator design. Three fairly widely used circuits are the RC oscillator, the Wien bridge oscillator and the crystal-controlled oscillator. The two first are used mainly for the generation of low-frequency sinewave signals.

A simple RC oscillator circuit is shown in Fig. 6.8. R_c is the collector load resistor and feedback from collector to base is via a series of RC networks— R_1, C_1 ; R_2, C_2 ; and R_3, C_3 . Since there is 180° phase change between the base and collector of a common-emitter amplifier stage, for the feedback to be positive the feedback networks must provide a 180° phase change between the collector and base. This they will do at a frequency deter-

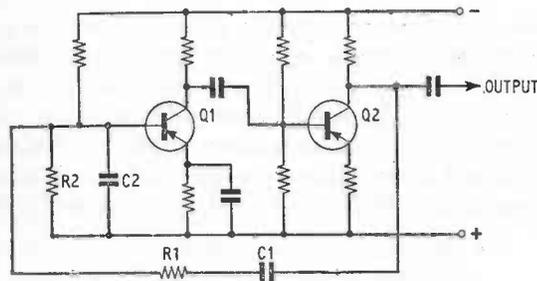


Fig. 6.9. The Wien bridge oscillator. Here the feedback is from the collector of the second transistor Q_2 to the base of the first Q_1 via the Wien network R_1, C_1 and R_2, C_2 . Positive feedback again occurs at one frequency only, depending on the values of R and C .

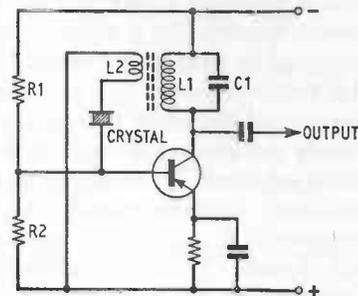
mined by the values of the RC networks, each RC network providing a phase shift of about 60° . At the frequency at which the feedback is positive, i.e. shifted by 180° in this way, oscillation occurs. R_4 with R_3 provide the initial base bias so that the transistor conducts. The RC oscillator provides an output that is relatively free of harmonics, and the output frequency can be made variable by the use of ganged capacitors in the RC feedback networks.

The Wien bridge oscillator (Fig. 6.9) uses two common-emitter transistors with feedback from the collector of the second to the base of the first via a Wien network comprising the series RC network R_1, C_1 and the parallel RC network R_2, C_2 . The values of R_1, C_1 and R_2, C_2 are the same. At one frequency only there

will be no phase change across the parallel RC network R_2, C_2 . At this frequency the feedback from Q_2 collector to Q_1 base, the 180° phase inversion in the first stage being compensated by the 180° inversion in the second, will thus be positive and oscillation will occur.

A simple crystal-controlled oscillator is shown in Fig. 6.10. It can be seen that this is simply a modification of the basic LC oscillator circuit shown in Fig. 3.10. The principle here is that

Fig. 6.10. A simple form of crystal-controlled oscillator. This is basically an LC -tuned oscillator. The crystal oscillates at one frequency only, its 'natural' or 'resonant' frequency, and at this frequency there is feedback from the oscillator 'tank' circuit (L_1, C_1) in the collector lead via L_2 and the crystal to the base of the transistor.



the crystal oscillates at one frequency only, its 'natural' resonant frequency, and at this frequency only there is feedback from the collector tuned circuit L_1, C_1 via L_2 and the crystal to the base of the transistor. While the simple LC oscillator is suitable for most purposes where a sinewave oscillator is required in electronics, it is nevertheless not all that precise due to the effects of heat, component tolerances and change in component values with age. The inclusion of a crystal in the circuit greatly improves the stability in the frequency of oscillation. Note that in this circuit the feedback is in parallel with the base bias network R_1, R_2 , and because of this R_2 cannot be decoupled (the signal at the base would otherwise be short-circuited by the decoupling capacitor). An alternative approach is to connect the crystal in the base-emitter circuit, with positive feedback to sustain oscillation provided by the collector-base capacitance of the transistor. Among other applications crystal-controlled oscillators are used

as the reference oscillator in colour television receiver decoders, where a very stable oscillator is required.

D.C. amplifiers

A number of direct coupled amplifier circuits have been described earlier in this book, mainly in connection with the amplification of a.c. signals. Additional problems arise when the signal to be amplified—as is the case in many applications in electronics—is a relatively slowly varying or d.c. one. Such a signal of course requires direct coupling between stages since a coupling capacitor or transformer would block it. But any change in the d.c. bias conditions—and, as we have seen, such changes do occur because of the effect of heat and also because of variations in supply voltages—will result in a spurious change in the signal being amplified. This change in signal level due to heat or other unwanted effects is termed *drift*, and the major concern of the designer of d.c. amplifiers is its reduction to the minimum possible figure. There are various techniques for doing this in use, including, as we have seen, the use of d.c. feedback. Since silicon transistors are less affected by heat than germanium ones, they are preferred for use in d.c. amplifiers. Another possibility is the use of manual resetting to a zero level, though this is obviously of use in only a limited number of applications. The inclusion of zener diodes in the biasing networks is a further common technique (see Chapter 7) in simple d.c. amplifiers. A more fundamental approach, however, is the use of an amplifier circuit that provides automatic drift rejection. Such a circuit is the differential amplifier which is very widely used, especially as an 'operational amplifier' in analogue computers.

Differential amplifier

The basic differential amplifier circuit—which is also known as the long-tailed or emitter-coupled pair—is shown in Fig. 6.11 (a) and consists of two transistors sharing a common emitter resistor R_E . With the transistors biased on equally by the equal base bias

networks R_1 , R_2 , a steady current will flow through R_E and divide equally between the two transistors and their two equal load resistors R_L . There will thus be no signal at the output since the voltage at each collector will be the same. Suppose now that due to the effect of heat the conduction of the two transistors increases. Several factors will tend to make the increase in conduction of the two transistors equal: first, the fact that the current

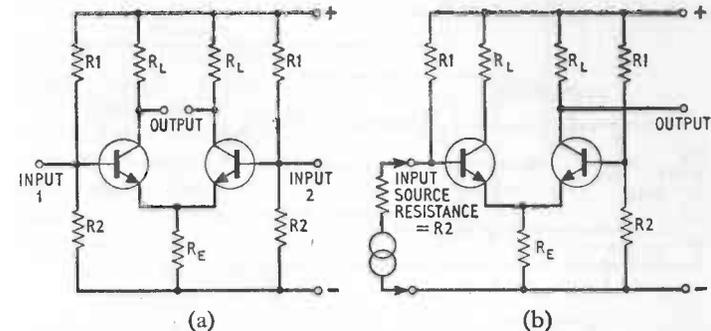


Fig. 6.11. Basic differential amplifier circuits: (a) 'push-pull' version, (b) single-ended input and output version.

flows through the same emitter resistor will mean that the emitter voltages move together; secondly, if the transistors are matched and thus have the same characteristics the change in current flowing through them will be equal; and thirdly, if they are mounted together and subject to the same heat change then they will be equally affected. This means that the change in voltage at the two collectors will be the same so that there will be no change of signal at the output. If, on the other hand, the signal to be amplified is applied across the two input terminals, then as the current through one transistor increases that through the other will fall, providing an output that is an amplified version of the difference between the potential at the two input terminals—hence the name differential amplifier. The input may alternatively, as shown in Fig. 6.11 (b), be applied between one input

terminal and the 'earth' side of the supply line (negative here since *npn* transistors are shown) if one of the base-emitter resistors is omitted and the source resistance of the input signal is the same in value as the other base-emitter resistor. The output may also be taken from one of the collectors as shown at (b).

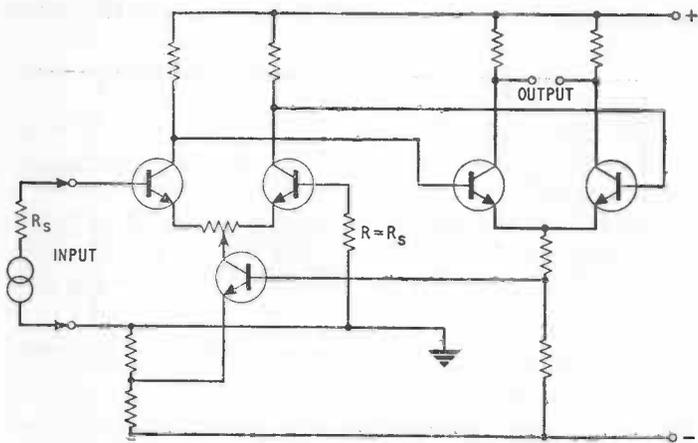


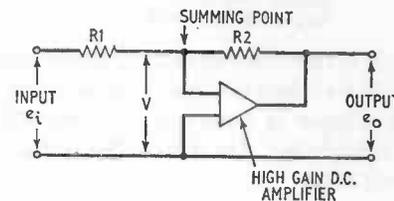
Fig. 6.12. Two-stage differential amplifier circuit with feedback from the emitter circuit of the second stage to the emitter circuit of the first to give improved immunity from drift. The feedback controls the series stabiliser transistor in the emitter circuit of the first stage. For a description of the operation of series stabiliser circuits see Chapter 7.

This circuit results in negligible drift. There is still, however, greater drift than is desirable for many applications, for example, strain gauge amplifiers, meter amplifiers and computer operational amplifiers, mainly because of the difficulty of precisely matching the transistors, and a number of techniques have been used to reduce drift further. These include the use of complementary pairs of *npn* and *pnp* transistors in each side of the differential amplifier, and the use, as shown in Fig. 6.12, of a two-stage differential amplifier with d.c. feedback from the emitters of the second

stage to the emitters of the first via a transistor stabiliser circuit. The principle of voltage stabilising circuits is described in the following chapter.

As previously mentioned, one important application of differential d.c. amplifiers is as operational amplifiers to perform mathematical operations such as integration, differentiation, etc. in analogue computers, and indeed the term operational amplifier is commonly used today for this class of amplifier. The basic operational amplifier configuration is shown in Fig. 6.13, from

Fig. 6.13. Basic operational amplifier configuration. This consists of a high-gain d.c. amplifier, e.g. differential amplifier, with series and feedback impedances— R_1 and R_2 respectively here—chosen to provide the required 'operation'.



which it will be seen that the input is applied via a series resistor and a feedback resistor is used. Provided that the gain of the amplifier is sufficiently high, its input and output voltages will be related as follows: $e_o/e_i = -R_2/R_1$, the minus sign indicating the signal inversion that occurs in a single stage common-emitter amplifier. To perform different mathematical operations, various combinations of resistive, reactive and non-linear series and feedback networks may be used. The junction of the series and feedback networks is called the summing point.

Chopper amplifier

An alternative approach to the problem of drift in d.c. amplifiers is to chop the signal to be amplified into a square-wave signal, amplify this in an a.c. amplifier, and then convert the output again to a proportional d.c. signal. Such an arrangement is termed a chopper amplifier. A transistor may be used, once again as a switch, to perform the chopping operation. A typical circuit is

shown in Fig. 6.14. The d.c. signal is applied across the collector and emitter of the chopper transistor Q1, and a square-wave signal is applied to its base to switch it on and off. The result of

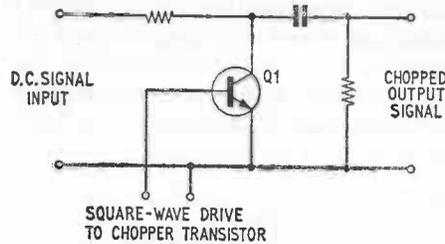


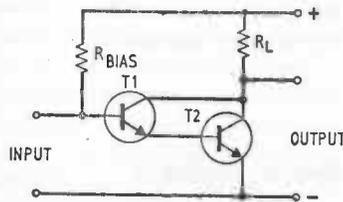
Fig. 6.14. Transistor chopper to provide a 'square wave' output proportional to the d.c. input. The chopped output signal can be amplified by an a.c. amplifier.

this is that the transistor short-circuits the d.c. signal input when it is switched on, and a square-wave output proportional to the d.c. input is obtained. A transistor may similarly be used to 'demodulate' the signal after it has been passed through the a.c. amplifier.

Darlington pair

Finally, a direct coupled transistor stage that is very widely used for both a.c. and d.c. amplification is the Darlington pair configuration, also known as the super-alpha pair, double-emitter

Fig. 6.15. The widely used Darlington pair configuration. In this two transistors are direct coupled, the base of the second being fed from the emitter of the first. The output is equal to the gain of the first transistor multiplied by the gain of the second one. The load may be connected in the collector or emitter lead of the second transistor.



follower and compound-connected pair. The circuit consists of a pair of transistors, often in fact supplied as an integrated pack, with the two collectors connected together and with the base of the

second transistor (see Fig. 6.15) fed from the emitter of the first. The input signal and bias is applied to the base of the first transistor, and the output may be taken as shown from a load resistor in the common collector circuit or from a load resistor in the emitter of the second transistor. This configuration provides a current gain equal to the product of the current gains of the two transistors and an input impedance that is far higher than that of a single common-emitter transistor amplifier stage.

POWER SUPPLIES

WHEN the d.c. power required for equipment is derived from the mains instead of from a battery it must be rectified. This process in its simplest form was outlined in Chapter 1 (see Fig. 1.10), using a junction diode to suppress one half of the input a.c. waveform and thereby obtaining an output in the form of a series of pulses as shown in Fig. 1.10 (c). For obvious reasons, this is termed half-wave rectification. Clearly, however, further measures must be taken to obtain a steady d.c. supply.

Practical half-wave rectifier power supply circuit

A practical half-wave rectifier circuit is shown in Fig. 7.1 (a). The input to the rectifier is applied via a transformer T1, which provides a convenient way of stepping down the 240 V of the a.c. mains supply to the order of voltage required by transistor equipment. The circuit is completed by the load resistor R_L , shown dotted, which represents the circuits supplied by the rectifier circuit. The addition of a reservoir capacitor, as it is called, C_R , across the rectifier output goes a long way to providing the steady d.c. output we require. When the rectifier, D, conducts, C_R charges; when D is cut off on the succeeding negative-going portion of the a.c. input waveform C_R discharges providing the output current during this time. The effect is illustrated at (b). Since C_R loses only a part of the charge it holds when the rectifier is cut off, the only time it fully charges is when the equipment is first switched on. This means that a fairly steady output is obtained across C_R though, as shown, there is still a slight 'ripple' left. This is largely removed by the filter network R1, C1 so that

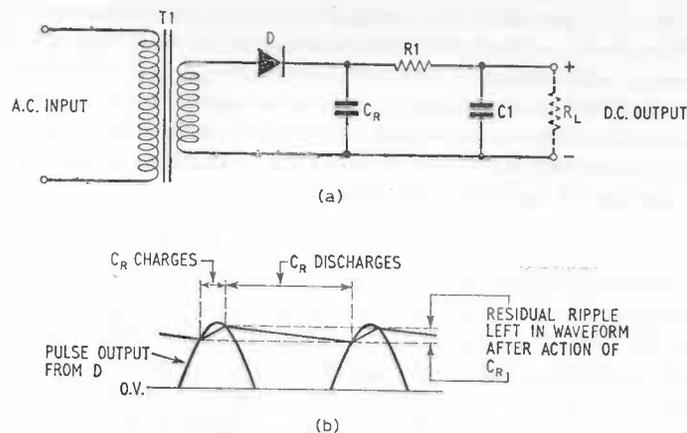


Fig. 7.1. (a) Basic half-wave rectifier circuit fed from a transformer, with reservoir capacitor (C_R) and RC filter circuit R1, C1 to smooth out the rectified a.c. (b) Waveforms showing the pulse output from the rectifier D and the action of the reservoir capacitor.

negligible ripple exists in the d.c. across the load. In many circuits an inductor is used in place of the resistor R1.

Full-wave rectification

In the arrangement just described, use is made of only half the input waveform. Slightly more elaborate circuits enable both halves to be used. Two methods of obtaining full-wave rectification are shown in Fig. 7.2: (a) makes use of an input transformer with centre-tapped secondary winding to feed two rectifier elements D1 and D2; (b) makes use of a bridge arrangement of rectifier elements. In practice D1, D2 in (a) and D1-D4 in (b) are generally physically a single component. Metal rectifier elements are generally used in these circuits.

Because of the centre-tapping of the secondary winding in (a), equal but opposite polarity voltages will appear across points A-B and B-C. This means that if A is positive to B on one half

of the a.c. input waveform so that D1 conducts, on the succeeding half of the a.c. input waveform C will be positive with respect to B and D2 will conduct. This will result in use being made of both halves of the a.c. input, and in a smoother output since variations across C_R will occur at twice the frequency and the current drawn by the load each time C_R discharges will be half what it would be in the case of half-wave rectification.

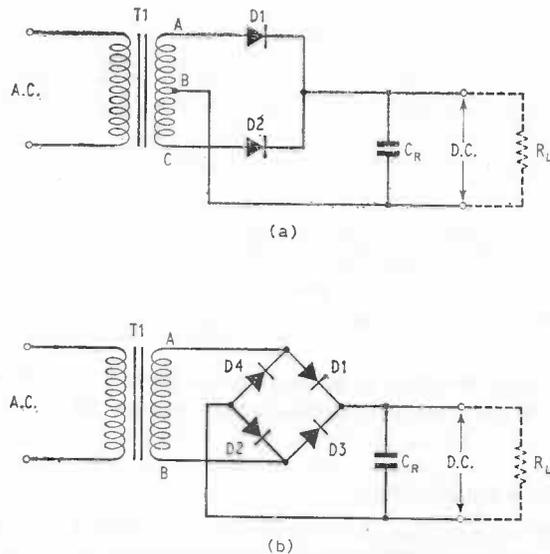


Fig. 7.2. The two basic forms of full-wave rectifier.
(a) Single-phase circuit. (b) Bridge rectifier circuit.

The bridge rectifier does not require a transformer with centre-tapped secondary winding. The operation of this arrangement is that when A is positive with respect to B—see Fig. 7.2 (b)—on one half of the a.c. input D1 and D2 will conduct, current flowing through T1 secondary winding, D1, the load R_L and D2, while when B is positive with respect to A on the following half-cycle of the a.c. input D3 and D4 will conduct, current flowing through

T1 secondary, D3, the load R_L and D4. This circuit is very commonly used in transistor equipment power supply circuits. It has the advantage that at any time two rectifier elements are in series with the supply so that each needs to have only half the voltage rating that would be necessary when only two rectifier elements are used.

Voltage regulation

A problem that arises in practice is that the load does not draw a constant current as this current must flow through the resistance of the supply circuit components this means that there will be variations in the supply voltage. This is not a problem with simple items of equipment. In more elaborate equipment with more stringent requirements, however, it may be necessary to make provision to compensate for this variation. There is also the problem that the mains supply itself is not constant but varies slightly. There are three principal ways of going about providing voltage stabilisation: the use of a voltage-dependent resistor; the use of a zener diode; or the use of some form of feedback stabilisation circuit.

Voltage dependent resistors (v.d.r.s) are generally used to provide voltage stabilisation in particular stages or parts of an equipment. If a v.d.r. is connected across the d.c. supply to be stabilised, variations in the supply will alter the resistance of the v.d.r. It will then pass a greater or smaller current to compensate for the supply fluctuation, stabilising the supply voltage. A v.d.r. will reduce a 10% supply voltage variation to a variation of about 3%.

The zener diode is a convenient way of providing stabilisation against fairly small variations in load current. The principle was mentioned briefly in Chapter 2, and is based on the fact that the zener diode operates in the region beyond the reverse breakdown voltage, where small changes in reverse voltage produce relatively large variations in reverse current (see Fig. 7.3). The zener diode is connected across the load, whose supply voltage it is to stabilise, in the reverse sense, i.e. so that with a slight fall in supply voltage

produced by increase in load current the reverse bias of the zener diode is decreased and the reverse current it passes falls. This decrease in the current it passes compensates for the increase in the load current, and since the supply voltage depends on the sum

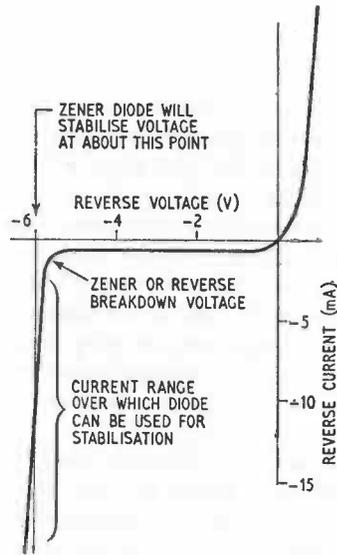


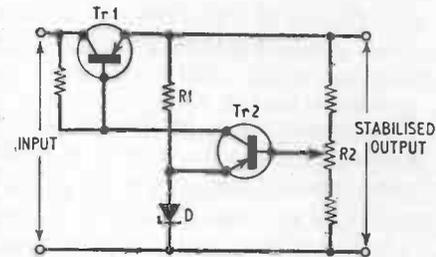
Fig. 7.3. Zener diode characteristic. The zener diode has a low reverse breakdown or zener voltage and if biased into this region can be used to stabilise voltage against variations in current.

of these currents, the variations in which tend to cancel out, the supply voltage is held constant. Likewise, of course, an increase in supply voltage will increase the zener diode reverse bias so that it passes a greater reverse current to bring the supply voltage down again.

The feedback series stabiliser circuit shown in Fig. 7.4 is commonly used where it is necessary to provide stabilisation against larger variations in load current. The feedback is between the emitter and base of the series transistor Tr1, which must of course be a power-type transistor. Variations in the voltage at its emitter are applied to the base of the d.c. amplifier

stage Tr2 via R2 and used to control the base biasing of Tr1. The zener diode D, biased by R1 into the required portion of its reverse characteristic, is used as a reference source to stabilise Tr2 emitter voltage, and R2 can be preset to give the required

Fig. 7.4. Simple transistor series stabiliser circuit. The series transistor Tr1 provides a stabilised output under the control of the voltage applied to its base via the voltage divider network and d.c. amplifier stage Tr2.



stabilised output voltage from the circuit. A change in the voltage at Tr1 junction is used by the feedback circuit to alter Tr1 base bias so that it passes a greater or smaller current. A common variation of the circuit shown is the use of a compound-connected pair of transistors in position Tr1.

INTEGRATED CIRCUITS

PLANAR transistors, as outlined in Chapter 2, are made in large numbers on a thin slice or wafer of single-crystal semiconductor silicon—typically 300 on a wafer of 0.8 in. diameter—which is then divided to obtain the individual planar transistor structures. The technique involves multiple diffusion to fabricate the various p and n regions, with surface protection after each diffusion followed by etching to define the areas to be subject to the next diffusion. Following diffusion, surface deposition is used at the appropriate positions to form ohmic (i.e. not rectifying junction) contacts to the emitter, base and collector regions. Clearly it is logical to take the process a stage further and form several transistors together on a slice with interconnections between them so that they will operate as a complete stage such as a bistable flip-flop, or as several stages so that we fabricate say a complete amplifier with d.c. coupling throughout on a single semiconductor slice. This is, in fact, how integrated circuit production, which is to some extent an extension of silicon planar transistor technology, commenced. Once again, a single wafer will yield a number of such circuits which can then be separated and provided with individual mountings.

An integrated circuit (i.c.) can thus be defined as a circuit capable of providing some electronic function such as amplification that is fabricated as a single indivisible structure rather than, as with conventional hand-wired or printed circuits, an assembly of separate components connected together, generally by soldered joints. The integrated circuit is then mounted and provided with lead-out wires to allow its use in equipment as a complete module or building block; it should be noted that integrated

circuits are often used in conjunction with other integrated circuits and/or additional discrete components.

Thin and thick film circuits

A different i.c. technique is the deposition of *thin films* on a suitable insulating and supporting substrate by vacuum, furnace, gas reaction or plating techniques. Thin films are usually defined as being about 1 micron thick. A further technology is the use of *thick films*, based on the deposition of viscous pastes or inks through a fine screen with subsequent firing so that the films become integrated with the supporting substrate. These thick films are of the order of 25 microns thick. Though it is possible to form thin film transistors by depositing layers of semiconductor material to produce unipolar devices, in practice both thin and thick film techniques are restricted at present to passive component networks—i.e. networks of resistors and capacitors with their interconnections. A special form of transistor, called a flip-chip transistor, is often used in conjunction with thin and thick film i.c.s. These transistors have contact pads on one face instead of connecting leads, this technique enabling them to be easily bonded to thin or thick film circuits. Flip-chip capacitors are also produced since there is a limit to the capacitance values that can be conveniently produced on a flat surface. A combination of thin film and semiconductor i.c. techniques enables the *active* devices—transistors and diodes—to be fabricated in the semiconductor wafer with the *passive* components and interconnections provided by thin film techniques.

Microelectronics

All these techniques come under the heading of *microelectronics*, which has been defined as techniques resulting in component densities greater than 50 per cubic inch. The initial motivation in producing microminiaturised equipment was primarily that of reducing the size and bulk of electronic equipment for space and military use. However, considerable success with these

techniques has now led to their being pursued on grounds of reliability, economy and convenience.

Semiconductor integrated circuits

Semiconductor integrated circuits fall into two main categories. In *multi-chip* or *chip* form, the active and passive components are produced on separate wafers of semiconductor material, and then assembled in a single package; on the other hand, the *monolithic* s.i.c. is entirely fashioned within a single slice of semiconductor material with all components, devices, interconnections and isolations between stages formed within one tiny wafer of the semiconductor. In practice, the monolithic s.i.c. is the most widely used, and is sometimes termed a *solid circuit*. Isolation may be achieved by using the high resistance represented by reverse biased *pn* junctions; resistors by the diffusion of impurities into a defined region; and capacitors by using the depletion capacitance of a *pn* junction.

Complete circuits can thus be fabricated by various planar diffusion and epitaxial processes, and since these techniques are easier with silicon this is normally the semiconductor material used. The only addition then needed is the leads used to connect the s.i.c. into circuit.

Integrated circuits are often referred to by the type of package in which they are enclosed. This may be a small TO-5 transistor-type metal can or the corresponding lower-cost plastics (epoxy) encapsulation, a small rectangular package termed a *flat-pack*, or the *dual-in line* pack in which the leads emerge from the two sides of a rectangular package similar to the flat-pack but are all turned in one direction.

An integrated circuit is normally a complete and unchangeable circuit module, having a finite number of connection leads (commonly up to about 22), so that there is a limited number of uses to which any one i.c. can be put. Since i.c. fabrication processes lend themselves best to the large-scale production of large numbers of identical units, most i.c.s are basic logic circuits—gates and flip-flops and so on—for digital operation in computers and

other data processing equipment for which a limited range of circuits is required in very large numbers. There are, however, also a number of devices intended for *linear* operation, that is for handling a.c. waveforms instead of the 'on-off' pulse trains used in digital equipment.

Circuit design

Because there are limitations in the range of component values, and also in component tolerance ratings, that can be produced by integrated techniques, a rather different approach to circuit design is required with integrated circuits than is the case for circuit design using conventional components. Since, however, integrated circuits are required to perform the same functions as the circuits described elsewhere in this book, their operating principles are much the same.

In the manufacture of a s.i.c., for example, a high value resistor or capacitor would require the utilisation of far more of the basic slice area of the semiconductor material than would a small signal transistor or diode. This means that it may be preferable to use several diodes or transistors if this allows the elimination of a single resistor or capacitor, unlike conventional circuit design where the active devices are normally of higher unit cost than fixed resistors or capacitors. Typically, it may be worth using say five transistors instead of one 47-k resistor, and capacitors impose an even heavier penalty in slice area.

In practice, the monolithic s.i.c.s marketed at the present time have a *limit* of roughly 20 k in resistor values and 200 pF in capacitor values; but many circuits are designed around passive components of much lower value. A considerably wider range of values and narrower tolerances are practicable with thin-film, thick-film and multi-chip circuits, but even here there are constraints which must be observed for economical design.

Advantages and disadvantages of integrated circuits

In general, integrated circuits can offer the following advantages over circuits based on conventional discrete components: smaller size and lower weight; low power requirements; low cost per function; high gains and low phase shifts; wide frequency response. The disadvantages include unwanted 'parasitic' couplings between components and the formation of parasitic transistors and diodes; limited range of resistor and capacitor values; low precision in resistor values and their high temperature coefficients; lack of inductors.

A common form of amplifier circuit used in these devices is the differential or operational amplifier which is particularly attractive because of the matched nature of transistors formed on the same silicon chip.

A problem with microelectronics is the dissipation of the heat generated by many circuits in a tiny volume, despite the low power levels of semiconductor circuits. There exists a requirement to reduce power levels still further (micropowers) if full advantage is to be taken of the extremely small unit volume of components.

SERVICING TRANSISTORISED EQUIPMENT

TRANSISTORS themselves are robust and if operated under correct conditions are seldom the cause of faults in transistorised equipment, though they may of course fail due to the breakdown of an associated component. Far more common than transistor failures are mechanical faults such as unsatisfactory switch contacts, 'tired' spring contacts in jack sockets, noisy volume controls, components being knocked off printed boards, badly soldered joints, short-circuits between adjacent wiring, broken aerial leads and cracked ferrite-rod aerials. The most common electronic faults are: battery failure, resistors going high-value, and capacitors becoming leaky. The first thing to check in battery-powered equipment should always be the battery.

Battery test

A simple method of testing the battery is to measure the voltage across its terminals (call this V_1), next to place a 100-ohm resistor across the terminals and measure the voltage across this (call this V_2), and then to calculate the internal resistance of the battery from the simple formula

$$\frac{V_1 - V_2}{V_2} \times 100 = \text{battery internal resistance (in ohms).}$$

A good battery will have an internal resistance of about 10 ohms. A battery with an internal resistance of up to 50 ohms is workable, but if the internal resistance is higher the voltage drop across it will be too great and the battery should be replaced.

Power supply check

Not all faulty battery conditions will be revealed by this simple test, however, so a further battery voltage check should be made with the battery on load, i.e. connected to the equipment and switched on. Faults that may be caused by a faulty battery and will not be revealed by the previous test include instability, distortion, and failure of the mixer oscillator section in radio receivers to oscillate, especially at higher signal frequencies. A second check should be made after the set has been working for a few minutes.

A further check that can be made is the equipment's current consumption. Connect a milliammeter in series with the collector supply line, i.e. in the case of *pnp* transistors, connect the meter negative lead to the negative terminal of the battery and the positive lead to the negative collector supply line. A convenient way of making this check is to connect the meter across the on/off switch (collector supply side) and to short across the 'earthy' side of the supply where a double-pole on/off switch is used. On first connecting up an above normal current will flow due to the various decoupling capacitors charging, so start off with the meter switched to a higher range than you would expect, then switch to a lower range. A reasonable current reading will indicate that the power supply and biasing networks are in order; no current suggests an open-circuit at some point in these circuits.

Initial checks

Some very simple checks can be made to help speedily locate a fault. In the case of a 'dead' set, for example, listen first to the loudspeaker on switching on. A slight audible thump will indicate that the set is taking power and that the loudspeaker and associated circuitry are in order. Then listen for background noise. Absence of this indicates a fault in the audio stages. Where background noise is heard and this can be increased by advancing the volume control, the audio stages are probably in order. This volume control check is most helpful in giving an

immediate indication as to whether the fault lies in the audio or r.f./i.f. sections of the set. More can be learnt by noting whether the background noise varies as the tuning control is altered. If it does the fault is probably in the aerial input tuned circuit, r.f. stage (if present) or mixer circuitry. If the background noise does not vary with alteration of the tuning control setting the fault is likely to be in the detector, i.f. or mixer stages.

Types of fault

Apart from 'no results', the main classes of fault are: (1) distortion; (2) a.f. instability (motor-boating); (3) i.f. instability (howling); (4) lack of sensitivity.

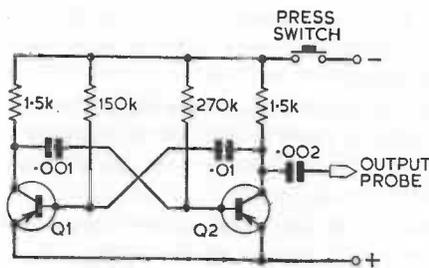
Distortion is generally due to a fault in the audio stages (assuming that the supply circuit is in order), due for example to a change in the bias conditions. Other possibilities are a faulty detector or a fault in the a.g.c. circuitry. A.F. instability is generally caused by unwanted feedback in the audio sections because of a faulty decoupling capacitor in the power supply line or biasing arrangements. I.F. instability is also generally due to faulty decoupling. Other possibilities are faulty capacitors, resistors or transistors in the i.f. stages, including the neutralising feedback networks if used, or a fault in the mixer stage. Low sensitivity is generally brought about by a faulty biasing component. If this is in the audio side, distortion will be present as well. Lack of sensitivity at one end of the tuning scale indicates a fault in the supply or in the mixer stage.

Fault location

The above notes should give some idea of where the fault may be. Follow this by visual examination of the set to see whether there are any obvious mechanical faults, dry joints, misplaced wiring/components, dirt causing short-circuiting on printed-circuit boards, cracks in the printed boards or wiring, or components that have obviously become defective through being subject to excess heat, i.e. are passing excessive current. Note in this

connection that in operation the transistors should—except for the output stage—not be warm, so that a warm transistor or hot output transistor indicates that there is a fault at this point. A damaged transistor can be very hot indeed, however, so care is needed in checking this point.

If the fault cannot be traced by these simple methods, a stage-by-stage examination must be undertaken. A simple method is



Supply voltage approx. 4.2V

Fig. 9.1. Simple transistor signal injector. The circuit comprises a multivibrator providing an output rich in harmonics extending up to r.f. When used for testing transistor circuits a resistor of about 47 k should be connected in series with the output probe to limit the output to avoid injecting signals that may be sufficient to damage the junctions of the transistors in the equipment being tested.

to use a low-voltage signal injector to check back from the output to the input, preferably with a resistor of about 47 k in series with the probe to limit the injector output to a safe level. A number of signal injectors is available; alternatively one can easily be made up. Fig. 9.1 shows a simple circuit that could be used. As can be seen, it consists of an astable multivibrator. Almost any small-signal type of transistor may be used in this circuit, and the whole unit can be built into an old pen-holder using three small mercury cells to provide the power. The circuit oscillates at audio frequency, but since its output is rich in harmonics which extend well into the r.f. range it can be used to inject signal 'noise' in both the a.f. and r.f./i.f. sections of a receiver. Check back through the equipment from the loudspeaker to the input,

i.e. aerial input or pickup: in this way the faulty stage is quickly revealed.

An alternative technique to signal injection is signal tracing. Again a number of instruments is available, and in this case the technique is to proceed from the input and move stage-by-stage to the output, checking of course not only the transistor inputs and outputs but the coupling networks as well. A signal tracer for a.f. may consist simply of a pair of headphones, with maybe a stage of amplification built in; for i.f./r.f. testing a diode is needed to detect the signal.

Without a signal injector or tracer the procedure is to use a high-resistance voltmeter (20,000 ohms/V or more) to check voltages stage by stage. A simple common-emitter amplifier stage is shown in Fig. 9.2, with voltage check points indicated.

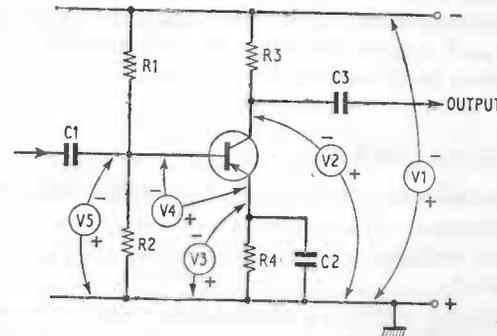


Fig. 9.2. Typical RC amplifier stage showing voltage measurements that can be made in systematically checking the stage.

Note the polarities of the meter leads indicated, for use as shown with supply and bias arrangements for a *pn*p transistor. In the case of an *np*n transistor the polarities are all reversed. First check the supply voltage V_1 : if this is much lower than the nominal battery voltage, either the battery, the supply line or its decoupling components, or a biasing network across the supply,

is at fault. The next step is to check that emitter current is flowing. This can be done by measuring the emitter voltage V_3 and working out the current with the aid of Ohm's Law. Say the emitter voltage is 1 V and the emitter bias resistor R4 is 1 k, then the emitter current in amperes is 1/1000 or 1 mA. If the emitter current is correct for the stage being tested, then almost certainly the stage is working correctly. If not, then further checks should be made. Measure the base voltage V_5 : this should be about 100 to 300 mV higher than V_3 in the case of a germanium transistor, or 600 to 700 mV higher in the case of a silicon one. If this is found to be the case, then the input circuit of the stage is in order. If not check the bias network R1, R2 and the coupling capacitor C1.

These checks for emitter and base-emitter voltages will not help with most output stages, however, since very low resistance emitter bias resistors are used. To check that collector current is flowing in such a stage disconnect the collector lead and connect an ammeter in series with it.

Detailed stage check

Once the faulty stage has been found, a fuller examination of it can be undertaken by checking the collector, base, emitter and base-emitter voltages. These, see Fig. 9.2, are V_2 , V_5 , V_3 and V_4 respectively.

If the collector voltage V_2 is found to be roughly the same as the supply voltage V_1 , and the emitter voltage is zero, collector current is not flowing. Then if the base voltage V_5 is very low or zero, either R1 is open-circuit or R2 short-circuit. If the base voltage is correct, R4 may be open-circuit. If R3 is open-circuit, V_2 and V_3 will both be low.

The collector current may alternatively be lower or higher than it should be. If the base-emitter voltage V_4 is found to be low, then the collector current will be low because there is insufficient emitter junction forward bias. This may be because the base voltage V_5 is too low or the emitter voltage V_3 too high. Too low base voltage may be due to R1 being high-value or R2 low-

value. Too high emitter voltage would be caused by R4 going high-value.

If the base-emitter voltage is too high then excessive collector current will flow. This means that either the base voltage V_5 is too high or the emitter voltage V_3 too low. Too great base voltage would be due to R1 being low-value, R2 high-value, or C1 leaky. Too low emitter voltage would be due to R4 being low-value or C2 leaky. If C3 is leaky, V_2 will be reduced and the following stage over-biased.

Note that in direct-coupled stages the d.c. conditions in a particular stage will be affected by the d.c. conditions in the stages to which it is coupled, and that direct coupling may continue over several stages with d.c. feedback paths also having to be taken into account.

Transistor tests

The transistors themselves may be tested by means of an ohmmeter—though the ohmmeter must have a testing voltage of less than 1.5 V and the equipment must be switched off before the ohmmeter is connected. More reliable checks are made with the transistors taken out of circuit, since otherwise the resistance

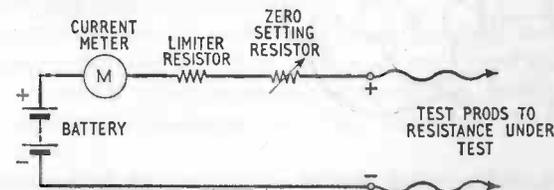


Fig. 9.3. Common ohmmeter arrangement.

readings obtained will be affected by the biasing networks in the circuit.

A simple ohmmeter circuit is shown in Fig. 9.3 (not all ohmmeters follow this configuration). The test prods are connected across the resistance to be measured. Current drawn from the battery then flows through the resistance under test, the zero

setting and limiter resistors, and the milli- or micro-ammeter M which is calibrated to give direct resistance readings. Note that a low-resistance reading is indicated by meter pointer deflection to the right, and a high-resistance reading by meter pointer deflection to the left. The zero setting preset resistor is used to adjust the meter reading to show zero ohms when the prods are short-circuited. The principle, therefore, is that the resistance being checked is supplied with current from the internal battery, and the meter is calibrated to show resistance readings which follow from the output voltage and the current passed by the resistance under test in accordance with Ohm's Law. The test prods are generally, as shown, given polarity markings which accord with the polarity of the battery, though this is not always the case.

The principle of transistor testing with an ohmmeter is that the meter's supply is used to reverse or forward bias the transistor junctions, and the junction resistance noted. The possible tests are shown in Fig. 9.4. Tests 1 and 2 forward and reverse bias the

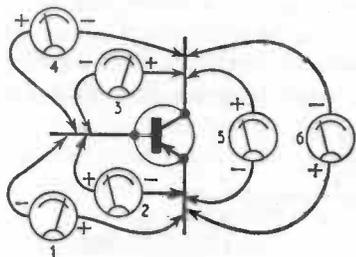


Fig. 9.4. Ohmmeter checks on a transistor.

emitter junction, which should result as shown in low and high resistance readings. Likewise tests 3 and 4 forward and reverse bias the collector junction, which should again result in low- and high-resistance readings. In the case of tests 5 and 6, since there are here two junctions back-to-back, high resistance readings should be obtained in both cases. With the meter connected as in test 6 and a resistor of about 220 k connected from the collector to base to forward bias the emitter junction, a fall in the meter resistance reading should be obtained indicating the passage of collector

current and the ability of the transistor to provide gain. Note that as in similar cases elsewhere all polarities shown would be reversed in the case of an *npn* transistor.

When using an ohmmeter to measure resistor values in circuit, note that misleading results can be obtained due to the meter output voltage forward biasing an associated transistor, e.g. when checking resistors in a transistor base circuit.

Precautions when servicing transistor equipment

While transistors are robust enough when correctly operated, they can nevertheless be easily damaged if mishandled. The following precautions are most important.

Make sure that the soldering iron and all test gear in use are connected to a good common earth.

Switch the equipment off or disconnect the battery before removing any component from the circuit or making any disconnection.

Before replacing a transistor, check the associated components in the stage.

Always switch off or disconnect the battery before soldering, and use a heatsink between the transistor and the joint: a pair of pliers gripping the lead is suitable.

Do not make resistance measurements or continuity tests with an ohmmeter that has an output voltage greater than 1.5 V.

Check transistor connections before connecting in circuit so that the replacement transistor—or diode—is correctly connected the first time.

It is most important when connecting the supply to observe correct polarity.

Make sure you do not short-circuit transistor leads—this can easily be done when using a screwdriver or probe. Do not short the base connection to chassis.

Do not bend transistor leads nearer than $\frac{3}{16}$ in. from the seal.

Remember that transistors are light sensitive and protected by an opaque coating which must not be damaged.

Do not short across the loudspeakers leads, or carry out tests with a lower resistance speaker load than correct.

Ensure that the operating conditions and circuit stability are such that thermal runaway cannot occur under the most adverse conditions likely to be encountered.

When fitting audio output transistors—and stabilising diodes—use a silicone grease and make sure that good thermal contact is made between the device and heatsink. Burrs and thickening at the edges of chassis holes must be removed and output transistors firmly bolted down on a flat surface.

Always replace mica insulating washers and bushes when refitting audio power devices.

Do not use carbon tetrachloride or trichloroethylene alone for cleaning switch contacts or potentiometers: these cleaning agents tend to be corrosive.

INDEX

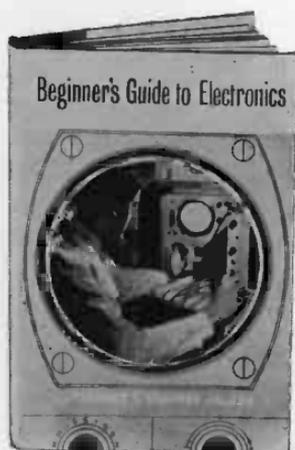
- A.G.C., 97, 98, 99, 106, 113
 Acceptor impurity, 15
 Aerial input circuit, 89, 93, 94, 99
 Alloy junction, 35
 Alloy junction transistor, 34, 35
 Alloy diffused transistor, 34, 38
 Amplification, 29, 32, 49, 50, 60, 63, Chs. 4 and 5, 132
 Amplitude modulation, 89, 97
 Astable multivibrator, 118
 Avalanche effect, 25
 Bandpass characteristic, 94
 Bandsreading, 103
 Base, 26, 28
 Battery test, 149
 Bias, 20, 57
 Bias stabilisation, 57, 61, 65, 76, 77, 79
 Bipolar transistors, 41
 Bistable circuit, 120
 Blocking oscillator, 114, 115, 119
 Bootstrapping, 74
 Bottoming, 126
 Bypass capacitors, 63, 66
 Capacitance, junction, 26
 Channel, 42
 Chassis line, 103
 Chopper amplifier, 135
 Chopper transistor, 135
 Class A operation, 67
 Class B operation, 68
 Collector, 26, 28
 Collector-bend detector, 99
 Collector clamping, 128
 Collector junction, 26, 28
 Common-base circuit, 29, 49, 54
 Common-collector circuit, 32, 49, 54
 Common-emitter circuit, 30, 49, 54, 55, 57
 Complementary-symmetry stages, 72, 75, 87, 99
 Compound connected transistors, 136
 Compound semiconductors, 48
 Crossover distortion, 69, 77
 Crystal oscillator, 129, 131
 Current transfer ratio, 50
 D.C. amplifiers, 61, 132, 142
 Darlington pair, 136
 Dead receiver, 150
 Decoupling, 63, 66, 98
 De-emphasis, 109
 Demodulation, 89, 94, 97, 107
 Depletion layer, 20
 Depletion-mode operation, 44
 Detection, see *demodulation*
 Differential amplifier, 132
 Diffused junctions, 38
 Digital circuits, 122, 127
 Diodes, 19, 45
 Direct coupling, 61
 Distortion, 56, 58, 68, 69
 Donor impurity, 15
 Doping, 11, 34
 Drain, 42
 Drift, 132
 Dynamic characteristics, 58
 Emitter, 26, 28
 Emitter-follower circuit, 49
 Emitter junction, 26, 28
 Enhancement-mode operation, 44
 Epitaxial junction, 34, 40
 F.M. detection, 107
 F.M. receivers, 104
 Fault finding, 151
 Feedback, see *negative and positive feedback*
 Feedback stabiliser, 142
 Ferrite-rod aerial, 89
 Field-effect transistors (f.e.t.s.), 41
 Flip-chip components, 145
 Flip-flop circuit, 122
 Forward bias, 21, 24
 Frequency conversion, 90, 94, 99, 105, 110, 112
 Full-wave rectification, 139
 Gain, 50
 Gate, 42, 46
 Gating, 128
 Gunn diode, 48
 H.F. bias, 82, 87
 Half-wave rectification, 138
 Heat, effect of, 12
 Heatsinks, 37, 76
 Holes, 12, 15, 16
 Hot carrier diode, 48
 Hybrid parameters, 51
 Impurity, 15, 34
 Input characteristic, 55
 Input impedance, 49
 Input stage, 81, 83
 Integrated circuits, 144 *et seq.*
 Intermediate frequency, 89, 91, 105, 107
 amplification, 94, 107, 113
 Insulated gate f.e.t., 41
 Junction diode, 19, 45
 Junction gate f.e.t., 41
 LC oscillators, 66, 129, 131
 Linear amplification, 32, 56, 58
 Load line, 58
 Logic circuits, 127-9
 Long-tailed pair, 85, 132
 Majority carriers, 17
 Mark-space ratio, 119
 Mesa transistors, 39
 Metal base transistor, 48
 Metal oxide semiconductor transistor (m.o.s.t.), 44
 Microelectronics, 145
 Minority carriers, 17, 18

- Mixer stage, 89, 93, 94, 99, 105, 110, 112
- Modules, 101
- Monolithic s.i.c., 146
- Monostable circuit, 123
- Multivibrators, 114, 116, 118, 152
- Negative feedback, 62, 74, 75, 78, 82, 85, 98
- Neutralisation, 95
- NOR circuit, 128
- n*-type semiconductor material, 15
- Ohmmeter tests, 155
- Operational amplifier, 135
- Oscillators, sine wave, 66, 90, 102, 106, 129
- Output characteristic, 58
- Output impedance, 49
- Pentavalent elements, 15
- Phase shift, 54
- Phase splitter, 71
- Photoconduction, 16
- Phototransistors, 16
- Planar transistors, 34, 39
- Playback amplifier, 82
- pn* junction, 19, 24, 26
- Point-contact diode, 45
- Positive feedback, 66
- Potential barrier, 20
- Power supplies, 98, 103, 138
- Power supply checks, 150
- Power transistors, 37
- p*-type semiconductor material, 16
- Push-pull amplification, 68
- Q* factor, 65
- RC oscillators, 129
- R.F. amplification, 63, 105
- Ratio detector, 107
- Recombination, 13
- Record amplifier, 84
- Rectification, 22, 138
- Relay operation, 126
- Reservoir capacitor, 138
- Reverse bias, 21, 24
- Reverse breakdown voltage, 25
- Reverse leakage current, 25
- Sawtooth waveform generators, 114 *et seq.*
- Schmitt trigger, 124
- Selectivity, 65, 94
- Self-oscillating mixer, 91, 93, 99, 105, 110, 112
- Signal injector, 152
- Signal tracing, 153
- Silicon controlled rectifier, (s.c.r.), 46
- Single-ended push-pull output stage, 70
- Solid circuits, 146
- Source, 42
- Speed-up capacitor, 123, 128
- Square wave generators, 118, 124, 125
- Steering diode, 122
- Step recovery diode, 47
- Super-alpha pair, 136
- Switching circuits, 125
- Tape recorder circuits, 81
- Television receivers, 109
- Thermal resistance, 52
- Thermal runaway, 57
- Thermistors, 48, 69, 77
- Thick film circuits, 145
- Thin film circuits, 145
- Thyristors, 46
- Time constant, 82
- Timebase generators, 114
- Tone control, 79
- Transformerless output stages, 70
- Transistors, basic operation of, 26 *et seq.*, 49 *et seq.*
- Transistor tests, 155
- Transition frequency, 52
- Trivalent elements, 16
- Tuned circuits, 63, 66, 89, 91, 94, 99, 106, 110
- Tunnel diode, 46
- U.H.F. reception, 109
- Unilateralisation, 95
- Unijunction transistor, 44
- Unipolar transistor, 41
- V.H.F./F.M. receivers, 104
- Varactor (variable capacitance) diode, 26
- Voltage checks, 153
- Voltage dependent resistors (v.d.r.s), 48, 141
- Voltage regulation, 141
- Volume control, 80
- Wien bridge oscillator, 129, 130
- Zener diode, 26, 46, 141
- Zener voltage, 25

This book will enable the reader to go on to more specialised works or courses with confidence.

BEGINNER'S GUIDE TO ELECTRONICS

*by Terence L. Squires,
A.M.I.E.R.E.*



**192
pages**
.
**128
line
diagrams**
.
**2nd
Edition**
.
15s

This book provides a "short-cut" for those wishing to obtain a quick acquaintance with modern electronics. It is not intended to replace "text" or "course" books on the subject that take the student through detailed theory and calculations: what it does is to describe as simply as possible the basic concepts of importance in electronic engineering, and the various components used in electronic equipment, so that the reader gains an understanding of the terms used and the practical side of the subject.

from all booksellers

— NEWNES BOOKS —

