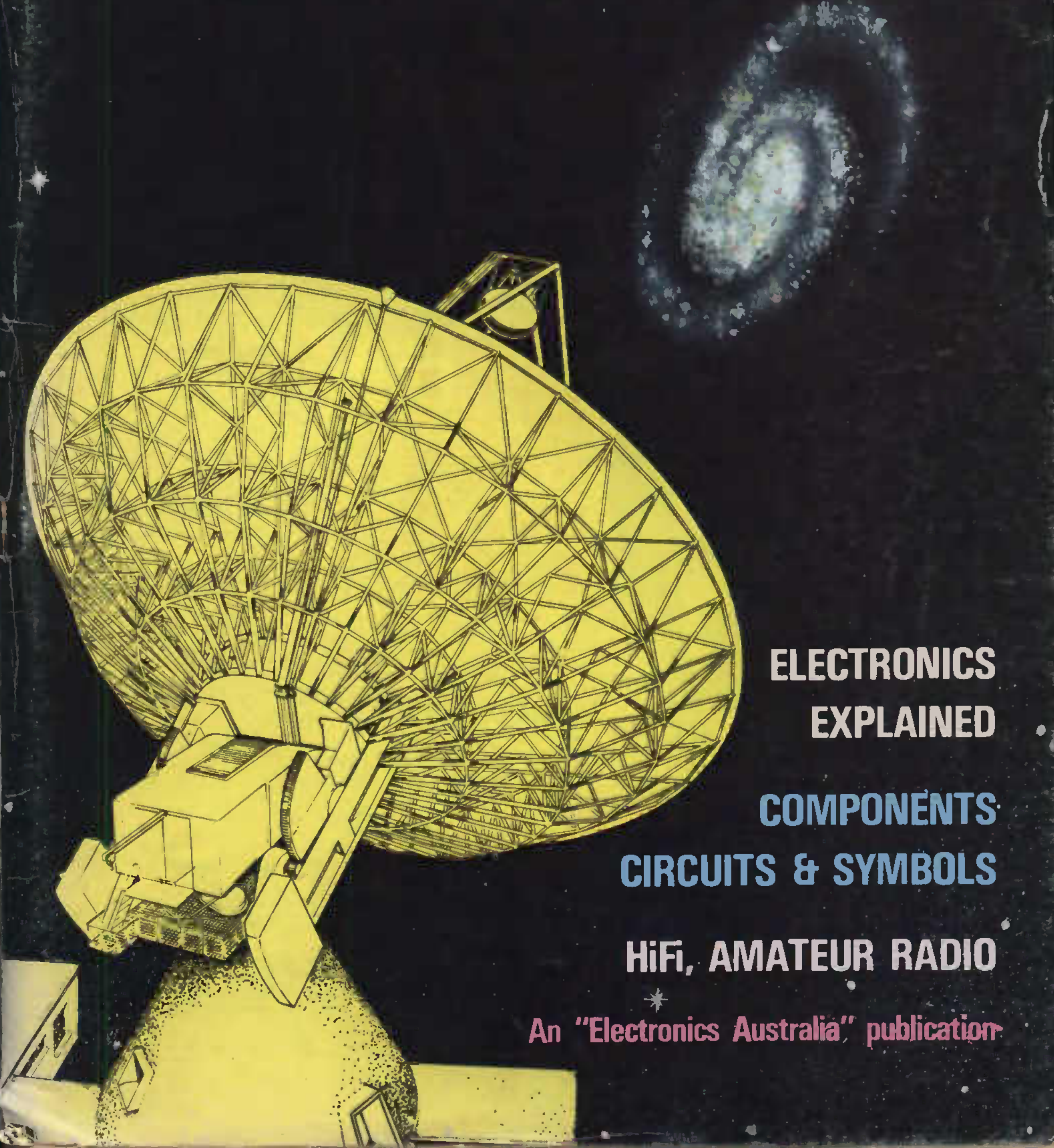
BASIC ELECTRONICS

Fourth edition
Price $2.00



ELECTRONICS EXPLAINED

COMPONENTS CIRCUITS & SYMBOLS

HiFi, AMATEUR RADIO

An "Electronics Australia" publication

BASIC ELECTRONICS

An introduction to electronics, radio, hi-fi

Fourth edition 1973

Co-Authors:

Neville Williams MIREE
Editor-in-chief "Electronics Australia"

Jamieson Rowe BA (Sydney) BSc (Technology NSW)
Editor "Electronics Australia"

Printed by Land Printers Pty Ltd of Lidcombe NSW
for Sungravure Pty Ltd of Regent St, Sydney, Australia.

Silicon Transistors

Choose the best technique to suit the application
☐ Alloy junction
☐ Diffused junction
☐ Mesa
☐ Planar
☐ Planar Epitaxial

Germanium Transistors

☐ Wide range of special and general purpose types.

Silicon Power Diodes

☐ Crest working reverse voltages to 1200V
☐ Average forward currents to 150A.

Power Transistors

☐ Germanium and silicon
☐ Dissipation to 125 watts
☐ Operation to 400 MHz

Integrated Circuits

☐ Digital DTL — CML — TTL (Interchangeable with SN74 Series) and MOS
☐ Linear Op Amps, Voltage Regs., Audio, Receiver types.
☐ DIL, Flat Pack, QUIL and TO-74/99/100
☐ Available to MIL Specs.

Silicon Zener Diodes

☐ Voltage range 3.3V to 75V
☐ Dissipation 400mW to 75W

Small Signal Diodes

☐ Silicon and Germanium
☐ Point contact
☐ Diffused junction
☐ Gold bonded
☐ Planar
☐ Planar Epitaxial
☐ Oxide Passivated

Silicon Controlled Rectifiers

☐ Repetitive peak off-state voltage to 1600V
☐ Crest working reverse voltages to 1600V
☐ Average forward currents to 250A
☐ Maximum junction temperatures to 125°C.

Transmitting Transistors

☐ VHF
☐ UHF
☐ SSB

High Voltage Diodes

with controlled avalanche breakdown characteristics.
☐ Crest working voltages to 150, 000V
☐ Current up to 50mA

PHILIPS

Australia's largest range of semiconductors. The brand - "Miniwatt"

Australia's most experienced manufacturer in Semiconductor techniques, provide for the Industry
☐ Devices suitable for all applications
☐ Comprehensive applications laboratory services
☐ Ready availability via branches in all States
☐ Stringent quality control

PHILIPS
38.2180

ELCOMA
Electronic Components & Materials Division
of Philips Industries Limited
Sydney • Melbourne • Brisbane • Adelaide • Perth • Canberra • Hobart

Preface

Of the many branches of modern science, electronics is almost certainly the fastest growing and the widest in its implications.

It includes radio, television and sound reproduction — the activities which nurtured it — but it extends nowadays into many other areas.

Electronics provides the basis for virtually all modern communication, from the once-simple telephone through to the sophisticated equipment which made possible colour television from the Moon, and which still links us with probes millions of miles out in space.

Electronics has provided the computers which are essential to the control and navigation of space ventures, as well as the control of more prosaic "earth-bound" vehicles, from supersonic jets to driverless trains. At the other extreme, the same computer technology is responsible for tiny electronic calculators, small enough to slip into a coat pocket.

In modern hospitals, electronic equipment monitors the condition of patients undergoing major surgery and, in many cases, provides the surgical "tools" for the operation itself.

Electronics has revolutionised manufacturing processes and methods of measurement; it has become a basic means in virtually every area of scientific research.

All this adds up to virtually unlimited career opportunities. As an electronics engineer, a person may find interest and employment in research situations, in the medical field, in astronomy or space installations, in aviation or other branches of transport, in computing and statistics, manufacturing, communications, light and power reticulation, printing and the graphic arts, entertainment, and so on.

Not everyone, of course, will want to become a fully qualified electronics engineer but, undoubtedly, a person who can think in terms of electronics — even at an elementary level — has an advantage over one to whom the subject is sheer mystery.

This book should serve as an introduction to electronics — we trust an effective one. When you have assimilated the information it contains, you will be in a good position to enlarge upon it by studying other more comprehensive textbooks and by embarking on more advanced constructional projects, with the help of magazines such as "Electronics Australia".

The Authors.

YOUR GREATEST SCOPE FOR MEASUREMENT COMES FROM THE B.W.D. OSCILLOSCOPE RANGE.

b.w.d

GL&C927BWD

509B

Ampl. DC to 7MHz.
10mV 50V/cm.
T.B. 200nS - 1S 7cm. 2Hz - 10MHz trigg

$209

506

Ampl. DC to 15MHz.
5mV - 50V/cm.
T.B. 50nS-10S/cm. 15MHz trigg. Isol grnd.

$304

539A DUAL TRACE

Ampl. DC to 10MHz.
10mV - 50V/cm. T.B.
200nS-2S/cm. Triggering 2Hz-10MHz.

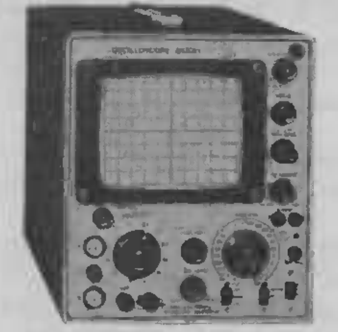
$375

502

Differential Ampl. DC to
3MHz. 20mV to 50V/cm.
T.B. 200nS-1S/cm. Isolated grnd.

$350

511



Differential Ampl. DC to
10MHz at 10mV 5Hz 4MHz
at 1mV T.B. 40nS-10S/cm. Isolated grnd.

$430

1722

17" Large Screen Ampl.
DC - 15KHz X & Y 500μV
to 50V/inch. T.B. 10μS-10S/inch triggered

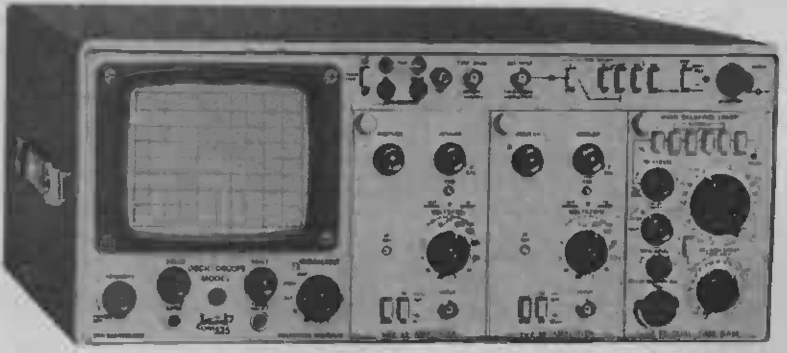
FROM **$975**

521 DUAL BEAM

Ampl. 10μV - 50V/cm.
Bandwidths to 30MHz.
T.B. 40nS - 10S/cm. Delayed sweep.

FROM **$835**

525 DUAL TRACE



Ampl. DC to 50MHz 5mV to 50V/cm. T.B.
10nSec - 10Sec/cm. 70MHz trigger delayed,
sweep, trigger and mixed display. 10KV EHT internal 8 x 10cm. graticule.

$1060

All prices plus sales tax and F.I.S. capital cities Australia (subject to change without notice).

THE MAJORITY OF OUR INSTRUMENTS CARRY THE FEDERAL NATO STOCK NO.

BWD ELECTRONICS PTY. LTD.

DESIGNERS & MANUFACTURERS OF THE FINEST ELECTRONIC INSTRUMENTATION
329-333 Burke Rd., Gardiner. Vic 3146. Ph. 25 4425. 127 Blues Point Rd, North Syd. N.S.W. 2060. Ph. 929 745
S.A. A. J. Ferguson Pty. Ltd. Ph. 51 6895. W.A. Cairns Instrument Services. Ph. 25 3230
Qld. Warburton Franki (Brisbane) Pty. Ltd. Ph. 52 7255 A.C.T. Associated Scientific Sales Pty. Ltd. Ph. 95 913
N.T. J.C. Electronics. Ph. 3605 and from B.W.D. Agents around the world.

WINNER OF AWARD
FOR OUTSTANDING
EXPORT ACHIEVEMENT

Chapter Index

CHAPTER 1

Background to Electronics

Electronics can lay fair claim to being the fastest-growing and most significant branch of science and technology in the present era. But where did electronics spring from? This first chapter sketches in the historical background, tracing events from early speculation, through the telegraph and the wireless, to the very sophisticated science we know today.

Around 500BC the Greek philosopher Thales made reference in his writings to the strange properties of "elektron" — a substance known to us as amber. When polished against a fur garment, elektron seemed to exhibit a mysterious attraction for tiny pieces of chaff or fluff.

The effect was treated largely as a curiosity for over 2000 years and it was not until around 1600AD that an English surgeon, William Gilbert, decided to make a study of the elektron effect. In the course of the study he identified more than twenty other substances which behaved in a similar manner. He made numerous observations about other substances and contrived simple instruments, describing the work in his historic book "De Magnete."

To establish a connection with elektron (amber) and, at the same time, to avoid confusion, Gilbert coined the word "electricity" as a name for the mysterious force of attraction. The amber-like materials he described as "electrics," the remainder as "non-electrics." By so doing he provided the first rudimentary division of materials into insulators and conductors.

Knowledge of electricity built up gradually over the next 300 years and many papers and books have been written setting out the contributions of research workers and experimenters.

Some of the names are better known than others because they have become part of the terminology of electricity. In most cases they relate to the nature of the research which the particular individual undertook. Here are some of them:

COULOMB, Charles 1785 (electric charge and force).
GALVANI, Luigi 1790 (electric current).
VOLTA, Alessandro 1800 (electric cell).
OERSTED, Hans 1819 (magnetic force).
AMPERE, Andre 1822 (force and current).
OHM, George S. 1826 (resistance).
FARADAY, Michael 1833 (electric "particles").
LENZ, Heinrich 1833 (induced electricity).
HENRY, Joseph 1834 (self-induction).
JOULE, James P. 1841 (electricity and heat).
KIRCHHOFF, Gustav 1847 (basic laws).
MAXWELL, James C. 1867 (electromagnetic waves).
HELMHOLTZ, Hermann von 1881 (electric charges).
HERTZ, Heinrich 1888 (radio transmission)

Particularly significant research was carried out by Joseph Thompson (1897) and Ernest Rutherford (1911) which substantially resolved long-standing arguments as to whether an electric current was a "fluid" flow or the movement of discrete particles — the so-called "aetherial" and "corpuscular" theories, respectively.

Not only did Thompson and Rutherford provide convincing support for the corpuscular theory but they were able to demonstrate that the particles involved had to be several orders of magnitude smaller than a hydrogen atom. From this came a new concept of current flow and of the structure of the atom itself: the atom was made up of much smaller particles carrying an electric charge (electrons) and a central positively charged nucleus.

The "electron theory," which was verified and developed by other workers, opened the floodgates of modern research and development. At the risk of oversimplifying matters, it could be suggested that this research has since followed three main courses:

1 — BASIC ELECTRICAL: The direct use of electrical energy for industrial, automotive and domestic purposes. This involved the design and development of power generating systems ranging from hand-size units to huge power stations. It involved the development of supply mains, sub-stations, street and house wiring methods, fusing, metering and so on. For the user's premises it involved the development of lighting, heating, cooling, power-drive and other such equipment. And, in the automotive

At left, a wall mounting Ericsson telephone in use around 1900. At top, an Audion valve available to US experimenters for $4.00 in 1910. Below it, an early type of solid-state detector with a pellet of lead peroxide between electrodes, one plated with platinum, the other with lead.

Reproduced from "Punch" for December 9, 1878, over the heading "Edison's Telephonoscope (Transmits light as well as sound)". Gazing hopefully into the future, the artist explains: "Every evening, before going to bed, Pater and Mater set up an electric camera-obscura and gladden their eyes with the sight of their children at the Antipodes."

area, electrical involvement included engine ignition, the battery, the generator and a variety of electrical ancillaries.

2 – WIRELESS COMMUNICATION: The outcome of observations which confirmed that electrical phenomena could be present in space, as distinct from wired circuits. This led to the development of "wireless" communication and broadcasting (now normally referred to as "radio") and to television. But the technology behind wireless opened up new ways of doing things in just about every field of human endeavour. Nowadays, we tend to group all this technology under the term "electronics" thereby establishing a verbal link with the electron and with the original Greek word elektron.

3 – MATTER & ATOMIC PHYSICS: While some workers devoted themselves to wired circuits (electrical) or to "wireless" circuits (electronic) others directed their energies to discovering more about the atom itself, its internal structure and the nature of matter. From this has come the modern science of atomic physics, atomic fission, fusion, atomic and hydrogen bombs, atomic power generation, and so on.

While these fields support their own research and their own industries there is, in fact, an enormous amount of common ground, with no clear lines of demarcation.

Traditional electronic components and techniques are being adapted freely into electrical environments, making it essential for electrical trainees to acquire a background in electronics.

Again, research workers probing the secrets of the atom and of matter rely very heavily on electronic techniques and instrumentation.

In fact, the science of electronics occupies a more or less central position, involved at one fringe with traditional electrical services and technology and, at the other, with research of the most advanced kind. We shall have more to say about this later on.

As we indicated earlier, the science of electronics has largely grown out of work aimed initially at achieving communication by electrical means, and especially without the use of intervening wires. In short, electronics has largely grown out of "wireless" and the most helpful thing we can do is probably to trace the course of events from the pioneering days to the present.

Electrical communication involving wires has a very long history, with the application pre-dating by far an understanding of the true nature of electricity.

As far back as 1753, a letter in "The Scot's Magazine," Edinburgh, over the initials C.M. proposed a form of electric telegraph with as many wires as there were letters of the alphabet. By applying electricity in sequence to particular wires, balls attached to the far end of the same wires would hopefully attract paper tabs marked with certain letters. By such means a sender could conceivably transmit messages to an observer at a remote point.

While there was a great deal of diffidence about the idea of communicating by electrical means, it was nevertheless pursued by many inventors in many countries. Most of their efforts were interesting but impractical but, from about 1832 onwards electric telegraphy made notable progress under the drive and initiative of men like Morse, Cooke, and Wheatstone. It became the new wonder of the age and an important adjunct to the operation of railways and newspapers.

Telegraph cables overland were followed by the first undersea cable between Dover and Calais, in 1850, by the first trans-Atlantic cable in 1858. It was just a year after the laying of the trans-Atlantic cable that a German schoolteacher Philipp Reis submitted a paper to the journal "Annalen de Physik" proposing a means of transmitting voice by electric telegraph (Reis coined the term "telephon"). He was refused publication on the grounds that voice transmission was clearly impossible. Even after he had produced and lectured on a practical prototype instrument in 1860, publication was still refused on the same grounds!

In due course the telephon – or telephone – was taken up and developed into a practical instrument by Alexander Bell, one of the best known inventors in the history of communications. Since then, continual development and refinement have produced the instrument and the facility that we know today.

ACROSS THE ATLANTIC 1901

Guglielmo Marconi (left) at the age of 27, with his assistant George Kemp from the British Post Office, receiving the first wireless signals across the Atlantic in 1901.

Research and observations on telegraphic apparatus served to confirm convictions that there was a relationship between electrical currents flowing in a metallic circuit and magnetic effects, the attractive qualities of amber, sparking in isolated circuits and atmospheric lightning. In short, electrical phenomena could fairly obviously occur in space as well as in metallic circuits. What was not so apparent was the distance through which such effects could be perceived and the chance of their being harnessed to convey any kind of information.

Michael Faraday had speculated about the existence of "lines of force" in space and the physicist James Maxwell had developed Faraday's theme mathematically, indicating that electrical waves in space were similar in character to light waves — but this was largely theory.

History records the efforts of quite a few contemporaries to solve the riddle of wireless waves, including men like Preece, Dolbear, Lodge and Crookes. However, it was the German Heinrich Hertz (1857-1894) whose work came to be the best known and who seems to have lodged the strongest claim to immortality.

In due course, the famous Italian Guglielmo Marconi (1874-1937) took up the mantle of Hertz' carrying on practical research and gaining a great deal of recognition in so doing.

At the age of 21, while still a student, he transmitted and intercepted wireless waves over a distance of 1200 feet. Two years later (1898) he covered a distance of 18 miles and confirmed Hertz' prediction that wireless waves would pass through solid structures, such as the walls of buildings. In 1901 he bridged the Atlantic with a wireless wave and, about the same time installed the first shipboard wireless equipment.

Within the next 5 or 6 years, a laboratory curiosity and experimenter's dream became the subject of speculation and rivalry on an international scale. Governments saw the military potential of wireless; maritime companies saw it as a link with their ships at sea. To cable companies it was a threat to their monopoly; to others it looked like a chance to break into the communications business.

Wireless had become a fact of life and it was a matter of improving techniques and capabilities and adapting to the realisation that, in terms of communication, no point on earth or at sea was now potentially more than a tenth of a second away!

One of the most significant developments was the invention of the wireless valve. In 1883, while experimenting with his early incandescent lamps, Thomas Edison had noted that current could be made to flow via a metal plate sealed inside a lamp envelope. Fairly obviously it was yet another manifestation of electricity traversing open space but Edison was content to note what seemed to be a purely academic observation.

In 1904 the English physicist J. A. Fleming showed that the combination of a heated filament and an adjacent metal plate could function as a rectifier of alternating current, with the ability to pass current in one direction and prevent current flow in the other. As a rectifier it offered the interesting possibility of being used as a detector of wireless signals.

The Fleming "diode" was important in its own right but the next step was even more so. In work which not everyone accepts as either original or unique, Lee de Forest added another element called a "grid" between the filament and the metal plate in a diode-like valve. This was in 1906.

The "triode" valve made it possible to generate wireless waves without resort to primitive and cumbersome spark gaps and rotating machinery. It could be done by what we would now describe as purely electronic means.

With valves, it was also possible to amplify very weak signals from a microphone and apply them to a transmitter, so that voice and even music transmission became much more practicable.

At the receiving end, valves made it possible to pick up and detect signals which would not have been discernible by other means, and to amplify those signals so that they could be heard clearly through earphones or even loudspeakers.

It would surely be true to say that the triode valve and its derivatives opened the way for wireless to became an everyday facility. In the process of evolution and acceptance, the old term "wireless" fell gradually into disuse and was replaced by "radio." Which word you use tends, in part, to be an indicator of your age group!

With the benefit of hindsight, it is possible

NEWS BY RADIO IN 1898

To the Editor of the "Financial News";
Sir,—

I fear that the recollections of "Midas" in your issue of the 20th instant, regarding the invention of wireless telegraphy, would not be accepted by everyone as strictly accurate. Few matters of fact seem to have excited so much diversity of opinion. In France the majority of people are firmly convinced that wireless telegraphy is a production of French genius, as exemplified in Dr Branly. In England "Midas" thinks the invention should be credited to Sir Oliver Lodge and Sir William Preece. In Russia, I daresay, Partisans of M. Popoff could be found. In Italy I believe that almost unanimously the decision would be in favour of my being the inventor, and I have reason to think that in the United States there is what "Midas" might consider a regrettable tendency to follow Italian opinion. In Sweden, however, which may be considered a neutral country, since no Swede has yet laid claim to be the inventor of wireless telegraphy, The Nobel Prize Committee, which gives its decisions on the strength of an international vote, unfortunately ignored the claims of Sir Oliver Lodge and of many others and made their award in 1909 to me and Professor Braun, of Strasburg. Perhaps "Midas" has never heard of the latter gentleman.

"Midas" is also rather at sea when he states that the first instrument for despatching messages was exhibited by me at Dover Town Hall. This took place in August, 1899, and over a year before, in July, 1898, I reported the Kingstown Regatta's by wireless from Dublin Bay to the Dublin Express. And before that wireless messages had passed between Osborne and the Royal Yacht, and before that again between warships of the Italian Navy.

Yours, etc.,
(Signed) G. MARCONI.
Marconi House, Strand, London,
W.C. 2 July 22, 1921.

to appreciate that the history of the science could have followed a somewhat different path. While some were concentrating on the development and application of valves, a few others were speculating about the possibility of making crystal detectors amplify. It went no further than speculation and, with valve technology forging ahead, the idea of "amplifying crystals" was pushed aside for about thirty years. Had it been otherwise, we might have had transistors decades sooner!

The history of wireless (or radio) in Australia mirrors fairly closely events in other parts of the world.

In 1905, the infant Marconi Company successfully demonstrated transmission across Bass Strait from Queenscliff in Victoria to Davenport in Tasmania, a distance of about 180 miles. The same company established an office in Sydney in 1911 and, by 1913, 19 radio-telegraph coast stations were in operation providing contact with ships at sea.

Up till 1920 most of the transmitters in use for long distance communication used spark gaps to trigger the tuned circuits into oscillation, principally because valves were not available to produce the amount of radio frequency power available. The detectors in receivers were mainly of the crystal type, valves again being still one step away from standard production equipment.

The changeover to valve-type wireless communication equipment dated from about 1920, with the last of the spark-gap ship transmitters being phased out in about 1929.

Also during the twenties international radio-telephone links were set up with Britain, New Zealand and across the Pacific to North America. By way of interest, there was a simultaneous trend away from the long waves (low frequencies) on which the pioneering work had been done, towards the use of shorter waves (higher frequencies).

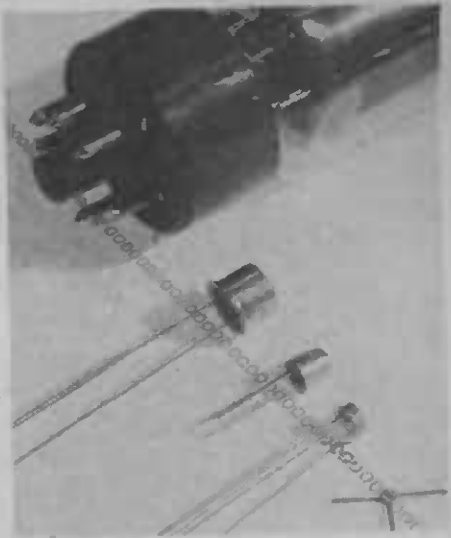
Strangely enough the pioneers of wireless telephony saw it mainly as a means of rapid point-to-point communication; very few saw it as a means of disseminating information, much less entertainment, to the population at large.

In 1920 Amalgamated Wireless (A / sia) Ltd gave a demonstration of broadcasting in Melbourne and, by the following year, the demonstrations had progressed to becoming a regular weekly feature. Stations 2BL and 2FC opened in Sydney towards the end of 1923, serving a very sparse audience made up mainly of enthusiasts and experimenters.

By 1928 broadcasting had become firmly established with a multiplicity of stations serving the major population centres and with radio sets finding their way into a significant proportion of Australian homes. By 1938 the number of medium-wave broadcasting stations in Australia had reached 97; the current figure is in excess of 200.

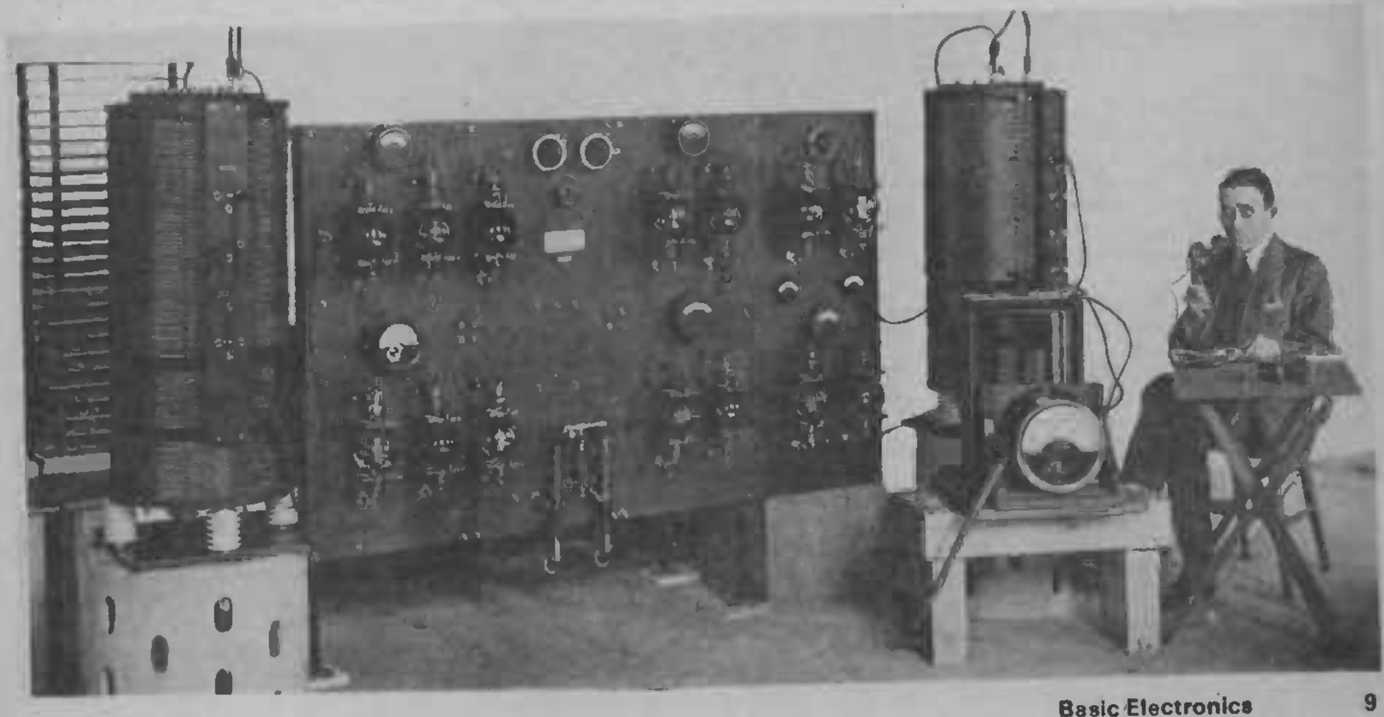
Television followed in the wake of radio, using the technology that radio made available. Work by John Logie Baird in England during the twenties, led to experimental tranmissions by the BBC in 1932, intended primarily for enthusiasts and experimenters. In 1936 the BBC conducted experimental test services from the Alexandra Palace, which led to the initiation of an official television service in 1937. This service was cut short by the war and, by the time hostilities had ceased, technology had advanced so far that Britain had to reappraise and up-date the earlier transmission standards.

In the United States, development roughly paralleled that in Britain but, due to rivalry and the more diffuse administrative situation, the distinction between experimental and "official" services is more difficult to draw. The war did not interfere to quite the same extent with television transmission but, even so, there were still only nine active television stations



Dramatising the trend towards miniaturisation: At the top is the base end of a medium-size radio valve. Ranged below it are a series of progressively smaller transistors. But even the smallest is huge by comparison with the transistor elements inside a modern integrated circuit, as pictured below.

The world's first wireless telephony news service was inaugurated on February 23, 1920 at the Chelmsford works of the Marconi Company. The 6kW transmitter pictured below was a valve type, operating on 2500 metres or 120kHz. At the microphone is W. T. Ditcham.



throughout the whole of the country in 1945.

Apart from such preliminary and rather small-scale activities, television has essentially been a post World War II development. It began in Australia in 1956 and, at present, the number of television transmitters in this country (including translators and repeaters) is getting close to the 200 mark.

Colour television, already operating in many overseas countries, is scheduled to begin service in early 1975.

This American designed Paragon RA10 receiver was ultra-modern in 1921. It covered from 160 metres to the "short waves" — 1000 metres, without loss of amplification due to the use of Armstrong regeneration. Price in the USA was $85.00.

In the meantime, of course, the radioman's techniques have spawned a variety of other entertainment and consumer items, many of which provide an industry in their own right. Among these could be listed sound films, audio visual devices, public address and mood-music amplifiers, hifi systems, tape recorders and electronic musical instruments — all of which are well known.

A truly historic photograph: The famous Australian prima donna, Dame Nellie Melba broadcasting from Chelmsford, England on June 15, 1920. The recital was heard as far afield as Persia.

However, the story of wireless — or radio — cannot be confined to entertainment or consumer items. This merely provides a convenient central theme of a much wider story.

- Medically orientated people quite early began to realise that the radioman's amplifier could be used to enlarge tiny electrical voltages in the body, so that they could be read on the radioman's meters and observed on his oscilloscopes.
- Structural engineers and aircraft engineers found that they could transduce pressures and stresses into tiny electrical signals which could be amplified, displayed and recorded.
- Industrialists found that they could use a light beam and a radioman's "electric eye" to monitor the density or level of a liquid, or count objects passing along a conveyer belt.
- Food manufacturers found that a radioman's coil, along with valves and a bell, could draw attention to metal particles in food.

One could go on adding to these paragraphs but the point has been made. The longer the list grew with the passing years, the more inadquate the word "radio" became as a collective term. The word that found universal acceptance was "electronics," derived from the most elemental component of electrical phenomena, the electron. So complete has been the acceptance of the word that it has no synonym in the English language.

The term came into its own very strongly during and after World War II. This was due, in no small measure, to the fact that wartime research accelerated the development of numerous electronic devices and techniques, ranging from radar and sonar to the forerunners of the modern computer.

The future of electronics in the postwar era revolved around the valve, as it had done thirty years before — but this was soon to change.

In 1920 the emergence of the thermionic valve had transformed wireless from a rather clumsy extension of electrical and telephone practice into a new and fascinating technology. But, by 1950, valves had become a virtual barrier to progress. Scientists and engineers could envisage important and complex roles for electronic equipment but the number of valves which would be required was horrifying!

Despite refinement to something of an ultimate, valves remained open to a number of criticisms, particularly in relation to highly complex equipment:

- Too costly per valve function;
- Too high a failure rate;
- Too large physically;
- Unnecessarily high power consumption;
- Excessive heat production.

These factors became more and more critical as engineers and scientists attempted to apply electronic techniques to more and more sophisticated tasks. They reached some kind of a climax when the first computers were constructed, using valves by the hundred and conventional components by the thousand. The equipment literally filled a large room; drew an embarrassing amount of current from the power mains; produced an embarrassing amount of heat — and was notoriously unreliable.

The breakthrough came with the transistor, pea-sized device which could take over many of the tasks performed by valves.

Developed by physicists Shockley, Brattain and Bardeen in the Bell Research Laboratories during 1947, the first functional transistor was the climax of speculation during the 1920s and of active research from the late 1930s onwards.

In a transistor, there is no heated filament, no grids, no anode and no space-wasting glass envelope. Electrons are still put to work, but their movement is promoted or restrained within tiny chips of germanium or silicon — chips measuring possibly one-twentieth of an inch square.

Instead of thermionic valves and vacuum techniques, scientists, engineers and technicians began to think in terms of "solid-state" technology — the manipulation and control of electrons within solid materials. Progress was rapid and far-reaching in its implications.

Spurred on by the success of transistor manufacturers and, in fact, using some of their new-found knowledge, other designers came up with a whole array of miniature components to go with them; capacitors, resistors and inductors smaller than had ever been seen before, and again more reliable.

Another vitally important development about this time revolutionised the method of connecting electronic components together. The new technique involved the use of a high-grade insulating board, clad on one side (sometimes both sides) with a thin layer of copper. By coating the copper with a photo-sensitive resist, exposing to light through a mask and then etching, the copper could be etched away where desired to leave a pattern of connections on

PEAK VALUE PACKED QUALITY MULTIMETERS

200-H. 90° quadrant meter. Pocket size.

AC/V	: 10V 50V 100V 500V 1,000V (10,000Ω/V)
DC/V	: 5V 25V 50V 250V 500V 2,500V (20,000Ω/V)
DC/A	: 50µA 2.5mA 250mA
OHM	:60kΩ 6MΩ
Capacitance	: 100pF to .01µF.001µF to 0.1µF
db	: −20db to +22db
Audio Output	: 10V 50V 100V 500V 1,000V AC
Approx. size	: 4½" x 3¼" x 1-1/8"

$15.50

A-10/P. Giant 6½" Meter. Inbuilt signal injector. Overload-Protected.

AC/V	: 2.5V 10V 50V 250V 500V 1,000V (10,000Ω/V)
DC/V	: 0.5V 2.5V 10V 50V 250V 500V 1,000V at 30,000Ω/V 5,000V (10,000Ω/V)
DC/A	: 50µA 1mA 50mA 250mA 1A 10A
AC/A	: 1A 10A
OHMS	: 10kΩ 100kΩ 1MΩ 100MΩ
db	: −20 to +62db
Signal Injector	: Blocking oscillator circuit with a 2SA102 transistor
Approx. size	: 6-2/5" x 7-1/5" x 3-3/5"

$60.00

CT-500/P. Popular, medium-size, mirror scale. Overload-Protected.

AC/V	: 10V 50V 250V 500V 1,000V (10,000Ω/V)
DC/V	: 2.5V 10V 50V 250V 500V 5,000V (20,000Ω/V)
DC/A	: 0.05mA 5mA 50mA 500mA
OHM	: 12kΩ 120kΩ 1.2MΩ 12MΩ
db	: −20db to +62db
Approx. size	: 5¼" x 3-5/8" x 1¾"

$19.25

M-650/P. Plastic dial cover. Mirror scale. Overload-Protected.

AC/V	: 6V 30V 120V 300V 1,200V (15,000Ω/V)
DC/V	: 3V 12V 60V 300V 600V 1,200V (50,000Ω/V)
DC/A	: 30µA 6mA 60mA 600mA
OHM	: 16kΩ 160kΩ 1.6MΩ 16MΩ
db	: −20 to +63db
Approx. size	: 5¼" x 3½" x 1¼"

$24.95

370W/P. Bench model. Overload-Protected.

AC/V	: 2.5V 10V 50V 250V 500V 1,000V (4,000Ω/V)
DC/V	: 0.5V 2.5V 10V 50V 250V 500V 1,000V (20,000Ω/V)
DC/A	: 50µA 1mA 10mA 100mA 1A 10A
AC/A	: 100mA 1A 10A
OHM	: 5kΩ 50kΩ 500kΩ 5MΩ 50MΩ
db	: −20db to +62db
Approx. size	: 7" x 5¼" x 3-1/8"

$57.50

OL-64D/P. Popularly priced model, rugged construction. Overload-Protected.

AC/V	: 10V 50V 250V 1000V (8,000Ω/V)
DC/V	: 0.25V 1V 10V 50V 250V 500V 1000V (20,000Ω/V) 5000V (4,000Ω/V)
DC/A	: 50µA 1mA 50mA 500mA 10A
OHM	: 4kΩ 400kΩ 4MΩ 40MΩ
db	: −20 to +62db
Approx. size	: 6" x 4-1/5" x 2"

$21.55

AS-100D/P. High 100,000 Ω/Volt sensitivity on DC. Mirror scale. Protected-movement.

AC/V	: 6V 30V 120V 300V 600V 1200V (10,000Ω/V)
DC/V	: 3V 12V 60V 120V 300V 600V 1,200V (100,000Ω/V)
DC/A	: 10µA 6mA 60mA 300mA 12A
OHM	: 2kΩ 200kΩ 20MΩ 200MΩ
db	: −20 to +57db
Audio Output	: 6V 30V 120V 300V 600V 1,200V AC
Approx. size	: 7½" x 6⅜" x 2¾"

$37.95

H. ROWE
& CO. PTY. LTD.

SOLE
AUSTRALIAN
AGENTS

ADELAIDE BRISBANE HOBART
MELBOURNE PERTH SYDNEY

All prices are suggested prices only.
All prices include sales tax.

Semiconductor Circuit Design ▶ INFORMATION

Semiconductor Data Handbook

New, first edition, bound handbook. Over 1,300 pages. Includes complete product specifications for GE's complete line of semiconductors, over 2,500 types arranged in alpha-numeric order. Interchangeability information, Selector Guides, and check lists make this a complete, comprehensive guide.

Transistor Manual

7th Edition
Written for circuit designers by circuit designers. Applications, characteristics, and theory for transistors, unijunctions, diodes, tunnel diodes, plus other discrete signal devices. Over 600 pages.

SCR Manual

5th Edition
New, updated edition. Covers the complete story on SCR's, plus expertise on Triacs, regulating circuits, motor controls, and other power semiconductor applications. Over 500 pages.

Experimenter's Manual

3rd Edition
A select collection of useful and intriguing circuits covering audio, automotive and marine, home, farm, and camping applications, plus other areas. Personal enjoyment for engineers. . . as well as hobbyists. Over 250 pages.

AUSTRALIAN GENERAL ELECTRIC LTD

86-90 BAY STREET, ULTIMO. PHONE 212 3711

DISTRIBUTORS: **Fairchild Australia Pty Ltd**, Melbourne, Tel 723 4131. Sydney, Tel 439 4355.
GEC-Elliott Automation, Sydney, Tel 439 1922. Adelaide, Tel 71 7971. Brisbane, Tel 32 0441. Melbourne, Tel 387 2811.
PB Components, Melbourne, Tel 53 2766.
Watkin Wynne Pty Ltd, 32 Falcon Street, Crows Nest, Tel 43 2107.
H. Rowe & Co Pty Ltd, Perth, Tel 28 5444.
The National Electrical & Engineering Co Ltd, Wellington, Tel 553 709. Auckland, Tel 599 089.

what came to be known as a **printed wiring board**. The technique increased uniformity, reduced the risk of error and lent itself to mechanisation with economies in both time and cost.

With transistors and the new miniature components, the room-sized, power hungry, hot, unreliable computer shrank to a couple of head-high cabinets — cheaper, more economical to run, cooler, more reliable and capable of tackling more complex computations.

It became possible for aircraft to carry an array of electronic equipment that would have been unthinkable earlier — multiple transceivers for communication, multi-purpose radar, multiple navigation aids, blind landing equipment and a variety of other devices to monitor, control and display internal functions in the plane.

Even the humble portable radio was transformed from a rather rare item, too heavy, unreliable and costly to run, to the modern transistor radio, small enough to slip into a coat pocket.

Tiny as transistors are in comparison with valves, they are still unnecessarily large for the job of manipulating electrons. As someone has remarked, electrons moving through a transistor are like toy cars on a super highway.

Appreciating this fact and having learned the new tricks of micro-photography, micro-etching, doping, diffusion, metallic deposition and such like, engineers quickly realised that they could shrink their dimensions to fit several transistors on to much the same semiconductor chip or "die" that formerly accommodated one. They could even interconnect them on the chip in the kind of circuit hook-ups that equipment engineers might require.

Variations of the techniques made it possible to add other components such as resistors to the circuitry on the chips. These component groups, concentrated on to tiny silicon dice, soon came to be known as integrated circuits, now commonly abbreviated to ICs.

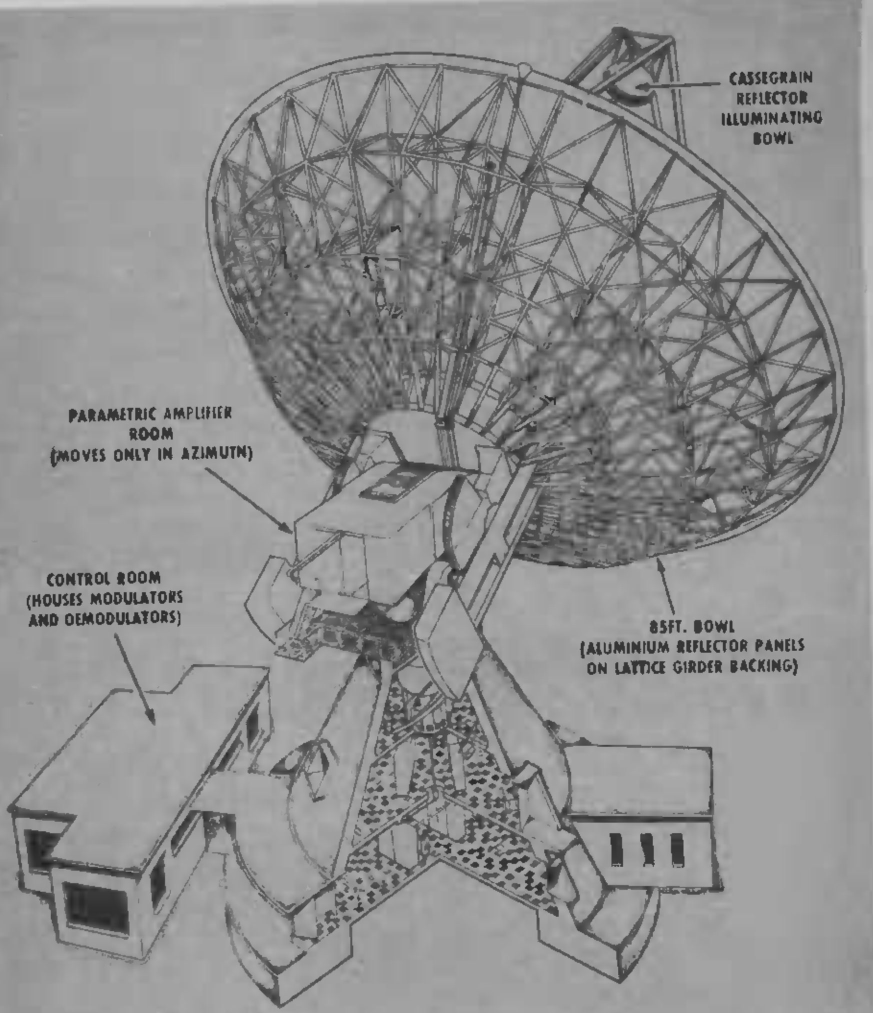
The execution of electronics with component groups as small as this has also given rise to the term "microelectronics."

Early integrated circuits, introduced about 1964, managed to cram about 10 components on to a single chip. By the following year the figure had risen to 50, and to about 500 by about 1968.

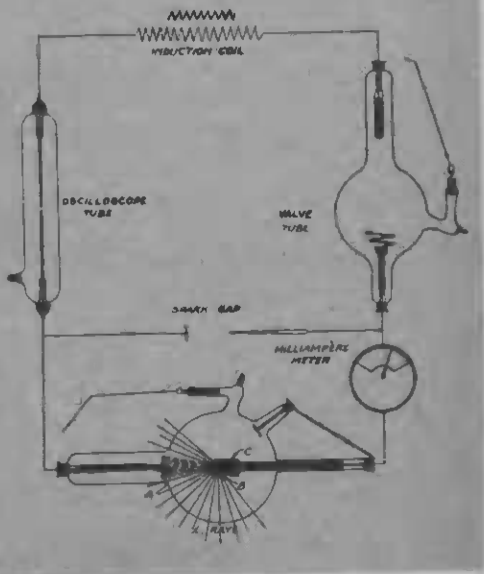
Nowadays, large-scale integration (abbreviated to LSI) crams on to a single chip not only many hundreds of diodes and transistors but the interconnections necessary to perform highly complex circuit functions. Devices that, a few years ago, would have required hundreds of transistors on multiple circuit boards are now being concentrated on to single chips and packaged in a moulded case no larger in area than a postage stamp.

Perhaps the most dramatic examples of this can be seen in the electronic pocket calculators that are currently flooding the market. They carry out calculations that would have involved a desk-top instrument with transistors, or a room-size monstrosity if anyone had bothered to attempt one using valves.

Nor is this last observation in any way exaggerated. Twenty years ago, electronic organs needed a mechanical generator system or a chassis-full of valves and other components to generate the basic octave of musical tones. Now it can been done in two tiny microcircuits, far more cheaply, accurately and reliably!



One of the developments which most dramatises the progress of communications and electronics generally is our present ability to communicate via satellites or with space probes many millions of miles from earth. Illustrated above is a British designed satellite terminal, involving a fully steerable antenna.



Fifty years ago X-rays provided the most obvious hint of developments that would ultimately change "wireless" into "electronics". This diagram is reproduced from a magazine of the period. The article indicated that X-rays were being used in Australia for medical purposes, as well as for examining welds and castings.

Electronics has now reached the stage where it is involved in virtually every aspect of life, art and science. It is monitoring space flight, controlling surface traffic and guiding submarines. It provides the measuring equipment for science and industry, helps combat disease and spots the counterfeiter and the saboteur.

The more you know about electronics, the better equipped you will be to live and move in this modern world. That is why we recommend that you take the time to study the chapters that follow.

CHAPTER 2

Basic Electrical Concepts

Electrons, protons and electric charge — atoms, molecules and ions — electric current and pressure — the Ampere and the Volt — conductors, semiconductors and insulators — the Ohm — fixed resistors: wire-wound, carbon and metal oxide film — variable resistors: rheostats and potentiometers.

Most basic electronic text books begin with a discussion of the electron theory — not because the authors lack imagination but simply because it is the logical place to start. One can never hope to understand the principles of modern radio and electronic apparatus, without at least a working knowledge of electrons, their nature, their behaviour and their relationship to what we have called electricity.

Once having grasped the idea of electrons in motion, electrical current will cease to be a mystery. You will be able to appreciate readily — and graphically — the significance of terms like conductor, insulator, resistor and so on.

Let's start, then, with a little bit of basic chemistry or physics — call it what you like. Either term would be justified.

Scientists have demonstrated that all matter, whether it be solid, liquid or gas, is composed of minute electric charges of which there are two main varieties. One of these is the electron, which may be defined as the smallest existent quantity of negative electricity. Its opposite number is the more massive proton which is an electric charge of the same magnitude, but of opposite or positive sign.

Just what electrons and protons and other uncharged components of atoms actually are need not concern us greatly here. Depending on their activities, some scientists prefer to regard them as basic particles exhibiting definite electrical characteristics; others prefer to regard them as "bits" or quanta of electricity which behave in the manner of particles.

Since we are concerned in this book with electrical and electronic theory, it is convenient and appropriate simply to think in terms of electric charges.

Electrons and protons do not normally exist alone, but are usually associated with one another in groups. Each group is thought to resemble in certain ways the planetary or "solar" system of which our world forms a part. If it were possible to peer into these sub-microscopic realms, we would behold a host of miniature "planetary systems." In each, we would see a proton, or a group of protons and other particles forming a core or nucleus, and moving in the space around them, a number of single electrons.

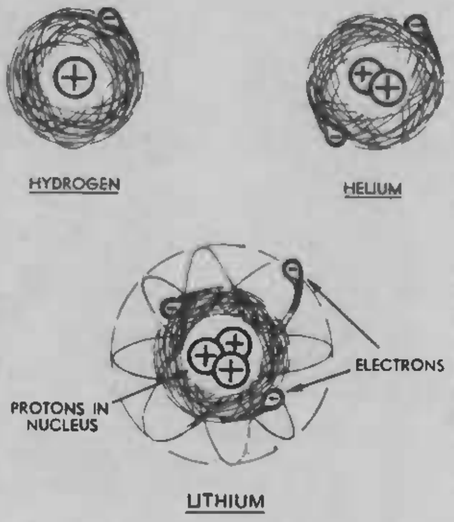


Fig. 1: Typical atoms, which consist basically of a positively charged nucleus surrounded by a "cloud" of electrons in motion.

In actual fact, each tiny "planetary system" constitutes an atom of a particular substance. And, from our chemistry, we know that an atom is the smallest particle of any substance obtainable by chemical separation, or capable of entering into chemical combination. Atoms are very small — one hundred million of them, arranged end to end, would just about equal in length three words of this type.

Although no one has ever seen a single atom, or the electric charges which compose them, scientists have been able, by roundabout means, to deduce quite a lot of information about them. They know, for example, that the simplest of the lot is the hydrogen atom, which has a single proton as a nucleus, with a single electron spinning around it. Only slightly more complex than this is the helium atom, which has two protons with two planetary electrons. An atom of lithium has three protons in the nucleus, two electrons whirling on an inner orbit and a single electron on an outer orbit.

These are shown in simplified form in Fig 1, at left.

Every atom is electrically balanced in its normal state. The three examples just quoted contain respectively one, two and three positive charges each, and the same numbers of negative charges. This state of affairs is maintained in even the very complicated atomic structures. An atom of copper, for example has 29 protons grouped in the central nucleus; but the excess positive charge here is exactly balanced by the 29 electrons which revolve in various orbits around the nucleus.

It is important to note that the electrons and protons comprising the various atoms remain simply electric charges. An electron associated with a hydrogen atom is identical to one in a copper atom, even though it may be placed differently in the structure. It is the particular combination of protons and electrons which determines to what chemical element the atom belongs. The number of natural chemical elements is believed to total around 100, which means

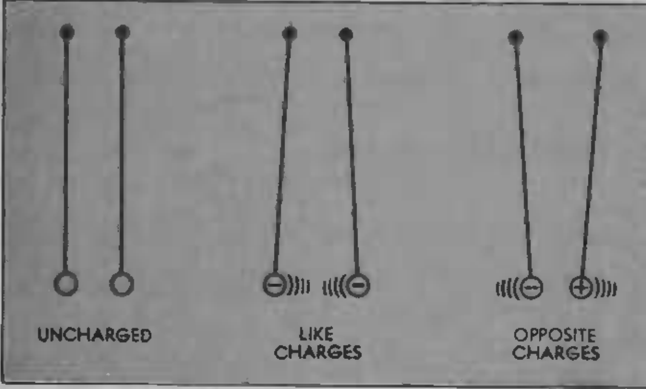


Fig. 2: The classic rule that "like charges repel, unlike charges attract" is illustrated by the behaviour of charged pith balls, suspended on fine silk or nylon threads.

that there are at least this many different atomic structures.

Theoretically, at least, some atoms can exist in the solitary state and exhibit all the chemical properties of the element in bulk; examples of this are the atoms of the rare gases helium and neon. Such atoms may also be referred to as **molecules**. In the case of other elements, the atoms, for chemical purposes, are normally found in groups of two or more, and these atomic groups are also known as molecules. A third variety of molecule is that belonging to a chemical compound, in which unlike atoms are found in combination.

In short, a molecule can be defined as: "The smallest portion to which any given substance can theoretically be divided, without altering its chemical properties."

The chemical compounds in existence are virtually numberless but they all result from different combinations of the same basic elements; in other words, they are all structured from two or more of the 100-odd basic atomic "bricks."

It is hard to realise that the most uninteresting, the most inert substance one can imagine is composed of countless atoms, with their numerous swiftly moving electrons. Indeed, the speed with which each individual electron revolves in its orbit is enormous, despite the fantastically small circumference of even the largest atom.

But there are other matters of more immediate interest to the reader of this article. The outer electrons in many types of atom are rather wayward under certain conditions and it is not uncommon for them to wander rather aimlessly into the structure of adjacent atoms.

This leaves the original atom short of an electron so that, for a very brief interval, it may have an excess positive charge. In other words, the loss of an electron has upset its electrical balance.

Atoms in this unbalanced state are known as **ions**. They are said to be positive ions if they have lost an electron, and negative ions if they have acquired an extra electron. The electrons which are wandering free of attachment to an atom may also be called negative ions. The word ion is, in fact, often simply used to signify a charged particle.

A fundamental principle of electricity is that like charges repel while unlike charges **attract** (see Fig. 2).

The operation of this principle soon cancels the state of unbalance created by individual meandering electrons. A positively charged depleted atom (ion) immediately attracts an electron from elsewhere, while negative ion formed by an atom gaining a surplus electron quickly disposes of one. Naturally, nearby atoms are disturbed in the process, so that there tends to be a continuous and random movement of electrons within the confines of particular substances. This movement can be greatly accentuated by the application of heat and the agitation can actually become so violent that some electrons momentarily jump off into space.

Such movement is purely random in character and the average movement of electrons in one direction is exactly balanced by electron movement in the opposite direction. In other words, there is no nett migration in any particular direction.

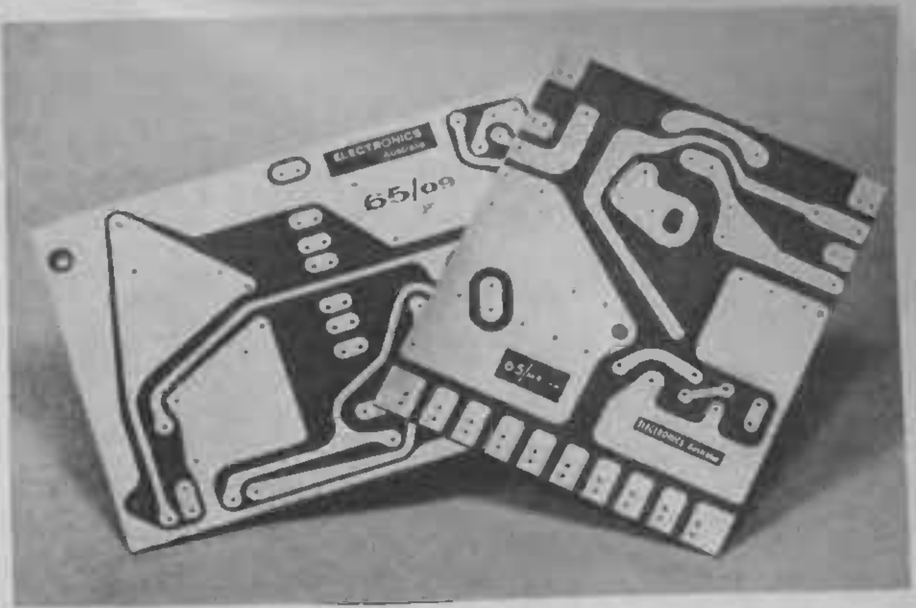


Fig. 4: Modern printed wiring boards serve to illustrate the contrasting roles played by conductors and insulators. The conductor, a thin layer of copper shown as the light areas, performs the circuit connections. The board (dark areas) to which the copper is fixed provides support and isolation; the board is usually of phenolic bonded paper or fibreblass.

IN SEARCH OF THE ELECTRON . . .

"Now the most startling result of Faraday's Law is perhaps this: if we accept the hypothesis that the elementary substances are composed of atoms, we cannot avoid concluding that electricity also, positive as well as negative, is divided into elementary portions which behave like atoms of electricity." (Helmholtz, in his Faraday lecture before the Royal Institution, 1881.)

"Thus on this view we have, in cathode rays, matter in a new state: a state in which the subdivision of matter is carried very much further than in the ordinary gaseous state; a state in which all matter — that is, matter derived from different sources such as hydrogen, oxygen, &c — is of the same kind; this matter being the substance of which all chemical elements are built up."(J. J. Thompson, in Philosophical Magazine, 1897.)

However, it is possible to alter this state of affairs so that there is a defined migration of electrons in a certain direction through the substance.

Any such clearly defined movement or migration of charges such as electrons, in a given direction through a substance is known as an electric current.

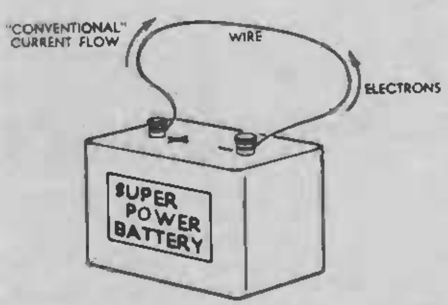


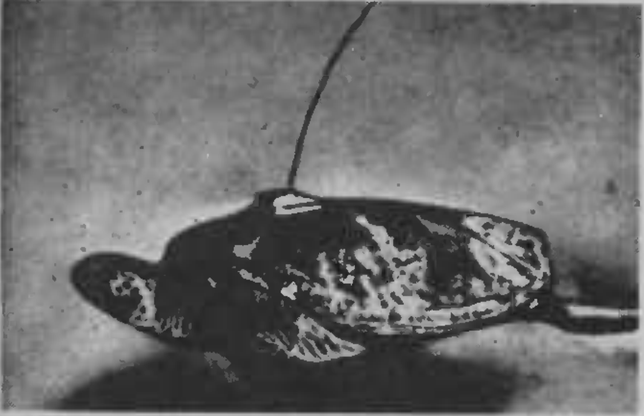
Fig. 3: A source of EMF such as a battery causes a net flow of electrons through a conductor, in one direction. Such a flow forms an electric current.

All readers will have handled an electric battery at some time or other, in one of the many forms. These devices will be treated in detail in a later chapter; it is sufficient just to state here that a battery is a device which, by electro-chemical action, can produce an excess of free electrons at one terminal, and a deficiency of electrons — hence positive ions — at the other.

Between its terminals the battery thus builds up an electrical unbalance or pressure which, in electrical language, is known as an electro-motive force (EMF) or, more simply, a potential. These terms actually have definite and slightly different meanings, but for our purposes they may be regarded as identical.

Fig 3 depicts in schematic form a piece of copper wire connected between the terminals of an electric battery. We can expect the positive terminal to exercise a strong attraction for the outer electrons in the nearest atoms of the wire. This, indeed, is the case and a definite movement of

ELECTRONICS ARE GOING PLACES



A radio-equipped turtle is tracked by orbiting satellites to check its migratory voyages in the Atlantic Ocean. One of the million incredible uses of electronics for commerce and industry.

EARN BIG MONEY as an ELECTRONICS ENGINEER

STUDY AT HOME IN YOUR SPARE TIME

You know as well as we do that electronics is the big new field that's here to stay. Industry is using electronics in fields many people hadn't dreamed of a few years ago.

TRAINED MEN ARE NEEDED. Australia's industries need, and must have, Electronics Engineers urgently. Salary scales are rising fast and electronic engineering specialists are making big money. Trained Australian Electronics Engineers can choose jobs anywhere in the world—the lack of these trained men is world-wide. Training is the key—qualifications are what matter.

Select the career that suits you best

Basic Electronics
A short course ideal for those wishing to obtain a sound knowledge of radio theory and the industrial applications of electronics.

TV Service Technician
Start a busy and profitable career. In this course you will learn how to install, calibrate and repair modern radio and TV cells. An ideal course for beginners.

Electronic Engineers
Designed to give a thorough knowledge of Electronic Engineering in both theory and practice. Instruction in radio theory is also included.

Colour TV
Colour TV is on the way and this course has been designed to give the trained technician a good knowledge of the principles. The NTSC and PAL systems are studied in detail.

Computer Technician
This course includes programming as well as instruction on the technical aspects of computers. Ideal for technical personnel concerned with installation and maintenance of digital computers.

BE HIGHLY PAID IN THE WORK YOU LIKE MOST. You already have the interest you need to be successful in electronics—you can get the training you need through International Correspondence Schools. You can train for your career in electronics at night, in your own time, with the help of a School of Electronics as close as your mail box.

ACT NOW! Fill in the coupon below and send it to I.C.S.—we will send you by return mail our Free Book "Your Career in Electronics."

TV Radio Electronics
Electronic Instrumentation and Control
Electronic Technology
Fundamentals of Electronic Computers
Industrial Electronics
Electronics Maintenance
Radio-Television Engineering
Radio Electronic Telemetry
Radio Technology
Television Engineering
Electronic Drafting
Solid State Electronics Technology
Radio, Hi-Fi, Stereo and Sound Systems Servicing
Computer Programming
Business Data Processing
Cobol Programming
Computer Systems Analysis
Fortran Programming
General Computer Programming
IBM 360 Computer
Systems and Procedures Analysis

POST THIS COUPON TODAY TO

ICS
DEPT. 560, 400 PACIFIC HIGHWAY
CROWS NEST, N.S.W. 2065

Please provide me with information about the Career on:

____

Name (Mr., Mrs., Miss)__________ Age____

Address____________________

____________ Postcode____

Occupation____________________

LB3.4353

electrons becomes evident in the direction of the positive battery terminal.

At the other end of the wire, the negative battery terminal is quite ready to part with some of its excess electrons, and these go to replace those lost to the positive terminal. The potential or EMF of the battery thus initiates a definite flow of electrons — an electric current — through the copper wire. There is a nett electron flow from the negative to the positive terminal.

Unfortunately, in making this statement we run up against an apparent contradiction. The early electrical pioneers knew little of electrons, although they could observe the effects of electric current and potential. So charges were simply labelled "positive" and "negative" for identification, after which they deduced that current consisted of positive charges moving from positive to negative.

According to modern theory, they were wrong, as electric current is considered to consist in most cases of a movement of electrons from negative to positive. This is certainly so in the case of a current flowing through a metal, such as copper. However the original concept of current flow still tends to persist, being known as the "conventional" direction of flow. It is not unusual, therefore, to encounter statements based on the original assumption.

In this book we have sought to avoid confusion by majoring on the term "electron flow," and avoiding use of the word current where direction is involved.

Electrons are far too minute to be useful directly as a measure of current flow. It would be rather like measuring sand by the grain!

The standard unit by which current flow is expressed is the **ampere**, generally abbreviated to "amp" or simply "A."

Some idea of the minute size of electrons may be gained by the following: When a current of one amp (which is quite a modest figure) is flowing in a circuit, it corresponds to something like 6¼ million, million, million electrons passing any given point in the circuit each second!

The precise definition of the amp need not concern us here, being mainly a matter for standards laboratories. The important point to remember is that the ampere is the basic unit for current measurement.

For purposes requiring a smaller unit than the ampere, we have:

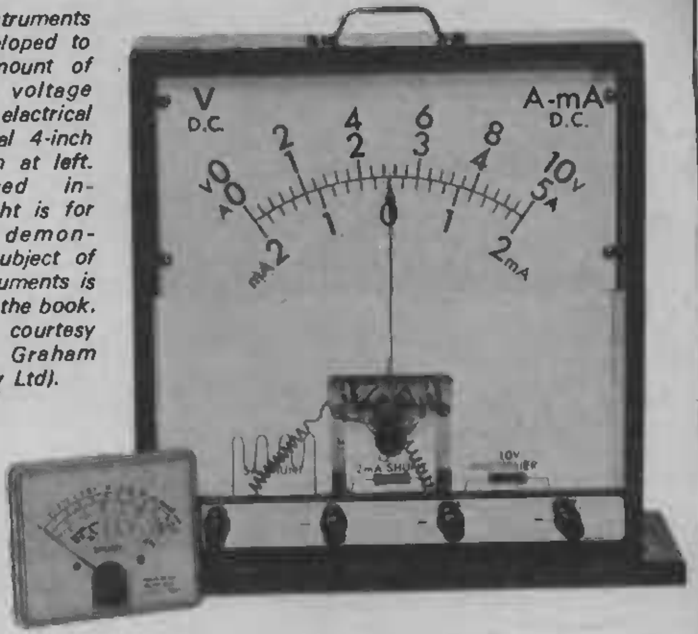
1 milliampere equals .001 amp.
1 microampere equals .000001 amp.

Conversely:

1 amp equals 1,000 milliamperes.
1 amp equals 1,000,000 microamperes.

These relationships should be memorised as you will find yourself coming up against them repeatedly in electronics and electrical theory.

The accepted unit of electro-motive force (EMF) or potential is the volt. A battery may typically be defined as having an EMF between its terminals of, say, 1.5 volts. Alternatively we may state that its voltage is 1.5, or any other figure which might apply.



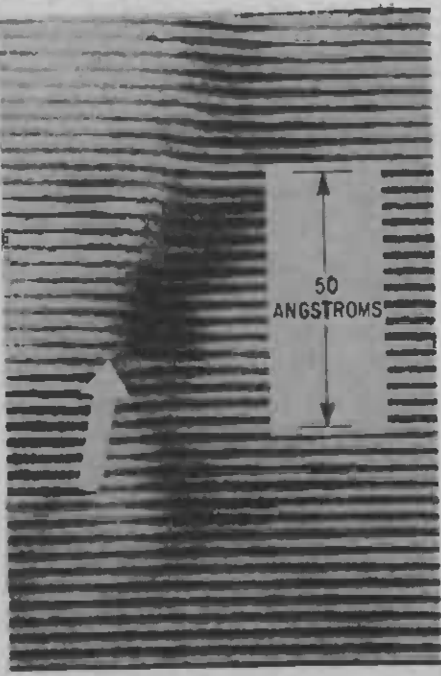
Special test instruments have been developed to indicate the amount of current or voltage present in an electrical circuit. A typical 4-inch meter is shown at left. The king-sized instrument at right is for class-room demonstration. The subject of measuring instruments is covered later in the book. (Photograph by courtesy of University Graham Instruments Pty Ltd).

Larger and smaller terms for defining voltage are:

1 kilovolt equals 1,000 volts.
1 millivolt equals .001 volt.
1 microvolt equals .000001 volt.

Alternatively:

1 volt equals .001 kilovolt.
1 volt equals 1,000 millivolts.
1 volt equals 1,000,000 microvolts.



While no one has yet succeeded in photographing an individual atom, this picture shows the planes or lines of atoms in a crystalline substance (e.g. silicon, germanium, etc.) The picture was taken with an electron microscope and subsequently enlarged. The arrow points to a fault in the structure where one plane ends abruptly. Other nearby planes are distorted as a result.

Coming back to figure 3, we must not imagine that the current flowing in the wire consists of individual electrons whizzing through the copper wire unimpeded, like meteors through outer space. Typically an individual electron may leave the negative terminal and collide with the first atom in its path. It may take its place in the planetary system of that atom, while the electron it displaced moves off toward the positive battery terminal. An instant later it may itself be displaced, allowing it to move on again. So the current may quite accurately be said to consist of many small inter-atom "jumps" by many different electrons.

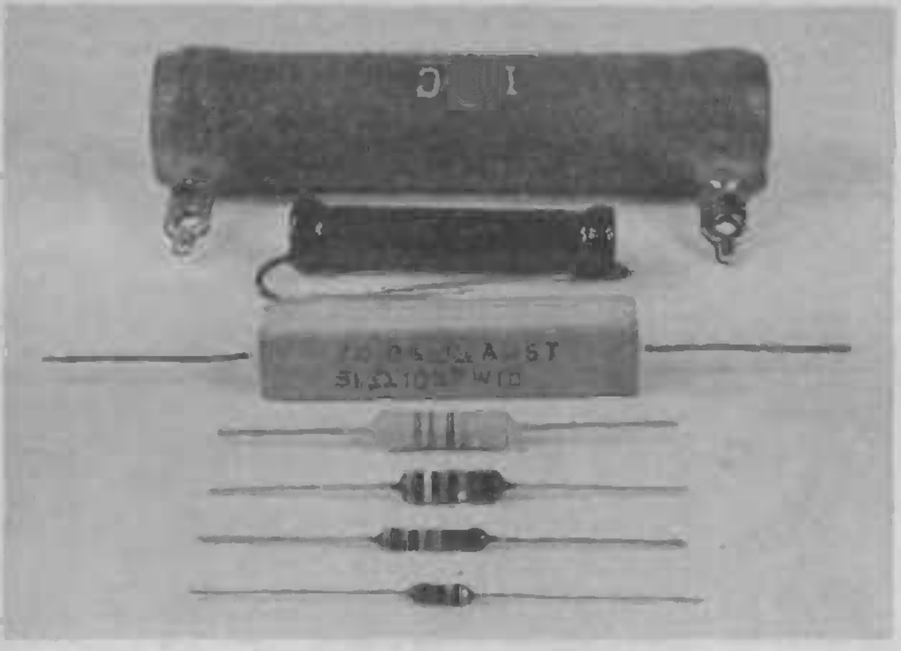
By its very structure, the copper wire offers some resistance to the passage of the current, because of electron-atom collisions, etc. For all that, the current passes easily enough, so that copper is one of the substances classified as a **conductor** of electricity.

Into this general classification **fall** the various **metals**, carbon and certain other substances and liquids. Some offer more resistance to current flow than others, but all are conductors. Current can flow relatively easily through them, or from one to another if two or more conductors are brought into direct physical contact.

In another group are a large number of substances in which there is very little transfer of electrons from atom to atom, under normal conditions, unless a very large external electric pressure is applied. Such substances, which under normal conditions conduct to a negligible degree only, are called **Insulators**.

Some substances are very useful in this respect, maintaining their insulating properties despite heat and high electric pressures. Others can withstand only moderate heat or applied electric pressure before they conduct appreciably, and may "break down" and char, due to permanent structural damage.

In a later chapter we deal with a third group of substances which are in some



respects midway between conductors and insulators. These are called **semiconductors** but, apart from mentioning their existence, we will not deal with them any further at present.

Typical insulating materials are ceramic, ebonite, mica, shellac, silk, oil, dry hair, and most modern plastics.

Insulating materials are used to support, isolate or contain conductors which are part of an electric circuit. Because no appreciable current can flow through the insulating materials, their presence does not — or should not — affect the operation of the circuit.

A good illustration of conductors and insulators and their respective roles is seen in Fig 4. Here are shown two of the so-called "printed wiring boards" which are nowadays used very frequently in radio sets, TV receivers and a lot of other electronic equipment.

Printed wiring boards consist of a "sandwich" of an insulator and a conductor. The insulator (dark areas in the picture) is usually either resin-bonded bakelised paper or epoxy-fibreglass, in a sheet about 1 / 16in thick. This is used as an insulating support for both the conductor and the various parts mounted on it. The conductor itself is usually copper, in a sheet about .0015in thick, bonded to the surface of the insulator. It is etched away to form thin strips and other areas (light in the picture), which serve to provide the "wires" connecting the parts together to form a circuit.

As we shall see in later chapters, there is a very definite place in electrical circuits for circuit parts made from substances which can be classed as neither good conductors nor good insulators; in other words, for parts which offer predictable resistence to the passage of current through them.

The connection of a short copper wire between the terminals of a battery, as in Fig 3, would result in a heavy flow of current, sufficient probably to discharge the battery. A longer connecting wire would tend to reduce the current by reason of the longer path the current would have to traverse. Substitution of a finer gauge wire would have the same effect, since a smaller cross-sectional area would be available for the electron movement.

By substituting an iron wire for the copper, still greater resistance would be evident, with a consequent reduction in current flow.

In other words, by deliberately introducing **resistance** into a circuit, it is possible to limit or control the current flowing, exactly as a tap controls the flow of gas in your domestic range. Actually, there are other uses for this property of resistance in an electrical circuit, but the above remarks should be sufficient for the time being.

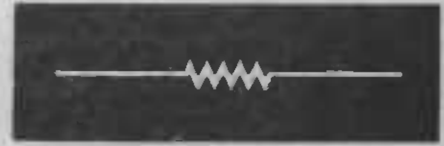


Fig. 5 (top left): An array of typical resistors, shown approximately normal size. The size is a guide to their power rating. Figure 6 (above): The usual circuit symbol for a resistor.

The basic unit of resistance is the ohm. which is that amount of resistance which will limit the current through a circuit to one amp, when the applied EMF is one volt. For very high values of resistance it is more convenient to speak in terms of **kilohms** and megohms, the symbols for which are k and M.

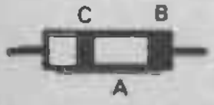
1 kilohm equals 1,000 ohms.
1 megohm equals 1,000,000 ohms.
1 megohm equals 1,000 kilohms.

Resistance values up to 999 ohms are generally expressed directly in ohms, but between 1000 and 999,000 ohms they are commonly expressed in kilohms, and above 1,000,000 ohms in megohms. It is obviously easier to say and write 120k and 12M, for instance, than 120,000 ohms and 12,000,000 ohms, although the two ways of showing the figures concerned mean the same thing.

Coming to the practical side of the matter, many substances exhibit sufficient opposition to the passage of an electric current to suggest their use in a circuit as a **resistance element**. However, their physical adaptability is also an important consideration. Certain metals and alloys, for example, exhibit quite high resistance compared to copper, and lend themselves well for the purpose. Carbon, too, is the basis for some resistance elements.

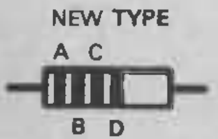
RESISTOR IDENTIFICATION — THE COLOUR CODE

OLD TYPE



A — BODY COLOUR
B — END COLOUR
C — DOT OR BAND

NEW TYPE



A — FIRST BAND
B — SECOND BAND
C — THIRD BAND
D — FOURTH BAND (OPTIONAL)

BAND	A	B	C	D
COLOUR	TENS	UNITS	MULTIPLIER	TOLERANCE
BLACK	0	0	$10^{0} = 1$	NONE = 20%
BROWN	1	1	$10^{1} = 10$	1%
RED	2	2	$10^{2} = 100$	2%
ORANGE	3	3	$10^{3} = 1,000$	
YELLOW	4	4	$10^{4} = 10,000$	
GREEN	5	5	$10^{5} = 100,000$	
BLUE	6	6	$10^{6} = 1,000,000$	
VIOLET	7	7		
GREY	8	8		
WHITE	9	9		
GOLD	—	—	$10^{1} = 0.1$	5%
SILVER	—	—	$10^{2} = 0.01$	10%

Examples: 47,000 ohms, 10%: Yellow, Violet, Orange, Silver.
1,000 ohms, 20%: Brown, Black, Red.
68 ohms, 10%: Blue, Grey, Black, Silver.
4.7 ohms, 5%: Yellow, Violet, Gold, Gold.
0.47 ohms, 5%: Yellow, Violet, Silver, Gold.

To facilitate the inclusion of a specific amount of resistance in an electrical circuit, component manufacturers produce small parts known as **resistors**. For resistance values up to a few thousand ohms, these units often consist generally of a length of so-called "resistance" wire, wound spirally on a tube of cardboard, glass, porcelain or other suitable insulating material.

The wire is joined to a terminal at either end, or to a lug or tinned copper lead, to facilitate soldered connections. The resistor is encapsulated and / or then coated with lacquer to protect and locate the individual turns of wire. It is finally branded with its nominal resistance value in ohms.

The actual resistance depends on the nature and the gauge of the wire used, and on the length of wire wound around the tube or former. Resistors made up after this fashion are known as **wire-wound** types.

For resistance values above a few thousand ohms, the wire-wound type of resistor presents manufacturing difficulties, owing to the fine gauge and the nature of the wire which would have to be used, and also to the bulk of the finished component.

As a result, an alternative type of resistor is generally used. Finely divided carbon or similar material is mixed with a non-conducting binding material and either moulded into small rods or deposited as a film on a small ceramic rod. The proportion of carbon and binder and the physical dimensions govern the ultimate resistance. This basic element may be moulded into a bakelite tube for further protection, or simply coated with lacquer or enamel. Two copper wire "pigtails" for soldering complete the assembly of these so-called **carbon resistors**.

Carbon resistors are not confined to high values, but may also be used for quite low resistances — down to about 10 ohms. Below this value, a similar type of resistor may be produced by depositing a metal film (typically nickel) on a ceramic rod, and machining this film into a spiral.

Another type of resistor element is made by depositing a thin film, of a metal oxide, such as tin oxide, on a small glass rod. The resulting **metal oxide resistor** is usually sealed in another glass tube, with tinned copper pitgails emerging at each end for the external connections. This type of resistor is more costly than either the wire-wound or carbon varieties, but also tends to be more reliable, being less susceptible to damage by moisture and heat.

Like wire-wound resistors, carbon and metal oxide elements may have their nominal resistance value marked directly on the outside, in numerals. However due to the small size of many of these parts, they are often marked alternatively using bands of coloured paint, according to the code shown.

When current flows through a resistor, the erratic movement of the electrons and the resulting electron friction absorb electrical energy from the source of current, and turn this into heat energy. Without being more precise just here the heat generated can be expressed in watts. The amount of heat which any resistor can dissipate is closely related to its physical size and to the ability of the associated insulating materials to withstand the effects of increased temperature.

Resistors commonly used in electronic work, and illustrated in Fig 5, range from less than one half to many inches long and carry dissipation ratings of from 1 / 10 watt to 20 watts. Larger resistors are available for special purposes.

In Fig 6 you will see the first of many rather mysterious symbols. The **zig-zag line** is actually the schematic circuit symbol normally used for a resistance.

When electrical or radio engineers are depicting an electrical circuit, solid lines are used to represent wires or connecting links. This is simple enough, but it would be a nuisance to have to sketch a resistor in detail whenever one had to be shown in a particular portion of a circuit.

Simplified or "schematic" symbols are thus used, to simplify both the drawing and the reading of circuit diagrams. The subject will be explained later in more detail; for the present just remember that a zig-zag line in a circuit normally represents a resistor.

There are various applications where it is necessary to vary the amount of resistance in circuit; this brings us to the term **variable resistor**.

A variable resistor normally involves a wire-wound or carbon resistance element, with provision to move a sliding contact along its surface. By making a connection to the sliding contact and another to one end of the resistor, the amount of resistance between the connections will obviously be capable of adjustment.

Where ease of adjustment is important, it is possible to have the resistance wire wound on a flat insulating strip, which is formed or bent into about a three-quarter circle. A spindle is located at the centre point of the arc carrying a radius arm. As the spindle is rotated, this radius arm moves over the resistance element, making contact, as it passes, with each individual turn. One terminal connects to one end of the resistance element and the other terminal to the contact arm. Any value from zero to maximum resistance may be effective between the terminals, depending on the setting of the contact arm.

In the earliest controls of this type, the wire-wound resistance element was left exposed. This is still done with large high power units, as may be seen in Fig 7, but present practice is to house most of the smaller units in a case of pressed metal or moulded plastic, for appearance and protection.

For resistance values in excess of a few thousand ohms, the resistance element is usually a carbon film, or a carbon compound moulded into the required shape.

The appropriate schematic symbol for a variable resistor is the upper one in Fig 8.

Variable resistors having just the two connections, are commonly known as **rheostats**. Modern practice is to provide a variable resistor with three connections, one to each end of the resistance element and one to the moving contact. Such a unit is generally referred to as a **potentiometer** and all three connections are often employed in modern circuit practice. The symbol for a potentiometer is the lower one in Fig 8.

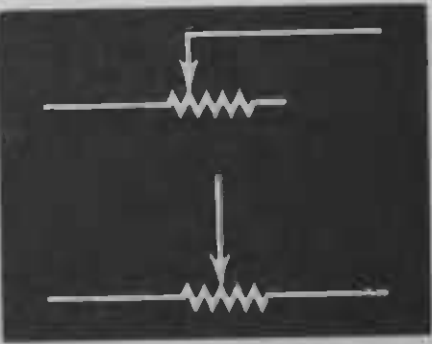


Fig. 7 (below): An array of typical variable resistors. Actually all of the units shown are potentiometers, having three connections. Fig. 8 (right): Circuit symbols for variable resistors.

CHAPTER 3

Batteries and Cells

The electro-chemical battery — primary and secondary cells — the Le Clanche cell — construction and operation — series and parallel combinations — the alkaline cell — the "mercury" cell — the lead-acid accumulator — the Ni-Fe cell — the nickel-cadmium cell.

A battery may be defined as a device which, by chemical action, produces an excess of free electrons at one terminal and a large number of positive ions at the other. The first terminal is negative; the second is said to have a positive charge, because it is deficient in free electrons.

When the two terminals are joined by an external conducting path, the electron pressure or EMF of the battery causes a movement of electrons along this path. In other words, the battery initiates the flow of an electric current through the external conductor.

To be precise, a single electro-chemical unit should be referred to as a cell; an interconnected group of cells makes up a battery.

Investigators have produced quite a variety of electric cells, and a lot of space could be devoted to a discussion of their chemical action and their characteristics. However, for the purpose of this chapter, it seems more appropriate to devote the necessary mental energy to a study of the types normally used with electronic equipment.

It will be time enough to worry about the more specialised varieties, if and when your studies lead you into specialised areas.

Cells fall into two broad classifications: primary and secondary.

Primary cells are intended to operate for a period, then be discarded when they are discharged, or no longer able to maintain the appropriate EMF across their output connections. This being the case, their cost has to be kept as low as possible. Familiar examples of low cost primary cells are those which you put into your hand torch every now and again. Your transistor portable radio operates from a battery made up, most likely from six very compact primary cells.

Secondary cells are normally much more expensive but they have the advantage that, when exhausted, they can be recharged. In fact, with due care, they can be cycled (ie charged and discharged) hundreds of times — a fact which offsets their higher cost. Charging normally involves the use of a special charger, intended to plug into the power mains. A variety of factors may influence the choice of primary or secondary cells for particular applications but secondary cells tend to be favoured where a large amount of current is required and where charging facilities can be arranged without undue inconvenience.

More recently, there has been developed a cell which is not strictly either a primary or secondary cell, but falls midway between. This is the manganese-alkaline cell which, although normally sold as a primary cell, is also available in a modifed form which can be recharged — within limits.

Most primary or secondary cells may be further classified as "wet" or "dry".

These terms are not strictly accurate, but are used to differentiate between those types in which the electrolyte is a free liquid and the container is not completely sealed, and those in which the electrolyte is a jelly or paste and the cell is effectively sealed.

The first really practical primary cell was produced by Georges Le Clanche in 1868. His cells used a zinc rod for the negative pole and a carbon rod, surrounded by a mixture of manganese dioxide and powdered carbon in a porous container, as the positive pole. The whole lot was housed in a glass jar containing a solution of sal-ammoniac (ammonium chloride).

The Le Clanche cell in this form, though an entirely practical arrangement, suffered from the disadvantage of having a liquid electrolyte, subject to spilling and "creeping". These problems were largely overcome when, in 1888, a "dry" cell was produced by Dr Gassner, using the basic Le Clanche principle but in a more convenient form.

The modern "dry" cell is essentially similar to that produced by Gassner. The glass jar is eliminated and the negative zinc electrode becomes the container for the chemicals. The carbon rod and manganese dioxide is retained as the positive electrode, but the sal-ammoniac is thickened to a non-

A range of modern batteries and cells. Large Le Clanche "dry" batteries (rear) are used in lanterns and higher power electronic equipment. The smaller batteries and cells (foreground) are used in low power applications. The two disc-type cells are of the nickel-cadmium type.

spillable paste. A coat of wax or other suitable compound over the top serves to seal the whole unit.

Such a cell is not truly dry and, indeed, could not be so. However, the cell may be used in any position without danger of the electrolyte spilling or doing damage to other apparatus.

Although dry cells have changed little in appearance for many years, various chemical refinements have trebled and even quadrupled their service life.

When first manufactured, the cell builds up a certain potential across its terminals and, thereafter, the chemical action becomes very slight. However, there is always a slow chemical action going on and a drying-out process, so that a cell will not remain fresh indefinitely. After a few months, its terminal voltage begins to diminish and the amount of electrical energy it can supply becomes very limited.

The ability of a cell to remain "fresh" for a long period of idleness after manufacture is expressed in terms of its shelf life. Generally speaking, large cells have a much longer shelf life than small ones.

The main factor controlling shelf life is the purity of the materials employed, particularly the zinc. If these were all 100 per cent pure, most of the shelf life problems would be solved. Unfortunately, this level of purity is impractical on a commerical basis and the battery manufacturer has to be content with something less.

However, against this background battery manufacturers have improved shelf life significantly over the years. In addition, they can design a battery to have either a long shelf life — at the expense of some other characteristic — or with less emphasis on shelf life where other characteristics are more important.

Again, for any one design, the temperature at which the battery is stored will have a significant effect on shelf life. High temperatures shorten it, low temperatures prolong it, so that it is customary to keep batteries in cold storage where the type of battery, kind of usage, and climatic conditions would otherwise create problems.

When, in actual service, a conducting path is provided between the terminals there is an evident electron movement from negative to positive along the path, which tends to relieve the electron pressure within the cell. However, the cell automatically tries to maintain the initial potential between its terminals, so the flow of electric current is invariably accompanied by increased chemical activitiy.

It is only comparatively recently that there has been general agreement on the chemical reactions in zinc-carbon cell, due to its extreme complexity. A popular concept for many years was that of a main energy producing reaction between the zinc and the ammonium chloride, with hydrogen liberated as a by-product.

The formation of hydrogen bubbles on the carbon rod is responsible for a very undesirable cell characteristic called polarisation. Previously, it was assumed that the manganese dioxide overcame this by combining with the hydrogen to form water. More recently, it has been shown that the role of the manganese dioxide is much more direct, in that it prevents the formation of gas, rather than simply reducing any already formed. In addition, it is now known that the manganese dioxide is not simply a depolariser, but an active electrode which contributes more than half the energy which the cell delivers.

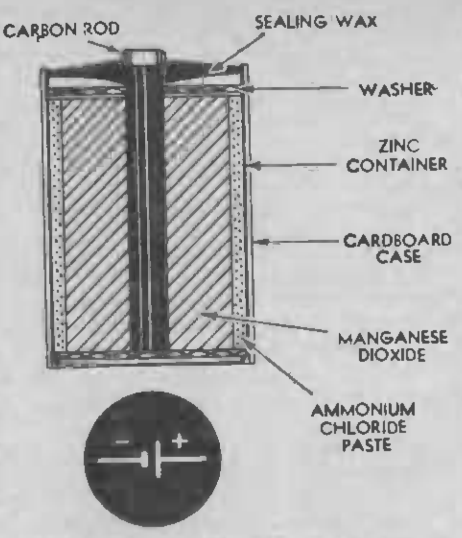


Fig 1. Cross-section of a typical "dry" Le Clenche cell, showing details of its construction. Note the paste form of the electrolyte.

All dry cells which correspond to the foregoing description have an initial EMF of 1.5 volts or thereabouts between their terminals, irrespective of their physical size. However, large cells have a much longer shelf life than small cells, and, in service, can supply much heavier currents without running down.

Because of this, some care should be exercised in selecting the best size cell for a particular application. In general, larger cells are more economical than smaller ones, even though the latter have the advantage of low initial cost and, possibly, greater convenience. On the other hand, it is unwise to use a cell which is so large, relative to the current drain, that its theoretical discharge life far exceeds its shelf life. There is, therefore, an optimum size for any application.

The larger a cell, the greater is the surface area over which chemical activity can be evident. Therefore, such a cell is better able to maintain its electron pressure — or EMF — when called upon to deliver relatively high current.

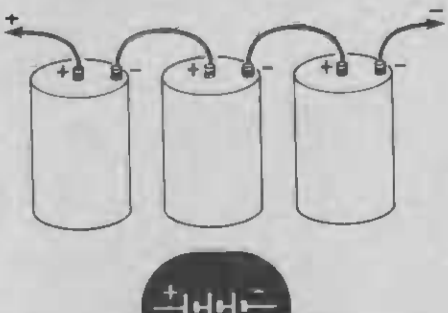


Fig 2. Cells connected in series deliver a voltage higher than that from a single cell. The total voltage is determined by multiplying the single cell voltage by the number of cells.

The loss of EMF due to output current may be regarded very conveniently as an effective internal resistance of the cell itself. This internal resistance is small with large cells, and increases as the size of the cell is diminished.

The largest dry cell in common usage measures about 6in in height and 2in in diameter, and has a shelf life of two or three years — sometimes much longer. It is designed for a normal intermittent load of between 200 and 300 milliamps and is frequently used in telephone circuits.

The cells used in the larger type of "radio battery", used in the valve-type portable radios of a few years back, were designed to handle normal loads up to only 15 milliamps or so, but gave a useful life on intermittent service of 500 to 700 hours.

The cells in the "energiser" type of batteries used in transistor equipment are designed along somewhat different lines to the usual type of cell, to suit them to the type of loading conditions imposed by transistor circuitry. When used with a piece of equipment which draws the current values for which they are intended, such batteries have about the same useful life as the "B" batteries mentioned above.

Small single cells of the "penlight" variety will deliver only 8 or 10 milliamps comfortably and would give only a small fraction of their service life if asked to supply greater currents for any length of time.

A sectional drawing of a typical dry cell is shown in Fig 1, together with the usual schematic symbol. As may be seen, a single cell is represented by two parallel lines of unequal length. As a rule the longer stroke signifies the positive side of the cell, but it is occasionally found the other way round. To be on the safe side, most circuit draughtsmen clearly mark the polarity with a plus and a minus sign, as shown.

For many purposes the EMF or voltage available from a single dry cell is often inadequate and a higher voltage is obtained by connecting two or more cells in series. That is to say, they are linked together with the positive terminal of one cell going to the negative terminal of the next, and so on. Any number of cells may be connected in this way, and the two terminals left over are the positive and negative connections to the bank of cells — or the battery, as it is generally called.

The most familiar illustration of this principle is that afforded by the ordinary two or three-cell torch or "flashlight".

In ordinary torch cells the tip of the positive carbon rod protrudes through the wax and is fitted with a brass contact cap. The negative connection is actually the bottom of the zinc can. When the cells are pushed into the torch case they are automatically connected in series and the resultant EMF is equal to the sum of their individual voltages. For two cells it is approximately 3.0 volts, for three cells about 4.5 volts, and so on. The connection is illustrated graphically and schematically in Fig 2.

For the sake of convenience, battery manufacturers often assemble a number of cells in a single package or carton, connecting the cells together internally and bringing out just the one negative and the one positive lead. The actual voltage depends mainly on the number of cells contained.

The method of representing such batteries schematically is worth noting. There is an obvious objection to drawing a large number of individual cells, so that circuit draughtsmen normally show a couple of cells at each end, with a dashed line between them and a figure signifying the total terminal voltage.

For certain applications, voltages may be required intermediate between the negative and the full positive EMF. Standard procedure, in this case, is to provide the battery with one or more intermediate terminals, connected to the appropriate point in the series network of cells.

In the days when valves were widely used in battery-operated equipment, multiple-cell batteries were produced which gave quite high terminal voltages — typically 45V and 90V. These were commonly known as "radio B batteries". There were also smaller batteries with intermediate terminals, used for grid biasing, and known as "C batteries".

To make them as compact as possible for valve type portable receivers, B-batteries were often made up using a "layer building" form of construction, in which the cells were formed by successive flat layers of zinc, carbon, electrolyte-soaked cardboard, and manganese dioxide.

It is well to note in passing that two or more cells may alternatively be connected in parallel, the positive and negative terminals being joined together as shown in Fig 3.

With this arrangement, the EMF between the terminals remains at 1.5 volts (approx), but the load current is shared between the total number of cells. Because each cell is then called upon to deliver only one half, one third, one quarter, etc. of the total load, the bank of cells can handle a proportionately higher current than any one of the cells singly. A similar current capacity, with possibly longer shelf life, could be obtained from a single very large cell, but the parallel arrangement is sometimes used for convenience or expedience.

A modifed version of the Le Clanche cell is the alkaline cell, which uses an alkaline electrolyte instead of the acidic ammonium chloride. This gives a better high-rate performance and a cell which can be more readily manufactured in smaller sizes. The smaller sizes are generally for low current drain applications only, although an "inside out" construction (ie, with the positive plate connected to the outer case) has been designed, with potassium hydroxide as the electrolyte, for higher current drains.

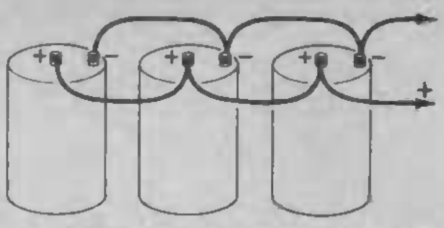


Fig 3. Cells connected in parallel deliver greater current with a less significant loss of terminal voltage. This assumes that the cells concerned all have the same nominal terminal voltage.

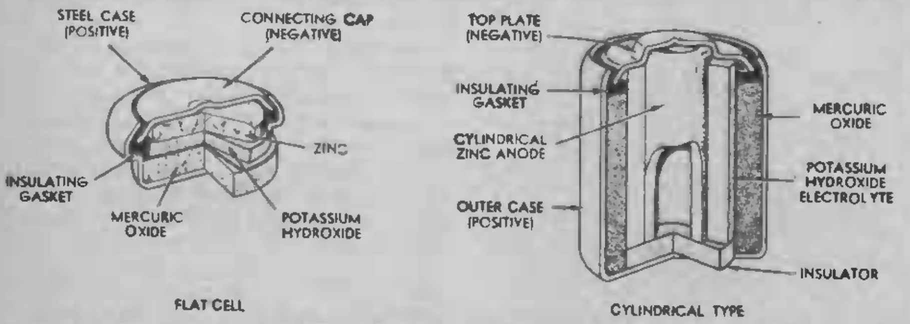
The energy volume ratio of alkaline manganese cells is higher than the acidic Le Clanche cell and they have a longer shelf life. They may well eventually replace the Le Clanche for general commercial use but, at present, are more costly to produce.

Another type of primary cell is the so-called "mercury" cell, which has been used for some years in hearing-aids and other continuous-use devices, and is finding use in some of the newer transistor equipment (Fig 4).

The mercury cell uses zinc as the negative electrode, as before, but the positive electrode is of mercury. Mercuric oxide is used both as a source of the mercury and as a depolariser. The electrolyte is a strong solution of potassium hydroxide and zinc oxide in water.

Mercury cells are commonly made "inside out" — that is, with the negative (zinc) electrode brought to the top as a cap. This enables a steel case to be used, both to connect to the positive mercury electrode and to act as a container.

They may also be made in the "flat" variety, for stacking in series to produce voltages higher than that of a single cell. The voltage of a single mercury cell is approximately 1.35 volts, slightly lower than the Le Clanche type, but the internal resistance is generally very low in comparison, and mercury cells have a very high capacity. This makes them more suited to applications where the battery cannot be replaced very often, and to circuits which draw high currents for short periods.



FLAT CELL

CYLINDRICAL TYPE

Fig 4. Cross-sectional views of two types of Ruben-Mallory mercury cells. These primary cells have very high capacity and low internal resistance. Note the "inside out" construction, compared with the familiar Le Clanche cell; the steel outer case is the positive terminal.

In all the cells we have described so far, the electro-chemical action necessarily ceases when the chemicals have been used up, and the cells are then of no further use.

These so-called primary cells are distinguished from secondary cells by the fact that the latter can be made to deliver current afresh by a process called recharging.

The most common type of secondary cell or battery is the lead-acid accumulator used in motor vehicles. These were also used as filament of "A" batteries in the earliest battery radio sets.

In the lead-acid accumulator, two sets of lead plates are immersed in a mixture of sulphuric acid and water. To "charge" this type of cell the plates are connected to an external EMF in such a way that a current of up to several amperes flows through the acid via the plates.

The flow of current causes the acid-water mixture to be decomposed, releasing atoms of hydrogen and oxygen at the respective plates. The hydrogen mostly escapes, but the oxygen liberated at the positive plate combines with the plate to form the dark brown peroxide of lead.

When the charging has proceeded for an adequate length of time, the battery may be removed from the charging circuit.

If a conductor is then connected between the two plates, a current is found to flow, and, as it does so, a chemical change becomes evident in both plates. The positive plate, coated with lead peroxide begins to take a coating of greyish lead sulphate, and much the same thing happens to the lead negative plate. The "sulphate" molecule comes from the acid, which is therefore progressively weakened as the discharge continues.

When next the cell is recharged, the lead sulphate on the negative plate is reduced to pure spongy lead, and the coating on the positive plate reverts to the dark coloured lead peroxide. At the same time, the strength of the acid is built up again, so that the cell is ready once more for a period of discharge. This charge and discharge cycle can be repeated hundreds of times until, after a period of perhaps several years the active electrodes begin to disintegrate.

The EMF of a lead-acid secondary cell is normally reckoned at 2.0 volts, irrespective of its size the actual current handling capacity is governed largely by the number, the size, the spacing and the nature of the plates.

For voltages greater than 2.0, it is usual to mount several cells, side by side, in the one large subdivided container. The cells are sealed across the top to prevent spilling of the acid and then connected permanently in a series arrangement, as already described for dry cells.

In the early days of domestic battery-powered radio receivers, accumulators delivering 2, 4 and 6 volts were quite common but nowadays the most common voltages are 6 for use in older style cars and 12, as used in most modern vehicles. The voltages mentioned involve the use of 1, 2, 3 or 6 cells respectively.

Because the strength of the acid is affected by the chemical condition of the

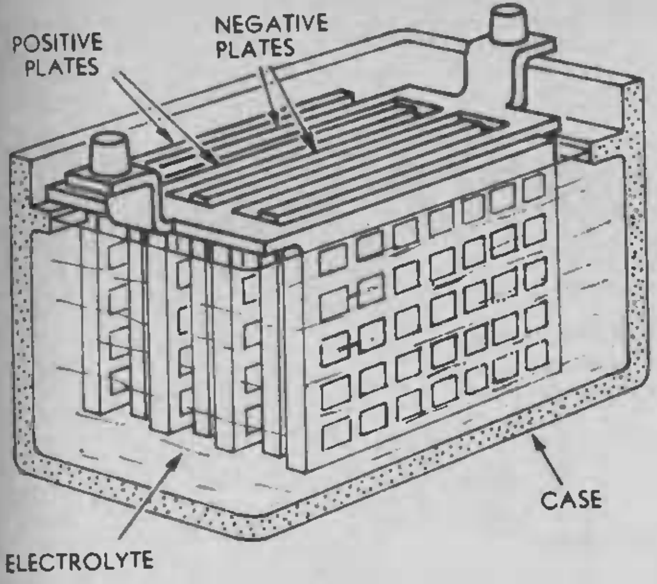


Fig 5: A single lead-acid secondary cell, showing the alternate lead plates supported in the acid electrolyta. Three or six such cells are used in the familiar automotive battery to obtain 6 or 12 volts, respectively.

Fig 6: An hydrometer, used for assessing the state of charge of a lead acid cell. Some of the electrolyte is drawn up into the barrel. If the battery is fully charged, the specific gravity of the electrolyte is high and the float rides high in the liquid. When the battery is discharged, the reverse applies.

plates it may be taken as a good guide as to whether the cell is fully charged or otherwise.

In new cells for car or radio use, it is usually arranged that the specific gravity of the acid be about 1.25, when fully charged (often quoted as 1250), 1.18 when half-charged and 1.11 when discharged. (Specific gravity has effectively the same meaning as "density compared with that of water, which is taken as 1.0.")

Where cells are intended for industrial or home lighting purposes, acid of somewhat lower specific gravity is often specified in the interests of long service life.

To measure the specific gravity, a device known as a hydrometer is used. This is rather like an oversize syringe with a small glass float in the barrel.

The acid is drawn up into the barrel and specific gravity is indicated by the depth to which the float submerges in the acid. The upper section of the float is usually calibrated directly in terms of specific gravity, with "Full" and "Discharged" zones indicated for general guidance.

The EMF of a fully charged cell is often as high as 2.2 volts to begin with but rapidly drops on load to almost exactly 2.0 volts. There it remains for a long period, falling to 1.8 volts at very low charge.

Generally speaking an accumulator should not be discharged further when the voltage per cell is less than 1.8, or when the specific gravity is less than about 1110 (1.11, strictly). Even if it is not needed immediately, it should be recharged to avoid it being left idle for lengthy periods in a discharged condition.

The current capacity of an accumulator is summed up in its ampere-hour rating, which gives a rough idea of the number of hours it will continue to deliver a specified amount of current after a full charge.

The rate at which an accumulator is discharged does not appear to be important as far as the life of the battery is concerned.

The same is not true of charging, and, in general the charging rate should never be such that it causes excessive temperature rise or gassing. When a battery is in a discharged condition it can safely be charged at a high rate, since most of the charging energy is employed in producing chemical change and very little is wasted in generating heat or gas. As the charge progresses there is an increasing tendency to generate heat and gas, and the charge rate should be decreased, or "tapered," to keep these within safe limits.

In a modern motor car the charging system is designed to behave in exactly this manner. If the battery is well discharged the system will charge it, initially, at a very high rate, up to 50 amps in a six-volt system. As the charge progresses, the rate decreases until, at full charge, it may be no more than 1 amp. Fast chargers used in service stations are — or should be — designed to observe the same precautions.

As well as the lead-acid accumulator, there are a number of other secondary cells used for special applications. One of these is the **Nickel-Iron (Ni-Fe)** cell, sometimes called the Edison cell, after Thomas Edison, who is credited with its invention.

The Ni-Fe cell is a complex structure, the positive electrodes being rows of perforated nickeled-steel tubes filled with alternate layers of nickel hydroxide and thin flakes of pure nickel. The negative electrode is a grid made from nickeled sheet steel containing pockets filled with iron oxide. The electrolyte is a 21 per cent solution of potassium hydroxide in distilled water, plus a small percentage of lithium hydrate.

The voltage of this cell is 1.37 when fully charged, dropping to 1 volt when discharged. A greater number of cells are therefore required for a given battery voltage. The specific gravity of the cell does not vary appreciably from charge to discharged conditions, and the voltage is usually taken as a guide to the condition of charge.

The cells are a good deal more costly than the lead-acid type, but offer a number of advantages in certain applications, particularly for electric traction vehicles and similar heavy-duty service. They are a very robust cell, and tolerant to badly arrenged charging procedures — even to accidental reverse charge — which would quickly ruin a lead-acid cell. The working life may be 20 years or longer.

A more recently developed secondary cell is the **nickel cadmium type**. Although initially deerer than the lead-acid cell, it has many advantages which make it attractive for modern, compact, electronic equipment.

Among its advantages are a very favourable power to weight ratio, and a far greater tolerance to neglect than the lead-acid cell. Also, it is designed to provide a completely closed chemical cycle, whereby gas evolved during charging is re-absorbed within the cell. Thus the cell can be completely sealed. It is finding increasing use in portable equipment, such as electronic photo-flash, transceivers, portable radio, etc.

Another application is in self-contained electrical appliances. Electric shavers end electric drills are two examples of this in which these cells make possible a compact, convenient device which can be used anywhere, independently of power lines, or where the use of high voltage equipment may be dangerous.

CELLS THAT BURN FUEL . . .

Readers may have noticed occasional references in technical literature to fuel cells. While involving chemicals they differ fundamentally from batteries.

In a battery, the production of electricity involves a chemical reaction between the electrodes and electrolyte, as explained in this chapter.

Basically, a fuel cell involves combination of a fuel (hydrogen, gases, alcohol, etc) with an oxidising agent. It is, in fact, a kind of flameless combustion, in which the electrodes and cell structure play only a secondary role. Basically the cell will continue to operate for as long as it is supplied with fuel and oxygen.

Despite intensive research, fuel cell development has been held up by various technological problems and the principle has had only limited application outside the laboretory.

CHAPTER 4

Magnetism, Inductance and AC

Magnetic fields and currents in conductors — conductors in coil form — coils with air and iron cores — inductance and mutual induction between two coils — direct and alternating current — the alternator — inductive reactance and its effects — phase angle — practical inductors.

At some time or other, most of us have toyed with a horseshoe magnet and noted the mysterious attraction it had for iron and steel. We have also seen how a bar magnet, when suspended so that it can swing freely in the horizontal plane and away from the influence of other magnets, comes to rest with one pole pointing to the North and the other to the South. The pole pointing North is called the North-seeking pole of the magnet — usually abbreviated to North pole — while the other pole is a South-seeking pole or South pole. This phenomenon is the basis of the magnetic compass.

If a pocket compass is held close to a wire carrying heavy direct current, the pointer is deflected from its normal due-north setting. From this simple experiment, the fairly obvious deduction is that the wire must be surrounded by a magnetic field.

The exact nature of a magnetic field is a matter of much debate, but it is convenient to think of it as so many **lines of force**. The stronger the magnet or the current in a wire, the stronger the associated magnetic field and the more numerous the lines of force. Alternatively, we may say that the **magnetic flux density** is greater.

The direction of lines of force can be seen in Fig 1. Every line passes out from the North pole of a magnet, makes a complete circuit through the surrounding medium and returns into the South pole and through the magnet to the North pole again thus completing the magnetic circuit. The direction of the lines of force surrounding a conductor or a coil is determined by a specific law but there is little point in the beginner committing it to memory at this stage.

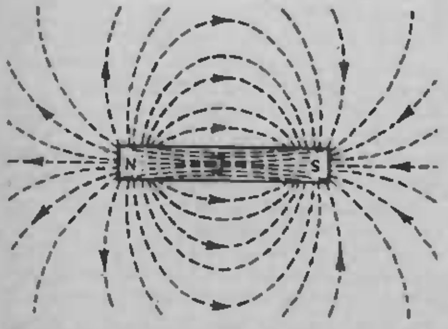


Fig. 1. A bar magnet showing the direction of lines of force.

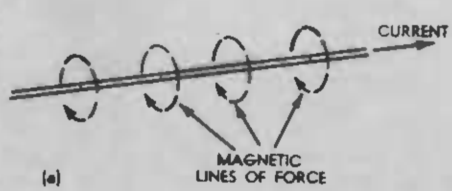
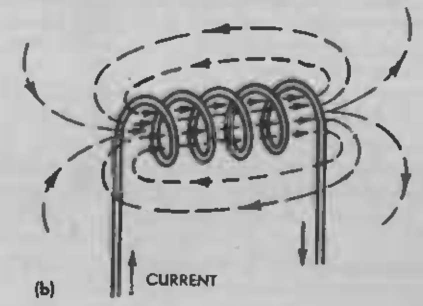


Fig. 2. (a) Lines of magnetic force about a current-carrying conductor. (b) When formed into a coil, the magnetic field becomes more concentrated.



In actual fact, the magnetic field surrounding a single, straight, current-carrying conductor (Fig 2a) is not very great, but it becomes much more evident if the wire is wound in the form of a coil. The magnetic field assumes the shape indicated in Fig 2b, the lines of force all passing through the centre and completing the magnetic circuit via the much greater space outside the coil. Obviously, the magnetic concentration is much greater within the coil than outside.

Due to its magnetic properties, a core of iron, nickel, cobalt or an alloy of these metals, when placed within a coil, can greatly increase the magnetic flux density for a given current through the coil. To be a little more technical, we may say that the core has greater **permeability** to the magnetic flux than air; the term "permeability" may be roughly defined as the flux multiplying property of a substance.

According to particular requirements, a core may consist of a simple iron rod, a bundle of iron wires, a stack of iron strips, or a piece of ceramic material (embedded with iron dust) moulded into any desired shape. If the core is straight, as shown in Fig 3, it assumes the characteristics of a bar magnet in the presence of a steady current through the coil. In other words, it has north and south poles; it will attract iron objects; and it will align itself with the earth's magnetic field if there is no mechanical hindrance.

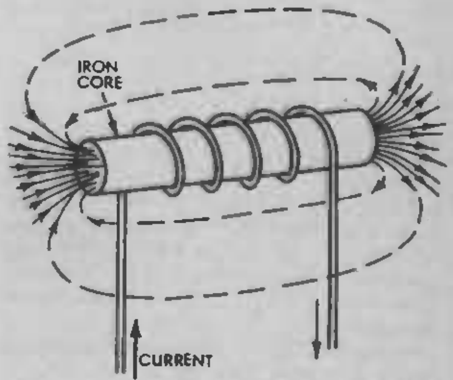


Fig. 3. A current-carrying coil wound about an iron core. The high permeability of the core increases the flux density for a given current through the coil.

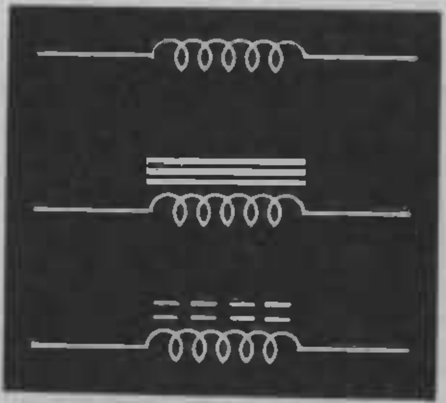


Fig. 4. Schematic symbols for air, iron and ferrite-cored inductors.

Note should be taken of the schematic symbols for air and iron-cored coils, shown in Fig 4. In the case of an iron core, the core lines are drawn either within or alongside the coil symbol, depending upon the "school" of the circuit draughtsman. The third symbol, with dashed lines, is for "iron dust" and ferrite (ceramic) cores.

With a soft iron, iron dust or ferrite core, the magnetic effect is dependent on the strength and the direction of the current through the coil. If the direction of the current is reversed, then the magnetic poles are interchanged. If the current ceases, nothing but a slight residual magnetic effect remains.

Magnetic effects in a core due to current through the associated coil are used in a variety of radio and electronic equipment, as we shall see in later chapters.

However, it is interesting to note that cores of steel and certain other alloys will retain their magnetic properties even after the energising current is switched off. In other words, they become permanent magnets.

Where a core is extended outside the coil but within the magnetic circuit, the lines of force tend, for the most part, to follow the easier path through the iron, and the magnetic flux in the free air surrounding the coil is correspondingly reduced. This is particularly the case if the core is extended to form a complete **magnetic circuit**.

If a small gap is left in the magnetic circuit, an intense magnetic field exists across it. This fact is used to effect in numerous electromechanical devices.

The important points to grasp in all this are: (1) That a magnetic field is evident around a conductor carrying an electric current, and (2) that this magnetic field can be controlled to serve particular purposes.

As a simple experiment, let us assume that we have a battery and a switch, so arranged that an electric current is passed through a coil. We can see this in Fig 5 which, by the way, is the first wholly schematic drawing we have used.

When the switch is closed, the current does not rise immediately to the value determined by the battery voltage and the resistance of the coil, but is found to build up to this value over a definite time interval. This rate of increase is determined by the size and shape of the coil and the number of turns in the coil. With a very large iron-cored coil, the rate is low enough to be seen on an ordinary meter (set up to read current) if one is connected into the circuit.

This effect cannot be duplicated with a purely resistive circuit, where the current rises to its ultimate value almost instantly.

Fairly obviously, the coil must offer some initial and temporary opposition to the flow of current by a property quite distinct from the resistance of the wire. To cut a rather long story short, this opposition is produced as part of the process of building up the magnetic field around the coil. The very action of building up the field produces in every turn of the coil what is known as a **counter EMF** — an induced voltage which is in opposition to that from the battery. This counter EMF is evident only while the magnetic field is in the process of building up, but while present it exercises a definite retarding effect on the action.

Once the current has reached its ultimate value — as determined by the battery voltage and the resistance of the coil — the magnetic field remains stationary and constant while the switch remains closed. At this point the counter EMF ceases to exist.

When the switch is opened, one might expect the EMF to disappear from across the terminals of the coil, but, for sheer stubbornness, the magnetic field must rival the classic mule. The opening of the switch is the signal for the magnetic field to begin collapsing, which is to be expected, but in the very process of doing so another EMF is generated briefly across the coil of the same polarity as the original battery!

This magnetic effect is just as much a property of the coil as its resistance. It is defined as **Inductance**, and is measured in units known as **Henries**. Thus the inductance of a coil may be defined as the property of the coil which causes it to oppose any change in the current flowing through it.

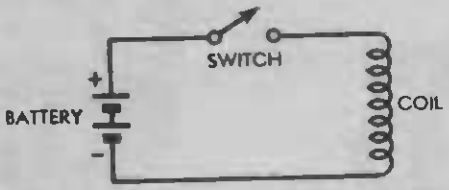


Fig 5. The schematic diagram of a simple circuit to direct a current through an air-cored coil.

A coil is said to have an inductance of one Henry if one volt is induced across its terminals when the current through the coil is changing at the rate of one ampere per second.

Smaller units of inductance are the millihenry (equal to one-thousandth part of a Henry) and the microhenry (equal to one-millionth part of a Henry).

In the form of a general statement, while resistance is the opposition to the flow of current, inductance is the opposition to a change in the flow of current. If there is zero initial current, then the coil offers temporary opposition to the building up of a current in either direction. If there is an initial steady current, then the coil opposes any tendency for it to increase or diminish. It would be wise to read those last two statements again. They are very important.

A coil of wire designed specifically to exploit these effects is known as an **inductor**.

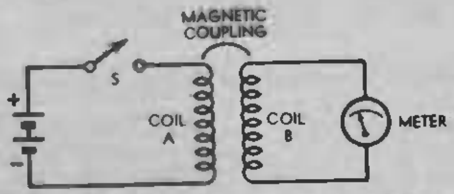


Fig 6. Changing currents in coil A induce EMF's in coil B, as a result of mutual magnetic coupling.

An interesting phenomenon can be demonstrated by mounting two coils in close proximity, so that the magnetic flux from one coil can link with the turns of the other.

Assume, as in Fig 6, that coil A is connected through a switch to a battery and that a sensitive indicating meter is connected across the terminals of Coil B.

On closing the switch, the meter pointer will be found to move slightly — say to the right — and then return to its normal position. If the switch is now opened, the pointer will show a slight movement to the left, demonstrating that, somehow or other, an EMF has been induced in the second coil.

The explanation is found in a general law that an EMF is always induced in a conductor, which is intersected by moving lines of force, or, yet again, in a conductor which itself moves through a stationary magnetic field. In the illustration of Fig 6, the closing of switch "S" causes a magnetic field to build up around coil A and, for a brief interval, lines of force are evident moving outwards from the coil. These moving lines cut through the turns of coil B and induce in it a voltage, which is made evident by the meter.

Once the magnetic field has become stationary, no further voltage is induced until such time that the circuit is broken and the field begins to collapse. The reverse movement of the lines of force then causes an induced EMF of opposite polarity to appear in coil B.

A similar effect could be produced by passing a powerful bar magnet rapidly through the centre of a coil, or by sliding the coil to and fro over a stationary magnet.

During the time the magnet and coil are in close proximity and there is movement of the coil cutting lines of force within the magnetic field, small voltages will be evident between the terminals of the coil.

An interesting condition arises if the switch in Fig 6 is opened and closed rapidly for a sustained period of time. Under these conditions, the magnetic field around the two coils is in continual movement, building up one instant, collapsing the next. The result is an induced EMF in coil B, and a resulting current through it (and the meter) which are continually changing in polarity. At one instant, current flows from right to left; the next instant from left to right, and so on.

Let us assume, for argument's sake, that the two coils are quite large and intimately coupled together, and that heavy currents are caused to flow through them; further, that the meter M is replaced by a lamp. Each time the switch is opened or closed, we can expect the induced current to cause a momentary glow in the lamp, so that it flickers off and on at a rapid rate.

Carrying our assumption a little further, let us step up the rate of opening and closing the switch to such an extent that the individual surges of current through the lamp can no longer be discerned, and it appears to glow steadily. We then have a state of affairs where useful work is being done — lighting of the lamp — not by a steady direct current from a battery, but by an ever-changing current produced by the interaction of a conductor and a changing magnetic field.

The electron movement which is capable of producing heat in the lamp filament is no longer a steady flow of electrons through a conductor, but an oscillatory or to-and-fro motion.

In this simple way, then, we have been introduced to **alternating current** which is of tremendous importance in the fields of electronics and electricity.

Before pursuing the subject further, it is as well to be quite sure of the difference between **direct and alternating** current. In the first case, electron movement is always in the one direction. It may be large or small, or subject to continuous variation

IADE IN
FAIRCHILD

MADE IN
FAIRCHIL

DE IN
AIRCHILD

MADE IN
FAIRCHI

- ✱ ICs—LINEAR AND DIGITAL IN CERAMIC AND PLASTIC PACKAGES
- ✱ TRANSISTORS—PLASTIC, METAL CAN, FETs AND POWER DEVICES
- ✱ DIODES—SMALL SIGNAL, HIGH CONDUCTANCE, HIGH SPEED AND ZENER DIODES
- ✱ MICROWAVE AND OPTOELECTRONIC TRANSISTORS, LIGHT EMITTING DIODES, PHOTOTRANSISTORS, DISPLAYS AND HYBRID RF AMPS
- ✱ ELCO CONNECTORS—AUSTRALASIAN DISTRIBUTORS FOR THE ELCO RANGE OF PRINTED CIRCUIT AND RACK AND PANEL CONNECTORS
- ✱ G.E.—VICTORIAN AND N.S.W. DISTRIBUTORS FOR THE G.E. RANGE OF SCRs, TRIACS, RECTIFIERS AND METAL OXIDE VARISTORS

PRODUCT
DATA
SERVICE
DELIVERY

FAIRCHILD AUSTRALIA PTY. LTD.
420 Mount Dandenong Road, Croydon, 3136. 723 4131
Sydney: 37-43 Alexander Street, Crows Nest. 439 4355
Adelaide: North Terrace House, 19 North Terrace, Hackney. 42 4482
Auckland, N.Z.: 1 Gordon Road, Otahuhu. 6 9299

Distributors & Stockists:
Victoria: Radio Parts Group, 562 Spencer St., W. Melbourne. 329 7888. J. H. Magrath & Co. Pty. Ltd., 208 Lt. Lonsdale St., Melbourne. 663 3731. Warburton Frankl, 220 Park St., S. Melbourne. 69 0151.
N.S.W.: Geo. Brown & Co. Pty. Ltd., 247 Sussex St., Sydney. 29 7031. General Accessories, 443 Concord Road, Rhodes. 73 0211. Stockists: Dick Smith Manufacturing Co., 10 Atchison St., St. Leonards. 43 1414. Martin de Launay Pty. Ltd., Cnr. Clarence & Druitt Sts., Sydney. 29 5834.
A.C.T.: Electronic Components Pty. Ltd., 3 Pirie St., Fyshwick. 95 6811.
Queensland: Douglas Electronics Pty. Ltd., 322 Old Cleveland Rd., Coorparoo. 97 8222.
S.A.: General Accessories, 81 Flinders St., Adelaide. 23 4022. Hills Electronics Telefix, 10 Port Rd., Hindmarsh. 46 3832. Gerard & Goodman Pty. Ltd., 192 Rundle St., Adelaide. 23 2222.
W.A.: Warburton Frankl, 41 Great Eastern Highway, Rivervale. 61 8888.
N.Z.: John Gilbert & Co. Ltd., Tasman Building, Anzac Ave., Auckland. 3 0839. Stockist: Tisco Ltd., Wellington, Christchurch, Dunedin, Hamilton, Hastings, Palmerston North.

FC5910

but, so long as it remains uni-directional, it is referred to as a direct current, or DC.

An alternating current or AC is one in which the actual direction of electron movement is changing continually, or in periodic fashion. The amplitude of the electron movement may, of course, be large or small at any particular instant, depending on the circumstances which initiate it.

An alternating current cannot be produced directly by electro-chemical action, as in a battery, but is always generated by relative movement of a conductor and a magnetic field. An elementary type of AC generator or alternator is illustrated in Fig 7.

Imagine that a loop of wire is mounted so that it can rotate between the poles of a permanent magnet; further, that the two ends are brought out to mutually insulated copper bobbins, which make rubbing contact with a pair of fixed metal or carbon rods known as **brushes**.

The magnetic field surrounding the loop of wire will be evident as lines of force between the faces of the two magnet poles. When the loop commences to rotate, the movement relative to the lines of force will immediately initiate a flow of current through the wire. Thus, as the left-hand side of the loop (as illustrated) moves upwards with respect to the South pole of the magnet, there will be a movement of electrons in one direction along the wire. At the same instant, the other side is moving downwards with respect to the North pole, so that the electron movement in this half is in the same general direction around the loop. As a result, an electron pressure — an EMF — becomes evident between the ends of the loop and is thence transferred to the two brushes.

At the instant when the plane of the loop is exactly parallel with the lines of force — that is, in the horizontal position — the loop cuts through the magnetic flux at the highest rate, so that the generated voltage reaches a peak value. As the loop continues to rotate, the relative movement becomes less, so that the generated voltage is also reduced.

At the instant when the loop is in an exactly vertical plane, the sides are moving parallel to the lines of force and no voltage whatever is generated. Beyond this point, the loop again approaches the horizontal position, so that the voltage or EMF rises once more. However, the electron movement through the loop is now found to be in the reverse direction; or we may choose to say that electrons are accumulated at the alternate brush. Whichever way we look at it, the simple result is that the generated voltage which appears between the brushes is reversed in polarity.

Following the movement further, the current reaches a peak in this direction, then falls away to zero, immediately building up again with the original polarity. This process goes on just as long as the loop continues to rotate.

The matter can be made clearer with the aid of a simple graph. Assume that we can observe the instantaneous value of voltage or current generated over a brief period of time, and that values of voltage (or current) are plotted against corresponding time intervals. On joining the plotted points we would have a curve similar to that shown in Fig 8.

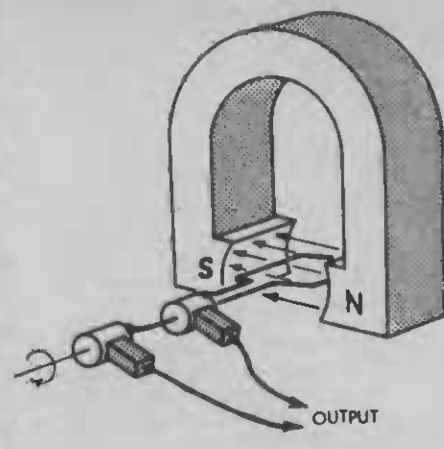


Fig 7. An elementary AC generator or "alternator".

Assuming a zero reference line, the voltage would rise smoothly to a peak value, then diminish and pass through zero to an equal peak negative value; then back to zero again and so on, "ad infinitum". The plane of the loop, relating back to Fig 7, is suggested by the small arrows beneath the curve.

A practical alternator is actualily a far more elaborate device than Fig 7 might lead one to believe, but the operating principles are the same.

In practically every country of the world, large banks of alternators are ceaselessly producing power for factories, homes, street lights and a dozen other purposes. Rotating machines of rather different type are employed to generate direct current for similar purposes, but AC has proved easier to handle and distribute on a large scale; the specific reasons need not concern us here.

Direct current finds its greatest application these days for trains, trams and other forms of electric traction, and a DC installation on a small scale is found in every automobile and aeroplane — in fact, anywnere a battery and generator are in direct association.

It is just as well at this stage to memorise certain common terms which have to do with alternating current. A wave pattern such as is shown in Fig 8 is referred to as a **sine wave**, because it has a definite relationship to the sine of an angle — something familiar to all who have studied trigonometry.

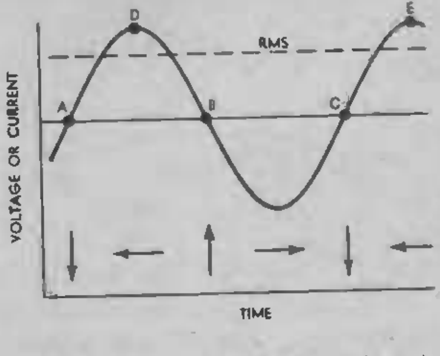


Fig 8. Output voltage plotted against the position of the coil in the magnetic field.

We will not interrupt your train of thought by enlarging on the matter, but remember that a sine wave represents alternating current or voltage in its purest form. If irregular peaks, dips or bumps appear in a wave pattern, it is no longer a pure sine wave.

The maximum value attained by the voltage or current in either direction is known as the peak value. However, this peak value is a purely instantaneous one and, at all other times, the current or voltage is of lesser magnitude. The peak value, therefore, can scarcely be regarded as a fair indication of the ability of the current to produce secondary effects, such as lighting the filament of a lamp or raising to red heat the element of a household radiator.

The ability of an alternating current to do useful work, in this sense of the term, is expressed by what is known as the **RMS** value of the current (or voltage). This RMS (or root-mean-square) value has a definite mathematical basis and, for a sine wave, it is equal to .707 times the peak value; in other words, it is a little under three-quarters of the peak value of the wave.

Unless otherwise stated, figures of AC voltage or current are always assumed to refer to the RMS value. Therefore, the application of an alternating potential of 6.0 volts RMS to the filament of a lamp will cause it to glow with exactly the same brilliance as 6.0 volts DC, as from an electric battery or DC generator.

To express the magnitude of an alternating voltage or current, we use the same units as for DC: amps, milliamps, microamps; volts, millivolts, microvolts.

It is necessary to bear in mind, however, that the stated RMS value basically summarises the heating or energy value of the alternating wave over a period of time. The value of the current or voltage at any particular instant may be greater or less than this, depending on which particular portion of the wave train we happen to pick.

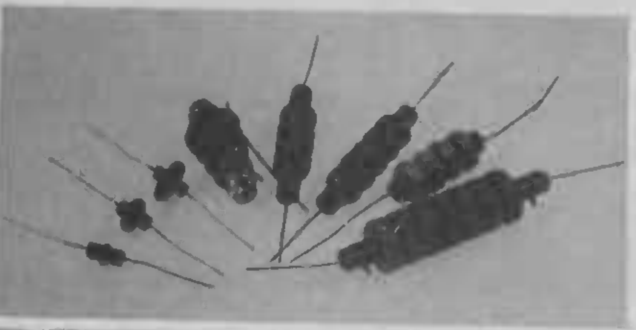
Referring again to Fig 8, the space between points A and C is known as one complete cycle. From point D to E is also one cycle, since they are exactly equivalent points on the curve and represent one complete rotation of the simple alternator.

By the same reasoning, from A to B, and from B to C is exactly one half-cycle, representing half the time interval of a complete cycle and a half rotation of the alternator loop. From A to D and from D to B are each one quarter cycle.

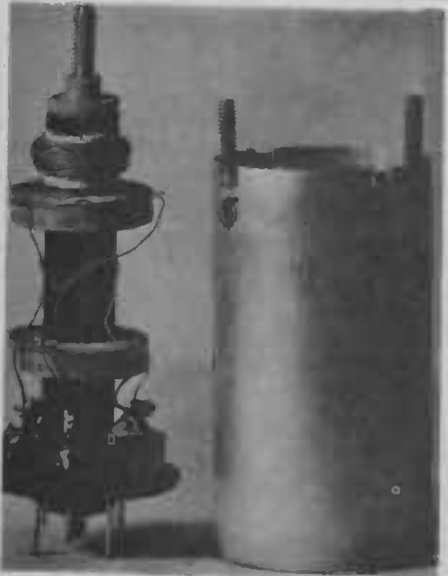
Of vital importance is the **frequency** of an alternating current or voltage, which is the number of complete cycles occurring in one second of time. The accepted unit of frequency is the Hertz, equivalent to one cycle per second. The usual abbreviation is Hz.

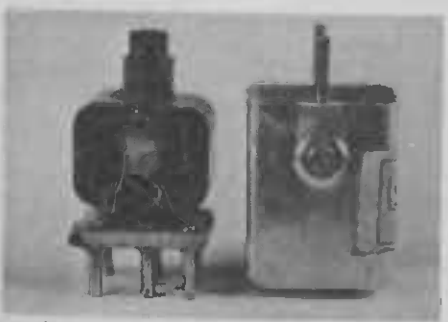
In the simple alternator of Fig 7 the frequency is obviously dependent on the speed with which the loop is rotated, or the number of revolutions per second.

The frequency of the alternating current



A group of radio frequency chokes or "RFCs", which are often used in radio receivers, TV sets and other equipment. Each is basically a coil of wire on a ferrite or iron-dust core.





Coils for "radio" frequencies (around 1 Megahertz) reproduced about normal size. The larger unit (left) is of early design, while the smaller (below) is a modern type. Both have adjustable ferrite cores and aluminium shield cans.

power mains is not standardised throughout the world but it is usually either 40, 50 or 60 Hertz, depending on the particular country or area in question.

In most cases, the adopted frequency is maintained very accurately by the electricity supply authorities, who maintain a constant check against standard time signals. The widely used synchronous clocks depend for their accuracy on the frequency of the power mains.

The vista of frequency is almost unlimited, as we shall see later in our discussions of sound waves and radio broadcasting. So while you are busy forming a mental picture of an alternating wave, leave room for an enormous range of frequencies and for the peculiar ways the higher frequency voltages are liable to behave.

The final step in this chapter is to correlate the earlier discussion of inductance and our more recent discovery of alternating currents. Remember that a coil of wire exhibits an initial and temporary opposition to the passage of an electric current but, once this current is maintained for a brief period of time, the coil offers just the same opposition to any tendency for the current to increase or decrease. As you will recall, the property is referred to as the inductance of the coil.

It is, therefore, not difficult to imagine just how an inductance reacts to the passage of an alternating current, which is continuously changing in value and direction. Fairly ovbiously, the coil is going to offer virtually a continuous opposition, quite distinct from and additional to that of the DC resistance of the wire.

Furthermore, since the opposition involves the building up and collapsing of a magnetic field, it is fairly obvious that the opposition will increase in proportion to the speed with which the current tries to build and collapse the field; that is, with rising frequency.

The opposition offered by a coil (ie, by its inductance) to the passage of current is referred to as its **reactance.** Like resistance, it is measured in ohms, but, as we have seen, its value for a given coil is directly related to the frequency of the current or voltage concerned. Thus, a given coil may have a reactance of 1000 ohms at one frequency and 2000 ohms at twice that frequency. In this respect, it differs from pure OC resistance which is a substantially fixed and definite quantity for DC or for AC of any frequency.

Inductive reactance differs from pure resistance in another vital aspect, stemming from the opposition it offers to the flow of an electric current. When a voltage is applied across a resistance, the flow of current commences instantaneously and ceases immediately the voltage is removed. Thus, when the voltage applied is alternating, the current reaches its peak value at the same instant as the voltage peak, and goes through the same variations in value.

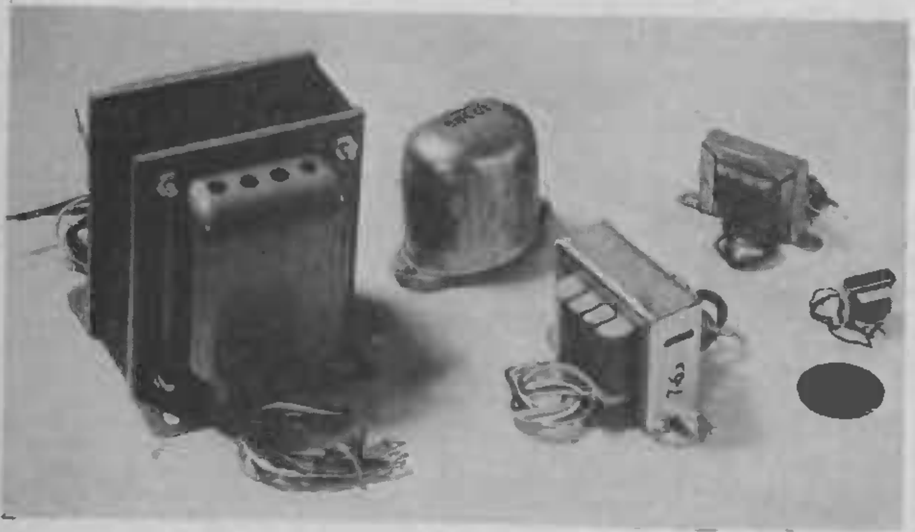
In the case of an inductor, the current builds up more gradually, taking a definite time to reach a peak value. Thus, when the applied voltage is alternating, the current does not reach its peak value until a measurable time after the voltage peak. It goes through the same cycle of alternations, but the peak and minimum values are attained somewhat later than the corresponding voltage variations. The current and voltage are said to be out-of-phase one with the other — the current lagging behind the voltage by a definite amount. Alternatively, it may be said that the voltage leads the current.

With a purely inductive circuit, the difference in phase actually amounts to a complete quarter-cycle.

Inductance plays a vital role in electronic circuits generally. Among the simplest examples are inductors involving just a few turns of stiff wire, wound to form a self-supporting coil, and looking not unlike a spring. In other coils, turns of insulated wire are wound side by side on a former of plastic or other such material, or layer upon layer to constitute what is often referred to as a pie-wound coil. Sometimes a powdered iron or ferrite core is provided inside the coil former, usually with the object of achieving a given order of inductance and efficiency from a component of smaller dimensions.

Inductance is also fundamental to the operation of transformers and chokes, as pictured at the top of this page.

Chokes have the property of passing direct current quite readily but, because of their inductive properties, they oppose the



A selection of iron-cored transformers and chokes of the type used in the power supplies of domestic electronic equipment. The windings are commonly on plastic bobbins, around which are assembled the laminated iron cores. Clamping brackets complete the assembly.

flow of alternating or rapidly varying current. This particular property of inductors was referred to earlier in the chapter.

A transformer normally has two or more windings, wound side by side or one over the other, in such a way that any alternating magnetic field associated with either coil reacts upon the other. The word "alternating" in the previous sentence should be especially noted, because transformers are devices which operate only in relation to alternating voltages and currents. (Look again at Fig. 6)

Transformers intended to operate at the frequency of the power mains (40-60Hz) or at audio frequencies (up to 15000Hz and beyond) are constructed with a laminated iron core and have the general appearance of the components as pictured at left. With frequencies higher than this again, adequate magnetic coupling between adjacent windings can take place without need for a heavy laminated core.

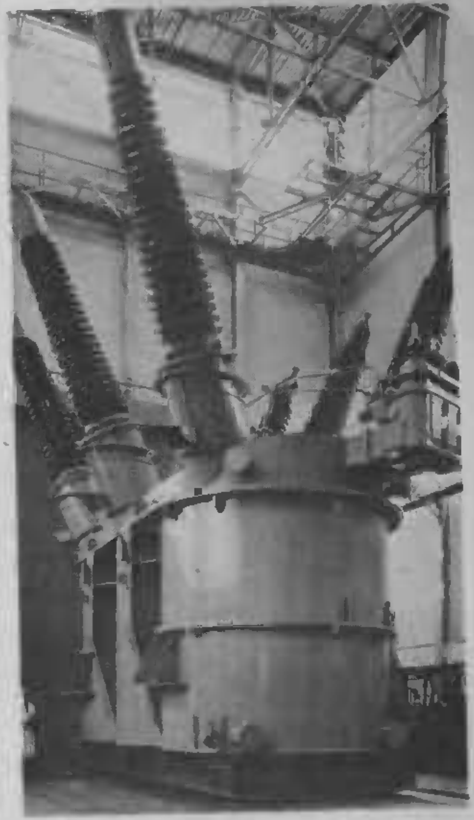
What happens in a transformer is basically this: when an alternating voltage is applied across one winding, alternating current flows through it. This sets up a magnetic field around the winding which alternates in strength and polarity, in sympathy with the current which produces it. This alternating magnetic field sets up corresponding voltages across any other windings present within the field. These induced voltages are greater or smaller than the original voltage according to the ratio of turns between the windings concerned.

Transformer action is probably best illustrated by visualising a transformer intended for connection to the power mains. The AC power main is connected to the primary or input winding, causing current to flow through it and creating an alternating magnetic field in the laminated iron core. The alternating field induces corresponding voltages in any other so-called secondary windings which are present. If the primary winding happened to have 1000 turns and a particular secondary 100 turns, the latter would deliver 100 / 1000 or one-tenth of the input voltage; therefore with 240 volts AC across the primary winding output from the secondary would be 24 volts AC.

Generally, power transformers used in electronic equipment include several secondary windings, each delivering a particular voltage that may be required, and which may be higher or lower than the original 240V AC input. Transformers which have to deliver high current from the secondary windings—and which therefore have to handle high power — need to have larger cores and be wound with thicker wire than transformers which have to handle lower power.

Transformers vary in size from tiny units — the size of a pea — to huge devices like that illustrated. They are particularly useful components.

We shall discover more about coils, chokes and transformers as this course progresses.



A very large high voltage power transformer capable of handling 2,500 amps at a voltage of 400,000V. Note the men at lower left, showing the size of the unit.

CHAPTER 5

Capacitance and Capacitors

Capacitance and capacitors — charging current, and storage of energy — the definition of the Farad, unit of capacitance — mica capacitors — tubular capacitors, paper and plastic types — ceramic capacitors — temperature coefficient of capacitance — electrolytic capacitors — variable capacitors and "trimmers".

To understand just what capacitance is, it is necessary to recall an earlier discussion on conductors and insulators. It was pointed out that an atom consists of a positive nucleus, around which revolve one or more planetary electrons. Since these outer planetary electrons are not always intimately bound to the nucleus, an externally applied EMF can cause a definite drift of electrons from one atom to the next along the length of the material. This is the basic theory supporting a "conductor".

The electron movement is always toward the positive pole of the applied EMF, in accordance with the general law that "unlike charges attract, like charges repel". In a vacuum, where there are no atoms present, this migration of electrons obviously cannot occur. In materials known as "insulators", it is difficult to move these electrons because they are strongly bound to their parent atoms. However, the application of an EMF to a vacuum or an insulator does produce an interesting effect.

Consider Fig. 1, where two plates of a metal conductor are placed near and parallel to one another and connected to the opposite terminals of a battery. Because of the chemical reaction within it, the battery tends to remove electrons from atoms of plate A and add electrons to plate B, making A positively charged with respect to B. A current thus tends to flow, despite the gap in the circuit.

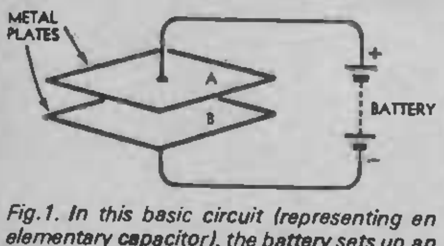


Fig.1. In this basic circuit (representing en elementary capacitor), the battery sets up an electron unbalence between the plates and an electric field in the gap.

This current, in turn, causes a "state of tension" in the gap between the plates, which tends to resist the electron flow. The "state of tension" is called an electric field, and it eventually sets up a voltage across the gap which opposes the EMF of the battery and causes the current to cease. In other words, even though the gap between the plates does not conduct a current, it permits a current to flow for a brief period in the rest of the circuit, until a state of equilibrium is reached.

This effect is known as capacitance and the circuit components involved are capacitors. An older term commonly used was condensers.

The initial current which flows in the connecting wires, when an EMF is applied to a capacitor, has a strength or magnitude which depends upon the value of the capacitor. In turn, the capacitance is dependent upon the surface area of the plates, the spacing between them, and any insulating material or dielectric which may be introduced between them. But more of this later on.

Capacitors, in a practical sense, consist of plates made from metal sheet, foil or coating; between them may be a vacuum or dielectric material, such as air, waxed paper, mica, plastic, ceramic or a metallic oxide.

When an EMF is applied to a capacitor, as we have seen, a current flows for a brief interval and an electric field is set up in the gap between the plates. If the capacitor is now disconnected from the battery, the electric field remains across the gap. Referring to Fig. 1 again, plate A is left positively charged with respect to plate B, having less electrons than normal, while B has more than normal. As far as the gap is concerned, nothing has altered with the removal of the battery, so that the gap maintains a voltage difference equal to the battery voltage.

A capacitor in this condition is said to be charged. There is a voltage across it and an electric field between the plates.

However, if the two charged capacitor plates are joined by a conductive path by touching the wires together, the excess electrons on plate B will move along the conductor to plate A and correct the state of unbalance. The electric field across the gap will collapse and the voltage will disappear. The capacitor is then said to have been discharged by this second flow of current.

From this it may be seen that a capacitor has the ability to store electrical energy, as it will retain a voltage across its terminals and can supply a current for a period if connected to a conducting circuit.

If a capacitor is connected across a source of fixed voltage, the current flows for only a brief interval of time and then ceases. However, if the applied voltage is varied at a given rate over a definite, interval of time, the current through the circuit is found to vary in sympathy with it. On this is based the definition of capacitance.

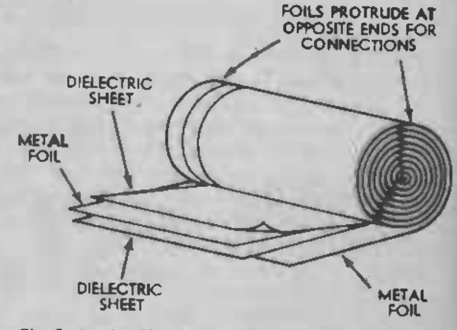


Fig.2. In the "tubular" type of capacitor, the plates are formed from two long strips of metal foil, rolled up between peper or plastic.

When a change of one volt per second in the EMF across a capacitor produces a current of one ampere, it is said to have a capacitance of one farad.

Just as resistance is measured in ohms, and inductance is measured in henries, capacitance is measured in farads.

In actual practice, the farad is too large a unit for ordinary radio and electrical work, and the unit most commonly used is the microfarad, which is one-millionth of a farad. Typical capacitors employed in radio receivers range from something like 2000 microfarads down to .00005 microfarads, although still larger and smaller units are not uncommon.

To shorten and simplify things, various abbreviations for the word microfarad are commonly employed. One is the purely Anglicised abbreviation "mfd". Another follows technical practice in using the

Greek letter mu(µ) to mean one-millionth, this being coupled to a capital F for Farad, giving the term µF. Typesetting difficulties may cause the English letter u to substitute for the Greek µ, giving the symbol "uF".

To save unwieldy decimals, small capacitance values are usually written in millionths of a microfarad, or picofarads. Thus .0001uF becomes 100 picofarads, and so on. The usual abbreviation for picofarad is "pF". An earlier term meaning the same as picofarad was "micromicrofarad", abbreviated "uuF" or "mmF".

The value of a capacitor depends upon several factors, as we have mentioned before. Firstly, the total area by which the plates actually overlap; capacitance increases directly with area. Secondly, capacitance varies inversely with the distance between the plates; the smaller the separation, the greater the capacitance. Thirdly, the dielectric or material in the space between the plates also has a marked effect on the capacitance; a vacuum is the reference standard, but air, impregnated paper, mica, plastic, ceramic or glass increases the capacitance substantially.

The ability of an insulator to multiply the capacitive effect between capacitor plates of given area and spacing is expressed by its dielectric constant. Thus, if the space between the plates is occupied by an insulator having a dielectric constant of 5, the capacitance between the plates is five times greater than if the intervening space was occupied only by a vacuum.

We have assumed, thus far, that only two plates are involved in a capacitor. However, in some capacitors a number of plates are used. Generally, the plates are interleaved, with alternate ones connected together. The total capacitance then becomes the sum of each of the individual capacitors formed by adjacent surfaces of the plates.

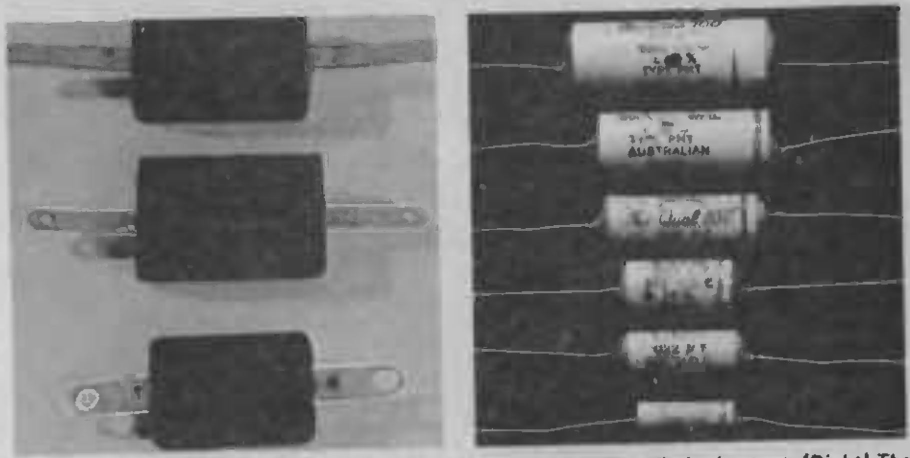
In producing a capacitor of any given value, it is necessary to consider the amount of space which the component will occupy, the mechanical rigidity of the unit and also the voltage which it can withstand before breakdown.

If the spacing between the plates is too small, or the dielectric material unsuitable, the stress set up when an EMF is applied may be so great that the dielectric could be punctured, allowing a flow of current and rendering the capacitor unserviceable.

In the past, capacitors rated up to about .001uF (1000pF) were usually of the so-called mica type. Small strips of thin metal were sandwiched between pieces of mica insulation, alternate metal strips being joined together at the ends. The whole assembly was held together with a metal clamp, or sealed with wax into a ceramic case. Many manufacturers moulded the completed assembly into a bakelite case. Connection to the two sets of plates was provided by terminals or short copper wire leads commonly called "pigtails".

Different values of capacitance were obtained by varying the size and number of the plates, and the thickness of the insulation between them.

Mica capacitors are still used nowadays, but to a limited extent. Plastic and ceramic dielectric units are often preferred in many of the applications where the mica type was formerly used. We will discuss the plastic and ceramic capacitors a little later on.



(Left) The older mica type of capacitor is still used today but to a limited extent. (Right) The paper capacitor, as the name suggests, uses impregnated paper as the dielectric which is then placed between two thin metal ribbons and rolled together in the form shown.

Ordinary mica capacitors are comparable in size to a postage stamp and up to ¼in thick. They will operate safely with applied voltages up to about 500, and their capacitance value is usually stamped directly on the case. Some brands, however, employ a colour code to distinguish one capacitance from another.

For capacitance values above about .01uF, the bulk of a mica assembly begins to multiply unduly and a more convenient and economical form of capacitor is usually employed; this is the paper or plastic "tubular" type. (See Fig. 2)

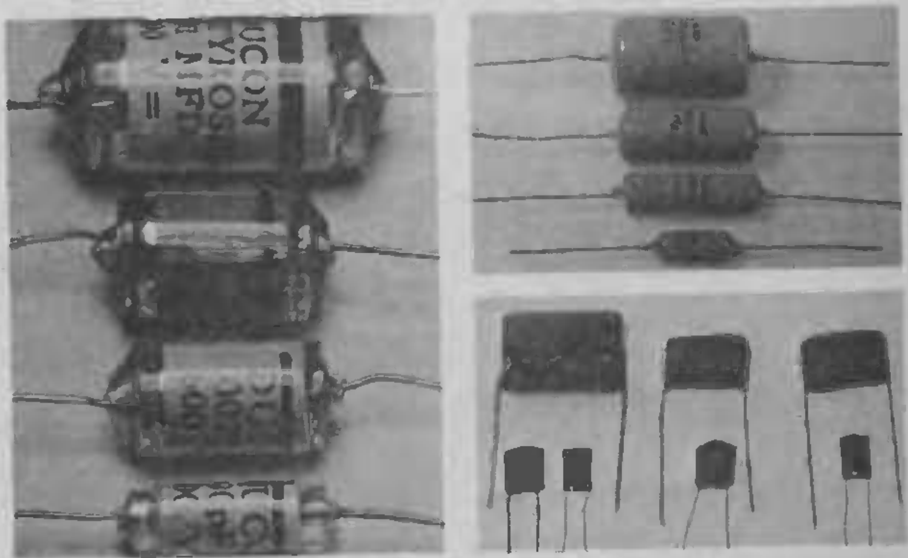
Two thin metal ribbons (usually aluminium) are rolled tightly together with a special impregnated paper or plastic film separating them. The two strips of foil serve the same purpose as the metal plates previously mentioned, while the impregnated paper or plastic film is the dielectric. The capacitor is usually wound with the respective foils overlapping the paper at each end. The overlapping portion of the foil is pushed into a compact mass and maintained in contact with a pigtail for connection to the external circuit.

Squeezing together the ends of each foil ensures a short path between each pigtail and all portions of its metal foil. If connections were made only to one end, each foil would resemble a flat coil and exhibit an undesired inductance, as well as the required capacitive effect.

The ultimate capacitance depends on the length of the foils, the distance they overlap and the characteristics of the paper or plastic dielectric.

Tubular capacitors are generally sealed with a special wax or plastic material into round bakelite tubes, or coated directly with plastic. In size, they may be up to two inches or so long and one inch in diameter.

They are stamped or printed with their nominal capacitance values and maximum DC voltages which they will withstand across their terminals. Most ordinary tubular capacitors are rated to operate with a maximum voltage of about 400 DC but ratings as low as 30 or as high as 2,000 volts are not uncommon. They are made up in a variety of values between about 100pF



Plastic dielectric capacitors (above left) are similar in construction to paper types but with a plastic film (in this case polystyrene) substituted for paper. At top right is a range of plastic capacitors with polyester as the dielectric. Those at bottom right also use polyester as the dielectric but, instead of metal foils interleaving the polyester film, the film is metallised.

and 0.5uF, the particular value usually being marked on the case.

In recent years, a particular type of tubular capacitor known as the "metallised film" type has been introduced. In these, the metal foils are not separate from the dielectric strips, but are actually thin metallic film deposits on one surface of each strip. When the strips are rolled together as before, the films again act as the "plates" being separated by the layers of dielectric. The main advantage of this process is that it allows a significant reduction in capacitor size for a given capacitance value and voltage rating.

It is possible to make metallised film capacitors with capacitance values up to 5.0uF at a voltage rating of 200V whose size is sufficiently compact to allow the normal "pigtail" construction. Formerly, a component of this value was so large that most were sealed into rectangular metal cans designed to bolt to the equipment chassis. Large paper and plastic dielectric capacitors intended for use at high voltages still use this construction. In some cases, a number of capacitors are mounted in a common case.

Plastic dielectric capacitors are of similar construction to paper types but, as the name suggests, they employ a thin plastic film as the dielectric material. They are often given names like "polyester", "polystyrene", "polycarbonate" or "mylar", to signify the particular type of plastic film actually used.

Plastic dielectric capacitors are made in approximately the same values as paper types. However, they are more stable and reliable than paper, owing to the chemical characteristics and moisture-resistance of the plastic dielectric. For a given value of capacitance and a given voltage rating, plastic capacitors also tend to be somewhat smaller than equivalent paper capacitors, particularly if the metallised film technique is used.

Ceramic dielectric capacitors are different again in construction from paper, plastic and mica capacitors. This is due to two factors. One is that fired ceramic material, of the type used by these capacitors as the dielectric, is rather brittle, and cannot be stressed, or made in the form of a thin film to be rolled up with the metal foil "plates".

The other factor is that the ceramic material can be made with an extremely high dielectric constant, which means that, in any case, it is not necessary to adopt the multi-plate or rolled-foil type of construction, in order to produce high values of capacitance. A flat ceramic disc with silver "plates" fired on to each side, or a hollow tube with a deposited silver surface inside and out is usually sufficient. (See Fig. 3.)

By using the same basic construction and varying both the size and the dielectric constant of the ceramic, capacitors ranging from 1pF up to 1uF may be produced, in voltage ratings of from 1 to 20,000 volts. The smallest ceramic capacitors are about 3/16in in diameter and 1/32in thick, while the largest may be more than an inch in diameter and ½in thick.

The high-capacitance types use a ceramic material which has a very high dielectric constant (around 6000). Two different types of high-capacitance ceramic capacitors are commonly available, being known as "HI-K" and "ultra-cap" units. The latter type is a comparatively recent development and is mainly suitable for circuit applications where the voltages are below 25 volts. The range offers a high capacitance in a very compact form.

Due to the effects of heat upon both the physical dimensions and the characteristics of the dielectric material, most capacitors exhibit a capacitance change with variations in temperature. The amount of capacitance change which occurs for a variation in temperature of one degree C, expressed as a fraction of the nominal value of the capacitor, is called the **temperature coefficient** of capacitance.

The temperature coefficient of mica, paper and plastic capacitors cannot be manipulated to any great extent and is usually kept as small as possible during manufacture. With ceramic capacitors, at least in the smaller values, it is possible to produce almost any desired value of temperature coefficient — over a fairly wide range of both positive (capacitance increases with temperature) and negative (capacitance decreases with rising temperature) values.

This makes ceramic capacitors extremely useful in a variety of circuits, where temperature variations in other components would otherwise cause the circuits to change their function with temperature. By using a ceramic capacitor which has the opposite kind of temperature coefficient, it is possible to arrange for a degree of cancellation so that the circuits as a whole are less affected by temperature variations.

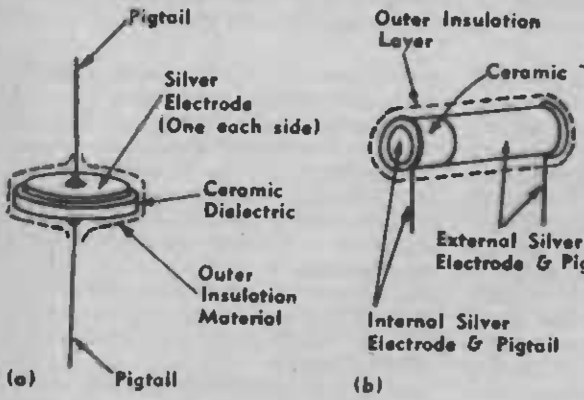
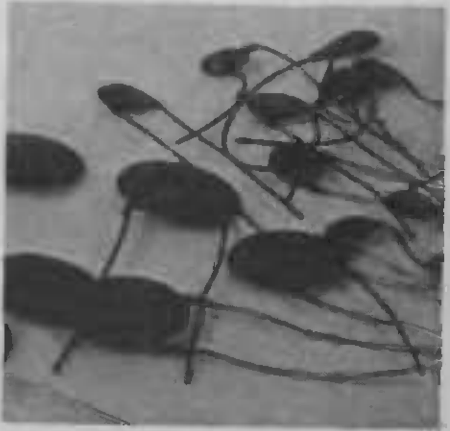


Fig 3 (Left): Construction details of typical ceramic capacitors. Basically, each consists of a two-plate capacitor formed by silver electrodes separated by ceramic material, either in flat disc or cylindrical form. (Right) A selection of disc-type ceramic capacitors.

Besides being classified in terms of capacitance and voltage rating, therefore, ceramic capacitors are also classified in terms of the value and polarity (positive or negative) of their temperature coefficient. Types having a positive temperature coefficient are commonly given a code number having as a prefix the letter "P", while the letter "N" is used to signify a negative coefficient. For example, "N750" signifies a negative temperature coefficient of 750 parts per million per degree C, or 0.075%. If the construction and ceramic material have been arranged to give a zero temperature coefficient (no variation with temperature) the capacitor is usually marked "NPO".

The ceramic used for very high value capacitors — "HI-K" types — is less suitable for this form of manipulation and normally has quite a high temperature coefficient. This fact, coupled with a natural tendency for individual units to vary considerably from the nominal value, usually upwards, means that these are unsuitable for use in any circuit where the value of capacitance is critical.

Modern circuit practice often calls for capacitors having very high values — up to 4000 microfarads or more — and capable of operating with applied voltages to 60 or 70 volts DC. With valve type equipment less capacitance is required — up to say 200uF — but at voltages up to about 600 DC.



To produce rolled-paper, plastic-film, or ceramic capacitors to these specifications would be a costly job, besides which the finished article would be very bulky indeed. This difficulty was overcome by the development of the so-called **electrolytic** capacitor, which makes use of the fact that aluminium and some other metals form extremely thin layers of an insulating oxide, when immersed in certain electrolytes.

An electrolytic capacitor, in its most common form, consists of two shaped aluminium plates immersed in a liquid electrolyte, itself a conductor of electricity. During the process of manufacture, the capacitor is connected to a source of direct current and formed. In other words, the aforementioned layer of oxide is caused to form on one aluminium electrode.

Once the forming process is complete, the capacitor ceases to pass appreciable current, since the oxide coating forms an insulating layer between the aluminium electrode and the electrolyte in which it is

A group of modern electrolytic capacitors. The largest unit has a capacitance of 8000uF and a voltage rating of 75. The smallest units at the front have much lower capacitance ratings — typically 100uF or less, as well as lower voltage ratings.

immersed. The aluminium is, therefore, one electrode, the oxide coating is the dielectric, and the electrolyte and the other aluminium plate together form the second electrode. The thickness of the coating and the voltage which the capacitor will withstand is determined by the method of manufacture.

Since the dielectric is produced and maintained by electro-chemical action, it is necessary to use an electrolytic capacitor with the electrolyte and unformed plate connected to the negative side of the voltage source, and the formed electrode to the positive side. This formed electrode is often referred to as the "anode". Ordinary electrolytic capacitors should not be used in reverse polarity, or connected to a purely AC circuit.

The oxide coating is actually a less efficient insulator than, say, ceramic or paper, so that a certain small leakage current usually has to be tolerated. The coating and the electrolyte are also likely to deteriorate after a few years' service, necessitating replacement. But for all their disadvantages, electrolytic capacitors are widely used, because they make possible large capacitance values with economy in both space and cost.

In old-style can-type electrolytic capacitors, the electrolyte was contained as a liquid within a sealed metal can — usually aluminium; this can provided electrical connection to the electrolyte. The positive aluminium electrode — or anode — was supported centrally within it, being roughened, fluted, crimped or otherwise treated to present the maximum possible surface area to the electrolyte.

Liquid-filled can-type electrolytics have long since given place to "tubular" types, employing an aluminium foil rolled with an absorbent material soaked in electrolyte. These "dry" electrolytic capacitors often resemble the larger paper tubular types, although others are assembled within small cans for mounting on the chassis. Electrolytic capacitors are distinguished usually by the description on the container, by the comparatively high capacitance rating and by the "plus" and "minus" markings at either end.

While the majority of electrolytic capacitors use aluminium as the anode material, tantalum has emerged recently as an alternative with some advantages. The anode of a tantalum electrolytic is typically produced in the form of a blob or rod by vacuum sintering. Being highly porous, it presents a large internal surface area.

By introducing an electrolyte into the porous structure and forming with the passage of an electric current, a microscopic layer of tantalum can be produced which forms the dielectric between the electrolyte and the metallic tantalum. When provided with connections and sealed in epoxy resin, finished tantalum capacitors look more like ceramic types than electrolytics. In fact, they are sometimes referred to as "solid electrolytics".

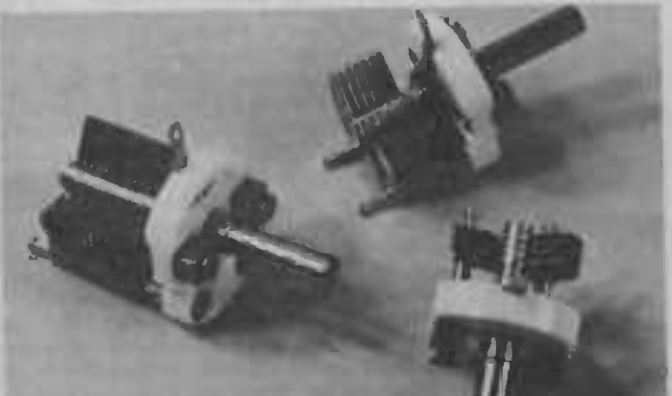
In general, tantalum capacitors can be manufactured to closer tolerances than aluminium types. They are more stable, exhibit much lower leakage and are more tolerant of reverse or alternating voltages. They do tend, however, to be more expensive.

Typical capacitance values range from about 0.1uF to 300uF, while operating voltages range from just over 1 volt to about 35 volts.

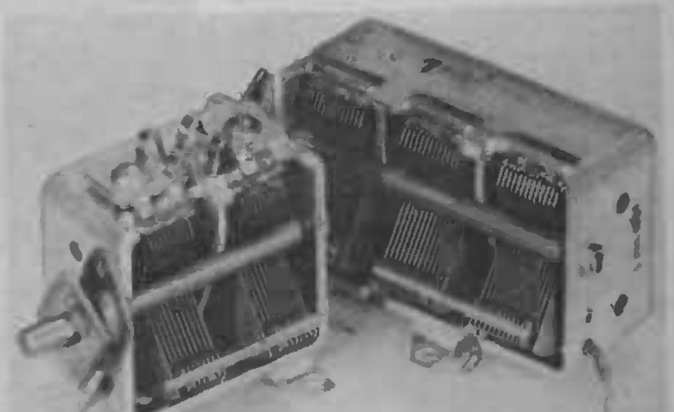
An almost universal feature in the tuning portion of radio receivers and in other pieces of electronic equipment is one or more variable capacitors (or variable condensers) whose capacitance value is capable of easy adjustment. The most common type employs a number of fixed and a number of moving vanes which mesh with one another to a greater or lesser extent.

The fixed vanes, which are rigidly mounted on a supporting framework, are referred to as the stator plates. The moving vanes, locked on a revolving shaft, are known as the rotors.

A variable capacitor must be well constructed mechanically, so that the two sets



Typical single-section variable capacitors, of a type often used in transmitters and other specialised equipment.



A 2-gang and a 3-gang variable capacitor similar to those used for station selection in broadcast band receivers.

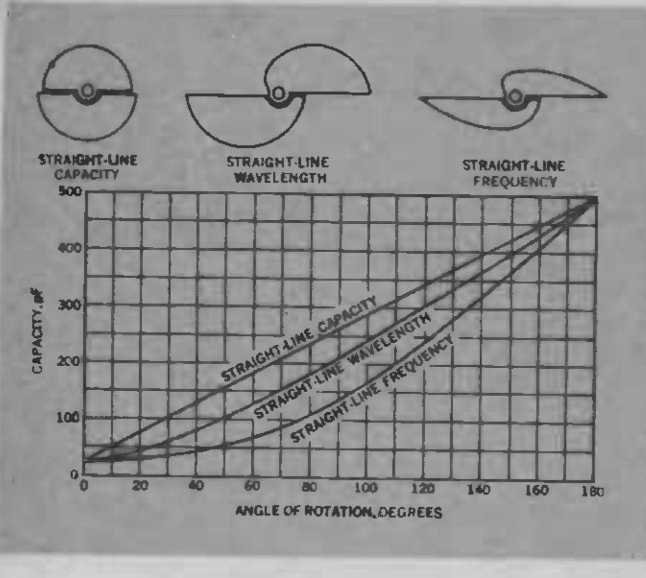


Fig. 4 (left): It has been common practice through the years to shape the plates of tuning capacitors so that stations will be spread more evenly across the tuning scale, either in terms of wavelength (centre curve) or frequency (bottom curve). With semi-circular plates, stations tend to seem crowded towards the minimum capacitance end of the rotation.

of plates will not foul one another at any position.

The plates of variable capacitors are shaped in various ways to obtain a specific relationship between capacitance and the number of degrees through which the control shaft is rotated. One purpose of this is to spread the stations more evenly across the tuning dial of a completed radio receiver.

Variable capacitors may be made singly, but two, three or four units are often mounted within the one framework and on the one shaft, so that they operate together. Multiple units coupled together in this way are commonly called "gang capacitors" or "tuning gangs".

Ganged capacitors commonly used in the tuning circuits of radio receivers generally have a maximum capacitance of about .0004uF (400pF) but only about .00002uF (20pF) in the fully-open or minimum position. This capacitance is achieved with anything from 17 to 23 plates in all, the ultimate capacitance depending on the number of plates, the area enmeshed, the spacing between them, and whether or not they use a dielectric other than air. Some of the smaller units have thin sheets of plastic between the plates, to give a higher capacitance.

Smaller variable capacitors of the same general pattern are used for special purposes. Some employ a smaller number of conventional-sized plates; others use a large number of small plates, comparable in size to half a 20-cent piece.

Some circuit applications call for capacitors which are capable of only occasional adjustment within certain relatively narrow limits. Variously described as semi-variable or semi-fixed capacitors, the most common are the mica compression type, the air-concentric type, and the ceramic type.

The mica compression types incorporate two or more metal leaves, insulated with mica and normally held apart by spring tension. Adjustment of a screw forces the leaves closer together, increasing the capacitance. These trimmer or padder capacitors are comparable in size to a postage stamp, and about one quarter-inch thick. They are widely used to effect adjustments made necessary by the use of ganged tuning capacitors.

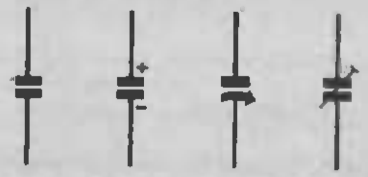
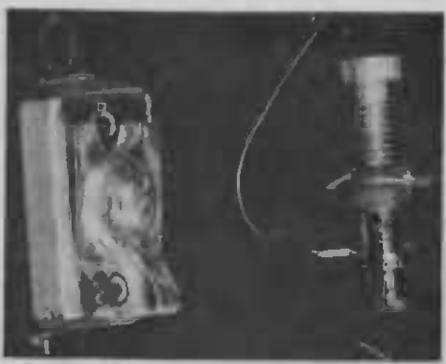
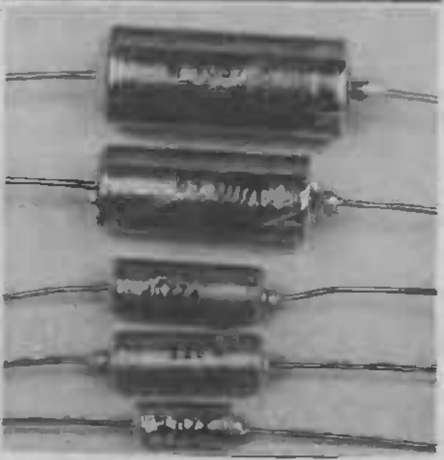


Fig 5(Above): Schematic capacitor symbols. From left, a fixed type, an electrolytic, a variable, and a trimmer.

Below: At left is a mica compression type trimmer, while at right is a ceramic tube type.





A range of tantalum capacitors is pictured above. Although more expensive, they offer advantages over the aluminium types.

For critical purposes, mica compression capacitors are not above reproach, which is why the various other types of trimmer capacitor have been devised.

The air-concentric type uses air as the dielectric, as the name would imply, and has two sets of concentric metal tubes — one set of which may be moved along a threaded shaft so that its tubes enmesh with, but do not touch, the other set.

Ceramic trimmers are made in a number of different forms, but the most popular consists of a hollow, internally threaded ceramic tube which has a metallic electrode fired on to the outside and a screw which is threaded into the tube. Adjusting the length of screw inside the tube varies the capacitance between it and the outer electrode.

Another type of trimmer which has recently appeared on the scene is a plastic-film dielectric type, whose construction is basically a miniaturised version of a normal variable capacitor. Small semicircular plates are used for both the stator and rotor, with thin discs of plastic film between them. The rotor shaft is slotted to permit adjustment using a screwdriver.

Before concluding the chapter, it is appropriate to mention one other very special kind of variable capacitor, defined in full as a **voltage dependent capacitor** or VDC. A coined word which is tending to be used as a generic term is **varicap**.

Voltage dependent capacitors have emerged largely as a spin-off from the technology surrounding semiconductor materials, (which have already been mentioned) and transistors (which have yet to be covered.)

In a voltage dependent capacitor, there is a **junction** or **interface between** two conducting materials which have a slight but very deliberate difference between their atomic structure. When a voltage is applied across the two conductors, free electrons within the atomic structure of each react differently to the electric field. What actually happens is that the free electrons tend to retreat from the junction or interface, so that it becomes a virtual insulating zone.

What results is a device with two conductors abutting one another, but with an artificially created insulating zone between them. The device, in fact, has the essential elements of a capacitor.

Since the movement of electrons from the Junction zone — and therefore the thickness of the zone — is sensitive to the applied voltage, one ends up with a device which behaves like a capacitor but which is variable in value according to the magnitude of the applied voltage.

In practical terms, voltage dependent capacitors can be made to cover much the same range as the type of variable capacitor just discussed and they are tending to be regarded for some applications as alternatives to the mechanical component.

Over and above this, they are particularly useful in certain items of test equipment where it may be necessary to have the tuning sweep back and forth over a range of frequencies.

TOP CAREERS IN ELECTRONICS

A top salary, job security and positive success.
That's what a career in the exciting, rapidly expanding electronics industry
can give you—if you have the right qualifications. An ICS course
can give you the training you need for the rewarding career you deserve.
Select your course then fill in and mail the coupon today.

COLOUR TV

With colour TV on the way, trained colour television technicians are in great demand. This ICS course will give you all the knowledge you need for a career in this field.

COMPUTER TECHNICIAN

In today's fast growing world computers are the key to the future. Every day their use is becoming more and more widespread. Step into a career of the future with an ICS computer technician course.

ELECTRONIC ENGINEERING

As a qualified Electronics Engineer you will find a career that's both fascinating and rewarding. This ICS course will equip you with all that necessary knowledge for a professional position in this field.

TV SERVICE TECHNICIAN

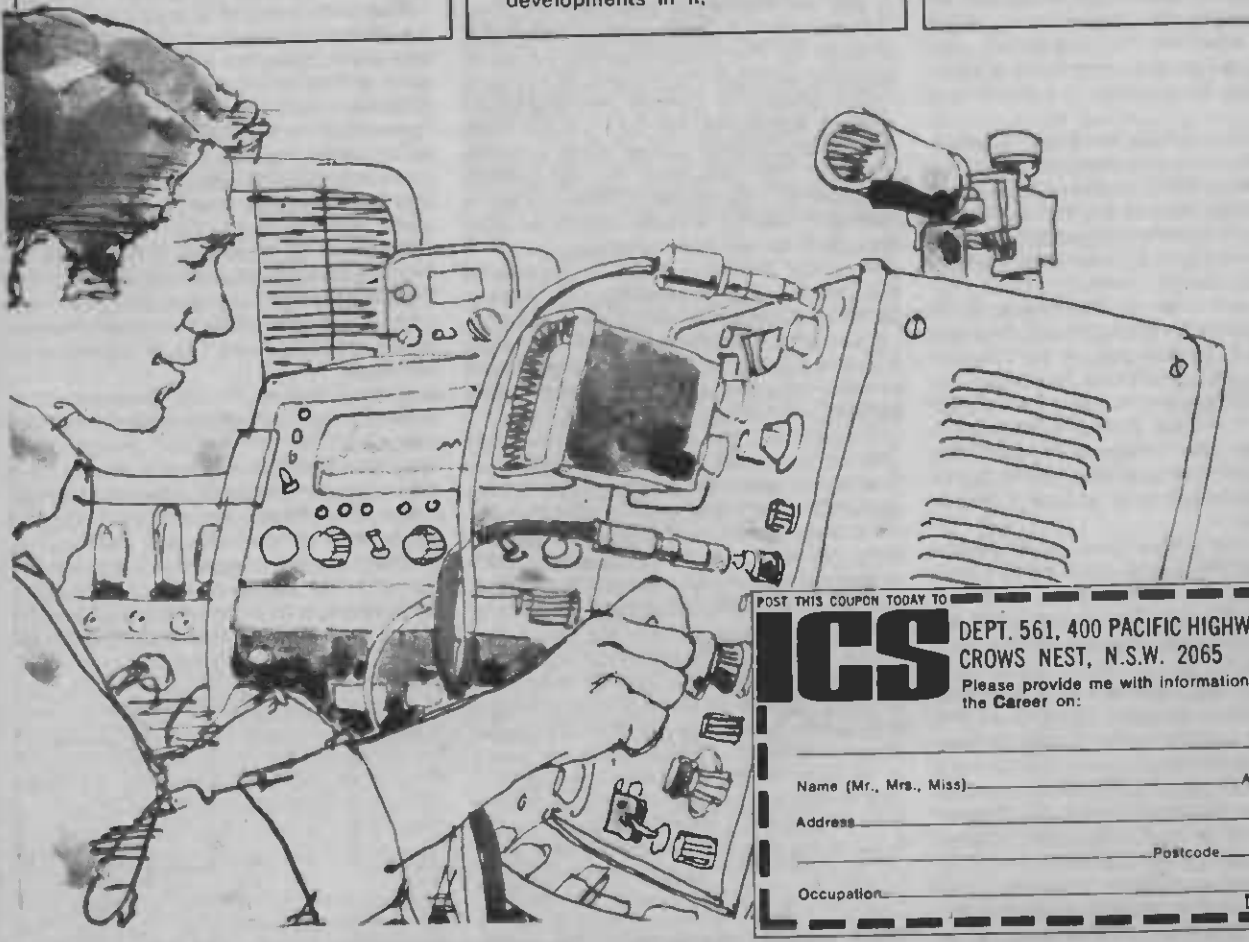
This comprehensive course could be the starting point of a busy and profitable career. The opportunities for you as a trained TV technician are boundless. Take up the challenge and reap the rewards.

SOLID STATE ELECTRONIC TECHNOLOGY

An advanced course for engineers and technicians who wish to advance their knowledge of solid state theory and hardware. The course will give a comprehensive knowledge in this area and the latest developments in it.

BASIC ELECTRONICS

A sound grasp of fundamental principles is the key to a successful career in the electronics industry. And that's just what this ICS course is designed to give you.



POST THIS COUPON TODAY TO

ICS

DEPT. 561, 400 PACIFIC HIGHWAY,
CROWS NEST, N.S.W. 2065

Please provide me with information about the **Career** on:

____

Name (Mr., Mrs., Miss)__________ Age____

Address__________

__________ Postcode____

Occupation__________ LB3.4262

CHAPTER 6

Basic Circuits

Connections and circuits — the effect of resistence — Ohm's law — resistors in series and parallel — inductors and inductive reactance — inductors in series and parallel — capacitors and capacitive reactance — capacitors in parallel and series — impedence — inductors and capacitors in parallel and series — resonance — quality factor or "Q".

The word "circuit" is quite broad in meaning, so that it is perhaps better that we try to explain and illustrate its meaning by example rather than try to define it. An electrical circuit is best regarded as a combination of components so designed that the movement of electrons within such an arrangement is controlled for a specific purpose; or alternatively, to cause an electric current to perform some desired task. The concept of controlling the movement of electrons is fundamental.

When a house is in the process of being built, the plumber installs a water circuit, which carries the water from the street mains to certain parts of the house. The pipes are made large enough to carry the expected flow of water, and appliances are designed on the assumption of a certain maximum expected water pressure. The taps at the end of each run provide a means of controlling the quantity and duration of flow.

An electrical circuit is designed just as deliberately. There are conductors to lead electrons along certain paths and insulators to prevent them flowing in other directions. Resistance, inductance and capacitance all have a definite part to play in controlling and utilising electron flow.

The natural urge at this stage is to present a series of mathematical formulae setting out in precise fashion the function of these circuit properties. However, not everybody is equally at home in the realm of mathematics and our aim here is rather to impart a clear mental picture of conditions in various electrical circuits, deriving the few basic formulae in as natural a fashion as possible.

Consider the simple circuit arrangement of Fig 1. in which a small lamp (torch globe) is connected by means of wires and a socket to a dry cell. The connections are illustrated pictorially and by means of a schematic circuit.

Remember that there is a surplus of free electrons at the negative terminal of the battery and positive ions at the positive terminal. Immediately the connections are made as shown in Fig 1, electrons move from the negative terminal of the battery along the wire "a" to the connection shown at the base of the lamp socket.

An insulating washer separates this bottom connection from the rest of the socket, so that there can be no direct movement of electrons across it. Instead, they move along the centre pin of the socket, which is bearing against a metal blob in the base of the lamp. Because there is still no alternative path, the electrons move along the leads within the globe, through the fine metal filament, thence to the metal shell forming the base of the lamp.

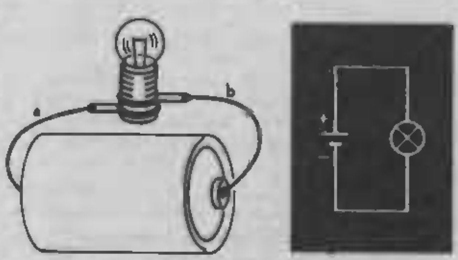


Fig.1. A very elementary circuit, shown both pictorially and as a schematic.

This base is in contact with the metal shell of the socket, so that the electrons pass on to this shell thence to the terminal, the wire "b" and back to the positive terminal of the battery. This forms a complete electric circuit, which is often described as being **continuous.**

Provided the lamp is able to withstand the battery voltage, the flow of electric current will cause the lamp filament to glow — the normal purpose of such a circuit. However, various circumstances could prevent this desirable effect from being realised.

For example, the washer in the base of the socket could conceivably split sufficiently to allow the two portions of the socket to touch one another directly; the electrons, instead of flowing through the lamp, would then take the easier path across the socket and the unwanted result would be described in technical parlance as a **short-circuit.**

If one of the leads were to be snipped through, or the lamp partially unscrewed from its socket, the electron path would be interrupted and the lamp filament would again fail to glow. This is described aptly by the term **open circuit.**

The switch on an electric torch, or on the wall of your home, is simply a device which makes and breaks the electrical circuit between the lamp and the source of power. When moved to the "off" position, it opens the circuit; in the "on" position it closes the circuit. It provides a simple illustration of how electron flow can be controlled at will.

Going a step further, consider the schematic diagram of Fig 2a which shows a lamp and a resistor connected **in series.**

The word "series" is important because it indicates that the components to which it refers are connected end-to-end, electrons moving first through one, then the other. Another combination of components "connected in parallel" indicates that the particular components referred to are connected side-by-side in the circuit, so that the electron flow divides between them. See Fig 2b.

As might be expected, it is possible to have a **series-parallel** circuit, in which the individual circuit components are connected, some in parallel and some in series. However, we are tending to get ahead of ourselves.

Coming back to Fig 2a, there will fairly obviously be a movement of electrons from the negative terminal of the battery, through resistor R, through the lamp, and thence back to the positive battery terminal.

We have already learned that a resistor, by its very nature, offers a definite opposition to the passage of an electric current. The effect may be compared in some respects to a ferry boat turnstile. On one side there is a crowd jostling to get

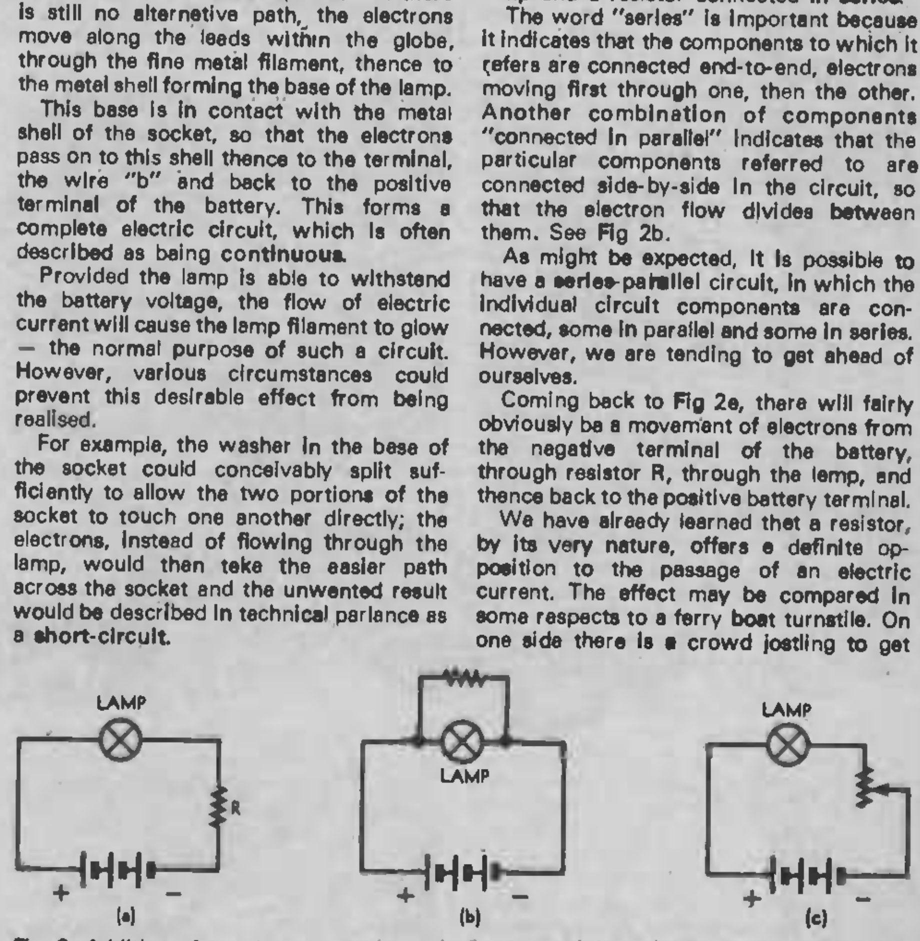


Fig. 2. Addition of a resistor to the simple circuit of Fig. 1. In (a) and (b) the resistor is fixed, while in (c) it is adjustable.

through, while a thin file issues from the other side. So it is with a resistor — an accumulation of electrons on one side and fewer on the other. Look again at that last sentence.

What we really have is a potential difference — or a voltage drop — across the resistor. The larger the electron flow or the higher the resistance value, the greater is the voltage drop across the resistor.

This observation is of fundamental importance, and deserves plenty of attention. The exact mathematical relationship between current, resistance and voltage is set out in what is known as Ohm's Law:

$$E = I \times R \quad \ldots(1)$$

where E is in volts. I is in amps and R is in ohms.

In plain words, the formula simply means that the potential difference developed across a resistor is equal to the product of the current flowing through it (expressed in amps) and the resistance (expressed in ohms). Thus, a current of 0.5 amps, flowing through a resistor of 6 ohms, would develop a voltage drop of 3 volts across the resistor.

Try it out for yourself with the formula.

By simple mathematical "juggling" formula (1) can be transposed into a second form:

$$I = E / R \quad \ldots(2)$$

Thus, if the voltage across a resistor of stated value is known accurately, formula (2) makes it possible to calculate the amount of current which is flowing through it.

For example, if a potential difference of, say, 12 volts exists across the terminals of a 4-ohm resistor, the current flowing through it would necessarily be 3 amperes.

Further transposition of formula (1) gives:

$$R = E / I \quad \ldots(3)$$

If the voltage across a resistor and the current flowing through it are both known it is therefore possible to calculate also the value of the resistor. This is the third form of Ohm's Law.

To elaborate further on these important and basic formulae would take up more space than is available, but the foregoing should give some idea of the essential relationship which exists in any circuit between voltage, current and resistance. It is important to note that all three formulae apply equally, whether one is concerned with direct or alternating current and potentials.

In Fig 2a, the voltage across the lamp filament must inevitably be less than that of the battery by the voltage drop across the series resistor. If the value of the resistor is small — just a few ohms — the lamp filament may still glow, but with reduced brilliance. With higher values of "R", the current flowing may be so limited that the filament will fail to glow at all, unless the battery terminal voltage is greatly increased.

By having a variable resistor in place of "R" (Fig 2c), it is possible to exercise control over the current flowing through the circuit and the voltage which is actually effective across the lamp. In many cars, such an arrangement allows the driver to adjust the brightness of the panel lamps. This is just one simple illustration of the actual effect of resistance in an electrical circuit. But let's take in a few more facts about resistors in a circuit.

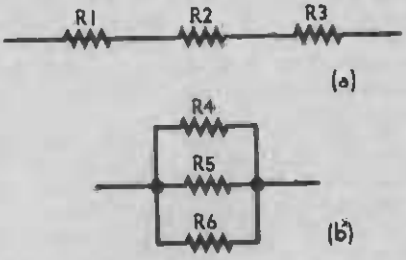


Fig. 3. The two important concepts of series and parallel connections, shown respectively in (a) and (b).

In Fig 3a, three separate resistors are shown connected in series — that is, connected end-to-end. The total resistance of the combination is simply the sum of their individual values. Thus, if R1 is 1 megohm, R2 is 50 kilohms and R3 is 7,000 ohms, the total resistance of the combination would be 1,057,000 ohms or 1.057 megohms. Call to mind our explanation of the terms ohm, kilohm and megohm and check back on our addition.

In Fig 3b, three resistors are shown connected in parallel — that is, side-by-side in an electrical sense. If the three occur in this fashion in an electrical circuit, they provide three suitable paths for the current. For a given electrical pressure (ie, applied voltage), more electrons can move through the combination than could move through, say, just R5 on its own; in other words, the parallel resistance of the combination is less than that of any individual resistor in it.

The general mathematical expression is:

$$1 / R_t = 1 / R1 + 1 / R2 + 1 / R3 \quad .(4)$$

in which Rt is the paralleled resistance of the combination, R1, R2, R3, R4, etc. These are individual resistors and all values are expressed in the same unit of resistance, be it ohms, kilohms or megohms.

Whether you can apply this formula immediately will depend on your familiarity with algebra. If your recollection of algebra is not too good, simply remember the general statement just ahead of the formula. But let us move on.

As explained in a previous chapter, the passage of an electric current through a conductor is attended by considerable atomic agitation and a certain amount of the electrical energy is dissipated in the conductor as heat. An extreme illustration is seen in the filament of an electric lamp, where the heat generated is so great as to cause the filament wire to glow brightly. The same heating effect is evident in resistors, but generally to a lesser extent.

The electrical energy dissipated in a resistor is expressed in watts, and is equal to the product of the voltage across the resistor and the current through it in amperes. That is,

$$W = E \times I \quad \ldots(5)$$

The above formula can be transposed in various ways and combined with the Ohm's law expressions given earlier, but it is not our intention here to pursue mathematical investigations. Just remember that there is a definite limit to the current or voltage which any given resistor or combination of resistors can handle. For this reason, electronic circuit designers often specify the minimum wattage rating required for certain resistors, as well as the resistance values.

If the constructor is careless enough to use too small a unit, the resultant overheating effect may easily cause it to char and fail in service. So much, then, for resistors.

Because of their special characteristics, inductance coils — inductors, to be more precise — find frequent application in circuits involving alternating currents. If necessary, they will carry direct current but offer marked opposition to the flow of AC; this much we have already seen.

The opposition to alternating current, that is the reactance of an inductor, is expressed by the formula:

$$X_L = 2\pi fL$$

In which reactance is in ohms, pi is equal to 3.1416, f is the frequency of the current in Hertz, and L is the inductance in Henries.

Without worrying too much about the formula, one point should be fairly clear, namely that the reactance of any given inductor varies with frequency. The higher the frequency of a current, the greater the opposition a given inductor offers to its passage. Plotted graphically, the result is as shown in Fig 4.

Imagine an inductor connected into a complex circuit arrangement requiring it to carry a direct current and two alternating currents of different frequencies. The inductor would offer little opposition to the direct, but definite and different degrees of opposition to the two alternating currents, in proportion to their frequencies.

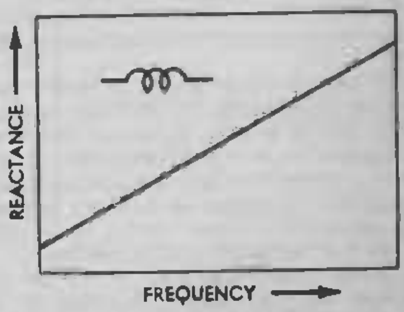


Fig. 4. A coil or inductor presents an "inductive reactance" to AC, the reactance increasing as the frequency rises.

This discriminating effect of an inductor is very widely utilised in electronic circuits, as we shall come to appreciate later.

If two or more inductors are connected in series, the total inductance value is equal to the sum of their individual inductances, providing there is no magnetic coupling between them; this is exactly the same as with resistors.

In like manner also, the inductance value of a number of parallel-connected noncoupled inductors is less than that of any one of them, and the expression is similar to formula (4):

$$1/L_t = 1/L_1 + 1/L_2 + 1/L_3 \ldots \text{ etc} \quad \ldots(7)$$

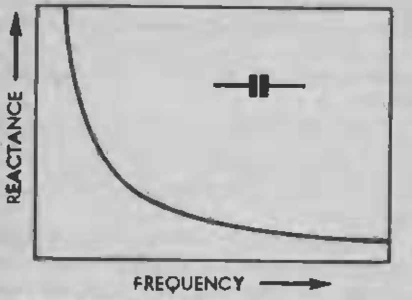


Fig 5. In contrast to an inductor, a capacitor presents a "capacitive reactance" to AC, the reactance decreasing as the frequency rises.

Capacitors are almost opposite in their characteristics to inductors. Since the component parts of a capacitor are necessarily insulated from one another, a capacitor cannot carry direct current. However, because of the phenomenon of electron flow during the charging and discharging processes, it will carry alternating current, although offering some opposition to its passage.

This opposition, known as capacitive reactance, is expressed in ohms and is given by the formula:

$$X_c = 1/2\pi fC \quad \ldots(8)$$

in which pi is equal to 3.1416, f is the frequency in Hertz, and C is the capacitance in Farads.

It is interesting to note that the reactance of a capacitor is again dependent on the frequency of the current flowing, but the relationship is just the reverse of that which characterises an inductor. With a capacitor, the higher the frequency of the current in the circuit, the lower is the opposition offered by the capacitor to its passage (Fig 5).

A capacitor in an electrical circuit thus offers no path for the flow of direct current, but allows alternating current to flow through it with an ease depending on its own capacitance value, and the frequency of the current. Here then is yet another discriminating effect, which is widely utilised in electronic circuit design.

When capacitors are connected in parallel, the total capacitance is equal to the sum of their individual capacitance values. That is:

$$C_t = C_1 + C_2 + C_3 \ldots \text{etc} \quad \ldots(9)$$

This is just the opposite effect to that stated for resistors and inductors. For capacitors in series, we revert again to the familiar reciprocal expression:

$$1/C_t = 1/C_1 + 1/C_2 + 1/C_3 \ldots \text{ etc} \ldots (10)$$

These expressions may be rather baffling to some, but don't worry too much about that. The whole purpose in this chapter is to give a general picture of how components behave and why they are used, rather than to get down to a precise mathematical basis.

Circuit requirements often call for a combination of resistance, inductance and capacitance, arranged in series, in parallel or in a combination of both. The opposition of such a combination to the flow of alternating current is generally neither wholly resistive nor wholly reactive. Because of this it is usual to speak of the opposition of such a combination as an impedance.

Providing certain mathematical rules are observed, the overall impedance of a combination circuit may be found by combining the resistance and reactance components, and expressing the answer in ohms.

Examination of the word impedance from a mathematical viewpoint is rather too involved for inclusion in this chapter, but the student would be well advised to follow up the subject in other texts which choose to cover it in detail. But remember the term and remember also the effect of resistance, inductance and capacitance in an electrical circuit.

- Resistance is a definite property, independent of frequency.
- Inductive reactance depends on, and increases with frequency.
- Capacitive reactance depends on, but varies inversely with frequency.
- Inductors will carry direct current, but capacitors will not do so.
- Circuits combining the above properties offer impedance to the flow of alternating current, the impedance depending on the frequency and the quantities involved.

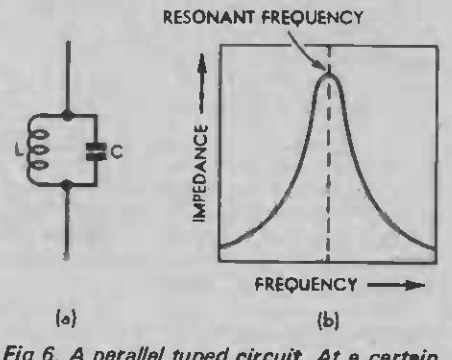


Fig 6. A parallel tuned circuit. At a certain frequency, the impedance of such a circuit rises to a peak. This is known as "resonance".

At this juncture, it is necessary to mention a matter of vital importance, namely the combination of an inductor (or coil) and a capacitor. We have seen that the reactance of a coil increases with frequency and that the reactance of a capacitor decreases with frequency. Referring then to Fig 6a it should be fairly clear that, for any given value of inductance and capacitance, there must be a particular frequency at which the reactance values of the two will have become equal.

The frequency at which this occurs is commonly known as the resonant frequency and the coil and capacitor are said to be tuned to this frequency. Because the two components are connected side-by-side (in the physical sense), the combination is referred to as a **parallel tuned** circuit.

In previous chapters we have learned something of phase; that the current through an inductor lags behind the voltage, but leads the voltage in the case of a capacitor. At the resonant point of a tuned circuit such as that of Fig 6a, the currents through the two components — which are opposite in phase, one lagging the voltage by one-quarter of a cycle, the other leading by the same amount — are equal in amplitude. They accordingly cancel each other, so that the circuit offers an infinitely high resistance to current.

The above assumes that the components are "lossless", or perfect — the coil having no resistance, and the capacitor no dielectric loss. In practice, however, this will not be true, so that at resonance the parallel tuned circuit will have a large but finite resistance. Note that the word used is **resistance**, for at resonance the inductive and capacitive reactance cancel each other, and the impedance of the combination becomes simply a resistance.

Because the reactive currents cancel at resonance, any current which still flows is in phase with the applied voltage, since the circuit behaves as a pure resistance. For this reason resonance is usually defined as that frequency at which the phase shift of the circuit is zero.

The resonance of a parallel tuned circuit may therefore be thought of in terms of two connected but different effects. One, the cancellation of reactances to leave simply an effective resistance (in other words, zero phase shift), and the other a rise and peak in impedance to current.

The relationship between circuit impedance and frequency is suggested by Fig 6b.

At its resonant frequency, a tuned circuit has an **oscillatory** quality. Assume, in Fig 6a, that the capacitor is disconnected, charged from a battery then instantly connected again to the coil. The charge across the capacitor would immediately commence to leak away, initiating a current through the inductor; in so doing, it would build up a magnetic field around it. Immediately the charge disappears, the initial current would cease, causing the magnetic field to collapse.

Even as it collapses, however, the magnetic field maintains the current flow to the point where the capacitor becomes charged again in the opposite polarity. Then the capacitor would begin to discharge in the opposite direction and the whole action would be repeated again and again.

In practice, some of the energy is lost in each **alternation**, and the oscillatory action ultimately comes to an end. The result, illustrated in Fig 7, is known as a damped train of waves. It is important to note that the natural frequency of the oscillation of a tuned circuit corresponds to the resonant frequency — or that frequency where the reactance of inductor and capacitor are equal.

For an inductor and capacitor of fixed values there can be only one definite and fixed resonant frequency. However, as we shall see later, electronic circuit design frequently calls for tuned circuits whose resonant frequency is capable of adjustment within certain limits.

In practice, this is achieved by varying either the value of the capacitor or the inductance of the coil. Of the two methods, the former is the more usual where a large variation is required (Fig 8a), and it is for just this purpose that the variable capacitor illustrated in the last chapter is employed.

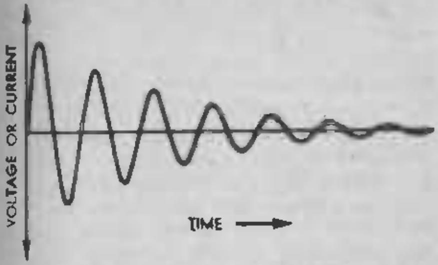


Fig. 7. A damped train of waves. Energy fed into a tuned circuit tends to oscillate between the inductor and capacitor at their resonant frequency.

Variation of inductance is used mainly where the resonent frequency is to be adjusted over a relatively small range. It can be achieved by altering the physical relationship between two halves of a coil, or by inserting or withdrawing a metal core through the centre of the coil former, thus acting directly on the magnetic circuit.

While tuned circuits are most commonly encountered in the so-called parallel configuration, as shown in Fig 8, it is appropriate to mention that a resonant effect is also apparent when an inductor and capacitor are connected in the series configuration shown in Fig 9a. When examined between the end connections, the reactance of the combination reaches a minimum value at resonance, as indicated in 9b.

The maximum value of impedance reached by a parallel tuned circuit at resonance, and also the minimum value reached by a series tuned circuit, are both determined by the so-called "losses" in the circuit. The losses are due to the resistance of the inductor, and any leakage resistance across the dielectric of the capacitor.

When the losses are relatively low, the impedance of a parallel tuned circuit rises quite markedly at resonance, while that of a series circuit falls to a similar extent. The circuits are then said to exhibit a high quality factor or "Q". On the other hand when the losses are relatively high, the peak of the parallel circuit and the dip of the series circuit both become less marked, and the circuits are said to exhibit a low "Q".

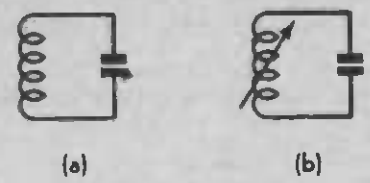


Fig. 8. The resonant frequency of a tuned circuit may be varied, if desired, by varying the value of either the coil or the capacitor.

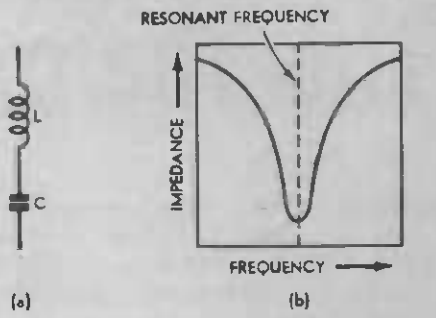
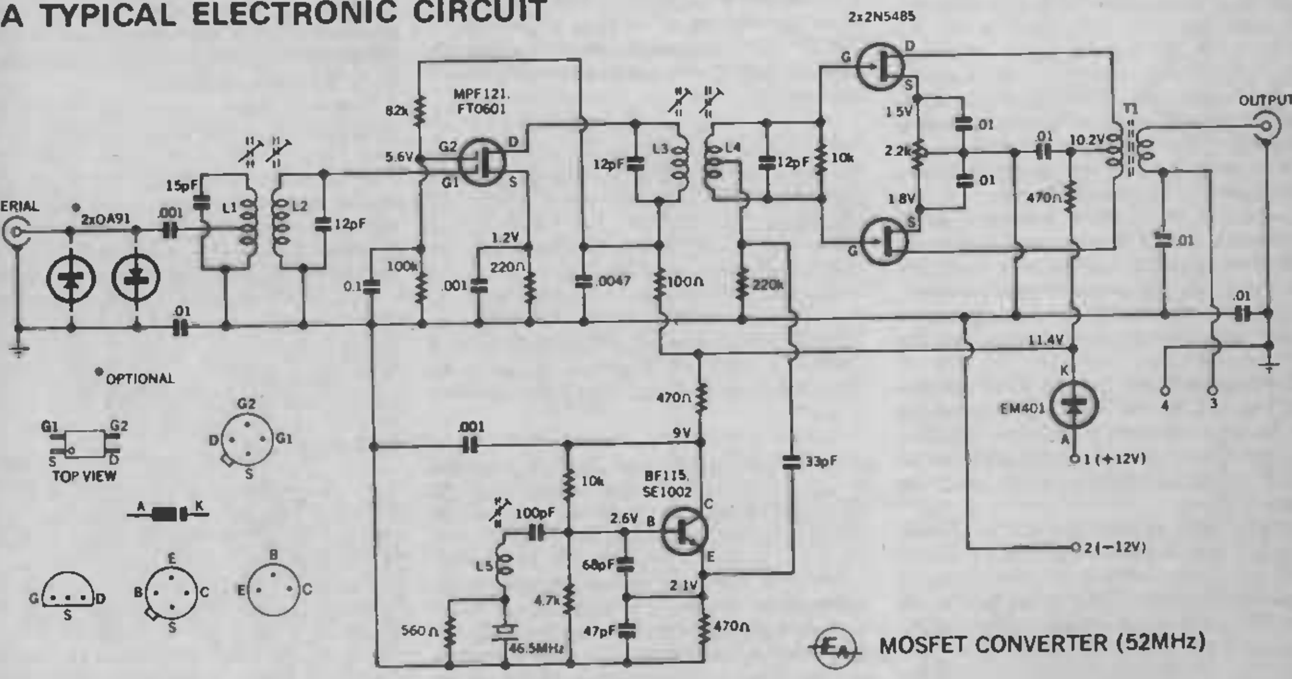


Fig. 9. A series tuned circuit. At its resonant frequency, the impedence of such a circuit falls to a minimum.

In general, the higher the Q of a tuned circuit, the "sharper" its resonant peak or dip. Low Q circuits have a "broad" resonance.

In later chapters we shall see just how these various combinations of resistance, inductance and capacitance are employed in practical circuits.

A TYPICAL ELECTRONIC CIRCUIT



Here is a typical electronic circuit, reproduced from the pages of "Electronics Australia". You may not be able to make very much sense out of it as yet, but don't worry; you'll develop that ability as you study further. Pick out the resistors, which range in value from 100 ohms to 0.22 megohm. Have a look at the values and see whether you can relate them to the information and the colour code set out in chapter 2. No batteries are specifically involved (chapter 3) but the circuit does call for a supply voltage of 12 which could come from an automobile electrical system.

Now pick out the inductors, based on information set out in chapter 4. Note that all except one are drawn to suggest inductive coupling between adjacent coils. Then there is quite a variety of capacitors shown in the circuit (see chapter 5) and you can logically spend some time looking at the values. In fact, certain of the coils and capacitors are obviously in parallel, constituting a parallel tuned circuit, as explained earlier in this chapter. What are those mysterious devices inside circles, in this and many other electronics circuits? You'll be able to answer that yourself after reading the next couple of chapters covering valves and transistors. Following that is a chapter on how to read circuits as a whole.

CHAPTER 7

The Thermionic Valve

Electron emission from a heated surface — filaments and indirectly heated cathodes — the plate or anode — the diode valve and rectification — the triode valve and amplification — amplification factor and transconductance — tetrode, pentode and beam-power valves — valve construction, multiple valves and symbols.

We have already seen that all matter, whether in solid, liquid or gaseous form, is made up of innumerable atoms. Further, that these atoms are composed of protons and electrons in a constant state of motion. In materials known as conductors, some of the outer planetary electrons can migrate from one atom to the next in more or less random fashion. We launch into our discussion of valves from this point.

This random movement of electrons can be accelerated by increasing the temperature of the conductor — usually one or other of the metals, or a metallic compound.

As the temperature is gradually raised, the velocity of electron movement increases and a condition may ultimately be reached where some of the planetary electrons acquire sufficient velocity to temporarily "escape" from the confines of the conductor. However, because of the attraction of the parent body and the outside blanket of air, the electrons so emitted return immediately to the parent body.

Conditions are different, however, when the emitting body is contained within an evacuated space.

With no gas molecules to bar the way, the electrons emitted from a heated conductor can move an appreciable distance from it. In the normal way, the loss of electrons leaves the parent body with a positive charge, so that it acquires a marked attraction for the minute negative particles. They shoot out from the surface of the heated conductor, then fall back again.

The net result is that the heated body is surrounded by a cloud of free electrons, like steam above the surface of boiling water. These free electrons can nevertheless be attracted from the immediate vicinity of the parent body simply be placing a positively charged metal plate elsewhere within the evacuated space.

Evidence of this so-called "electron emission" was first noted by Thomas Edison in about 1880. Edison discovered that an electric current could be made to pass through the ostensibly empty space between the heated filament of an electric lamp and another metallic conductor contained within the same glass envelope.

The effect was duly noted but was unexplained until about 1899, when the electron theory was first propounded.

When the reason for the current was appreciated, and something of its characteristics investigated, its application to the new science of wireless communication was envisaged. Electron emission ceased to be a mere incidental effect and became a phenomenon to be controlled and utilised.

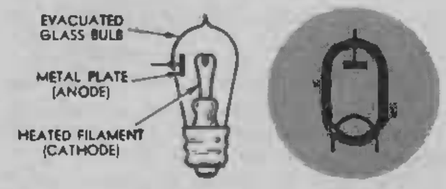


Fig. 1. An elementary form of thermionic valve — a diode — with its schematic circuit symbol.

The essential features of a valve are, therefore, a heated electrode to supply free electrons, and one or more additional electrodes to control and utilise their movement — the whole mounted within an evacuated envelope. (See Fig. 1).

The term "thermionic" is often used to describe such a valve, drawing attention to the fact that it depends on the heating of a material ("thermo") to provide a source of electrons (negative charges, and therefore "ions").

The electrode supplying the electrons is commonly referred to as the **cathode** and is usually heated by the passage of an electric current, in precisely the same manner as the filament of an electric light globe. The shell enclosing the evacuated space may either be of metal or glass, the latter being the more common.

Before going on to investigate just how the emitted electrons are controlled and utilised, it is as well to become familiar with the emitting electrode — the cathode.

For many years, the cathode was a tungsten or similar wire, heated by the passage of an electric current; because of its similarity to the element in a light globe, this type of cathode was commonly referred to as a **filament**. Since this current is used to raise the temperature of the emitting surface, the filament is called, in technical language, a **directly heated cathode**.

In the early valves, the filaments were generally of pure tungsten, which emitted a goodly supply of electrons when heated to a brilliant white heat. Subsequent experiment showed that similar emission could be obtained, with greater economy in heating power, by impregnating the tungsten with thorium and operating the filament at a bright gold colour.

Constant research has produced the more recent oxide-coated filament, which gives a copious supply of electrons at a temperature sufficient only to heat the filament wire to a dull red colour.

The current utilised to heat the filament in a sense represents a waste of power, since it does not contribute further to the basic function of the valve; thus, for most purposes, there is every reason to reduce it to an absolute minimum.

Filament-type (or directly-heated) cathodes have chiefly been used in valves intended for battery-operated equipment, their particular advantage being economy of filament heating power. (It should be noted, in passing, that transistors have almost completely displaced valves nowadays in battery operated equipment.)

A typical filament or directly-heated cathode is shown in Fig 2a.

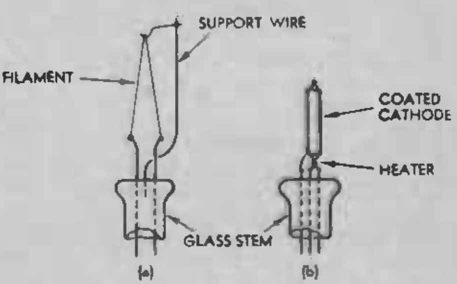


Fig. 2. Directly and indirectly heated cathodes. (a) Directly heated, (b) Indirectly heated.

Although this type of valve was used in the early days of radio, there was a definite need for a valve that could be operated from a low voltage AC supply, provided by a simple stepdown transformer. Conventional directly heated valves could not be used in this way because the AC introduced severe hum into the output circuit. This was due both to the heating and cooling of the filament during the AC cycle, and the change in polarity across the filament, which formed part of the signal circuit.

One approach to this problem was to use a very heavy filament, so that it would store considerable heat, and to make the signal connection to an electrical centre tap of the filament. This idea was only partly successful. It was satisfactory for the last stage (output valve) of a receiver, but introduced far too much hum if used in earlier stages. Such valves are still used to a limited extent in this and similar applications, such as high power stages of transmitters.

However, the real breakthrough came with the development of the **indirectly heated cathode**. In this the cathode is a nickel sleeve coated with electron emitting oxides. Inside the sleeve is a filament or heater, insulated from it electrically, but able to heat the cathode indirectly by conduction. Since the filament no longer forms part of the signal circuit, it can be operated from AC without introducing hum, while the cathode stores enough heat to ensure constant emission throughout the AC cycle. An indirectly heated cathode is illustrated in Fig. 2b.

The fact that the emitting surface is isolated electrically from the heater also confers other advantages which are much appreciated by the designers of electronic circuits.

Many thousands of different valve types have been sold within the short history of electronics and there is a wide variation in the voltage at which their filaments or heaters are intended to operate. In battery valves, the filaments have variously been rated to operate at 1.4, 2.0, 4.0, or 6.0 volts, according to the particular type. Mains-type valves have had heaters rated to operate at various potentials from 1.6 to over 100 volts.

When installing a valve in a circuit, this matter has to be kept in mind, as the application of an excessive voltage to a filament or heater may completely ruin a valve.

Coming back to the original line of thought, let us assume that we have a metal plate mounted within the evacuated envelope and in proximity to the heated cathode. This second electrode can quite correctly be referred to as an **anode** or a **plate**. (Fig. 1).

If it is connected externally to the cathode, a few high velocity electrons may strike it and return ultimately to cathode via the external connecting circuit. However, the vast majority of the emitted electrons remain in the immediate vicinity of the cathode, forming what is known as a **space charge**.

Think of this space charge as a restless cloud of negative electrons — and therefore a negative charge.

If a battery is connected between cathode and plate in such a way that the plate is negative with respect to cathode, it naturally tends to repel any electron which happens to approach it. Remember our earlier axiom — like charges repel. Reversing the polarity of the battery so that the plate is made positive has just the opposite effect, the plate tending to attract electrons to itself; the higher the applied plate voltage, the greater the attraction, and therefore the greater the electron flow across the space between cathode and plate.

A point is ultimately reached where the plate attracts every electron the cathode is capable of emitting, and when this stage is reached, the valve is said to have reached **saturation**.

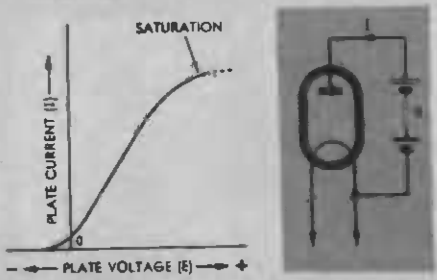


Fig. 3. The basic relationship between plate-cathode voltage and current for a diode, or its "characteristic".

The whole relationship is most easily illustrated with the aid of a simple graph, as shown in Fig. 3. Plate current is plotted in a vertical direction, and plate voltage, positive and negative, along the horizontal axis. There is no plate current while the plate is negative with respect to cathode, but a small flow becomes evident in the region where plate and cathode potentials are similar.

Beyond this, an increase in positive plate voltage is accompanied by an increase in plate current until ultimately the region of saturation is reached.

A valve having only a plate and a cathode — either directly or indirectly heated — is commonly known as a **diode**.

Just what happens when an alternating voltage is applied to a diode is shown in Fig. 4. The voltage / current curve has been redrawn and the quantities are designated by the common abbreviations "E" (for voltage) and "I" (for current).

The added curve at the bottom suggests the application of an alternating voltage which swings the plate alternately positive and negative. The diode passes current only on the positive swings, the negative swings having no effect.

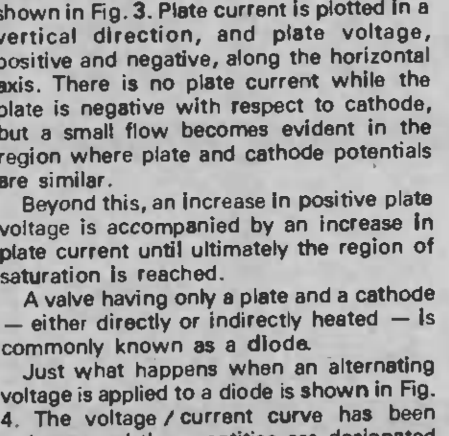


Fig. 4. Illustrating the way in which a diode is able to perform as a "rectifier" of alternating current by conducting only positive half-cycles.

The ability of a diode to suppress the negative peaks of an alternating input wave is an extremely valuable feature, the process being described as **rectification**.

Diode valves, with heavy cathodes and large plates, were formerly widely used to convert alternating current to direct current in power supplies for electronic equipment, large and small. (Their place has now largely been taken by semiconductor devices of similar function.)

Other diode valves, small enough to conceal in the palm of the hand still play a vital part in television receivers, helping to make the signals from a broadcast station visible on a cathode-ray tube. The whole subject of rectification is a big one, but no more can be said of it just now.

Diode valves of various patterns were used quite extensively in the early days of radio, but, in 1907, Dr Lee De Forest patented a type of valve rather like previous diodes in construction, but with a third electrode interposed between the cathode and the plate.

Commonly called the **control grid** (or simply the grid), this third electrode proved to have a pronounced effect on the number of electrons passing from cathode to the positive plate. De Forest's valve, known as a **triode**, is shown pictorially and schematically in Fig. 5. The grid usually resembles in form a short length of open mesh helical spring, provided with one or more side rods for support.

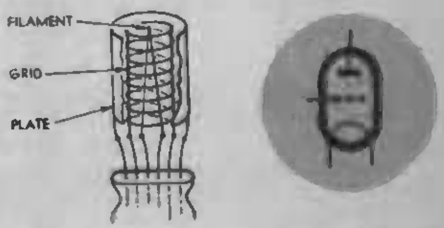
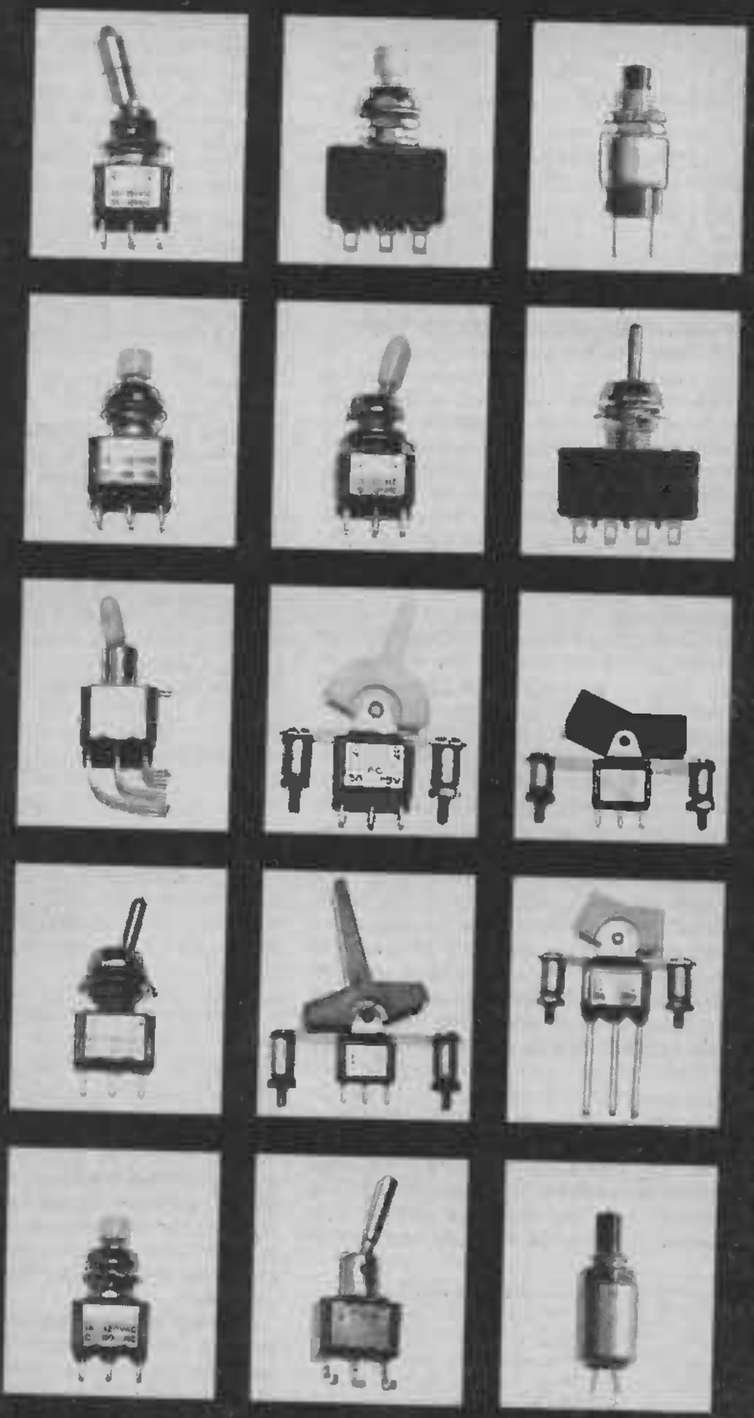


Fig. 5. A directly-heated triode in its basic form; alongside, its schematic circuit symbol.

If the grid is unconnected, the electrons pass through its meshes from cathode to plate in the manner already described for a diode. However, if the grid is made slightly negative with respect to cathode, it exercises a repulsive effect on electrons moving outwards from the cathode, and some of them are turned back. If the grid is made progressively more negative, the repulsive effect becomes stronger and a condition can ultimately be reached where the electron current to the plate is cut off altogether.

On the other hand, making the grid positive with respect to cathode has an accelerating effect on the electron movement, so that the plate current is made higher than would normally be the case. In fact, the positively-charged grid is likely to attract a certain number of electrons itself, giving rise to grid current. This is a condition which is generally avoided. Instead, the valve is designed to operate normally with a certain initial negative voltage, known as a **bias**.

Making the grid negative by a greater or smaller amount than this initial bias voltage



all shapes and sizes of sub miniature switches

Toggle, rocker, pushbutton, paddle, lever lock . . . every practical alternative lever action combined with maximum performance, minimum weight and size and available in one, two, three and four pole models. Subminiature switches manufactured by C & K Components Inc. U.S.A. are readily available in just about every shape, size and configuration imaginable and they are readily adaptable to a multitude of uses where space is at a premium. All switches feature rugged construction and simple mounting . . . long-term, trouble-free operation is ensured. Toggle switch contacts are rated 2 amps @ 240v. AC and 5 amps @ 28v. DC resistive load. The full C & K range is now available from the Professional Components Department, Villawood, NSW or Plessey Ducon Interstate agencies. Catalogue containing full specifications, options, information on hardware, panel layouts, mounting, etc. is available on request.

PLESSEY
Ducon

Plessey Ducon Pty. Limited
Box 2 PO Villawood NSW 2163
Telephone 72 0133 Telex 20384
Melb: Zephyr Products Pty. Ltd. 56 7231
Adel: General Equipments Pty. Ltd. 63 4844
Bris: Douglas Electronics Pty. Ltd. 97 8222
Perth: H. J. McQuillan Pty. Ltd. 21 4821
Everett Agency Pty. Ltd. 84137
Henderson (N.Z.) 64189

AC68

decreases or increases the plate current with respect to its initial value.

The degree of control which the grid exercises over the plate current depends very largely on the physical construction of the valve. However, De Forest's triode demonstrated conclusively that a small voltage applied to the grid could exercise control over a plate circuit in which comparatively large currents' were flowing.

In modern triode valves of quite ordinary design, a change of 1.0 volt on the grid can effect a change of 10 or more milliamps in the plate current. Here is a chance to recall and apply Ohm's Law, as mentioned in an earlier chapter.

In Fig. 6, the plate of a triode is shown connected through a plate load resistor to the positive terminal of a battery. This is commonly referred to as the high tension supply or "HT".

The grid, on the other hand, returns to an adjustable resistor across a bias battery. No heater supply is shown, it being taken for granted that this requirement is attended to as a matter of course. Note that both batteries are connected, at one end, to the cathode of the valve. In equipment operating from the power mains, the various batteries are supplanted by electronic power supply equipment.

Assume that the grid bias of the valve is set at −4 volts, and that it is drawing a plate current of 3 milliamps through the load resistor, equal to 50,000 ohms. The voltage developed across resistor RL by this current would be:

$$E = I \times R$$
$$= .003 \times 50{,}000$$
$$= 150 \text{ volts.}$$

If the grid bias voltage is now increased to −5 volts, we may find that the plate current has decreased by, say, 1 milliamp, so that the voltage drop across resistor RL becomes:

$$E = .002 \times 50{,}000$$
$$= 100 \text{ volts.}$$

In other words, a change of 1 volt on the grid has effected a change of 50 volts on the plate potential.

It follows that, if an alternating voltage of 1 volt were applied to the grid, an alternating voltage of no less than 50 volts would appear in the plate circuit. Thus, the triode, with its associated components and power supply, is capable of **amplifying** a signal.

Not all triodes can achieve an amplification of 50 times, as our example might suggest, but the importance of this development is not hard to imagine. Diode valves could only alter the nature of an incoming radio signal; the triode could do this, if necessary, and amplify it as well!

This fact was of tremendous importance in the development of early radio receivers. Signals from radio stations, too weak to be discerned with equipment using a simple diode valve, were amplified to complete audibility with triodes. Nowadays, one-valve sets using a single triode valve can receive stations from all over the world under favourable conditions.

Valve manufacturers have provided us with special high amplification triodes, and with others intended to deliver enough power to operate a loudspeaker. There are also general purpose triodes — valves intended for a variety of applications.

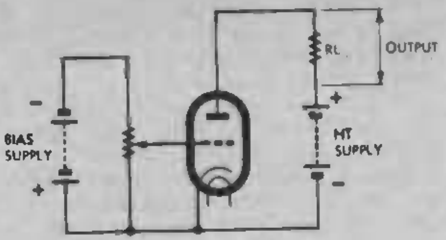


Fig. 6. A basic circuit used to demonstrate the ability of a triode to amplify a small change in voltage.

An important characteristic of a triode valve is, in fact, its **amplification factor.** This is the ratio of a voltage change at the plate, to a voltage change at the grid, necessary to effect a given small change in the current flowing through the valve.

Valves which are intended to give a fairly high stage gain, as in the example quoted above, are usually designed with a high amplification factor, typically in the range 50 to 150. General purpose triodes are designed with an amplification factor in the range 20 to 50. Triodes intended to deliver high power output may have an amplification factor almost anywhere between 5 and 100 depending on how they are intended to be used; they are normally distinguished, however, by larger than normal electrodes and a construction aimed at coping with the considerable heat which they tend to generate internally.

For our present purposes, it might be worthwhile at this stage to introduce the reader to the concept of **transconductance,** which can be applied not only to triode valves, but also to the other types which we will meet shortly. It is also equally applicable to the various types of transistor, to be described in the next chapter.

The transconductance is simply a description of the relationship between the output current — ie, the plate current, in this case — and the input voltage, here the grid voltage. More specifically, the transconductance is the ratio between a change in plate current, and the change in grid voltage which produced it. This is usually expressed in terms of so many "milliamps per volt" (mA / V).

In the foregoing example, the plate current of the valve concerned changed by 1 milliamp when the grid voltage was changed from −4 to −5 volts — a change of 1 volt. The transconductance of the valve would therefore be **1mA / V.**

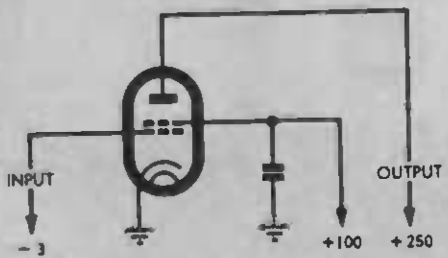


Fig. 7. With a screen-grid or tetrode valve, the second or "screen" grid is usually connected to a positive DC supply and tied to the cathode via a capacitor.

A term sometimes used loosely as an equivalent to transconductance is "slope". One can therefore speak of a valve with a slope of 5mA / V.

With the rapid advance in knowledge, it soon became evident that, notwithstanding all its advantages, the triode and definite limitations. And, as valve designers sought to overcome these limitations and to keep step with progress in circuit design, there came into being one new class of valve after another. Nor has the process finished even now despite the emphasis on transistors and other related devices which we have yet to meet. However, whereas attention at first was on new electronic principles, it is now more along the lines of achieving the most favourable physical and electrical characteristics.

The chief single difficulty with the triode was the considerable capacitance between the plate and control grid. A little thought will show that a capacitance effect must surely exist, because the grid and plate are both fairly large physically; they are mounted quite close together and insulated from each other.

In some circuit applications this grid-plate capacitance effect did not matter, but in others it caused a lot of trouble and seriously limited the usefulness of the valve.

If we can venture a little ahead of ourselves it may be said that the grid-plate capacitance caused instability and oscillation in radio frequency amplifier stages, and also loss of gain at high frequencies.

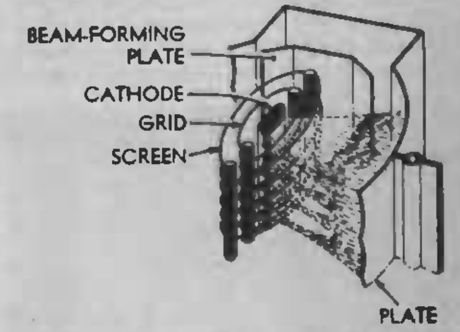


Fig. 8. The internal structure of a beam-power tetrode showing the form and relationships of the various electrodes.

The ultimate answer to this problem was the introduction of another grid, between the control grid and plate. The resulting four-electrode valve became known popularly as a "screen-grid" valve or "tetrode".

Valve designers arranged that this extra grid should be operated at a positive potential sometimes equal to, sometimes less than that applied to the plate. A bypass capacitor was — and is — connected between this screen-grid and cathode (or earth), the capacitor value being chosen so that its reactance is low at the signal frequencies at which amplification is required. Thus while the screen grid is positively charged with respect to the cathode for DC purposes, it is virtually at cathode potential as far as the AC signals are concerned. (See Fig. 7.)

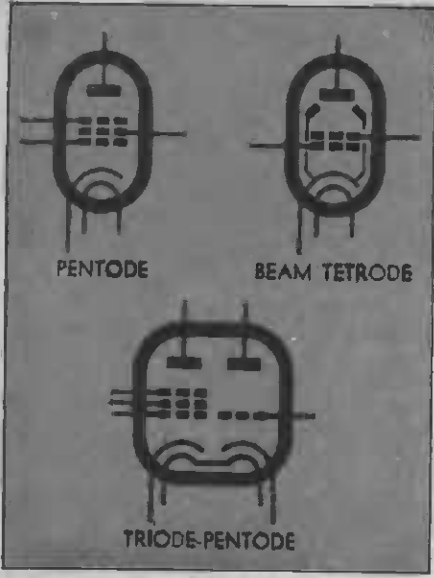


Fig.9. Circuit symbols for a pentode, a beam tetrode and a triode-pentode — the latter a multi-section valve.

It thus acts as an electrostatic shield between grid and plate, and reduces the capacitance effect between them to a very small figure. Whereas the capacitance averages 8pF in a triode, the grid-plate capacitance with a screen grid or tetrode valve is seldom more than 0.01pF.

Further, to reduce the grid-plate capacitance it became common practice to provide a connection for one or the other in the form of a cap on the top of the glass envelope, instead of connecting it to the usual base pin. American practice was to connect the grid to the top cap, while in Europe the plate was often brought out on top.

The reduced grid-plate capacitance, together with other characteristics peculiar to the screen-grid construction, allowed much higher amplification to be achieved in radio frequency amplifier stages without danger of instability.

Stability was further improved by installing a metal screen to enclose the valve, or alternatively spraying the envelope with a metallic coating which could be connected via a base pin to cathode or to the chassis.

In modern miniature valves, careful design and small size has largely obviated the need for top-cap connections and metal shield cans. Plate and grid connections are brought out to pins on opposite sides of the base socket and shields are built into the basic electrode structure.

The screen grid was found to have interesting effects on the electrical characteristics, quite apart from the reduction in grid-plate capacitance. Being operated at a substantial positive potential with respect to cathode, it tends to attract electrons from the latter in just the same way as the plate. Thus, some electrons are collected by the positive screen, but the majority simply pass through its open mesh and proceed outward to the plate.

However, the screen assumes a marked control over the plate current and, speaking generally, we may say that the plate current in a screen-grid valve depends very largely on the applied screen voltage, and to a lesser extent on the plate voltage. This is reflected in changed values of amplification factor, and transconductance.

Although the screen-grid represented a distinct advance, it introduced certain difficulties, chief of which was the so-called secondary emission effect. When moving electrons strike the plate of a valve at high velocity, they tend to penetrate its molecular structure and dislodge other electrons, which are thus literally "knocked" off into space.

Because this action is an incidental effect of the main electron stream from the cathode, it is described by the general term of secondary emission. The effect is evident in triode valves, but it does not lead to ill-effects because the secondary electrons are immediately attracted back to the positively charged plate.

In a screen-grid valve, however, the positively charged screen tends to attract some of the secondary electrons so that an electron flow is evident in the reverse direction between plate and screen. It is most noticeable when the screen voltage is equal to or greater than the plate voltage. This can occur easily in normal operation, when the control grid receives a positive-going signal impulse and the instantaneous plate voltage is reduced sharply.

Hence, secondary emission, by producing undesirable effects in the plate current, limits the amount of plate current swing and offsets partially the advantages of the screen-grid structure for some circuit applications.

The ultimate solution to the secondary emission problem was offered by the addition of yet another grid — this time located between screen and plate. The resultant five electrode valve came to be known as a pentode or penthoda. The suppressor grid is usually a very open spiral and is generally connected internally or externally to the cathode.

Being negative with respect to both plate and screen, it tends to reduce the velocity of electrons approaching the plate in the normal manner, and diverts back to the plate the slow-moving electrons produced by secondary emission.

Apart from the pentode, there is also the beam power tetrode valve which makes use of a different principle to suppress secondary emission. There are the four major electrodes — cathode, grid, screen and plate, but they are so shaped that secondary emission from the plate is prevented without an actual suppressor grid.

Because of the way the electrodes are spaced, electrons travelling to the plate slow down to a comparatively low velocity in a certain region between screen and plate. In this region the electrons form a dense cloud — a space charge. The effect of this space charge is to repel secondary electrons emitted from the plate and thus cause them to return to the plate.

Two additional electrodes known as beam forming plates, which are connected internally to the cathode, are so placed as to concentrate the electrons emitted to two beams, on opposite sides of the oval cathode.

Another feature common to beam power valves has the effect of reducing greatly the screen current. The helical coils of the screen and the grid are wound so that each turn of the screen is shaded from the cathode by a turn of the control grid. This alignment of the screen and grid causes electrons to travel in sheets between the

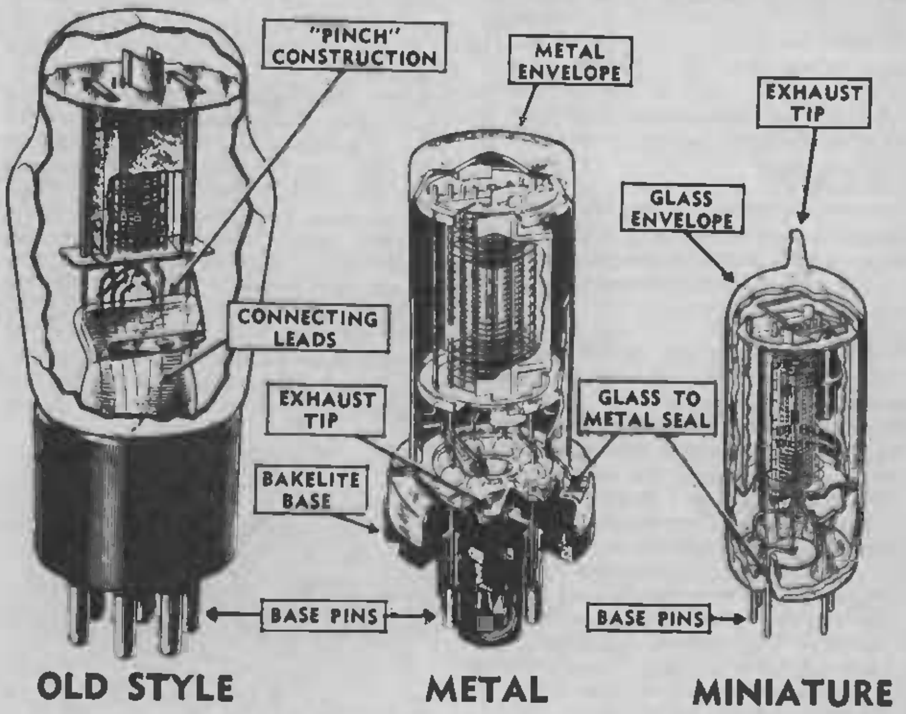


Fig. 10. Construction details of three types of thermionic valves. At left, an old-style "G" glass type with a construction adapted from an ordinary light bulb. In the centre, a metal-envelope "octal" (8-pin) type; and at right a more recent miniature all-glass type. Most valves still in use are of the last-named type.

RECEIVING AND AMPLIFYING TUBES

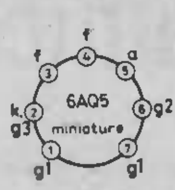
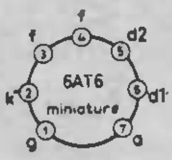
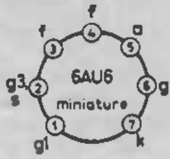
Type and application	V_f *(V)* I_f *(A)*	*Characteristic data*			*Limiting values*	*Base connections*
		Operating class A				
6AQ5 Output pentode	6.3 0.45	$V_a = 250$ V $V_{g2} = 250$ V $V_{g1} = -12.5$ V $I_a = 45$ mA	$I_{g2} = 4.5$ mA $S = 4.1$ mA/V $R_i = 52$ kΩ $R_{a\sim} = 5$ kΩ	$I_a = 47$ mA $I_{g2} = 7$ mA $W_a = 4.5$ W $V_i = 8.8$ V_{rms} $d_{tot} = 8\%$	$W_a = 12$ W $V_{kf} = 100$ V	
		Typical (triode)				
6AT6 Double diode Triode	6.3 0.3	$V_a = 250$ V $V_g = -3$ V $V_a = 100$ V $V_g = -1$ V	$I_a = 1$ mA $S = 1.2$ mA/V $I_a = 0.8$ mA $S = 1.3$ mA/V	$R_i = 58$ kΩ $\mu = 70$ $R_i = 54$ kΩ $\mu = 70$	$W_a = 0.5$ W $V_{kfp} = 90$ V	
		Typical				
6AU6 Sharp cut-off pentode	6.3 0.3	$V_a = 250$ V $V_{g2} = 150$ V $V_{g1} = -1$ V	$V_{g3} = 0$ V $I_a = 10.6$ mA $I_{g2} = 4.3$ mA	$S = 5.2$ mA/V $R_i = 1$ MΩ	$W_a = 3.5$ W $V_{kf} = 100$ V	

Fig. 11: While manufacturers make available to designers very detailed data about individual valves, abbreviated information like that above is published for the guidance of electronic servicemen and enthusiasts.

turns of the screen so that comparatively few of them flow to the screen.

The ultimate effect of forming the electron stream into horizontal and vertical beams, together with the reduction of screen current and other refinements, results in beam power valves being very efficient in operation.

It is possible to make quite a long story about valve design and development, for the evolution of the pentode structure has not by any means satisfied all requirements. Special circuit arrangements have brought the need for pentodes with a wide variety of characteristics. For example, "super-control" or variable-mu pentodes were evolved to permit smooth control to be achieved over the gain of a receiver.

Furthermore, numerous valves have been developed and released which contain in the one envelope two or more distinct electrode structures. Most rectifier valves contain two diode structures but twin triodes and triode-pentodes are quite common.

Physically, there is a like variety in design. Envelopes may be of either glass or metal, and fitted with bases having anything from three to 14 or more pins. Some glass envelopes are sprayed with a metallic coating, and overall size ranges from 6 to 7 inches in height to less than an inch for special types. The beginner can do no more than keep a very open mind on the whole subject, becoming acquainted with the characteristics of particular types, as they are met with in practice.

There is, unfortunately, little standardisation in the manner in which valves are depicted schematically. Generally, however, the filament or heater is shown as an arc at the bottom of a circle or rounded oblong representing the envelope.

A larger arc or a straight line in proximity to the filament indicates the presence of a cathode sleeve.

A solid or open rectangle at the top of the circle depicts the plate, while the various grids are shown by dotted or zig-zag lines between cathode and plate, generally in the same progressive order as they are found in the actual valve.

Typical valve symbols are shown in Fig. 9, but only familiarity will enable the student to interpret all circuits without hesitation, because of small departures from the usual practice.

Physical and electrical details of the various types are given in valve data charts released by most valve manufacturers. The various type numbers are listed in alphabetical or numerical order followed by details as to type, dimensions, base connections and electrical characteristics. It is a good idea to obtain one of these charts and spend an hour or so studying its columns, to become familiar with the presentation of valve data.

At the left is a fairly large transmitting type, about half size. At the top are two receiving valves to the same scale. Above is a miniature "nuvistor" with its socket.

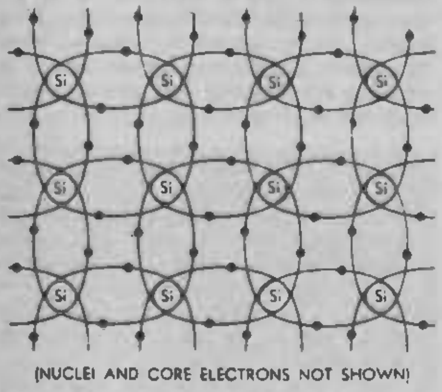
CHAPTER 8

Semiconductor Devices

Semiconductor materials — germanium and silicon — intrinsic semiconductors — "doping" with impurities — "P-type" and "N-type" material — the P-N junction — the junction diode and its properties — the bipolar transistor and its operation — leakage and gain — the field-effect transistor or FET.

In an earlier chapter, we learned of the atomic theory of the structure of matter, and how the atoms of the various elements are composed of differing numbers of electrons, protons and other particles. We saw that all atoms consist of a central nucleus of protons and other particles, around which "orbit" a number of electrons — like a miniature planetary system.

The orbiting electrons of every atom are imagined as being grouped in energy levels or "shells". In atoms of a particular element, the number of electrons in the outermost shell of each atom is of particular interest, since it determines many aspects of the behaviour of the element, both chemically and electrically. It provides the basis, in fact, for the whole discussion which follows.



(NUCLEI AND CORE ELECTRONS NOT SHOWN)

Fig. 1: The silicon atoms in a piece of crystal are bound together into a "lattice" by their valence electrons.

Considering solid materials, broadly, an element will generally allow current to flow through it in the form of a passage of electrons **if it has a small number of loosely-bound electrons in its outer shell.**

Elements which exhibit this characteristic are those which we have already defined in an earlier chapter as "conductors."

If, on the other hand, an element has an outer shell containing a number of electrons either near or equal to the normal maximum number (usually eight), it is generally much harder to produce a current of electrons through the material. Such elements are the "insulators."

The disinclination of insulators to support current flow may be pictured as being due to an increase in the "binding" of the outer electrons when there are close to the optimum number of them in the outer shell. The more there are, the more force must be exerted, in the form of an applied EMF, before they can be persuaded to move.

It so happens, however, that certain elements, such as germanium and silicon, exhibit electrical properties about midway between conductors and insulators, as well as behaving in other interesting ways.

These elements, as one might perhaps expect, have four electrons in their outer shell. They are called **semiconductors**, a name which includes not only elements but also certain chemical compounds having a similar electrical behaviour.

In a lump of a substance, the atoms or molecules of the element or compound comprising it, tend to be cohesive rather than to drift apart. There are a number of different ways in which they may be "bound together" but the type of bond which concerns us here is that which occurs in the semiconductors. This type of bond is called "covalent bonding" and is represented diagramatically in Fig 1.

Here the loops linking the "Si" symbols represent the outer or "valence" shells of each of 12 atoms of silicon (the inner shells of each are omitted for clarity). Note that each atom shares each of its four valence electrons (the round blobs) with the four neighbouring atoms. Thus any two adjacent atoms are effectively "sharing a pair" of electrons and each atom, in one sense, has eight electrons in its valence shell. This produces a stable bond and a disinclination to pass current.

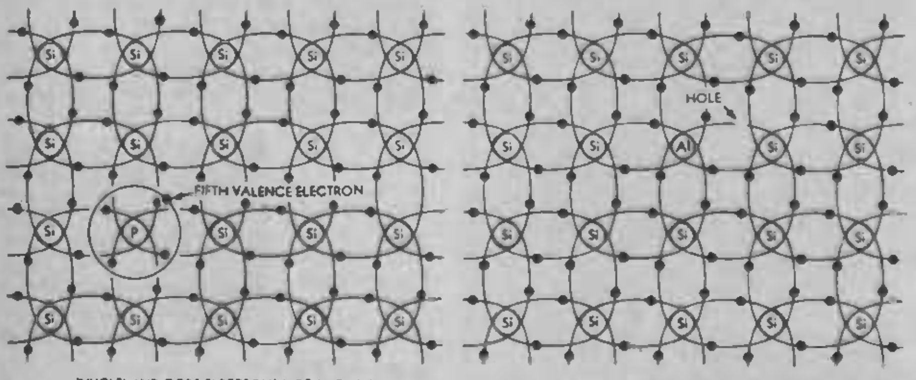
Silicon or germanium atoms, arranged as in Fig 1, in a typical pure crystal, are said to be in a "lattice," as the diagram would suggest.

As already mentioned, pure silicon or germanium having this lattice structure has a rather poor conductivity (high resistance), as all the outer shell electrons are employed in the lattice bonding and are fairly strongly bound.

If heat is applied, the added thermal energy causes the atoms and electrons to become energetic, so that a certain number of electrons are freed and made available for current flow. At low temperatures, however, the conductivity of pure silicon or germanium is poor.

Pure silicon and germanium are known as **intrinsic semiconductors**, because of this behaviour. The thing to remember about such intrinsic semiconductors is that they are virtually insulators at low temperatures, and only become conductors if energy is applied to their crystal lattice, for example in the form of heat (thermal energy).

If small amounts of other elements are added to the silicon or germanium as controlled "impurities", their conductivity at normal temperatures may be greatly increased. This may be made to occur in two different ways, by adding one of two groups of impurity elements. This process is



(NUCLEI AND CORE ELECTRONS NOT SHOWN)

(NUCLEI AND CORE ELECTRONS NOT SHOWN)

Fig. 2 (left), and Fig. 3 (right): The effects of adding to the crystal "impurities" having either 5 or 3 valence electrons.

known as "doping" the semiconductor.

Semi-automated transistor assembly at the Fairchild plant at Croydon, Victoria. Working with the aid of a binocular microscope, an operator is attaching individual transistor elements or "dies" to a punched 50-unit frame which will later be separated into 50 separate transistors.

For instance, if a small amount of an element having not four, but five electrons in its outer shell (such as phosphorus, arsenic, or antimony) is added, conditions such as that shown in Fig 2 may arise.

The impurity atom joins in the lattice structure and forms four covalent bonds with the surrounding atoms, but is still left with a "spare" electron. Thus this type of impurity is called "donor" impurity, as its addition results in there being free electrons available for current conduction through the crystal lattice.

Silicon or germanium crystal with a small amount of donor impurity is called N-type germanium, the "N" suggesting additional negative charges. Semiconductor materials in general having donor impurity are known as N-type semiconductors.

The other type of impurity which may be added is one which has only three electrons in its outer shell, such as boron, aluminium, gallium or indium. When small quantities of any such element are added to the germanium, conditions such as those shown in Fig 3 appear in the lattice.

Here again, the impurity atom joins in the lattice structure, but in this case it is only able to enter into covalent bonding with three of the surrounding atoms. There is no fourth electron to share with the remaining atom, so a defect in the form of an electron vacancy or hole is set up.

It happens that such holes are able to "travel" through the lattice from atom to atom, in much the same way as the free electrons do in the N-type material. Thus, under the influence of thermal (heat) energy or light energy or the electric field of an applied EMF, they move around the lattice and effectively constitute a current of positive charges.

The type of impurity which produces holes in the lattice is termed an "acceptor" impurity for, when the hole moves away from the impurity atom, the atom is regarded as having "accepted" a fourth valency electron from the atom which then finds itself with the hole.

A semiconductor material which has been doped with an acceptor impurity is termed P-type, to distinguish it from N-type material by indicating that the conduction is by holes rather than electrons. While hole movement may be interpreted as movement in the opposite direction of an electron or electrons, the holes concept is a convenient and simple way of imagining the different types of conduction in N-type and P-type materials.

It is interesting to note that the impurity atoms scattered throughout the parent material occupy normal positions in the lattice structure, just as if they were atoms belonging to the parent material (ie, atoms of silicon). However, the fact that the valence bond system associates each with one outer electron too few, or one too many, means that the orbiting electrons do not exactly balance the charge of the nucleus.

Thus, although the impurity atoms occupy normal positions in the lattice, they are actually atoms with an electrical charge — ionised atoms, in other words, or simply ions. Atoms of a donor impurity, having been deprived of one of their normal outer electrons, become positive ions. Atoms of an acceptor impurity, on the other hand, with one more than their proper number of outer electrons become negative ions.

Note that the ionised donor and acceptor atoms are fixed in the crystal lattice, in contrast with the conduction holes and electrons. Their polarity is opposite to that of the current carriers which they have introduced — the donor atoms in N-type material are positively charged, while the acceptor atoms in P-type material are negatively charged. And as there are equal numbers of fixed ions and current carriers in the material, the charges cancel out.

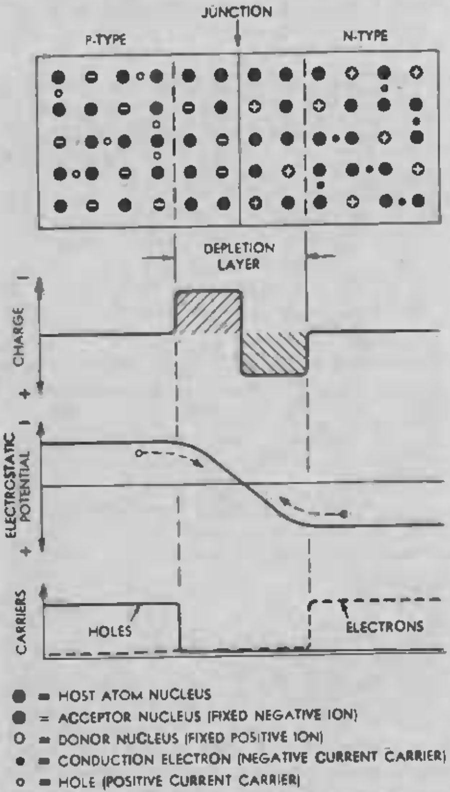


Fig. 4: When P-type and N-type regions are formed adjacent to one another in a crystal, a "depletion layer" is formed.

Thus the two types of impurity semiconductors should be visualised as (1) N-type, having fixed positively charged donor ions in the lattice and normally equal number of "free" (for current conduction) negative electrons, and (2) P-type, having fixed negatively charged acceptor ions in the lattice and a normally equal number of positive holes able to conduct current by moving through the bonding system of the crystal lattice.

The presence of impurity ions in the lattice structure of, say, silicon, with the attendant "mobile" (movable) electrons or holes, greatly modifies its conductivity.

Now let us see what happens when a crystal is arranged so that an area of P-type and an area of N-type are next to one another. Such a state of affairs is shown in Fig 4, which is a diagram representing what we can now call a P-N junction — simply a junction between a P-type and an N-type lattice. As we shall see, such a junction behaves in a very similar way to the thermionic diode valve which we examined in the last chapter, and thus is used as the basis of the semiconductor junction diode.

At normal temperatures, due to thermal energy, the current carriers — electrons in the N-type, holes in the P-type — wander freely throughout the lattice. Some may wander across the junction in the course of their travels.

If a current carrier from one side wan-

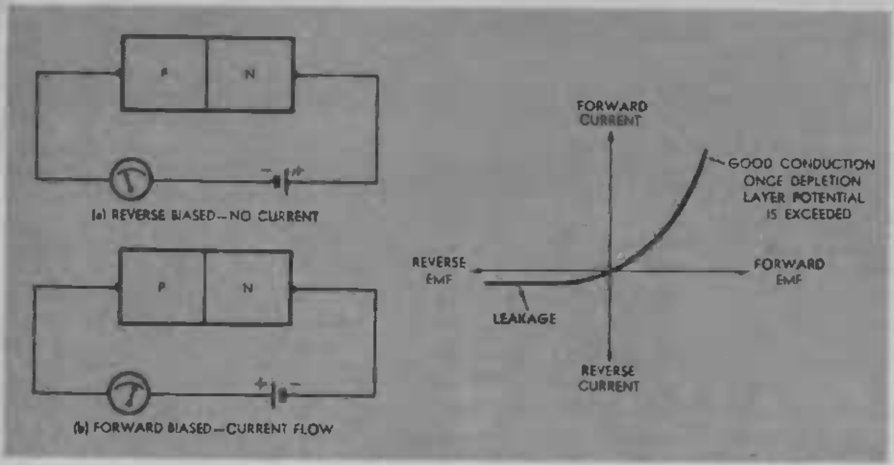


Fig. 5 (left), Fig. 6 (right): Because of the behaviour of the depletion layer, a P-N junction acts as a "one-way" conductor.

ders across the junction, there is a good chance that it will meet a carrier of the other type, each neutralising the other. For instance, if a conduction electron from the N-type side wanders into the P-type, it may easily meet a hole and "fill" it, the lattice in that immediate region reverting to normal. Similarly, when a hole appears in the N-type region, there is a very good chance that it will meet a conduction electron and the two will cancel one another as before.

Due to this mutually neutralising action, at normal temperatures there is a section of the crystal immediately surrounding the junction which has virtually no carriers. This is called the **depletion layer**, signifying that it is effectively a "no-man's land" as far as free current carriers are concerned.

In the P-type section of the depletion layer there are no free holes present, and there is thus a net negative charge — due to the fixed acceptor impurity ions. Similarly, the depletion layer in the N-type region acquires a net positive charge, due to the fixed donor impurity ions and the lack of free electrons.

As the small graph in Fig 4 (labelled "charge") shows, this means that the section of the crystal occupied by the depletion layer differs from the rest of the crystal in having net charge "humps," one each side of the junction. There is a negative charge "hump" on the P-type side, and a positive charge "hump" on the N-type side (this is shown as a "dip" because it is of the opposite polarity).

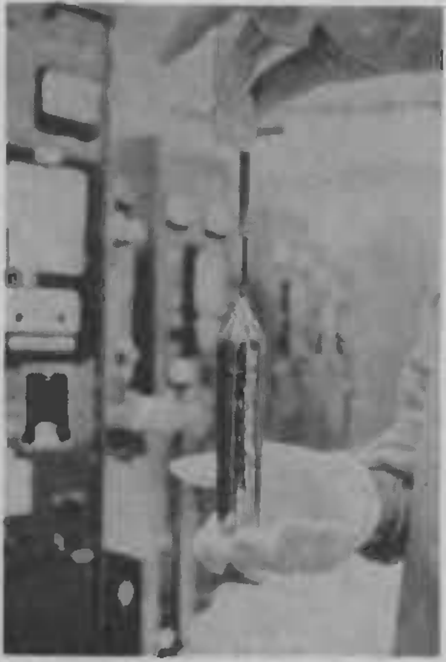
Because of this setting-up of opposite charge concentrations, each side of the junction, the P-type and N-type sections are effectively shifted in potential with respect to one another. For instance, the P-type region is effectively negatively-charged with respect to the N-type, because electrons in the N-type are prevented from passing to the P-type region due to the "hump" of negative charge on the far side of the depletion layer, (remember, like charges repel one another.)

The same is true of holes in the P-type region, which are discouraged from passing to the N-type region by the net positive charge N-type part of the depletion layer.

As the second curve of Fig 4 shows, we can represent the potential difference which is set up by the depletion layer by showing a gradient or slope in the depletion region, with the P-type negative with respect to the N-type. Or the N-type positive with respect to the P-type, whichever you prefer.

Looking at this potential diagram one can see the potential "hill" that carriers would have to surmount if they were to cross the depletion layer and reach the other region. Right-way-up, the hill for electrons may be seen, travelling from right to left; inverting the page will show the hill seen by the holes.

It is the potential "hills" set up at the junction as part of the formation of the depletion layer which under normal conditions cause the layer to remain at a certain constant width, and with a certain potential difference between the P-type and N-type sections of the crystal. For a given amount of thermal energy, electrons and holes cross the junction only until they set up the depletion layer and its potential hills to an extent which discourages any further movement. A balance or **equilibrium** is thus reached, with the width of the depletion layer and the potential difference between the two sections determined at least partly by the amount of thermal energy present.



First step in the manufacture of most semiconductor devices is to grow large single crystals of pure silicon or germanium. This picture shows a single crystal of pure silicon, from which many tens of thousands of devices will be made.

With us so far? If you're not too sure, it would be a good idea to read some of the foregoing material again, as the concepts involved are quite important.

Now if we try to pass a current through a germanium junction diode or "crystal" having a P-N-junction, by applying an EMF across its end, we find that interesting things happen. And it all depends upon the way we connect our EMF, for different things happen for each of the two possible connection polarities.

If we connect an EMF so that its positive polarity connects to the N-type region and its negative polarity to the P-type region, as in Fig 5(a), we are simply increasing the potential difference which the depletion layer has already set up between the two regions. What happens is that we simply increase the width of the depletion layer (and increase the steepness of its potential difference gradient) by causing more carriers in the crystal structure to cancel.

After a very brief and tiny flow of these carriers no current flows apart from a leakage current which is due to various lattice imperfections and effects which need not worry us just at present. The initial flow amounts to a capacitive charging current, just as if the junction was a capacitor — which, in fact, it is.

In this condition, which is called reverse biasing, the junction may thus be regarded as an open circuit. Or more exactly, as a low value capacitor with a small leakage current.

If on the other hand we connect an EMF to the crystal so that its positive and negative polarities connect to the P-type and N-type regions respectively, as in Fig 5(b), the situation is quite different. In this case the external EMF opposes that set up by the depletion layer, rather than assists it.

If the external EMF is greater than that of the depletion layer (which is in practice only a few hundred millivolts) there will thus be a net EMF acting in the direction of the external EMF, and the potential "hills" of the junction will be compensated by larger "downhill" sections. Current is thus able to flow, in terms of both electrons and holes within the crystal and electrons alone in the external circuit.

A junction in this condition is said to be **forward biased**. Forward biasing thus results in high conductivity (low resistance) and easy current flow.

Fig 6 shows a typical current-voltage curve for a semiconductor junction — note the low current with reverse biasing (due to leakage) and the sharp rise in current with forward biasing as the depletion layer potential is exceeded.

As this curve shows, the semiconductor junction is thus very similar to the diode valve which we examined in the last chapter — it is effectively a "one-way" path for conduction.

Sealed in a small glass shell or plastic or metal package for moisture protection, and fitted with two connection leads, it is the familiar junction diode which is used in "crystal" sets, in radio and TV receivers for detection and other jobs, and in power supplies for rectifying AC into DC. We see some of these uses in later chapters.

Many modern silicon diodes are made with the P-type and N-type regions formed together in the same tiny "chip" of crystal, rather like the structure of Fig 4. The regions are formed as part of the same crystal structure either by a controlled doping process when the crystal is being grown, or by growing a crystal which is uniformly doped with one type of impurity, and then changing certain areas into the other type of material by "diffusing" atoms of the opposite type of impurity into them from a surrounding vapour at high temperatures. The structure of a modern silicon diode of this type is shown in Fig 7. Note that the P-type and N-type regions of the chip are merely two electrically different regions within the same physical piece of material.

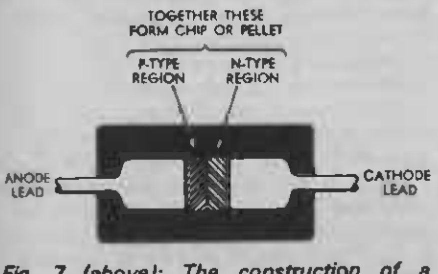


Fig. 7 (above): The construction of a modern silicon rectifier diode. Fig. 8 (right): The construction of a low-power glass package germanium detector diode.

Fig 8 is a diagram of a typical small germanium junction diode, showing its slightly different construction. During manufacture the piece of N-type germanium is soldered to the cathode lead, and the fine wire "cat's whisker" contact welded to the anode lead so that it is in contact with the surface of the germanium. When assembly is complete, a short but heavy pulse of current is fed through the device, whereupon a small amount of the wire material melts and dissolves in the germanium near the contact.

The composition of the wire and the techniques used are such that this produces a very small area of P-type germanium in the immediate vicinity of the wire contact, to which the wire is firmly welded. Thus a P-N junction is formed which is small in area — and thus low in capacitance, which is desirable for many applications — yet quite robust.

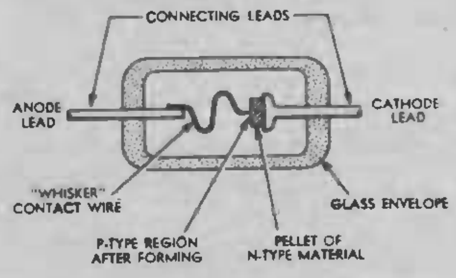
The usual schematic symbol for a junction diode is shown in Fig 9.

Having discussed the junction diode we can now progress to the bipolar transistor. This is simply the junction diode carried one step further — a semiconductor crystal having not one but two P-N junctions. They are near one another and, as we shall see, they interact in a way which makes the transistor able to amplify small signals. The bipolar transistor is thus the semiconductor version of the triode valve, in the same way that the junction diode corresponds to the diode valve.

There are two different types of bipolar transistors, as there are two different ways in which one can arrange two P-N junctions to be adjacent. One type is the PNP type, where, as the letters suggest, the two junctions share a common N-type region and have separate P-type regions; the other is the NPN type, where the two junctions share a common P-type region.

Both types are quite practical and, in fact, they are often used side by side in circuits — one type being best suited for some jobs, the other type for different jobs.

Fig 10 shows a diagram of an elementary PNP transistor. The common N-type region in the centre is thin, and is called the base region. The left-hand P-type region which is called the **emitter** is normally forward-biased with respect to the base by means of the battery E1, so that the base-emitter junction conducts freely.



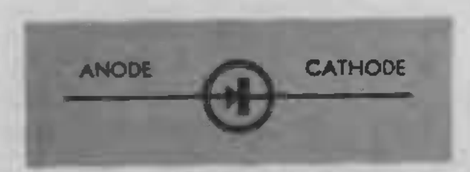


Fig. 9 (above): The usual schematic circuit symbol for a semiconductor diode.

The right-hand P-type region, termed the **collector**, is biased in the opposite direction by battery E2, so that the base-collector junction is reverse-biased and would normally be considered to be a high resistance. However things are not the same as they would be with a single junction.

As we mentioned before, the base region is made quite thin. It is also made weakly N-type by only doping it with small quantities of a donor impurity, whereas the emitter region is made quite strongly P-type by strongly doping it with an acceptor impurity.

This "differential" or unbalanced doping has one major effect: it makes holes play the major part in the current conduction of the emitter-base junction. In other words, most of the current passed by the junction is in the form of holes passing from emitter to base, rather than in the form of electrons passing in the reverse direction.

Because the base is thin, the base region part of the reverse-biased collector-base junction depletion layer extends almost to the emitter-base junction. This would not normally have much effect on the base, as the normal base carriers are electrons and would not be inclined to "climb" the potential "hill" to the negatively-charged collector. But it has quite an effect on the holes which reach the base from the emitter.

To these, the "hill" is not an uphill grade but an inviting downward slope. Thus, unless they are met and cancelled by one of the free electrons in the base as soon as they arrive, there is a good chance that they will "roll down" the potential hill into the collector region.

Confused? Then consider it from another viewpoint. The base-collector junction is reverse-biased, meaning that carriers would normally not pass from one side to the other because of the fixed-ion concentrations and potential "hill" of the junction depletion layer. But the "hill" is only an upward grade for the particular carriers normally present on each side.

What the emitter-base junction does, because of its forward bias, is to inject the "wrong sort" of carriers into the base region. Not the usual electrons, which see an upward grade to the collector region, but holes — which see the potential gradient as a downward grade, and may therefore pass straight through.

The holes "haven't been told" that the collector junction is meant to be an open circuit, as it were, and they accordingly have no objection to crossing over!

Not all of the holes which cross the emitter-base junction pass through into the collector. A few are cancelled by electrons before they can do so, and the base region accordingly receives a corresponding number of electrons from the external base connection; in other words, a small amount of the original emitter current flows "out" via the base lead, although most of it passes through to the collector.

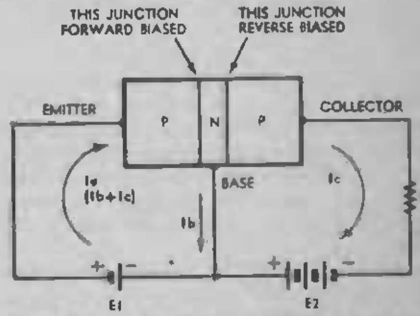
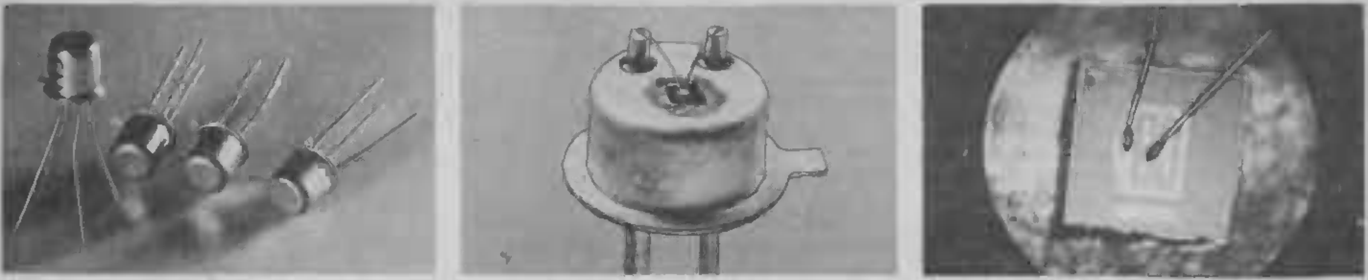


Fig. 10: An elementary bipolar transistor, whose operation relies upon the interaction between two P-N junctions.

The currents in the transistor can be visualised in terms of the base lead current Ib and the collector lead current Ic. In Fig 10 Ib can be imagined as flowing in the small lefthand circuit loop, while Ic flows around the large outside loop. Ie may be considered as simply the sum of the two, as they both flow through the emitter lead.

Due to the action of the transistor, there will be a definite ratio between I b and I c, the collector current normally being many times Ib. It happens that if we change Ib by a small amount, thereby modifying the base-emitter depletion layer, Ic will be found to change by a proportional and much greater amount.

In other words, the transistor will amplify small current changes. If we superimpose a small signal on the base current Ib, a larger



Three pictures which further illustrate the construction of a silicon planar transistor. At left are completed devices; in the centre, a device without its metal can; and at right, a close-up of a chip with its connection wires.

replica of the signal will appear in the collector current Ic. If we include a resistor Rc in series with the collector, it will develop a voltage drop which will be proportional to the current changes, to produce an amplified output voltage.

Like the triode valve, then, the bipolar transistor can amplify tiny signals so that they may be used to perform such tasks as operate earphones, loudspeakers, television picture tubes and so on. It is thus a very useful device in many fields of electronics — particularly as it does not need to be supplied with heater current as does the triode valve.

Before we end this discussion of the bipolar transistor, there are one or two points which we should look at. The first is the matter of the amount of amplification obtained.

It should be fairly obvious from the foregoing discussion of the operation of the transistor that the amplification or gain will be governed by the number of holes which pass straight through the base from emitter to collector. The fewer the number of holes which are intercepted by electrons in the base to produce Ib, the higher will be the gain.

In manufacturing the transistor the gain can thus be manipulated by two methods: (1) making the doping ratio between the emitter and the base high so that the current is mostly holes going from emitter to base, and (2) making the base thin so that there is less opportunity for the holes to meet electrons before crossing to the collector. There are also a number of other techniques which need not concern us at present.

There are a number of ways of expressing the gain of a transistor, but many need not concern us here. For the present, it is probably sufficient to know that there are two common terms used, both expressing gain in terms of current ratios.

The first is alpha (α) which expresses the gain in terms of the ratio between the collector and emitter currents:

$$\alpha = Ic/Ie \qquad (1)$$

Since the Ic is always less than Ie, *a* is always smaller than one. It is very close to one for high gain transistors, and lower for low gain transistors.

The other current gain ratio is beta *(β)*, which expresses gain in terms of the ratio between collector and base currents:

$$\beta = Ic/Ib \qquad (2)$$

Beta is almost always more than one, and generally much higher. Most modern transistors have a beta of between 50 and 500, although some may go as low as 20, or as high as 10,000.

Since alpha and beta are merely different expressions of the same basic effect within the transistor, knowing one always allows the other to be calculated.

$$\beta = \alpha/(1 - \alpha) \ldots\ldots\ldots (3)$$

and conversely

$$\alpha = \beta/(1 + \beta) \ldots\ldots\ldots (4)$$

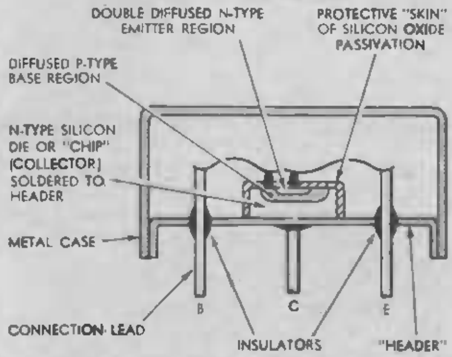


Fig. 11 (above): The construction of a modern silicon planar transistor. The size of the chip has been exeggerated to show its various regions more clearly.

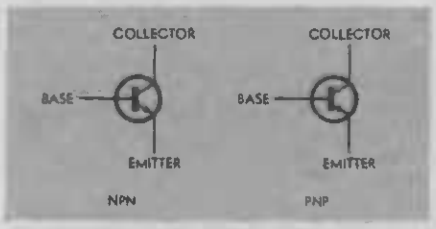


Fig. 12: The usual schematic circuit symbols for bipolar transistors.

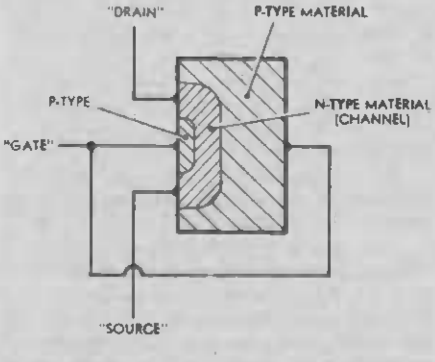


Fig. 13: The basic structure of a field-effect transistor or FET, whose operation is rather like that of the triode valve.

It should be realised that due to thermal energy at normal ambient temperatures, some electrons and holes are generated in small quantities in all types of semiconductor — intrinsic, N-type and P-type. Thus, in the bipolar transistor, even when an internal input current is not applied, there are small currents flowing because thermally generated carriers "of the wrong sort" are present in small numbers and are able to cross the various junctions.

For instance, due to heat energy, there will be some holes present in the N-type base region even in the absence of holes being injected from the emitter, and these will be able to "roll down the hill" into the collector.

These thermally generated "minority carriers," as they are called, are the major reason for the small leakage currents which are observed to flow in the transistor even when no input is applied. As the number of minority carriers generated depends upon the temperature, the leakage currents are likewise proportional to temperature, both externally applied and internally generated. This is one of the things which tends to complicate transistor circuits.

Although in the foregoing description of bipolar transistor operation we have considered only the PNP type, the operation of the NPN type is very similar, except that the operating potentials and currents are reversed. The main difference is that the principal conduction carriers are electrons, rather than holes. The reader might find it worthwhile to trace the operation of an NPN transistor out for himself, by working through the description again and substituting electrons in place of holes.

Fig 11 shows the construction of a modern silicon planar bipolar transistor, illustrating the form taken by a practical device. The collector, base and emitter regions of the transistor are formed within a single tiny chip of silicon, by the impurity diffusion process mentioned earlier.

Fig 12 shows the usual schematic symbols used to represent NPN and PNP bipolar transistors in circuits.

Although the bipolar transistor may be regarded as the semiconductor equivalent to the triode valve, there is one vital difference. The valve controls output (plate current by input voltage (grid bias), and has a very high resistance input. It draws microscopic input current, if any at all. On the other hand, the bipolar transistor for its very operation must draw at least some base input current — and may hence present an input resistance as low as a few hundred ohms.

Using suitable circuit configurations it is possible to arrange for a bipolar transistor to present an effective input resistance as high as several megohms; however, in order to obtain higher figures, it is generally necessary to employ a different type of device.

Such a device is the **field-effect transistor or FET**, which consists basically of a thin "channel" region of one type of semiconductor material surrounded on both sides by regions of the opposite type.

Like the bipolar transistor, the FET may be made in two possible versions. It may be made in a P-channel version, consisting of a thin channel of P-type material surrounded on both sides by N-type regions; or alternatively in an N-channel version, in which the channel is of N-type material and the surrounding regions are P-type material.

The construction of an N-channel FET is shown in Fig 13. It may be seen that the channel is a U-shaped region of N-type material formed into a chip of primarily P-type material. The P-type regions on each side of the channel are connected together to form the "gate" electrode of the device, while connections made to each end of the channel region form the "drain" and "source" electrodes.

The operation of the FET depends upon the fact that, as we saw earlier in this chapter, electrons and holes wandering across a P-N junction tend to neutralise one another and cause a carrier-free depletion layer to be set up on either side of the junction. Because this depletion layer region is exhausted of current carriers, it is virtually an insulating region unable to conduct current.

The P-N junctions forming the sides of the channel of a FET will have such depletion layer regions, and the depletion layers will extend into the channel. Because they are virtually insulating regions, they therefore tend to make the effective electrical width of the channel somewhat narrower than its physical width. And if an external voltage is applied between the gate and channel regions so as to reverse bias the junctions, the depletion layers will expand further into the channel to make it narrower still.

The effect of this narrowing of the channel is to reduce its ability to conduct current between the source and drain electrodes connected to its ends. In fact, if the reverse bias connected between gate and channel is increased sufficiently, the depletion layers will extend right into the centre of the channel and meet, virtually turning the whole channel into an insulator. Naturally this effectively "blocks off" the channel as a conducting path as far as the source and drain are concerned.

If a potential is connected across the length of the channel by means of an EMF applied between drain and source, it is found that the width of the gate-channel depletion layers may be used to control the resulting current. The drain-source current may be varied simply by adjusting the reverse bias applied to the gate-channel junctions, and a relatively small change in gate bias can produce a relatively large change in drain-source current.

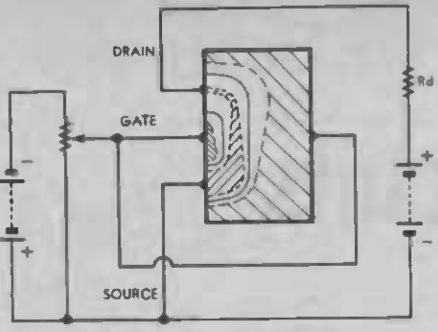


Fig. 14: Because a small negative bias on the gate of a FET controls the channel current, the device can amplify.

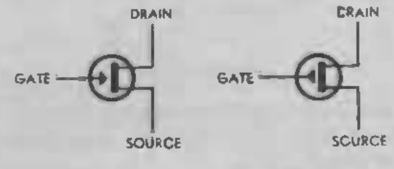


Fig. 15: The usual schematic circuit symbols used for FET devices.

The FET is thus capable of amplifying small signals, in much the same way as a triode valve. In fact, it is a much closer equivalent to the triode valve than the bipolar transistor, because like the valve its input control electrode — the gate — does not draw significant current. Thus the gate of the FET corresponds to the grid of the valve while the source and drain correspond to the cathode and anode or plate, respectively. This may be seen by comparing Fig 14 with Fig 6 in chapter 7.

Like the thermionic valve, the amplification of a FET is expressed in terms of transconductance, and measured in milliamps / volt. The usual schematic circuit symbols for FET devices are shown in Fig 15.

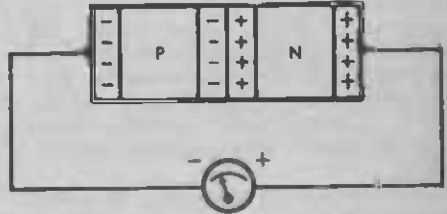
In passing, it should perhaps be noted that the type of FET we have just described is only one of a variety of practical types. Another type which the reader may meet is the MOSFET, in which the gate and channel are insulated from one another by a thin layer of oxide.

There are many other types of semiconductor device besides those which we have looked at in this chapter, but unfortunately space does not permit dealing with these here. Happily the reader will find that the junction diode, bipolar transistor and FET devices are those which he will tend to meet most often, so that the basic grounding provided in the foregoing should be found very useful. More information on semiconductor devices is general may be found in our companion handbook, "Fundamentals of Solid State".

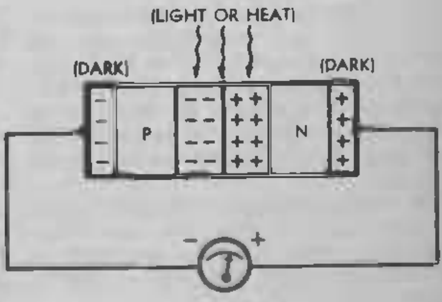
A final note before we leave the topic of semiconductor fundamentals. In reading the section earlier discussing a P-N junction "depletion" layer, it might have seemed that one should be able to measure the depletion layer potential of a junction using a sensitive voltmeter. This is not the case, however, because as soon as an external circuit of any description, such as a measuring circuit is connected to such a diode, another depletion "layer" forms. It forms in two parts, one at each end of the crystal where the wires connect (see Fig 16), and the two "half depletion layer" potentials equal and exactly cancel that of the central junction. With no net circuit voltage, the meter will show nothing.

While we cannot measure the depletion layer voltage of a normal diode, we can measure a voltage across a special type of diode known as a "solar cell." In this type of diode, the "main" P-N junction is made so that it can be exposed to light or heat radiation while the rest of the diode remains unexposed.

Because greater radiation energy is given to the carriers near the exposed junction, the depletion layer which is set up has a greater voltage than that at the connections, which are "dark" and at ambient temperature only. Thus there is a net voltage produced, which can be used to operate small radio receivers or other devices.



NORMAL DIODE — NO CURRENT



SOLAR CELL

Fig. 16: Diagrams used to explain the operation of the solar cell.

WHEN DID THEY START?

The first bipolar transistor was produced in December, 1947, at the Bell Telephone Laboratories in New Jersey, USA. It was produced by researchers Dr William Shockley, Dr John Bardeen and Dr Walter H. Brattain, who were awarded the 1956 Nobel Prize in Physics for their discovery.

Although it did not become a practical device until much later, the field-effect transistor was actually invented between 1925 and 1928, by an American engineer named Julius E. Lillienfeld.

CHAPTER 9

Reading Circuits

The language of electronics — schematic diagrams versus wiring diagrams — junctions and crossovers — standard circuit symbols — polarity and current flow — drafting practice for valve and transistor circuits — DC power line configurations — interpreting schematic diagrams.

Circuit diagrams, like algebra and shorthand, are a means of putting complicated ideas down on paper in the simplest possible form. They are usually called schematic diagrams, indicating they communicate the plan or "scheme" or a circuit but leave out the physical detail.

Also like algebra and shorthand, schematic circuit diagrams are a bit mysterious to those who don't understand the symbols.

But if you take the time to familiarise yourself with the symbols, which are reasonably standard throughout the world, and if you get some practice using them, you will find that circuit diagrams begin to make sense.

Schematic circuit diagrams are the language of electronics, and as with any other language, the beginner has to start with simple constructions and gradually learn to arrange and rearrange these "building blocks" into complex electronic devices.

Newcomers to the electronics field often wonder why we bother with schematic diagrams when wiring diagrams seem to be so much easier to follow, particularly for the newcomer to electronics.

Admittedly, wiring diagrams are easier to follow, in the purely physical sense, but they are not always available. They are usually supplied as part of a magazine article or a simplified construction kit when it is assumed that many of the people building the device are not familiar with schematic diagrams.

One of the principal drawbacks of wiring diagrams is that they are very time-consuming (and therefore expensive) to draw, particularly for more complex equipment.

Another reason why schematic diagrams are preferred is that they communicate the function of the various circuits — they tell us what's going on. A wiring diagram tells us what a part looks like and where to solder it into the circuit, but it does not readily tell us what that part's function is in relation to the other parts of the circuit.

On the other hand, if a circuit is wired by following a schematic diagram, the very reasoning required to wire the circuit correctly also teaches something about how it operates. The same type of reasoning is required when the circuit is not operating properly and you are trying to puzzle out why.

A schematic diagram is essential for rapid troubleshooting of electronic circuits; a wiring diagram is not.

A very important reason for learning the schematic symbols is that, as with any other language, you will sooner or later want to be able to express yourself in electronic terms. As you get more familiar with circuits you will want to modify them and maybe design a few of your own. You will soon be reaching for a pencil and scribbling down a circuit so you can study it, modify it and show someone else how it works.

A wiring diagram, being basically a drawing of the outward appearance of a circuit, uses solid lines to represent actual wires. In a schematic diagram, however, a solid line connecting two symbols means only that the two points are to be electrically identical, that is, they must somehow be connected together, but exactly how they connect is not shown and is usually unimportant. Different people could wire them together in different ways, with all of them being correct.

There are cases where the exact length or position of the wiring is important, especially in high frequency circuits, but even in these cases the schematic usually does not show the proper wiring method. Instead, the schematic will refer you to a partial wiring diagram or to written instructions, or sometimes rely upon your own experience.

Don't expect schematic diagrams to tell you everything you need to know about constructing an electronic device. Always refer to any additional information available if you want a device to work right the first time.

Where solid lines intersect on a schematic diagram, the draftsman must use a symbol to show whether the lines are connected or not. Unfortunately there is more than one method in common use. There are two alternative systems used by the majority of large industrial companies. Both are "simplified" systems intended to save drafting time, sometimes at the expense of confusing the reader.

With one system, shown in Fig. 1a, lines to be joined simply intersect at right angles and lines not to be joined are shown crossing over one another by means of a small semicircle. The other common industrial system, shown in Fig 1b, indicates joined lines by a large solid blob over the point of intersection. Unjoined lines are shown as simply intersecting with no symbol, exactly as with the first system, but with opposite meaning.

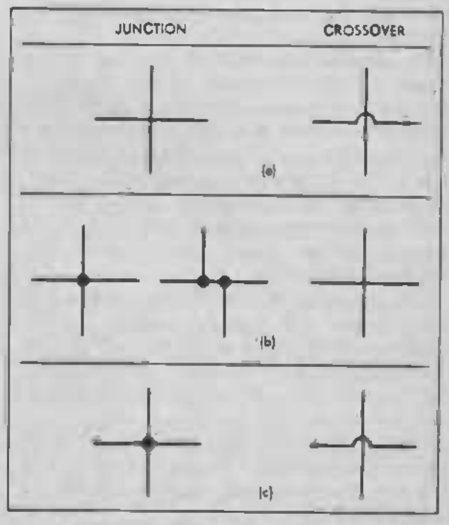


Fig 1. Three different systems in common use for distinguishing junctions from crossovers.

A failing of the second system is that copying and printing machines often add blobs of their own sometimes causing serious wiring errors. To avoid this, some companies offset each joined intersection and use two blobs to show a connection point (Fig. 1b).

Most engineers, technicians and hobbyists, when hand-sketching a circuit, use a combination of the two methods just discussed. They use the blob to show a connection and the semicircle crossover to show no connection. Almost all electronics hobby magazines also use this method (Fig 1c) as it avoids confusion and is relatively safe from "gremlins" in the printing presses.

Don't be too dismayed by the fact that several systems are in use (there are other variations of the ones we have discussed as well). Only one of the systems will be used in any particular schematic diagram you are looking at; it will have to be consistent within that diagram. It doesn't take long to recognise which method is being used.

Before we discuss how the various circuit symbols are used, let's take a look at how they are drawn. A collection of commonly

Some commonly used circuit symbols

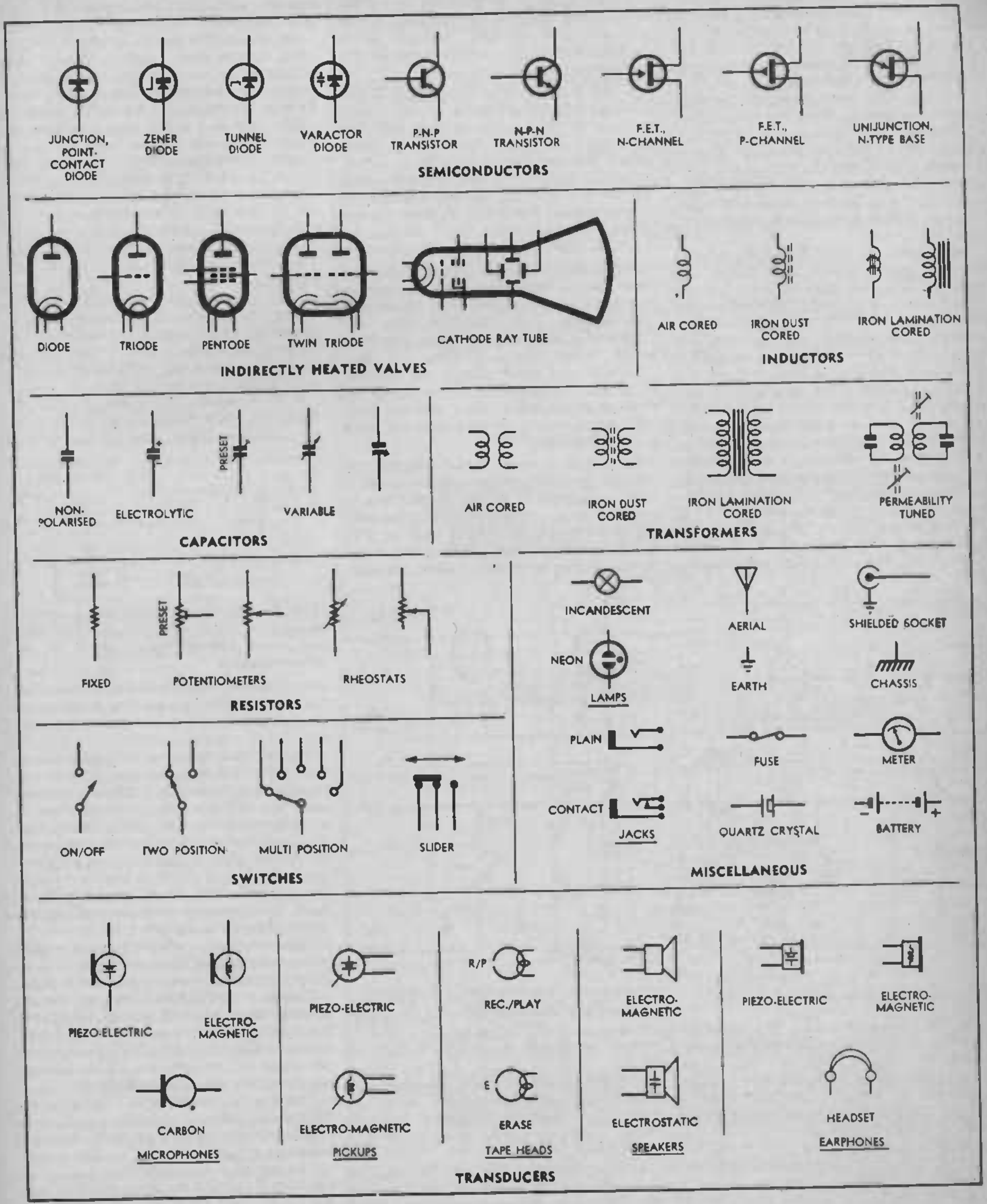


Fig. 2: Schematic circuit symbols vary somewhat not only from one publication to the next but also from one country to another. The ones shown above are a fairly representative cross-section of the types of circuit symbols that a beginner could expect to encounter. Although variations do exist, the electronics student should have no difficulty in recognising graphic representations. These tend to convey the functions of components rather than their physical appearance.

used symbols is shown in the full page chart in this chapter. You will find these symbols to be, with minor variations, extensively used in the electronics industry. But don't be confused if you see symbols drawn a little differently from those shown here. A valve, for example, may be drawn within a circle rather than an oval and its cathode may be a straight line rather than an arc, but it will functionally look similar.

You will find that schematic symbols usually try to communicate the function of a component rather than to copy its physical appearance. Notice that a resistor symbol attempts to show that electrons would have a harder path through it than through a wire; a capacitor looks like two plates with a gap between; a diode indicates one-way current flow, etc.

An exception to this is a recently adopted European system which uses simple rectangular shapes for most symbols, but the system is seldom used in English language schematics.

Components which are variable are shown with either an arrow or a long-tailed "T" symbol drawn at an angle through the standard symbol. The arrow generally means that the component is an accessible adjustment, such as a front panel volume control. The "T" symbol is usually used to indicate that the component is adjustable, but is a preset screwdriver adjustment inside the case and is not normally tampered with. In some circuit diagrams the arrow or "T" is drawn against the side of the symbol rather than crossing it at an angle.

Resistors, inductors and transformers which have connections (called "taps"), elsewhere than the normal connections at the ends, usually have either an arrowhead or a solid blob at each tap to indicate the junction. Symbols for variable or tapped components are not standardised, but as with the wire junctions, you can readily decide which system is being used on any particular schematic.

If the polarity of a component is important, it is shown by a + or — sign adjacent to the symbol, such as with the electrolytic capacitor shown in the chart. Where direction of current flow is important, it is shown by an arrowhead within the symbol pointing in the direction of "conventional" positive to negative current flow. The arrowhead in a diode, for example, points toward the cathode.

These arrowheads bring up a problem of word useage. When describing current flow in a circuit and talking about its relationship with voltage and resistance, the term "current" is not usually used in a directional sense. When someone uses the term "current" you often cannot tell whether they mean conventional positive-to-negative flow or negative-to-positive electron flow. And in many cases, it really does not make much difference.

As explained in an earlier chapter, the negative-to-positive electron flow concept is now considered to be the correct one in most instances, and is widely used to explain the behaviour of electrons in a circuit. This apparent contradiction in terms and arrow directions may confuse you when you use a schematic to figure out how to properly connect a semiconductor.

The principal thing to keep in mind is that electrons should be imagined as flowing **into the point** of an arrowhead inside any semi-conductor circuit symbol.

This means that the arrow should always point towards the negative side of the circuit. Since the negative supply has a surplus of free electrons, and electrons flow through the device toward the positive supply, this is a logical and easy way to remember correct polarity.

With an NPN transistor, for example, the arrow points toward the emitter, which must be connected to the more negative side of the circuit. This is because an NPN transistor has N-type collector doping, so its collector must be connected to the more positive side. The reverse is true of the PNP transistor.

Valve symbols are not marked for polarity, as there is little chance of connecting them the wrong way around. The cathode, which is easy to recognise, is always connected to the more negative side of the circuit and the plate to the more positive side.

It is, however, sometimes difficult to determine which base pins are attached to the cathode, plate and other internal elements of a valve. Many schematics list the base pin numbers adjacent to the valve symbol, but there are also many which do not.

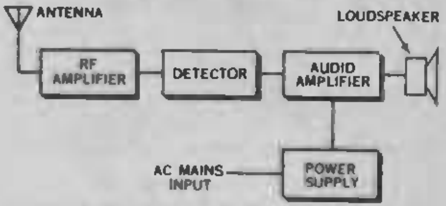


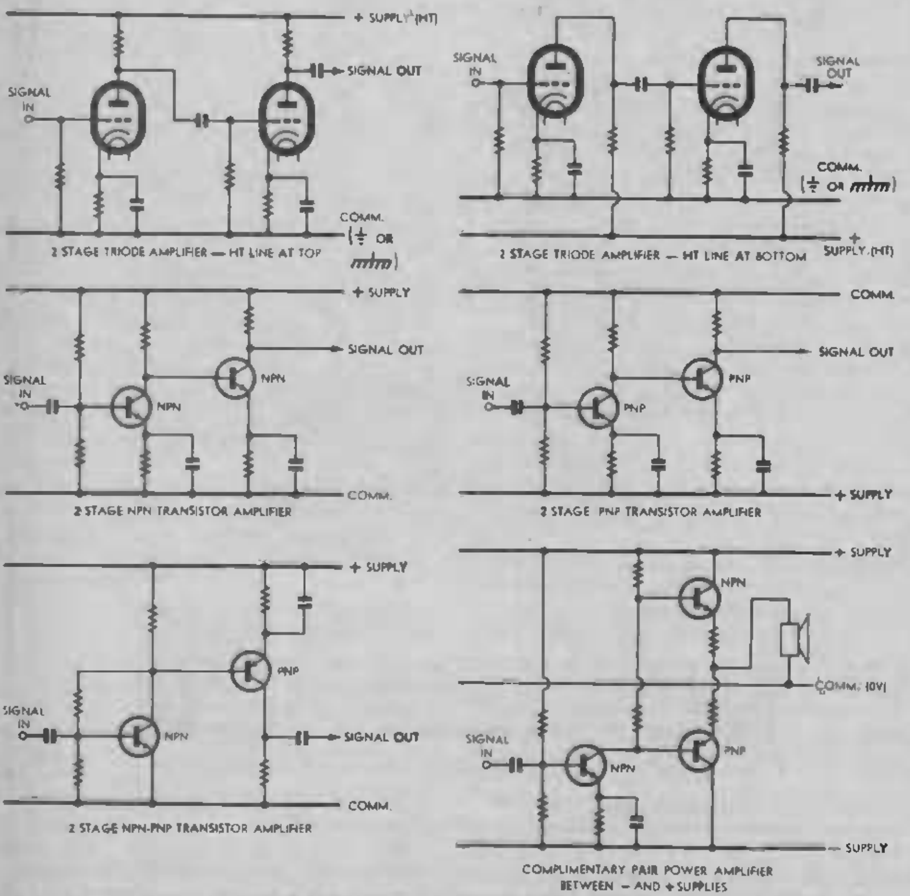
Fig 4: A block schematic diagram, as above, shows only the sequence of stages. It does not show components or detail of interconnections.

If base pin numbers are not supplied, they can be found in manufacturers' data sheets or valve data handbooks published by many valve manufacturers. If you do not have a valve data handbook you can probably find one in a library, but if you plan to make electronics your hobby or business, you would be wise to obtain at least one handbook of your own. Many are published in paper back editions at reasonable prices. Another useful paperback book is a valve substitution guide, which is especially valuable for finding modern valves to use in place of obsolete ones in older equipment.

Transistor base connections are easier to identify, especially with the newer types, but a transistor data handbook would also be useful. Substitution guides for transistors are helpful too, as some circuits you will see use types not available in Australia.

This may be a good time to bring up some differences between traditional ways of drawing valve circuits as compared to transistor circuits.

You will find that with both types it is common practice to lay out a circuit so that the signal which flows through it enters at the left and exits on the right. In a receiver for example, the RF (radio frequency) stage of the circuit will be at the left, followed by the stages which modify and amplify the



2 STAGE TRIODE AMPLIFIER — HT LINE AT TOP

2 STAGE TRIODE AMPLIFIER — HT LINE AT BOTTOM

2 STAGE NPN TRANSISTOR AMPLIFIER

2 STAGE PNP TRANSISTOR AMPLIFIER

2 STAGE NPN-PNP TRANSISTOR AMPLIFIER

COMPLIMENTARY PAIR POWER AMPLIFIER BETWEEN — AND + SUPPLIES

Fig 3: Typical ways in which valves and transistors are connected to the DC power supply lines in electronic circuits.

signal, and the output to the speaker or headphones will be at the right. Power supplies are usually drawn below and somewhat separated from the signal handling circuits.

There is a certain consistency in circuits using valves, in that valves are always drawn with their plates at the top. The HT line to which the valve plates are connected will usually either run across the top of the schematic or parallel to the negative line beneath the valve symbols. Since valves are most sensitive to voltage change, the stages are usually connected to one another by a high impedance coupling circuit based on either a capacitor or a transformer.

Some transistor circuits follow the same pattern, if, for example, NPN transistors are used, and they are drawn with their collectors at the top, they may resemble a series of triodes. But since NPN and PNP types can be intermixed in the same circuit, sometimes in close interaction with one another, you will find a much greater variation in drafting practice with solid-state circuits.

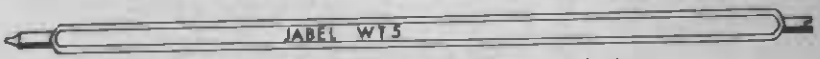
You may find the negative supply line at the top of the circuit and the positive line at the bottom. Sometimes there will be a line at zero potential between two supply lines with transistors across both pairs, that is, with some transistors using zero potential as a positive supply and others using it as negative supply. These variations will undoubtedly be confusing at first, but with experience you will learn to sort them out.

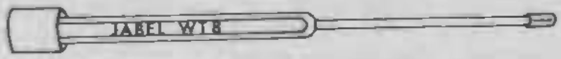
By the way, don't feel discouraged if you are not able to interpret a schematic diagram you are unfamiliar with. Many diagrams are very difficult, even for experienced people, unless they are accompanied by a written or verbal explanation.

The first step in interpreting a schematic diagram is to sort out the DC supply lines and identify the various valves, transistors and other components in relation to the DC supply in their "steady state" condition, that is, without a signal applied. Valves and transistors must, of course, be connected across a potential and they must have various load and biasing resistors in the circuit simply to make them conduct within the proper range of their transconductance curves.

Once you have identified the components concerned with simply "energising" the circuit, you can mentally eliminate them while you attack the main problem, which is, how do the various stages affect the signal as it travels, stage by stage, through the circuit. At first it is important to practice on circuit diagrams for which you have a good written explanation available, such as the project circuits in "Electronics Australia".

With enough experience, you will be able to simply look at a stage within a larger circuit and say "that's an oscillator because it has positive feedback" or "that transistor stage is biased as an amplifier and not a switching circuit". While you learn, take every opportunity to follow basic explanations of how different circuits operate and remember in a general way how they look. After you have built up a small memory bank of individual circuit ideas, you will find that schematic diagrams are no longer as mysterious as you thought at first.





CHAPTER 10

Radio Transmission

Electromagnetic radiation — radio waves — the escape of energy from charging and discharging reactance — the transmitting aerial — the use of a high-frequency carrier — frequency and wavelength — generating the RF carrier — RF oscillators — RF amplifiers — keyed CW transmission — amplitude modulation and AM transmitters.

So far in this course, we have been introduced to the basic components or "building blocks" of electronics — resistors, capacitors, inductors, valves, transistors, and so on. We have also seen something of the various ways in which these components may be connected together to make elementary circuits.

It is important that one has a firm grasp of these matters before one attempts to delve into the more practical aspects of electronics. However, there are no doubt many readers who have so far been wondering how all these components and circuits are used to send messages and music — or even pictures — to a distant place, without any physical connection.

With this chapter we begin to answer such questions, for we are now in a position to start examining how electronic components may be put together to transmit intelligence (whether it be messages, music, or pictures) from one point to another — without wires. In other words, we are now going to look at the basic principles of radio transmission and reception.

The whole of radio depends upon the fact that a certain form of energy, called "electromagnetic radiation," can travel from one place to another practically instantaneously, and even through a vacuum or the near-vacuum of outer space.

You are already familiar with at least two types of electromagnetic radiation — light and heat. You are also aware that these two forms of energy radiation can travel through the near-vacuum of space. Step out into the sunlight, and you have proof that energy is traversing the 93 million or so miles between the sun and you. It warms your body, it can stimulate the retina of your eye, it can be used to evaporate water, and so on.

Radio waves are simply another sort of electromagnetic radiation, along with light and heat. We shall see in a moment how these three forms differ from one another. But first we must learn just what electromagnetic radiation really is. To explain this fully we would need to delve into lots of mathematics, but we're deliberately going to simplify the story so that you will be able to form a mental picture of just what is going on.

In an earlier chapter, we saw that the application of a voltage or EMF to a capacitor caused the capacitor to "charge up." We saw that this was a process whereby the space between the two capacitor plates became "strained" or in a state of tension. We called this state of tension an electric field, and we said that it was stored energy which could be returned to the circuit when the capacitor was discharged.

In another chapter, we saw that passing a current through an inductor sets up a magnetic field around the inductor. The magnetic field, like an electric field, is a state of tension in space, but it is a different type of tension. It represents another sort of stored energy, which we said could be returned to the circuit when the field was allowed to collapse.

Now, in implying that all the energy stored in electric and magnetic fields could be returned to the circuit, we were simplifying the situation slightly. We did it to emphasise the difference between the basic energy storage behaviour of reactance (capacitance and inductance), as opposed to the energy dissipation (conversion to heat) behaviour of resistance.

In actual fact, however, not quite all the energy stored in an electric or magnetic field is returned to the circuit. Some is lost — it escapes, and flows or radiates away from the capacitor or inductor like the ripples from the surface of a pond disturbed by a stone. It can be picked up at a distant spot, by a suitable detecting device.

It happens that the form in which it radiates is the same in both cases — it doesn't escape from the capacitor as electric field alone, nor from the inductor as magnetic field alone. In both cases, the energy is radiated as combined electric and magnetic fields — hence the name, "electromagnetic" radiation.

The reason why the energy radiated is in the form of a combined field is that a changing field of either type is always accompanied by the other type. One cannot have a changing electric field without a magnetic field along with it, nor can there be a changing magnetic field without an

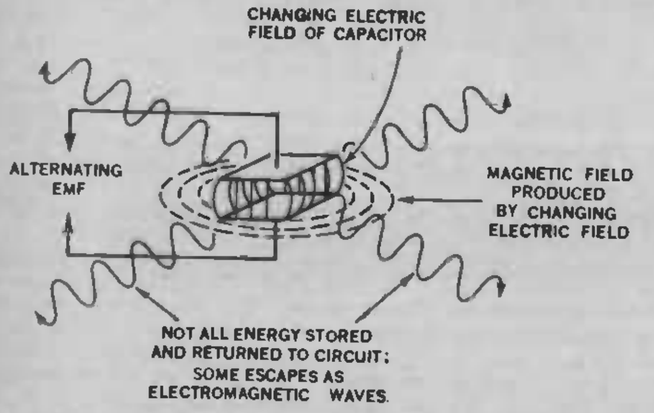


Fig 1 (left): The alternating electric field produced when a capacitor is connected to an alternating EMF is accompanied by a similarly alternating magnetic field.

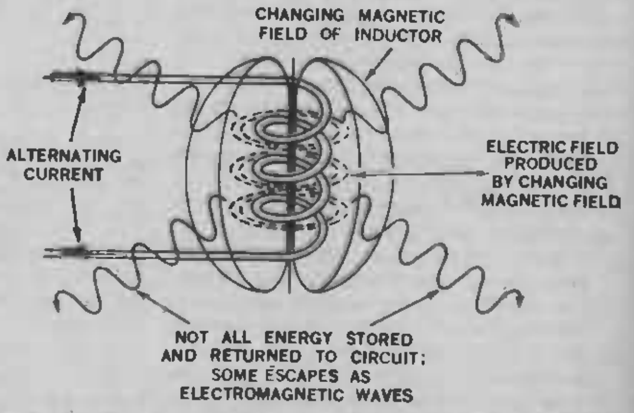


Fig 2 (right): With an inductor the converse occurs. In both cases some energy escapes as radiation, and is capable of being detected at some distance from its origin.

electric field. This is just a fact of life; no one knows why, nor do we know just why the energy "radiates."

Let's just summarise these ideas about electromagnetic radiation before we go any further: the total energy "stored" in a capacitor or in an inductor, in their respective types of field, can never be fully returned to the circuit, because some of it escapes or "radiates" away. It escapes as a combined electromagnetic field, which is produced during the charging and discharging processes when the fields are changing (building up or collapsing) because whenever one type of field changes it produces the other type.

When the field of a capacitor or of an inductor is built up or allowed to collapse, then a small "wave" of electromagnetic energy radiates away in all directions. Now if we apply an alternating EMF to the capacitor or inductor, a continuous series of electromagnetic waves will be produced.

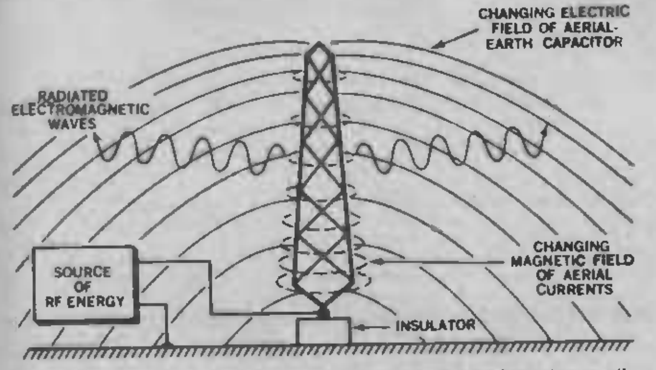


Fig 3: A transmitting aerial is a device specially designed to radiate energy. It combines the features of a capacitor and an inductor.

The continuous build-up-decay-reversal nature of the field in the capacitor or inductor will produce electromagnetic energy which will radiate away in waves, the waves having the same frequency as that of the alternating EMF.

What we know as radio waves are electromagnetic waves produced by currents so that they have a frequency of from about 10,000 Hertz to about 100,000,000,000 Hertz.

Incidentally, when discussing the frequency of radio waves or of the currents which produce them, the simple unit of frequency, the Hertz (Hz), often becomes unwieldy. Things are simplified by using the Kilohertz (kHz) which is equivalent to 1,000 Hz, the Megahertz (MHz), which is equivalent to 1,000,000Hz, and the Gigahertz (GHz), which is equivalent to 1,000 Mhz.

The radio frequency spectrum thus extends from about 10 kilohertz (10kHz) to about 100 Gigahertz (100GHz). The alternative descriptions in terms of wavelength (long-, short-, medium-, microwaves, etc.) describe the length of one cycle of the electromagnetic waves concerned.

The lengths of electromagnetic waves are inversely proportional to frequency, which, means that the higher the frequency, the shorter the wavelength, and vice-versa. Wavelength is measured in metres, and the length of a wave in metres is given by its frequency in Megahertz (MHz) divided into 300. Waves of a frequency of 100 MHz thus have a length of 3 metres, and so on.

We said before that light and heat were electromagnetic radiation, but that they differed from radio waves in some way. In fact, they differ in terms of frequency. Heat radiation is in effect super-high-frequency radio waves or "Extra-short" microwaves, while light radiation is a higher frequency again. Light waves are so short that their wavelength is measured in Angstrom units (an Angstrom is a ten-thousand-millionth of a metre).

But let us return to radio waves and their generation. When people first observed that energy was radiated from a changing electric or magnetic field, and saw that it could be picked up at a distant spot, they started thinking.

Surely, they reasoned, this effect could be used to transmit intelligence from one point to another. Thus was born the idea of using radio waves as a means of communication.

Experiments showed that radio waves could be radiated in more efficient ways than from a simple capacitor or inductor. As a result various types of special radiating devices have been developed which you will probably be familiar with as aerials. A properly designed aerial stores very little energy fed into it — it lets most of it escape, as radiation.

At this stage, it might be thought that to transmit messages by radio, one need simply speak into a microphone, amplify the resulting voice-frequency voltages with a valve or transistor amplifier, and feed the amplifier to an aerial. Then, it might be reasoned, one would only need another aerial and an earphone to receive the message at a distant spot.

Now while messages can be and have been sent in this way, it actually proves quite difficult to satisfactorily transmit electromagnetic waves with frequencies as low as those we can hear (between about 30Hz and 16kHz). Transmitting aerials miles in length are needed if practical amounts of energy are to be radiated. There are also other difficulties associated with the transmission of such low frequencies, but these need not concern us here.

It so happens that higher frequency waves are easier to radiate. Efficient aerials may be made in convenient sizes, which will radiate suitable amounts of energy if high frequencies are used.

In practice, then, we do not radiate voice-frequency radio waves. We radiate at considerably higher frequencies, called radio frequencies (RF), by supplying the aerial with alternating current generated by an RF oscillator and amplified by an RF amplifier. These may use valves or transistors, as we shall see a little later on.

The broadcast radio stations radiate waves with frequencies in the range 550 kHz-1500kHz. Long distance communication stations operate from about 2 MHz to 30MHz, in what is called the "short wave" or high-frequency (HF) band. Television stations transmit waves at frequencies between about 50MHz and 250 MHz (the VHF band), and so on.

Radiating at radio frequencies is desirable from the ease-of-transmission-of-energy point of view, then, but it complicates the procedure of sending messages. Unfortunately, human beings can neither talk nor hear at radio frequencies!

Means must, therefore, be used whereby our RF waves can be used as a vehicle or "carrier" for the information to be transmitted. This is called "modulating" the RF carrier.

The simplest way of doing this, and the way that was first used, is to arrange that the alternating RF currents fed to the transmitting aerial are turned on in bursts or pulses. The pulses are arranged to be either long or short in duration, and various combinations of long and short bursts made to correspond to letters of the alphabet and numerals.

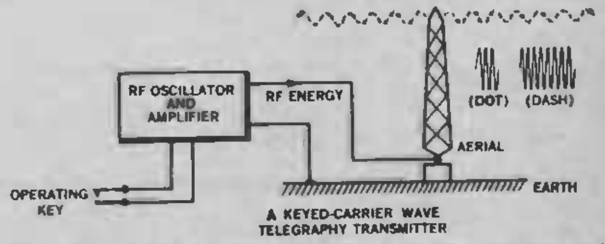


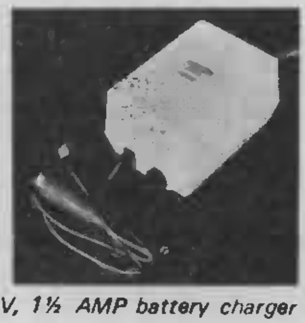
Fig 4: A keyed-carrier or CW transmitter is a system for radiating messages as long or short bursts of RF waves usually in the form of Morse Code.

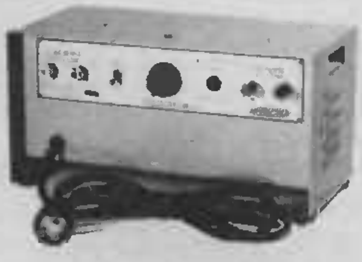
This type of transmission is known as keyed carrier wave transmission, or just "carrier wave" (CW) transmission. And the code used to pulse the carrier wave in short ("dots") and long ("dashes") bursts is, of course, the familiar "Morse" code.

With CW transmission the operator is provided with a "key," which is a switch connected to the transmitter. The key is arranged so that in its rest position the RF oscillator and RF amplifier send no energy to the aerial. When the key is pressed down, however, the oscillator and amplifier are turned "on", and by pressing the key briefly or for slightly longer the operator can send short or long bursts of RF energy to the aerial — to be radiated as dots and dashes.

A long burst followed by three short bursts means "B", for instance, while a short burst followed by a long means "A." Each letter of the alphabet and numeral is represented by a particular combination of short and long bursts.

Keyed carrier wave transmission is quite satisfactory as a means of transmitting simple messages, but it obviously lacks






Tape Questions—
Tape Answers
compiled and answered by
Heinz Ritter
Tape Questions — Tape Answers

something where speech, music or pictures are concerned. Who would be able to recognise their favourite piece of music translated into dots and dashes?

Fortunately, there are other ways of modulating the RF carrier in order to send information, besides the simple on-off modulation of keyed CW transmission. Although there are quite a large number of alternative modulation systems, we will confine ourselves here to the discussion of only one — that used by all medium-wave "radio" broadcasting stations.

The broadcasting stations amplitude modulate (AM) the RF carrier. Rather than switch the carrier through only two steps of amplitude (off and on), they vary its amplitude continuously. In this way, the continuous variations of the human voice or music can be transmitted faithfully, as similar variations in the strength of the radiated waves.

In the remainder of this chapter, we will see how this is done. Following chapters will be devoted to the operation of the receiving end of the system, to show how the receiver is able to recover the original transmitted voice or music from the amplitude modulated waves.

Before we examine how the signals to be transmitted are made to amplitude modulate the RF carrier, we should have a look at the way in which the RF carrier is generated in the first place. In other words, we should look at the RF oscillator, which perhaps can be regarded as the "heart" of any radio transmitter.

You may remember that in an earlier chapter in this course, we saw that a capacitor and an inductor may be connected in parallel to form a parallel tuned circuit. We saw that when such a tuned circuit is fed with a short burst of energy, it tends to oscillate, producing a decaying or "damped" alternating voltage.

The frequency at which the circuit oscillates, which is the frequency of the alternating voltage, is determined by the resonant frequency of the tuned circuit. This, in turn, depends upon the values of the capacitor and the inductor, as one might expect.

In fact, the frequency of the voltage produced is given by

$$F = \frac{1}{2\pi\sqrt{LC}}$$

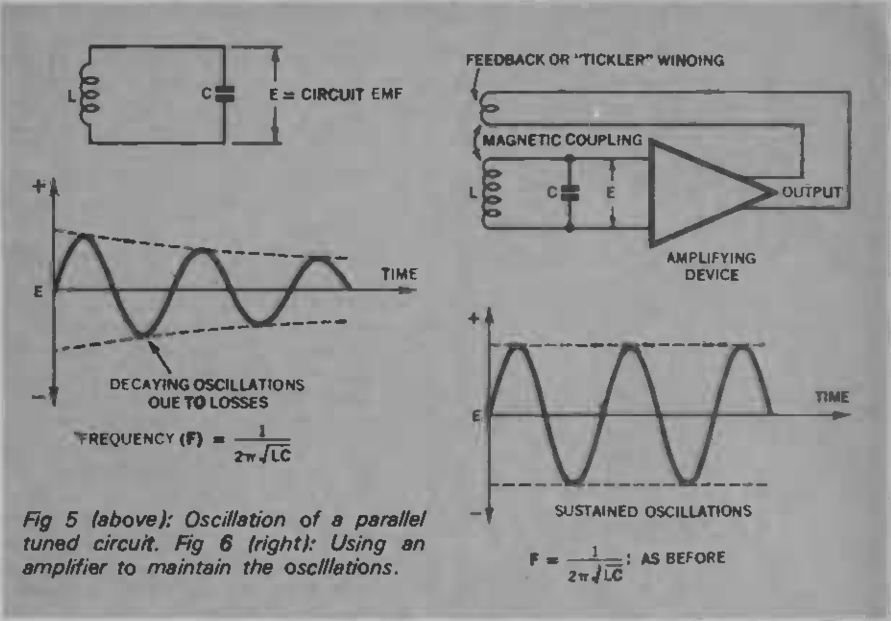


Fig 5 (above): Oscillation of a parallel tuned circuit. Fig 6 (right): Using an amplifier to maintain the oscillations.

where F is the frequency in Hertz, pi is 3.1416, L is the coil inductance in Henries, and C is the capacitor value in Farads.

A tuned circuit can thus be used to generate an alternating EMF of any desired frequency, by suitable choice of the inductor (coil) and capacitor. So if we want to generate an RF carrier of a certain frequency, we can select a capacitor and coil to resonate at this frequency.

But a tuned circuit alone is not sufficient, for it has coil resistance and other losses which make the alternating voltage decay and die away. To produce a continuous, steady supply of alternating EMF at our carrier frequency, we must arrange for the tuned circuit to be continually fed with energy, to overcome its losses and keep it oscillating.

Here is where valves or transistors or other amplifying devices enter the picture, for by means of a valve or transistor we can keep the tuned circuit oscillating steadily.

By the way, note that wording carefully. It is always the tuned circuit which oscillates, not the valve or transistor. The amplifying device simply keeps the tuned circuit going.

Fig. 6 shows the basic operation of a simple tuned circuit oscillator. The amplifying device is connected so that it picks up the oscillatory voltage E appearing across the tuned circuit. The output of the amplifying device is then connected to a feedback or "tickler" winding which is placed close to the inductor of the tuned circuit.

The feedback winding is arranged so that it can magnetically induce voltages into L which re-enforce the voltage E, when fed with an amplified version of E by the amplifying device. In this way, the tuned circuit is fed with energy which keeps it oscillating steadily.

The amplifying device may be a valve, a transistor, or anything else capable of doing the same job. Fig. 7 shows simplified circuits for tuned oscillators using a valve and a transistor.

In the valve circuit, the tuned circuit voltage is fed to the input of the valve, which passes the corresponding amplified plate current oscillations through the feedback winding to supply energy back to the tuned winding.

The transistor circuit does the same thing in a different way. It connects the tuned circuit in the collector (output) circuit of the transistor, and uses a small feedback winding to supply the input circuit of the transistor. Thus small oscillatory voltages induced in the feedback winding are amplified by the transistor and fed directly to the tuned circuit.

In all oscillator circuits of this type, the amplifying device not only supplies the tuned circuit with enough energy to overcome losses and keep it oscillating; it supplies more than enough, so that a small amount of the oscillatory energy of the tuned circuit can be picked off for external purposes — in our case, for amplification and supply to the transmitting aerial.

This "output" of the oscillator can be obtained in a number of different ways. A third winding may be used, magnetically coupled to the tuned and feedback windings, to produce an induced EMF, as shown in the valve circuit. Or a connection may be made directly across the tuned circuit, as shown in the transistor circuit. Or various other methods may be used, depending on the sort of oscillator actually used and the amplifier circuit which is to be connected to the output.

In our discussion of oscillators so far, we have been talking in terms of L-C parallel

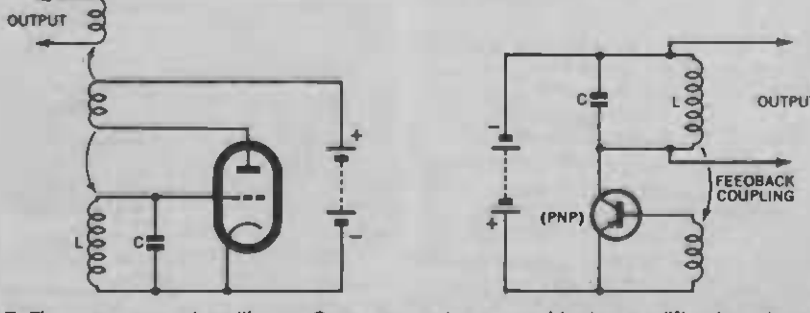


Fig 7: Elementary tuned oscillators. One uses a valve to provide the amplification, the other a transistor. The circuits illustrate different ways of providing feedback and output coupling.

tuned circuits. However, oscillators using such tuned circuits tend to waver or "drift" in frequency: only to a small degree, if the circuit is well designed, but generally enough to make them unsuitable as a source of RF carrier energy in a broadcast transmitter — for transmitters must radiate on a fixed frequency, or one would never quite know where to find them on the receiver dial!

Actual radio transmitters do not use L-C tuned circuits in the RF oscillator, for this reason. They use instead a carefully prepared wafer of quartz crystal, which has the property of resonating mechanically when an EMF is applied to opposite sides of the wafer. When it is made to oscillate, it does so with very much less frequency drift than a normal tuned circuit, particularly if it is kept at a constant temperature in a thermostatically controlled oven.

The frequency of such crystal-controlled RF oscillators is set by the dimensions and preparation of the quartz crystal. To change the frequency, the crystal must either be replaced by another, or taken out and altered in size.

So much, then, for the source of the RF carrier energy in our transmitter. But the output of the oscillator is seldom strong enough to be fed direct to the transmitting aerial. Usually, it must first be amplified by one or more valve or transistor stages in the RF amplifier, as we mentioned before.

An RF amplifier using a pentode valve is shown in Fig. 8. It has a tuned circuit in both the grid and plate circuits, and both tuned circuits are made to resonate at the oscillator frequency. Other types of RF amplifier stage called "multipliers" have the plate circuit tuned to a multiple of the oscillator frequency, and the stage is arranged to multiply the frequency. This is used where the required carrier frequency cannot conveniently be generated directly by the oscillator.

For instance, multiplier-type RF amplifiers must be used with crystal-type RF oscillators if very high carrier frequencies are required, as it is impractical to make quartz crystals to oscillate at very high frequencies.

Link windings couple the tuned circuit of the RF oscillator to the input of the amplifier, in this case. If the RF oscillator used a crystal rather than an L-C tuned circuit, one of the other types of coupling would generally be used.

Negative bias is applied to the grid of the valve to ensure that it operates at a convenient point and amplifies efficiently. The amplified RF carrier which appears in the plate tuned circuit (the so-called "tank"

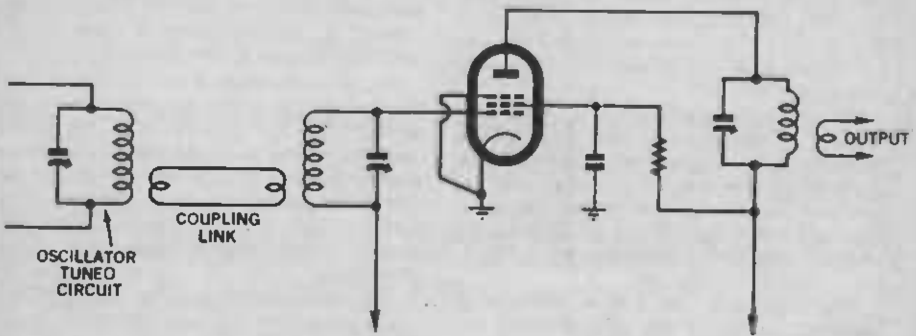


Fig 8: The basic form of an RF amplifier stage, using in this case a pentode valve. Transistors are also used for this purpose.

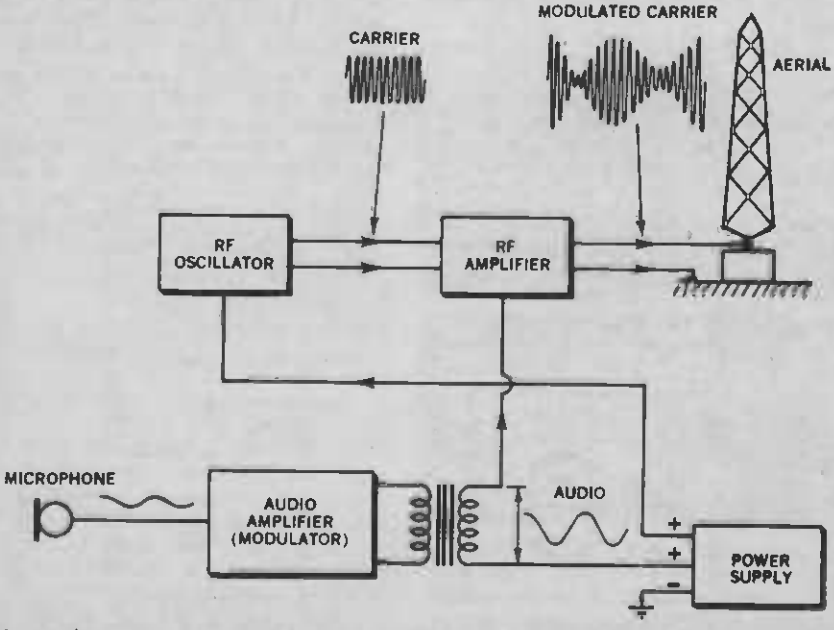
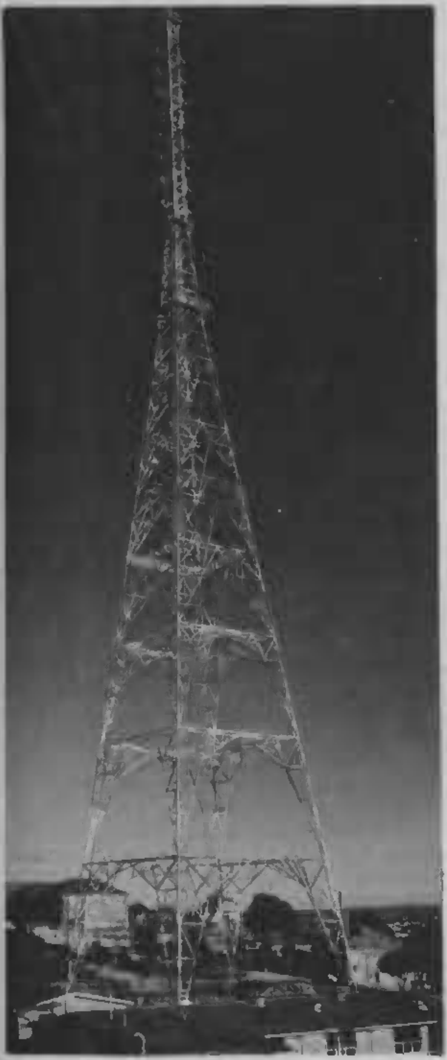
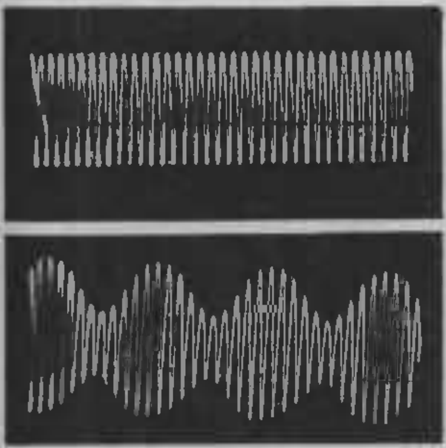


Fig 9: An elementary AM transmitter, showing how the RF signal is made to function as a "carrier" of the audio information.



A typical television transmitting tower. The actual aerial is the parallel-sided assembly on top of the tapered supporting tower. With medium-wave broadcast stations, operating on a much lower frequency, it is more usual to employ a slim guyed tower, resting on a huge insulator. The entire tower acts as the aerial (or antenna).



Two photographs taken from the screen of an oscilloscope, an instrument which allows us to "look at" electrical EMFs and currents. The upper pattern shows an alternating RF "carrier" signal, and the lower pattern the effect of modulation.

circuit) is coupled to the next stage — or to the aerial if this is the last stage — via another coupling loop.

We have now seen something of those parts of a radio transmitter responsible for the generation of the RF carrier energy. By adding a Morse key to this, we would have a keyed-carrier or CW transmitter, but let us progress a little further and see how the carrier may be varied in strength so that it can be used to transmit voices, music, or even pictures. In other words, let us see how continuous amplitude modulation is performed.

The strength of the RF carrier fed to the aerial depends on a number of things, but one of these is the supply voltage of the final RF amplifier stage. The output is, in fact, proportional to the plate voltage, with a circuit like that of Fig. 8.

Because of this, to vary the strength of the RF carrier — to amplitude modulate it — all that need be done is superimpose the audio (sound) frequency signals on the plate supply voltage. In this way the audio signals add to and subtract from the plate voltage, and vary the strength of the RF carrier in sympathy with the sound waves reaching the microphone.

There are other ways of amplitude modulating the carrier, but they all produce much the same effect and need not concern us here. The basic idea of a plate-modulated AM transmitter is shown in Fig. 9.

There is an RF oscillator and an RF amplifier, as with the CW transmitter, in order to generate the RF radiation energy. However, added to this section is the microphone, and audio amplifier (the "modulator") and a transformer used to superimpose the amplified audio frequency signals on to the plate voltage of the RF amplifier.

The audio amplifier may use either valves or transistors, and builds up the strength of the tiny voice-frequency voltages generated by the microphone. The output of the audio amplifier is fed to the winding of the modulation transformer.

The other winding of the transformer is connected in series with the plate circuit of the final RF amplifier, so that the amplifier receives its plate current through the transformer winding. In this way, the amplified alternating audio voltages induced in this winding of the transformer add to or subtract from the supply voltage, and can vary the strength of the carrier fed to the aerial.

The small waveform sketch shows what the modulated carrier would look like if we could see it. In fact, we can see it if we use an instrument called an oscilloscope, as the two photographs show.

Instead of the microphone, we can use a gramophone pickup, a tape recorder, and so on. In television transmission, we would use cameras, film scanning machines and video tape recorders instead.

And with the description of a basic AM transmitter, we must end this chapter. Now that we are reasonably familiar with the nature of radio waves and at least two of the ways in which information can be transmitted, we are ready to look at the way in which the radiated radio waves are used. In the next chapter, then, we will start at the "other end" of the radio system — the receiver.

CHAPTER 11

Radio Reception

Revision of basic principles of radio transmission — the use of capacitors and inductors in various forms as receiving aerials — tuning the desired signal by means of a parallel tuned circuit — demodulation or detection — the diode detector — the simplest radio receiver — earphones and their operation — practical construction.

In the previous chapter, we described the process of radio transmission; how telegraph code, speech and music can be carried by radio waves over great distances. In this chapter, we see what happens at the "other end". We find out how the radio waves are intercepted and the required signal selected and changed back into the information originally transmitted.

When we talked of radio waves in the last chapter, we said that they were waves of energy in the form of an oscillating electromagnetic field. Energy in this form radiates from any oscillating electric or magnetic field.

It therefore radiates to a certain extent from any capacitor or inductor in a circuit carrying alternating currents and EMFs, but it radiates most profusely from devices which are specially designed for the purpose. These radiators or aerials combine in varying proportions the characteristics of a capacitor and an inductor, and when properly designed will radiate most of the electrical energy fed to them.

Now it so happens that the process is reversible. If capacitors, inductors or aerial systems combining capacitive and inductive elements are placed in the path of radio waves, energy will be extracted from the electromagnetic field to appear as induced alternating EMFs or currents in the circuit. And, if the RF carrier voltage fed to the original transmitting aerial system was amplitude modulated, the voltages or currents induced at the receiving end will similarly be varying in amplitude.

Fig. 1 illustrates some of the ways in which energy may be extracted from the transmitted waves to obtain induced EMFs.

A capacitor C (right) placed in the path of the waves interacts with the "oscillating electric field" aspect of the electromagnetic waves, and generates an EMF, as shown. In a similar fashion, an EMF will be generated by the "capacitor" formed by a suspended aerial wire A and the earth beneath it.

The latter example is the familiar "inverted L" aerial, which often becomes a simple "length of wire strung around the picture rail" when the transmitter is not too far away. The connection to the earth side of the "capacitor" may not be an actual physical one, but may be made indirectly via inter-winding capacitance in the mains transformer and thence to earth via the power mains system.

Another version (not shown) of the capacitor aerial is seen on cars, where the insulated whip forms the "suspended wire" side of the capacitor and the car body itself forms the other side.

If we place an inductor L in the path of the radio waves, it will interact with the "oscillating magnetic field" aspect of the electromagnetic waves, and will also produce an EMF. Examples of this type of receiving aerial are the "loop" coils used in older-style portable sets and the ferrite-rod aerials used in modern transistor portables.

There are many more types of receiving aerials, such as the folded dipole shown in Fig. 1 as D. This is used in many aerials of the higher frequency type, and will be familiar to many as part of a TV antenna. The dipole combines capacitor and inductor action, in such a way that it responds mainly to waves of only one frequency. Aerials of this type are said to be resonant.

In general, we can say that any capacitor or inductor or device which combines the two types of reactances will produce an EMF when placed in the path of electromagnetic waves; the more the capacitor or inductor is "pulled apart" to embrace more of the waves, or the higher and longer the wire in an "inverted L" aerial, the better, while the greater the area of a loop aerial coil the better, and so on . . .

So far, we have "intercepted" the radio waves emitted from the transmitter and used an aerial to extract a small amount of energy from them and produce a corresponding EMF.

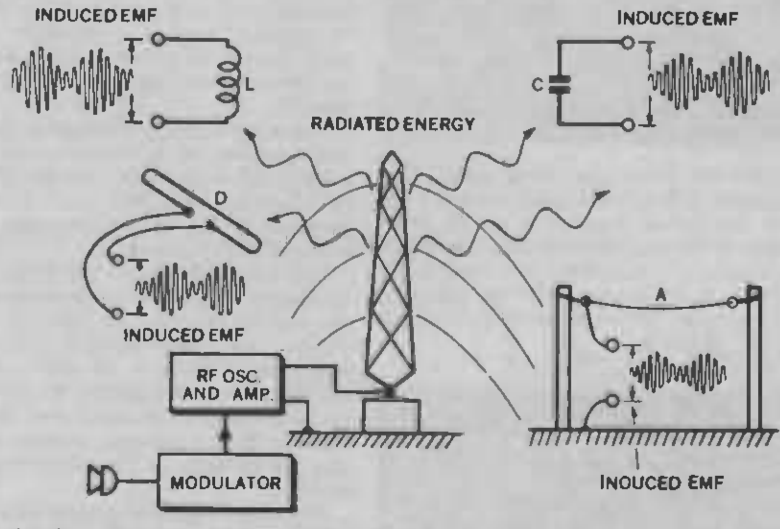


Fig 1: An illustration of some of the ways in which a small amount of the energy in radio waves may be extracted to produce an EMF for a receiver.

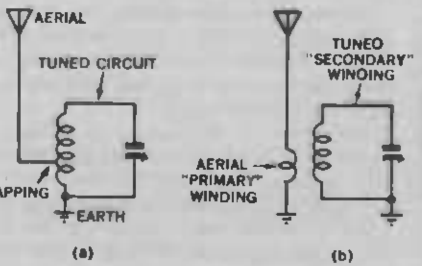


Fig 2: Two ways of feeding the induced aerial-earth EMFs to the resonant circuit used to reject unwanted signals.

The next step is fairly obvious when we consider that most aerials will produce EMFs corresponding to all the waves which stream past them. Before much more can be done, therefore, it is necessary to select the desired signal from the jumble. This is usually done by means of a parallel tuned circuit.

As we saw in the last chapter, a parallel tuned circuit will oscillate when it is excited into doing so by energy arriving as an alternating EMF corresponding to its resonant frequency. The resonant frequency is governed, you may remember, by the size of the capacitor and of the inductor.

It is most important that the rate of

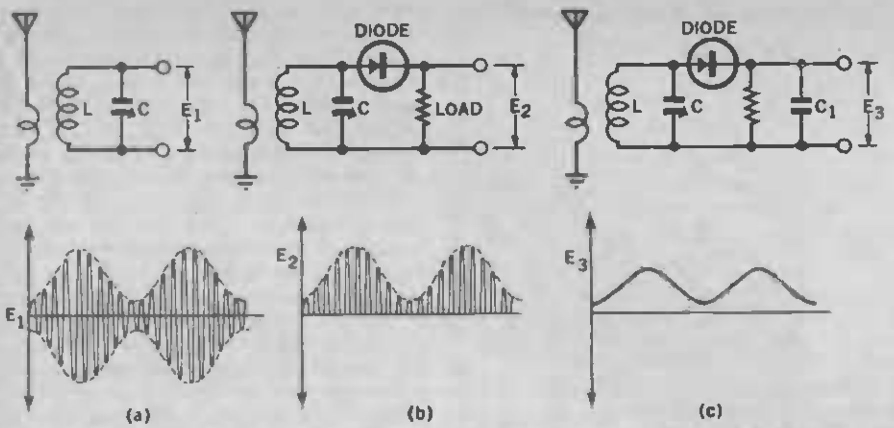


Fig 3: Steps in the demodulation or detection of the modulated RF carrier EMFs present across the tuned circuit. A diode and additional capacitor are used to separate out the modulation information as indicated in (b) and (c).

alternation of the incoming energy corresponds to the resonant frequency of the tuned circuit, to keep it oscillating. If it is not of the same frequency, it will not have a sustaining effect and the tuned circuit will not keep oscillating. It will be like trying to make a child's playground swing oscillate by giving it pushes at a rate faster or slower than that dictated by its size and weight; the energy will be almost completely wasted for a swing will only oscillate at its natural frequency.

To select a particular radio signal from the jumble of EMFs supplied by the aerial system, then, the aerial signals are fed to a parallel tuned circuit. The capacitor and inductor of the tuned circuit are arranged to be resonant at the desired frequency, so that the only aerial EMFs which can cause it to oscillate are those of the signal concerned. (It is very unlikely that the aerial will receive signals from two transmitters having the same carrier frequency, because regulations ensure that such stations are long distances from one another.)

If it is desired to receive signals from a number of transmitting stations, the tuned circuit is arranged to have its resonant frequency adjustable over the band of frequencies concerned. This can be done by making either the capacitor or the inductor a variable unit.

The control knob fitted to the variable component in the tuned circuit is the "tuning" knob familiar to all who have operated a radio receiver.

There are a number of ways in which the aerial EMFs may be fed to the tuned circuit. Two such methods are shown in Fig. 2. In one case (a) the EMFs appearing between the aerial and earth are directly connected to part of the inductor; in the other case (b) they are connected to the small primary winding which is magnetically coupled to the inductor.

In neither case is the aerial-earth circuit connected directly across the tuned circuit. Among other reasons, this is because the aerial-earth capacitance of the aerial would tend to upset the tuned circuit if it were connected directly across it. The same aerial "capacitor" would also tend to re-radiate some of the oscillatory energy as soon as the tuned circuit started to oscillate – remember, a capacitor connected across an alternating EMF will radiate!

We have now obtained across the tuned circuit of our receiver a tiny replica of the modulated RF carrier voltage which was fed to the transmitting aerial system of the station in which we are interested. The selective action of the tuned circuit has effectively discarded all the unwanted EMFs fed to it from the aerial circuit, to leave only the desired signal. The task is now to turn this tiny modulated RF carrier voltage back into the original speech and music.

There would be no point in feeding the signal from the tuned circuit directly to an earphone or to a loudspeaker, because these devices would not be able to vibrate mechanically at radio frequencies. And even if they could do so, the air would not transmit the vibrations more than a tiny fraction of an inch. Furthermore, we wouldn't be able to hear the "sound" anyway, as our ears do not respond to such high frequencies.

Before we can feed our signals to an earphone or loudspeaker for conversion into sound waves, we must separate the modulating information (the speech and music) from its "transport" – the RF carrier. This process is known as demodulation or detection.

The first of these terms is perhaps to be preferred, because it indicates that the process is the opposite of the modulation which occurs at the transmitter. The second term is quite commonly used, however, being a heritage from the pioneering days of radio.

There are a number of ways in which an amplitude modulated RF carrier may be demodulated, but the simplest is the diode detector, which takes advantage of the "one-way" conduction behaviour of any one of the various types of diode. Thermionic diode valves, germanium or silicon semiconductor diodes, or even a "cats-whisker" wire touching a suitable spot on a crystal of galena – all will serve as diode detectors.

Detector action is quite complex, unfortunately, even in the simple diode detector. However, we can form a simplified understanding of the way in which they work by reference to the waveforms of Fig. 3. above.

The waveform in (a) should be familiar, as it is a replica of the modulated RF carrier voltage fed to the transmitting aerial system; it is the signal developed across our receiver tuned circuit.

What the detector does is to let only part of this signal through, as shown in (b). The "one-way" action of the diode is used to allow only half of the RF oscillations to pass through – only the upper half-cycles of carrier, in the waveform shown.

The resistor marked "load" represents in simplified form the effect of an earphone or loudspeaker were it connected across the output of the detector.

As may be seen, the output of the diode detector consists of half-cycle RF pulses which are all of the same polarity and which vary in amplitude with the modulating information.

Were we to feed these pulses directly to an earphone or loudspeaker – and assuming them to be strong enough – we would hear faintly the original sound. We would hear the sound, not because the earphone or loudspeaker would be vibrating at the high frequency of the pulses, but because it would tend to regard them as "instalments" or a succession of "little pushes," which tend to displace the diaphragm.

In other words, it would be responding to the average value of the pulses. As the pulses are modulated, their average effect on the diaphragm will be similarly modulated and thus the earphone or speaker would produce faint sounds.

Before we go any further, make sure that you can see why an earphone or speaker would respond – even if faintly – to the waveform of Fig. 3(b), when it would not do so to the "whole" tuned circuit output shown in (a). The reason is simply because the complete waveform has an average effect of zero; there are as many negative pulses as there are positive.

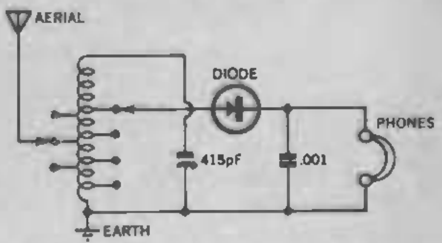
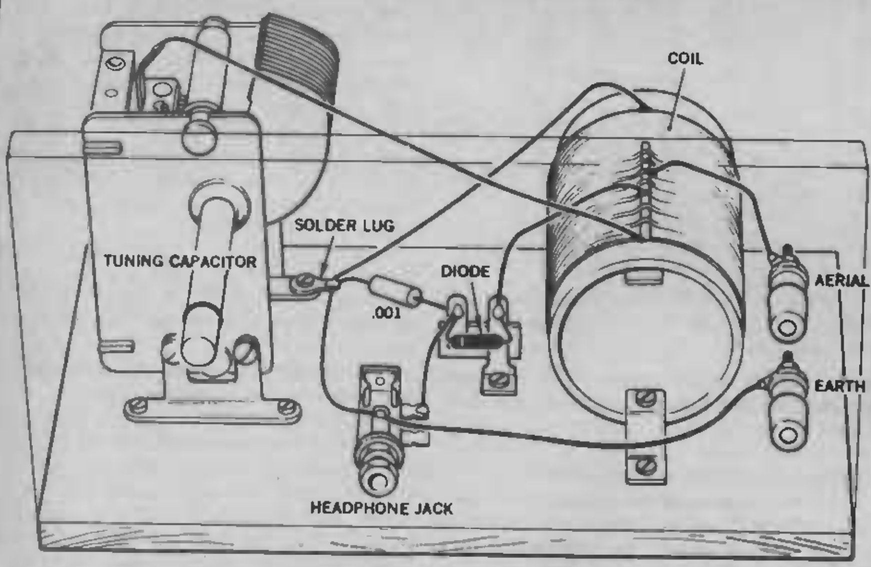


Fig 4: The circuit for a practical "crystal" receiver, using a germanium diode as the detector (see overleaf).

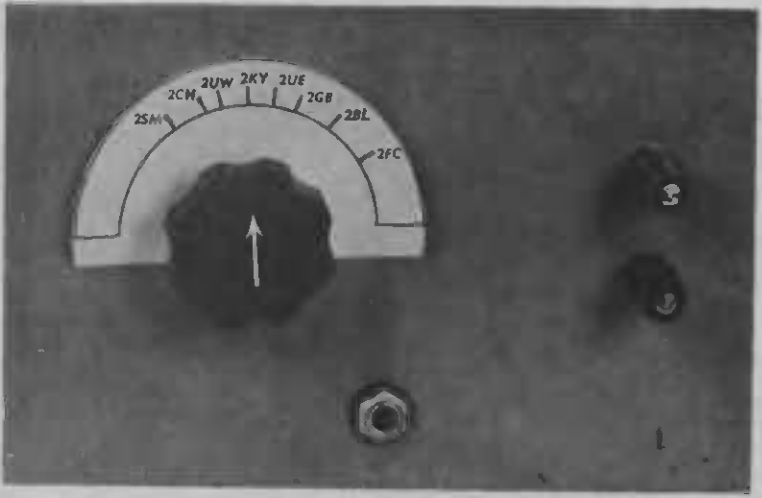
It is only by either removing one or the other set of half-cycle pulses, or by somehow "up-ending" them so that they add to the first set, that there is a net effective voltage to which the earphone or loudspeaker can respond. The diode detector uses the "remove one set" method, which is called half-wave detection. Most detectors use this method, but one or two use the "up-end one set" method, which is known as full-wave detection.

By using a diode detector to "chop off" one side of our tuned circuit EMF we can feed the result to an earphone or loudspeaker with the possibility of being rewarded with faint sounds. Just how faint they will be will depend upon our aerial

DETAILS OF A TYPICAL "CRYSTAL" SET



A wiring diagram showing how the "crystal" set of Fig 4 might be assembled. The parts are shown as they would be seen from the front of the set if the front panel were transparent. The coil consists of about 80 turns of enamelled copper wire (22 B & S or similar), close wound and with a tap at every five turns for the 35 turns nearest the "top" end. This is the end connected to the fixed tuning capacitor plates. Below is a front view of the receiver with a simple tuning dial, cut from card.



system and the distance from the transmitter, because these factors will control the strength of our aerial and tuned circuit EMFs.

In some cases, we can improve the efficiency of this "simplest" radio receiver by adding another capacitor.

Fig. 3(c) shows where this additional capacitor is connected — right across the load. If the capacitor is arranged to be a suitable value, it can help in the pulse averaging operation carried out by our earphones or loudspeaker. It does this by charging up during each tiny radio frequency pulse and supplying its stored energy to the earphone or loudspeaker until the next pulse arrives.

In other words, the added capacitor "smoothes" out the pulses into a direct voltage which varies only with the modulation. This is illustrated by the waveform in Fig. 3(c).

The circuit above this waveform represents what may be regarded as the "simplest practical radio receiver," and is, in fact, the basis of the "crystal" set which we will discuss in a moment. It involves no amplification of the received signals, but is simply a means whereby a small amount of energy may be drawn from the electromagnetic waves from a radio transmitter and used directly to reproduce the sound which has been superimposed on it.

In the more elaborate receivers which we will be discussing in following chapters, the signals are amplified to make them more powerful and better able to operate a loudspeaker or earphone under poor signal conditions. In most types of receivers, the signals are amplified both before and after they are demodulated in the detector.

For the remainder of this chapter, however, let us see how a typical "crystal" set carries out the tasks which we have been discussing. Incidentally, the term "crystal" is still used for sets of this type, even though the old galena crystal-and-cat's-whisker-wire detector is rarely ever used nowadays. As long as the set does not amplify the signal, it is usually called a "crystal set" despite the fact that it might use a diode valve or a germanium diode.

The circuit for a simple receiver of this type is shown in Fig. 4. As you can see, there is not really much difference between this circuit and that of Fig. 3(c). The earphones are shown connected to the output of the germanium diode, and the small smoothing capacitor is connected across the phones to ensure proper detection efficiency.

The EMFs induced in the aerial-earth capacitor system by the radio waves are magnetically coupled to the inductor L of the tuned circuit as before, and the capacitor C is variable so that the tuned circuit may be adjusted to resonate at any frequency within the broadcast band (550 to 1600 KHz).

A point of difference is that the detector diode connects to a tap on the inductor, so that the diode and earphone circuit are connected across only a small portion of the tuned circuit rather than right across it.

This is done because practical earphones load the tuned circuit too much if they are connected in series with the diode across the whole inductor. As this prevents the tuned circuit from properly rejecting all the unwanted signals — upsets its selectivity, in other words — the detector and earphone circuit are connected across only part of the inductor.

How far the detector is tapped down the coil depends also upon the strength of the available signals, and hence upon the distance from the transmitter. This is because tapping down the inductor reduces the proportion of tuned circuit voltage fed to the detector and earphone, and reduces the loudness of the sound heard.

The position of the tap is therefore a compromise between good sensitivity and good selectivity. The higher the tap is moved up the inductor the greater will be the volume, but the poorer the selectivity; conversely, lowering the tap will generally lower the volume but give better rejection of unwanted stations.

There is only one component in Fig. 4 which we have not as yet said much about — the earphones. Most readers will be aware that these somehow convert a varying voltage into sounds, but some may not be familiar with their construction or operation. It might be as well if we spent a few moments clarifying these matters, particularly as we will be meeting with earphones again in later chapters.

Fig. 5 shows the basic construction of an earphone of the type used in most "headsets" and in many telephone handsets.

Basically the earphone consists of a permanent magnet which attracts a thin "stalloy" metal diaphragm via two soft iron pole-pieces. The diaphragm is stressed inward due to the attraction and thus assumes a slight concave shape under normal conditions. The housing and cap are arranged so that the diaphragm almost touches the tips of the pole-pieces, but never quite does so.

A coil of fine wire is wound on each pole-piece, and the coils are connected in series so that when a current is passed through them they both have the same effect upon the magnetic field produced by the magnet.

When a current is passed through them in one direction, their magnetic fields add to that of the magnet and attract the diaphragm still further. Conversely, when a current is passed through them in the opposite direction their magnetic fields tend to cancel the magnet field and the attraction is reduced.

If a varying or alternating EMF is applied to the coils, a correspondingly varying or alternating current will flow through them. This will cause the diaphragm to move in sympathy with the variations or alternations in the EMF, as its attraction will vary with the current variations.

And when the diaphragm vibrates in sympathy with the voice — or music — frequency variations in the applied EMF, it sets up sound waves in the air. If the applied EMF is the output of our detector diode, the sound will be a reproduction of the sounds which were picked up by the microphone at the transmitter.

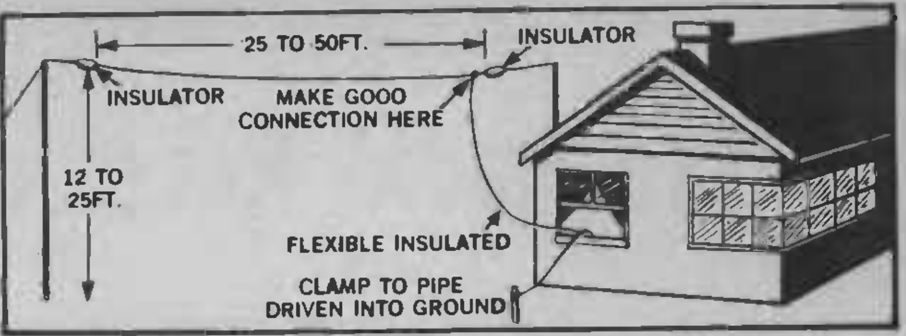
As we shall see in later chapters, earphones and their "big brothers" the loudspeakers are an important part of every radio receiver. Without them to reconvert the electrical signals back into sound waves, radio would scarcely be practical.

The simple receiver circuit shown in Fig. 4 is quite practical, and interested readers may care to build it up for the valuable experience which this will give. For these readers we give the following comments for guidance.

The construction and wiring of the receiver should be fairly obvious from the wiring diagram. The components are mounted on a piece of plywood or flakeboard some 6½in x 5in, to which is nailed a wooden or composition board front panel about 8in x 6in.

Two terminals on the front panel are used to allow the aerial and earth lead-in wires to be connected easily. The aerial should be as long and as high as is practical, and should preferably be a length of 7 x 0.022in copper "earth" wire. It must be insulated from its supports, and this may be done with "egg"-type insulators of plastic or porcelain.

The earth wire may be clamped to the house water-pipe near where it leaves the ground, or if this is not practical, it may be a connection to a yard or so of galvanised pipe driven into fairly damp ground. Don't connect to a gas pipe, though, for this is frowned upon by the authorities.



How to erect an aerial / earth system suitable for use with a crystal set or other small receiver. Stranded copper wire is excellent for the aerial and earth but use good quality plastic covered stranded wire for the lead-in.

The tuning capacitor need not be a new one. A tuning gang salvaged from an old radio set may be pressed into service providing it is still working satisfactorily. Make sure that none of the moving plates touch the fixed plates, and connect to only one set of fixed plates if the unit happens to be a multisection or "gang" capacitor.

The shaft of the capacitor is taken through a hole in the front panel and fitted with a large knob. A dial scale pasted on the front panel can then be marked with the station names when the set is completed.

Below the wiring diagram we give coil winding details for those who would like to wind their own. The former may be a tube of plain or waxed cardboard, or even of wood. To make the taps, wind the appropriate turns over a piece of match-stick. The raised wire at each tap can then be scraped free of enamel and soldered to (or connected to using an alligator clip) as required.

If desired, of course, a commercial coil may be used in place of the homewound one. The most suitable type would be a high-gain transistor radio aerial coil. Connect aerial, earth, and the two sets of capacitor plates to the coil exactly as originally intended, using the "base" tapping for the detector in the crystal set.

The germanium diode may be almost any of the common types, such as the OA91, OA90, OA85, OA81, OA80, OA79, OA5, GEX35, GEX34 or GEX33. If readers have an old "cats-whisker" detector they could try it, but results generally will not be as good as with one of the modern diodes.

For this particular application, it does not really matter which way round the diode is connected into the circuit.

The 0.001uF capacitor may or may not have much effect, depending upon the type of earphones used. Since it will only cost a few cents, it may as well be put in.

Low impedance earphones are not suitable for this type of receiver. For best results, use high impedance 1000-4000 ohm units. Piezoelectric crystal earpieces are in general rather too insensitive, despite their high impedance, although they may prove satisfacotry in strong signal areas.

The earphones are connected to a jack-plug, which plugs into a jack mounted on the front panel of the set.

There is no need for an on-off switch, for there is nothing to turn on and off. Simply connect up the aerial and earth, plug in the earphones, and you should be able to tune in to one or more stations, provided you are not further from them than about 15-20 miles.

How many stations you will actually be able to receive without mutual interference will depend on relative strengths at the listening site, some sites being much more favourable for crystal sets than others.

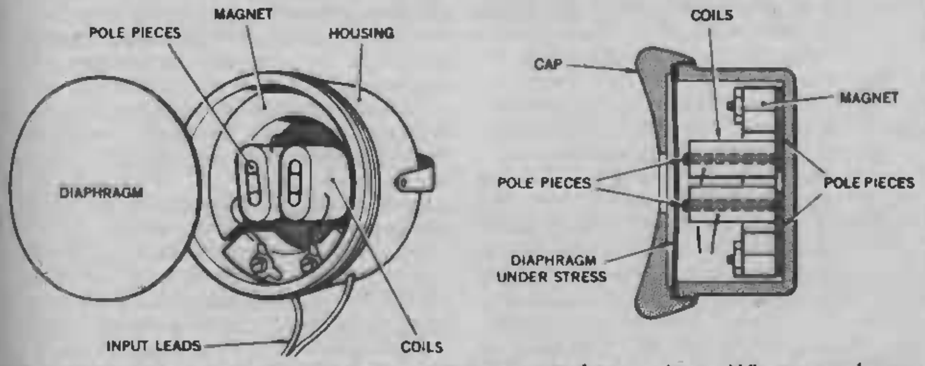


Fig 5: Two diagrams illustrating the construction of a typical earphone. When a varying or alternating current passes through the coils, the thin metal diaphragm vibrates to generate sound waves.

CHAPTER 12

Simple Radio Receivers

Limitations of a diode or "crystal" set — amplified crystal sets and audio amplifier stages — active detectors and their operation — reaction or regeneration — two stage receivers with transformer and resistance-capacitance coupling — additional audio stages — output power considerations — overload and gain or volume control.

Having used the crystal receiver to learn some of the basic facts about radio reception, we are now in a position to discuss simple transistor and valve receivers.

As we saw in the last chapter, a crystal set is a very useful and interesting device. It is simple to make, it costs nothing to operate and it demonstrates, in a practical way, many important radio principles.

For all that, however, a crystal set has very serious limitations. The only energy available to it is the radio frequency energy picked up by the aerial and earth system from the desired transmitter. This is selected, demodulated and made available to the earphones as an audible signal.

As the distance between receiver and transmitter is increased, the energy available becomes less and less until, at a distance which may be as little as 25 miles, the signal becomes inaudible. Only in very exceptional circumstances are the signals from a crystal set ever strong enough to operate a loudspeaker.

Yet another serious problem is that of poor selectivity, a crystal set often being unable to separate the wanted signal clearly from other strong signals in the receiving area.

In the face of such limitations, it is not surprising that engineers, very early, sought to improve the performance of crystal receivers or, alternatively, to supplant them altogether. Nor is it surprising that they have been relegated, in this modern age, to the role of a "beginner's set".

As you have probably guessed, the answer was found in a device we have already discussed — the thermionic valve. If you've forgotten this earlier discussion, we suggest you turn back to chapter 7.

Strangely enough, the very first valve receivers were no more ambitious in their performance than crystal sets — in fact, there were plenty of early radio operators of the day who claimed that they were not as good.

These early valve receivers were just like crystal sets, in fact, except that they used a diode in place of the metallic crystal and "catswhisker" detector.

As we explained in the earlier chapter, diodes exhibit the same rectifying properties as a crystal, being able to pass current only in one direction. They make signals audible in the phones by the same process as explained for a crystal set.

The main advantage of the valve or thermionic diode was that it needed no critical adjustment. This advantage was very real in a day when the surface of crystal diodes had to be probed with the "catswhisker" contact to discover a sensitive spot.

Against this, of course, the diode valve needed a filament battery, which was something of a nuisance. Hence the arguments of the day as to which was the better proposition.

The development of the triode valve settled such arguments, because it brought with it the ability to amplify the incoming signals. Instead of being utilised to operate the phones directly, the signals were applied to the grid to control plate current flowing from a B-battery. The resultant and larger plate current excursions, dependent on the grid signal, produced much louder signals in the phones.

Fig 1 shows a type of receiver which was quite popular in its day — the combination of what is virtually a crystal or diode receiver and a triode amplifier stage.

The incoming signal is selected by the tuning circuit and applied to the detector. This latter may be either a semiconductor diode or a thermionic diode, which suppresses half the incoming carrier and delivers to its output circuit what we described, in the last chapter, as a series of unidirectional pulses proportional in strength at each instant to the modulated carrier.

Instead of being passed directly through the phones, to produce an audible sound, these pulses are passed through the primary winding of an audio transformer. Perhaps we should pause here to explain these terms at least in brief.

The word "audio" comes from the Latin verb "to hear" and is used in electronics to describe any circuit or component which handles signals at a frequency within or adjacent to the range of sound frequencies. Thus an audio amplifier stage is one which amplifies signals at audio frequencies.

By the same token, an audio transformer is one which is designed to handle, or transfer, or couple signals at audio frequencies.

The principles of transformers generally have been discussed in an earlier chapter and obviously cannot be repeated here. An audio transformer is usually wound on a core made up from iron laminations. It normally has two windings, and each may comprise many thousands of turns of fine wire.

The input signal is fed to the winding and normally referred to as the primary. Because of the mutual coupling of the two windings (see Fig 1) a resultant signal is developed across the secondary winding and applied to the grid of the triode.

It is possible to secure a step-up in signal voltage from a transformer by winding more turns on the secondary than on the primary. Old-style transformers, which often come into the hands of experimenters, typically have a turns ratio of 1 to 3 or 1 to 5 from the primary winding to the total secondary winding. More modern transformers intended for transistor circuits may have less turns on the secondary than on the primary. This gives a step-down in voltage, but a step-up in current.

Now back to Fig 1.

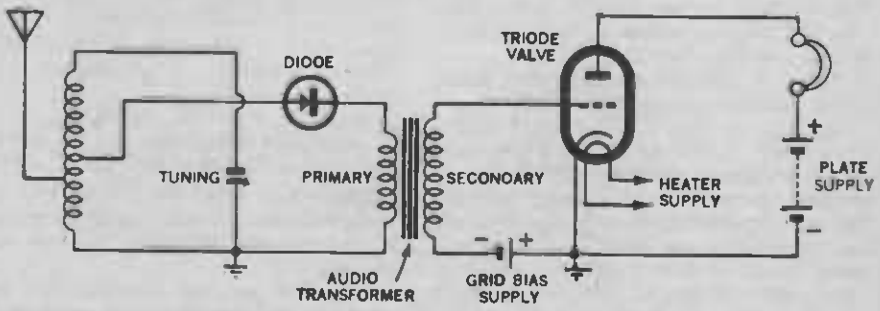
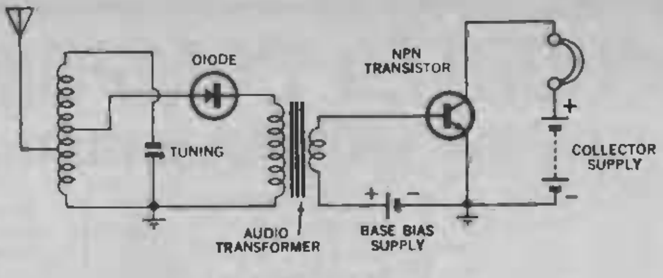
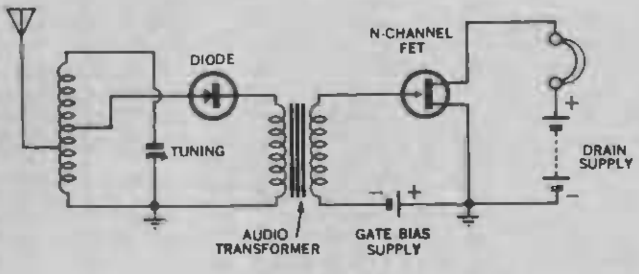


Fig. 1: An early type of receiver which used a triode to amplify the demodulated signals produced by a diode detector circuit.



(a) NPN BIPOLAR TRANSISTOR



(b) N-CHANNEL FET

Fig. 2: Bipolar transistor and FET versions of a basic receiver as shown in Fig. 1. Note the similarities between the three types of audio output amplifiers. A step-down transformer is desirable for the transistor, as explained in the text, but not for the FET.

The signal currents from the diode detector flow through the primary winding of the transformer. Now the transformer is no more able to respond to individual unidirectional carrier pulses than the earphones referred to in the last chapter.

However, the current through the primary winding, and therefore the magnetic field it produces in the core, tends to merge into a pattern, which follows the rise and fall of the incoming carrier with modulation.

The changing magnetic field, due to current through the primary winding, induces current in the secondary winding and a corresponding signal voltage between its two ends.

These two ends are connected respectively to the grid and cathode circuit of a triode amplifier valve and the audio voltage between them therefore constitutes a grid signal controlling the flow of electrons through the valve from cathode to plate.

For the reasons explained in chapter 7, a bias voltage is normally provided to keep the grid slightly negative with respect to filament, the optimum bias depending on the type of valve and its other operating conditions. When the incoming signal carries the grid more negative than the standing bias, current through the valve is reduced. Conversely, when the signal makes the grid less negative, current through the valve is increased.

This ever-changing current, flowing from the High Tension or B-battery through the phones produces much more output from the phones than could the small current pulses available from the detector.

Fig 2 shows the circuits for receivers which are the transistor and FET counterparts of that in fig 1. In (a) an NPN bipolar transistor is used in place of the triode valve, and is supplied with base bias voltage and collector supply voltage with the polarities shown. The audio transformer used with the transistor usually has a step-down primary-secondary turns ratio rather than a step-up ratio, because as we saw in chapter 8 transistors have a relatively low input resistance and are current amplifiers rather than voltage amplifiers.

While fewer turns in the secondary of the transformer than in the primary give a step-down voltage ratio, as mentioned earlier it actually gives a step-up current ratio, and this suits the transistor admirably. It also "matches" correctly the relatively high-resistance detector circuit in the primary and the low resistance transistor input circuit connected to the secondary.

Fig 2(b) shows how an N-channel FET would be used in a similar type of simple receiver. The circuit is very similar to those for the valve and bipolar transistor versions, as may be seen. Note that the polarity of the gate bias supply is opposite in polarity to the drain supply, pointing to the similarity between FETs and thermionic valves.

Simple receivers along the lines of Figs 1 and 2 are capable of substantially better performance than an ordinary crystal set. Sound volume from near-by stations is increased. Range is effectively improved because signals which might otherwise be inaudible are amplified to listenable strength.

Even the effective selectivity can be improved because amplification from the audio stage allows the tappings on the coil to be moved closer to the earthed end than would otherwise be the case. Selectivity is improved as a result.

Still further improvement would be possible by providing two or even three audio stages after the crystal detector. In practice, however, this is seldom done because better overall performance can be obtained by following different circuit principles, at least for simple beginner's type receivers.

What is involved, primarily, is the elimination of the crystal or diode detector and the substitution of an "active" detector stage using a transistor, FET or valve.

Fig 3 shows the basic circuit for a simple receiver using an N-channel FET as an active detector.

To understand its operation, one must remember that the gate and channel regions of a FET are separated by a P-N junction which is virtually identical with the P-N junction of a normal semiconductor diode. Normally, the junction of the FET is reverse-biased, and its depletion layer is used to control the source-drain current flowing through the channel. But if the gate-channel junction is forward biased, it will itself conduct current, just like a normal diode.

Now in the circuit of Fig 3, the input signal selected by the tuned circuit is fed to the gate of the FET through a coupling capacitor marked "Cg" (this kind of notation is often used to facilitate discussion of electronic circuits, by the way; thus "Cg" is short for "capacitor connected to the gate", and so on). In a typical circuit, the value of the capacitor might be between about 100pF and 470pF. There is no fixed reverse bias on the gate of the FET, so that the gate-channel junction will tend to be forward biased during each positive half-cycle of the RF input signal, and reverse biased during the negative half-cycles.

During the half-cycles in which the junction is forward biased, it conducts current. The pulses of current which flow in this fashion tend to charge up capacitor Cg to the peak value of the RF voltage developed across the tuned circuit. There is thus built up across the capacitor a DC voltage whose size is proportional to the peak value of the RF input voltage, and whose polarity is such that it tends to reverse-bias the gate-channel junction of the FET — ie, in this case with negative connected to the gate.

The gate-channel junction of the FET naturally cannot conduct current during the half-cycles of the RF input voltage which tend to swing it in the direction of increased reverse bias. In fact, during these half-cycles the charge on capacitor Cg tends to "leak away" through resistor Rg. The discharge current flows down through Rg and back up to the capacitor via the low resistance of the tuned circuit coil.

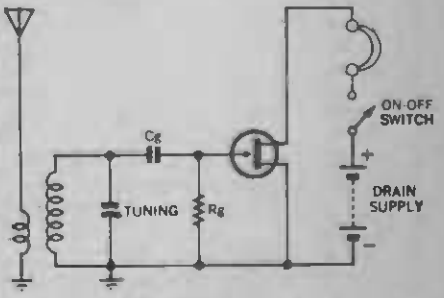


Fig. 3: The basic circuit of a simple receiver using a FET detector.

Far from being undesirable, however, this discharging action is actually necessary if the circuit is to be used to detect or demodulate any audio or other information impressed on the RF signal. Without resistor Rg, the charge on Cg would simply rise to a value corresponding to the highest positive peak of the input RF voltage, and stay at this value. Even if the RF signal were then completely removed, it would tend to remain at this value, dropping only very slowly due to leakage through the FET and through the dielectric of Cg itself.

In a nutshell, the "gate leak" resistor Rg is necessary to ensure that the charge developed by capacitor Cg due to gate

ALL PRICES
INCLUDE
TAX AND POSTAGE

IMPORTED COMPONENTS

P.O. BOX 1683P, MELBOURNE 3001

SPECIALS

2N3819	.75
2N3055	1.60
EM404	.22
OA5	.21
UA741 I / c's	1.00
S6M3 (3 amp, 600v)	1.50
S1M3 (3 amp, 100v)	1.20
1N914	.47
2N174	5.21
2N278	4.72
2N301	2.79
2N301A	4.02
2N406	.99
2N417	1.54
2N443	2.78
2N456A	3.51
2N489B	12.60
2N591	2.26
2N649	2.17
2N696	.99
2N697	2.02
2N706A	1.56
2N930	1.55
2N1038	3.35
2N1046	15.28
2N1073B	6.42
2N1302	1.14
2N1303	1.14
2N1305	1.24
2N1306	1.41
2N1307	1.41
2N1308	1.67
2N1309	1.80
2N1546	5.01
2N1639	1.41
2N1908	16.48
2N2102	3.02
2N2147	4.31
2N2160	3.34
2N2188	2.41
2N2270	2.25
2N2646	1.90
2N2647	3.17
2N2669	5.58
2N2926	1.95
2N3054	1.56
2N3525	3.18
2N3563	.75
2N3564	.60
2N3566	.73
2N3567	.67
2N3642	.79
2N3691	.53
2N3692	.60
2N3692	.52
2N3693	.60
2N3702	1.50
2N3703	1.23
2N3704	1.53
2N3705	1.50
2N3706	1.44
2N3707	1.44
2N3708	.70
2N3716	4.50
2N3790	9.60
2N4121	.66
2N4250	.73
2N4354	1.09
2N4355	1.30
2N4356	1.30
2SB186	1.30
2SB474	1.30
2SF28	4.81
AA119	.38

AC107	1.97
AC126	1.08
AC128	.90
AC132	.99
AC172	1.56
AD149	1.91
AD161 / 162	3.27
AN1104	.54
AN1105	.54
AN2001	.42
AN7102	.79
AN7105	.60
AS147	.88
AS148	.84
AS208	1.49
AS301	.84
AS306	.81
AS307	.81
AS308	.80
AS310	.93
AS11	.99
AS312	.99
AS313	.95
ASY73	2.05
ASY76	1.55
ASY77	1.87
ASZ16	2.42
ASZ17	2.05
ASZ18	2.32
ASZ20	.86
ASZ21	3.32
AT323	.54
AT324	.57
AX1101	1.34
AX1103	1.32
AX1104	1.34
AX1108	1.46
AX1127	.66
AX1130	1.28
AX1144	1.30
AX1166	.60
AX6168	.71
AY1102	.81
AY1108	1.44
AY1113	.62
AY1119	.54
AY6108	1.44
AY6109	1.44
AY8135	4.62
BA102	.53
BA114	.38
BC107	.57
BC108	.63
BC109	.68
BC147	.45
BC148	.42
BC149	.49
BC157	.48
BC158	.44
BC159	.52
BC177	.59
BC178	.55
BC179	.65
BC208	.56
BC212	1.25
BCY10	1.61
BCY11	3.74
BCY12	3.21
BCY39	5.97
BCY71	.78
BCZ210	1.95
BCZ211	2.37
BCZ212	2.16
BD139	2.04

BD140	2.09
BDY20	2.63
BDY38+	2.23
BF145	.64
BF173	1.20
BF177	1.63
BF178	1.20
BF179	1.29
BF180	1.22
BF185	.56
BF194	.44
BF200	.98
BFY51	.95
BSX19	1.50
BSX20	1.68
BT100A / 500R.	2.32
BT101 / 500	3.53
BTY79 / 100R	2.90
BTY79 / 300R	3.56
BTY79 / 500R	5.04
BY127 / 800	.63
BYX21L / 200R	1.17
BYX38 / 300+R	1.06
BYX38 / 600+R	1.15
BYX38 / 900+R	1.24
BYX38 / 1200+R	1.37
BYX39 / 600+R	3.21
BYX39 / 800+R	3.77
BYX39 / 1000+R	4.34
BZX70Scvies	.97
B2Y88C303 to C11	.49
B2Y88 C12 to C30	.82
BZY95 Scvies	1.84
BZY96 Scvies	2.16
BZZ15 to 29	1.58
C20D+	5.64
C106 / Y1	1.81
C122D	3.24
D13T1	1.93
DTG110F	6.96
Dtg1010	8.40
EM410	.66
FLV100	2.83
FPT100	1.56
H35	6.88
MB1	1.75
MB3	2.29
MB6	3.40
MB8	3.67
MJE2955	3.92
MJE3055	2.63
OA10	2.09
OA47	.90
OC20	5.45
OC22	4.92
OC23	7.54
OC24	6.12
OC44N	1.56
OC45N	1.56
OC74N	.86
OC141	2.50
ORP12	.99
OC201	2.33
OC202	3.70
OCP70	2.60
OCP71	4.32
ORP60	1.75
PA40	4.83
PB40	7.26
SC45D	11.10
SC50D	13.00
SE1001	0.42
SE2001	.98
SE2002	1.20

SE3001	1.12
SE4001	0.38
SE5001	2.10
SE5002	2.10
SE5003	2.48
SE5020	4.05
SE5023	3.15
SE5025	1.35
T1 / 40 / A2	3.30
T13027	2.61
TIC44	1.68
TIC45	1.88
TIC46	2.05
TIC47	2.30
TIP31A	2.10
TIP32A	2.70
TIP33A	2.98
TIS34	2.46
TIS43	2.36
40360	2.48
40361	2.70
40407	2.52
40408	2.98
40409	3.15
40410	3.30
40411	7.00

FETS

MPF102	1.00
MPF104	1.10
MPF105	1.50
MPF121	1.50
2N3820	4.55
2N4360	1.55
2N4889	2.50
2N5245	2.50
2N5485	1.50
2N5486	1.60

I.C.'S

uA703	1.94
uA709	1.30
uA710	1.30
uA716	4.80
uA723	2.58
uA739	4.91
FUL900	1.05
FUL914	1.05
FUL923	1.94
LM301A	1.94
LM305	6.21
LM309K	4.04
LM370	5.76
LM372	4.61
MC1303L	4.50
SN7400 series exstock	P.O.A.
TAA300	3.70
TAA570	3.68
TAA840	4.53
TA10B	6.25
TA10AB	6.50
TA20B	12.75
TA25B	15.00
TA50	35.00
TT797	1.32
TT798	1.33
40411	6.00
BC182L	1.25
TT800	1.95
TT801	1.95
3N81	5.50
TIL209 (LED)	.65
40362	2.74

Prices subject to change without notice.

PLEASE ENQUIRE FOR ANY DEVICES NOT LISTED

current flow can leak away fast enough to follow any downward changes in RF signal strength. Resistor Rg is thus given a value such that it can discharge Cg fast enough to correspond to the "downward slopes" of the waveform of the highest modulating frequency which it is desired to detect. This value usually lies between about 220K and 2.2 megohms.

Capacitor Cg, resistor Rg and the gate-channel junction of the FET thus act as a detector circuit very similar to that in the "crystal" receiver described in chapter 11. The flow of gate current during positive peaks of the RF voltage developed across the tuned circuit results in the building up of a unidirectional charge across Cg, and because of the action of Rg this charge is able to vary in size to follow faithfully any variations in the RF signal corresponding to modulation.

But this is only half the story. While the gate-channel junction of the FET is thus arranged to function as a detector, the FET as a whole is still able to function as an amplifying device. Because the drain-source current is related to the gate-channel bias by the transconductance, as we have seen in chapter 8, the varying reverse bias developed across capacitor Cg by the detector action results in magnified variations in the average drain current flowing through the phones. The sounds produced by the phones are thus considerably louder than they would be if the phones were connected directly into the detector circuit.

It is because this type of circuit performs both the functions of detection and amplification together that it is called an "active detector".

Although the circuit of Fig 3 uses an N-channel FET to illustrate the operation of an active detector, other devices can be made to operate in the same way. Thus a thermionic valve may be substituted for the N-channel FET, and would operate in virtually identical fashion provided it was supplied with filament or heater power and a suitable plate supply. A P-channel FET could similarly be used, simply by reversing the polarity of the drain supply.

NPN and PNP bipolar transistors may also be used as active detectors, although these require a slightly different biasing circuit to work efficiently.

Although an active detector circuit of the type shown in Fig 3 is quite interesting from a technical viewpoint, as it stands it cannot boast any special order of performance. The performance is not markedly different from the circuits of Figs 1 and 2, in fact, so that the main advantage offered is that it eliminates the need for a separate detector diode and an audio transformer.

That is not the end of the story, however. A simple addition to the circuit can make an enormous difference to the whole performance. It involves the use of reaction or regeneration or positive feedback, terms which all mean much the same thing. Fig 4 shows a FET detector incorporating what is probably the best-known reaction circuit.

It must be emphasised that this is not by any means the only possible arrangement for a receiver using reaction. It is a popular and typical arrangement but it would be possible to produce a quite imposing article on the many circuits which have been evolved during the last 30 or 40 years around regenerative detectors.

The tuning, detection and amplifying action are basically the same as for Fig 3. However, advantage is taken of the fact that, over and above the detected audio voltage, there is present on the gate of the detector some of the original RF input signal. This is amplified and the signal at the drain contains the audio component, which operates the phones, plus an amplified RF signal.

When reaction is employed, this amplified RF signal is coupled back into the tuning coil in such a way that it adds to the signal energy already present. This involves placing a reaction winding close to the tuned winding and so arranging the connections to it that the signals tend to add rather than to cancel.

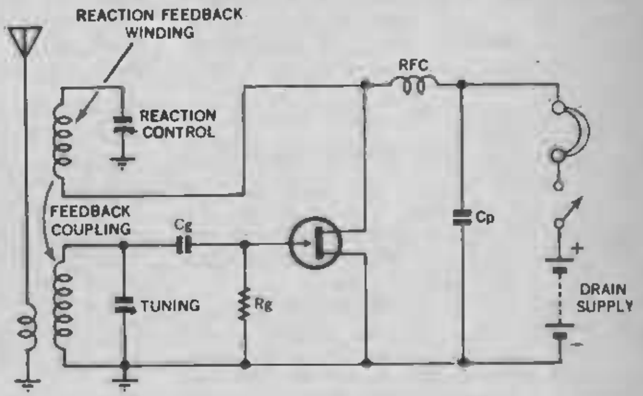


Fig 4: The way in which the performance of a simple type of receiver may be improved by applying positive feedback or "reaction". The reason for the improvement is explained in the text.

Assume, for example, that there is a positive signal pulse at the gate at a particular instant. This increases the drain current and causes a negative pulse at the drain. By impressing this pulse across the reaction winding and suitably arranging the connections, its phase can be reversed; ie, coupled into the tuned winding as a positive signal. This augments the original signal and produces a far greater total effect on the drain current than would the original signal without the feedback.

The effect of this type of feedback, therefore, is to make every positive signal excursion much more pronounced than it would normally be and every negative excursion likewise. The changes in signal level due to modulation are made much more evident and therefore the audio signal delivered to the phones is greatly increased.

The letters "RFC" in the circuit stand for "radio frequency choke". This component, which is usually a honey-comb-wound coil, is inserted between the drain and the phones to ensure that RF energy at the drain is not bypassed to earth by capacitance of the phone cords. The RF is therefore retained for use by the reaction feedback winding and control.

At the same time, RF energy is undesirable in the phone cords, because it can radiate into space and back into the aerial tuning circuit, causing the reaction adjustment to be upset by random movement of the phone cords or even by the person wearing the phones.

The radio frequency choke (inductor) is intended to prevent this trouble, its effect being augmented by the capacitor Cp shown in the circuit. This bypasses any RF energy to earth which may still be present but it does not bypass the audio components, which have a much lower frequency than the RF carrier.

Again, although Fig 4 shows an N-channel FET in the circuit, the principle of regenerative or reaction feedback can just as easily be appled to active detectors using other devices. And active detectors using thermionic valves or bipolar transistors, show an equally marked improvement in performance when this is done.

For a regenerative detector of this type to operate correctly, it is most important that the amount of feedback be properly adjusted. If there is insufficient feedback, only limited benefit is obtained from the scheme. If there is too much feedback, the detector will oscillate of its own accord and begin to act as a generator of RF energy, exactly as described in the chapter on radio transmitters.

To give the necessary control over reaction it is customary to connect a small variable capacitor in series with the reaction winding, as shown in Fig 4. When this is fully meshed, maximum RF feedbeck current can flow from plate, through the reaction winding to earth. As the capacitor plates are opened, the impedance of the circuit rises and less feed-back energy can flow through the coil.

To adjust the reaction in bipolar transistor circuits, it is sometimes more convenient to place a potentiometer across the feedback winding.

When the reaction control of the circuit in Fig 4 is set so that the detector is just below the point of active oscillation, the gain and selectivity of the detector and its tuning circuit are increased enormously. Used with an efficient aerial and earth, a one-FET or one-valve or one-transistor reaction set can receive signals under favourable conditions from transmitters thousands of miles away.

From the foregoing description, it might possibly be assumed that a one-stage regenerative set is all that should even be necessary to receive radio signals. But such is not the case.

Compared with a crystal receiver, a one-stage set has an enormous advantage in terms of sensitivity and selectivity — terms which relate to its ability to pick up a wanted signal and separate it from other signals. For all that, however, its performance is still capable of substantial improvement.

For example, the signals heard in the

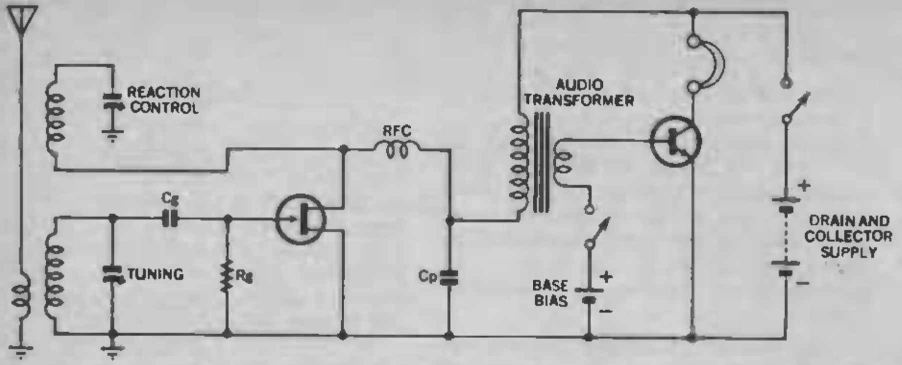


Fig 5: The addition of an audio amplifier stage to further increase sensitivity. In this case, the amplifier is shown using a bipolar transistor. It is coupled to the detector by means of an audio transformer, which would normally have a step-down ratio, primary to secondary.

phones from a distant station may be quite weak, requiring a good deal of concentration to follow them. The usefulness of the set can be increased greatly by adding an audio amplifier stage after the detector, exactly as already described in Figs 1 and 2 for a crystal set.

This gives a basic circuit such as that shown in Fig 5.

As before, the FET is used as a regenerative detector but, instead of its output being fed directly to the phones, it is passed through an audio transformer and fed to the base of an NPN bipolar transistor acting as an audio amplifier.

The amplified signals appearing in its collector circuit are then applied to the phones. Because of the extra amplification or gain, weak signals can be heard with less effort. Furthermore, the reaction control may not have to be set so critically to obtain adequate sound level, making operation and adjustment of the receiver that much easier.

The use of audio inter-stage transformers as shown in Figs 1, 2 and 5 was commonplace many years ago mainly because of the step-up they could give in the signal voltage. This supplemented, very usefully, the rather limited gain that was available from early valves.

As a component, however, audio transformers have always been rather bulky and expensive, prone to breakdown and liable to introduce distortion of one type or another. As a result, the growing tendency through the years has been to avoid them by using alternative coupling methods. One such method is a resistor-capacitor or R-C coupling, which is illustrated in Fig 6.

A resistor Rd, normally called the drain load resistor, is connected between the FET drain and the battery in place of the audio transformer primary winding. With no input signal, a certain drain current flows through this resistor and produces a corresponding voltage drop across it. The actual voltage at the drain of the FET is thus somewhat less than the battery voltage.

The base of the following audio amplifier transistor is fed with its appropriate bias as before, but not in this case by means of a separate battery as in Fig 5. It is necessary to supply the bias from a source having a reasonably high impedance, and although this can be done in a variety of ways, the method shown in Fig 6 is that most often used with bipolar transistors. Here the bias is derived from the main drain-collector supply battery using a voltage divider formed by resistors Ra and Rb. The relative value of the resistors determines the proportion of the battery voltage applied to the transistor, so that the bias is adjusted by altering the resistor values.

Between the FET drain and the transistor base is the coupling capacitor, Cc. Since the capacitor is connected between the drain, a point in the circuit at relatively high voltage, and the base, a point in the circuit at relatively low voltage, it will initially acquire a charge equal to the voltage difference between the two. And the capacitor is always made large enough with respect to the resistors Rd, Ra and Rb that it cannot alter this charge appreciably at an audio rate.

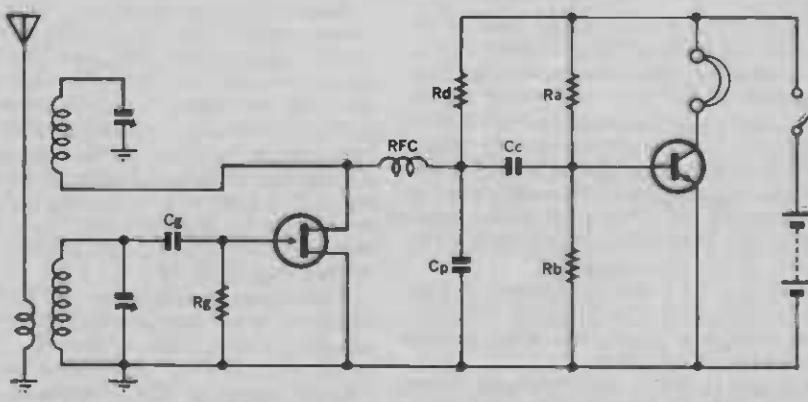


Fig 6: A two-stage receiver similar to that of Fig 5, but in this case using resistor-capacitor or "R-C" coupling between the detector and audio amplifier stages. This avoids the relatively costly audio transformer but, in general, offers somewhat less overall gain.

Now, when an audio component swings the FET drain current up and down, the voltage drop across the drain load resistor varies. As a result, the drain voltage itself varies at an audio rate.

Since the capacitor cannot alter its charge at an audio rate, it simply transfers the variations in voltage to the following base, the variations appearing at the base as an alternating audio signal. The signal is then amplified by the transistor in the ordinary way.

In other words, the coupling capacitor transfers the AC audio signal from the drain to the base, while at the same time preventing the relatively high voltage at the drain from upsetting the somewhat lower bias voltage at the base.

Much more could be said about resistance-capacitance coupling, but the foregoing should convey the general idea. Needless to say the technique is equally suitable for coupling between transistors, FETs and valves, or any combinations of these devices.

Just as the addition of one audio stage to a detector makes for a more sensitive and versatile receiver, so can further improvement be obtained by using two audio stages, with either transformer or resistance-capacitance coupling. In point of fact, many domestic receivers in the early days of radio were designed around a detector and two audio stages.

In such a case the amplification can be of such an order that the use of a loudspeaker can be considered, rather than headphones. The convenience of a loudspeaker is obvious but it does need to produce a great deal more sound output than phones, if it is to be heard properly.

This raises a special difficulty. If a loudspeaker has to produce a lot more sound output or acoustic power, it has to be supplied with a lot more audio power in the form of electrical energy.

If we can cut a lot of corners to make the point clear, we can say that most loud speakers and, of course, earphones operate by virtue of a changing flow of current through their windings. Therefore a lot of acoustic output requiring a lot of audio electrical power can also be thought of as requiring a large change of current flowing through the windings.

Now if a transistor, FET or valve is 'to amplify without distortion, its output current cannot swing beyond the limits of zero to twice the standing or "no signal" current. Therefore, if the last transistor in a receiver is intended to draw only 1 milliamp of standing collector current, the maximum current change it can effect through phones or a loudspeaker is plus and minus 1 milliamp — that is, from zero to 2 milliamps.

Such a change might be plenty for phones but it certainly would not be enough to produce much output from an ordinary loudspeaker. To operate a loudspeaker, therefore, it is necessary to use in the last stage of a receiver a device which can draw a higher standing current. With a signal, the current can then swing through wider limits.

Transistor and valve manufacturers in fact provide devices expressly designed for use as power amplifiers. Such devices are designed to be capable of passing relatively large currents, and dissipating relatively large amounts of power, without damage.

It is beyond the province of this chapter to discuss the many circuit arrangements possible using transistors, FETs and valves, and, from the beginner's point of view, individual designs have to be accepted and constructed on their merits. As knowledge increases, the general ideas conveyed by this chapter will gradually be supplemented by other knowledge.

It should however be mentioned that the provision of high gain or amplification in a receiver can introduce the problem of overload. The word is almost self-explanatory.

On weak signals, the amplification available in a receiver may be just enough to raise their level sufficiently to operate the phones or loudspeaker.

If the same amplification is applied to signals which are already fairly strong, they will be amplified so much in, say, the first stage that they are too great for the second stage to handle. As a result, the stage overloads and produces a very distorted output signal — sounding rough and harsh to the ears.

To avoid this difficulty, it is often necessary to include in a receiver some means of varying the amplification. To use another phrase, some method of volume control or gain control must be included.

A certain amount of gain or volume control effect can be obtained by varying the setting of the reaction control. The more nearly this control approaches the position for oscillation, the louder will the signals become, and vice versa.

The big difficulty with this method is that the setting of the reaction control also affects selectivity and it may easily happen that a position which gives adequate signal level may not give enough selectivity to select the wanted from the unwanted signals.

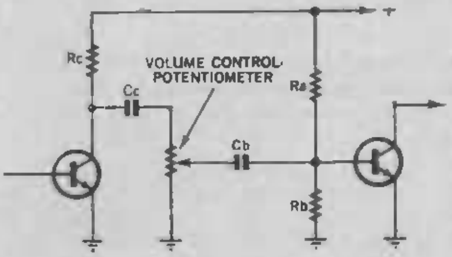


Fig. 7: A method of gain or volume control showing the way in which a potentiometer or "pot" is used. Overall resistance of the potentiometer depends on the circuit.

Ideally, the reaction control should be operable for best detector performance, with an entirely separate control for gain.

Over the years many methods of gain control have been devised, including in valve circuits the variation in filament voltage with a rheostat, variation in plate voltage or grid bias or variation of screen voltage in a pentode or tetrode. All of these schemes are open to criticism because, in reducing gain, they also limit the valve's ability to handle strong signals, thereby introducing distortion in many cases.

Nowadays the method almost universally adopted in audio circuits using any of the normal amplifying devices — transistors, FETs or valves — is that illustrated in Fig 7. A potentiometer is connected between the output of one stage and the input of the next in such a fashion that it may be used to adjust the proportion of output coupled between the two.

The audio AC developed across the load resistor Rc is fed by the coupling capacitor Cc across the whole of the potentiometer. The position of the moving arm of the potentiometer then determines the proportion of this AC voltage which is coupled through the second capacitor Cb into the following base. Thus by moving the arm of the potentiometer up and down, the resistance element varies the volume of sound heard in the phones or loudspeaker, and allows the volume to be adjusted to a convenient level.

In designing a radio receiver or audio amplifier, it is usual to connect the volume or gain control ahead of the first stage in the circuit which is likely to be overloaded in the event of a strong input signal. In simple radio receivers of the type which we have looked at in this chapter, the control would generally be connected between the detector stage and the first audio amplifier.

In the next chapter we will show you how to build a simple receiver, with ideas on the type of circuit, selection of components and some constructional details.

CHAPTER 13

Building Simple Receivers

After reading the last chapter, no doubt many readers will want to try their hand at building simple radio receivers based on the concepts which they have met. The circuits presented here have been designed for just this purpose.

Three basic simple receiver circuits, with variations, were described in Chapter 12. They were active detector, active detector with regeneration, and the latter with an audio amplifier stage.

In this chapter we shall describe simple receivers based on each of the circuits and discuss the level of performance to be expected from each. Fig 1 shows the first circuit, an active detector using an N-channel FET. This circuit is the same as Fig 3 in the last chapter but component values have been added.

The active detection function is performed by an N-channel FET, the 2N5459 (formerly called the MPF105) from Motorola. The coil assembly is a time-honored "valve" type Reinartz coil, still made by Aegis Pty Ltd and designated type M.12. The tuning capacitor can be any tuning gang with a maximum capacitance of about 400pF.

Other commercial Reinartz or "aerial with reaction" coils could be used, new or old, provided the windings are still intact. It will, of course, be necessary to identify the connections and relate them correctly to the circuit. Alternatively, you can wind your own coil to the specifications in the accompanying diagram.

The accompanying photograph shows a nine-volt battery but the circuit will give better performance with a 12 volt battery. Note however, that voltages over 12V should not be used otherwise the FET may be damaged. An old 12V car battery may be used to power the circuit, if one is on hand. One can really raid the junk box for a project like this.

There is no need for a metal chassis. Ours was made from a piece of particle-board and tempered hardboard (Masonite or Burnieboard) for the front panel. Wiring layout is not critical but beginners should follow our wiring diagram of Fig 2 to avoid mistakes. Cross check it with the circuit diagram, Fig 1. At any rate, keep all wires as short as possible, consistent with neatness.

The Reinartz coil was mounted on the board with the aid of an electrolytic capacitor mounting clamp, but there is no reason why it must be mounted in this attitude. It may be mounted on its side, if convenience dictates. Just don't let it float around. When making connections to the coil, take care not to overheat the coil terminals otherwise the plastic former will be melted.

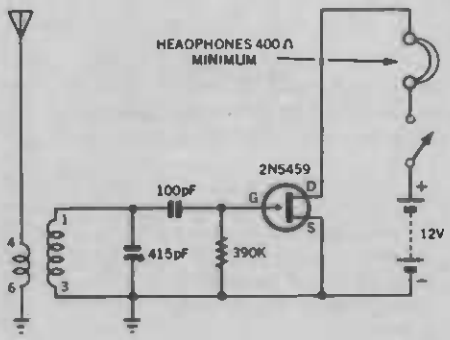
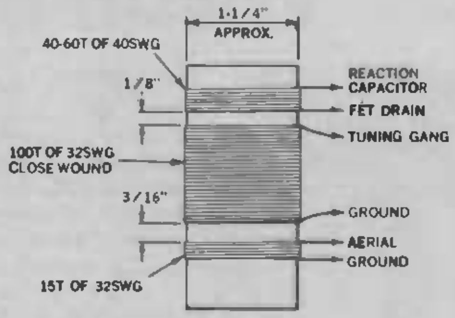


Fig 1: One FET and a handful of components make up the simple active detector circuit. The corresponding wiring diagram appears below.



Shown above are the winding details for those who wish to make their own "Reinartz" coil. Use a cardboard or plastic former.

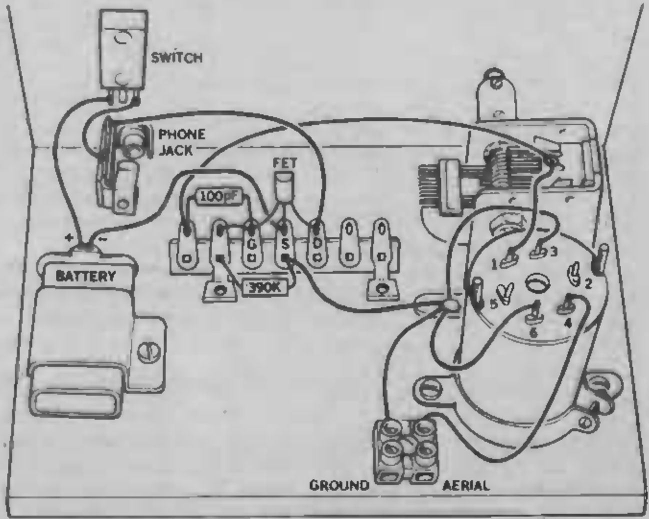


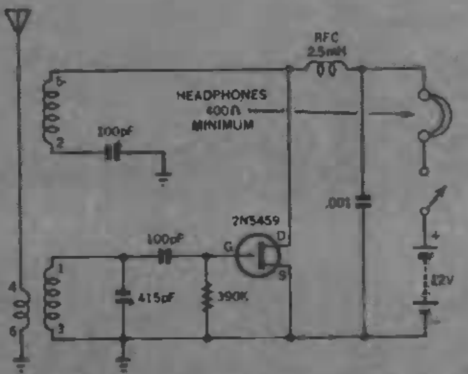
Fig 2 (right) is the wiring layout for the active detector circuit of Fig 1. Follow the diagram carefully to ensure correct connections.

If you use a tuning gang salvaged from an old radio set, make sure it is clean and that the moveable plates are not shorting to the fixed section. This can be easily checked with the aid of a multimeter switched to the "ohms" range.

If you are going to make a dial for the set remember that stations at the low frequency end of the broadcast band are received when the tuning gang capacitance is relatively high, ie, when the plates are approaching full meshed.

The FET and its gate components, 390K resistor and 100pF capacitor, can be mounted on a tagstrip as shown in the diagram. If you wish, it is not necessary to install an on / off switch for this circuit — disconnecting the phones from the jack socket effectively disconnects the battery.

Again, take care not to overheat components when soldering. This applies particularly to transistors and polystyrene capacitors. It is wise to tin the leads and to use a crocodile clip or a pair of long-nosed pliers as a "heatsink", if you are a beginner at soldering. The idea is to use the clip or pliers to grip the component lead between the solder joint and the body, so that most of the heat is conducted away from the lead before it has a chance to heat up the component itself and cause possible damage.



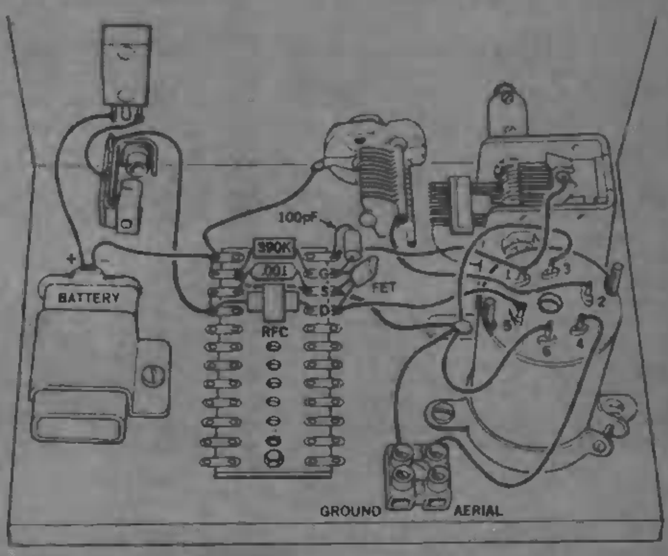


Fig 3, above, is the circuit for the active detector with regeneration while at right, Fig 4, is the wiring diagram. Note the extra components compared with Fig 2. For those wishing to wind their own Reinartz coil, winding details are given on the preceding page.

The headphones should be "medium" to "high impedance" — ie, they should have a nominal impedance of at least 400 ohms and preferably higher.

An insulated terminal block is mounted on the board for connection of earth and aerial wires. Strictly speaking, the earth connection should be made to a water pipe or a metal plate buried in the ground.

For best performance, the aerial wire should be as long and as high above ground level as possible. But many readers in the metropolitan areas may obtain adequate performance with just a few feet of wire.

Having checked all connections for errors, you are ready to apply power (connect the battery) and put on your headphones. As readers will remember from Chapter 12, the FET operates without bias when no signal is applied to the gate. This means that the current drain is relatively high at around 10 milliamps, depending on the impedance of the headphones.

When signal is applied to the FET gate, the FET develops a negative gate bias which is proportional to the peak value of the RF input signal. The mechanism by which this happens is fully discussed in Chapter 12. The effect of the negative gate bias is to reduce the current drain, although with this circuit unless the stations tuned are particularly strong, any reduction is likely to be minimal.

With this basic receiver you should be able to tune several radio stations, but this will depend very much on the area in which you live — whether or not it is a strong signal area, and on the parameters of the particular FET used.

Having determined the level of performance available from the simple active detector circuit, you are now ready to add components for regeneration. The circuit is shown in Fig 3. The additional components are a 100pF variable capacitor for controlling the level of regeneration, a 2.5mH RF choke and a .001uF capacitor. This is the same as Fig 4 of Chapter 12.

Again, we have provided a wiring diagram to facilitate connections — see Fig 4. Pins 2 and 5 of the Reinartz coil are now connected into circuit. The FET and its associated components are mounted on a short length of tagboard. Leave sufficient terminals available to add the audio stage to be described in the next step.

The reaction or regeneration capacitor we used is a small variable air dielectric type as used in transmitters but there is no reason why other types could not be used. One could, for example, use a smaller capacitance tuning gang or even a mica "compression" trimmer for the basis of the experiment.

Having connected the regeneration components, apply power to the circuit again and don your headphones. With the regeneration capacitor set for minimum capacitance, ie, with capacitor plates unmeshed, the circuit behaves very similarly to the active circuit tried previously.

Increasing the capacitance of the regeneration capacitor increases the loudness of the signal but also causes the signal to become distorted. This is because the loading effect of the regeneration coil and capacitor causes detuning of the main resonant circuit. Consequently, the tuning and regeneration controls interact and have to be adjusted in conjunction with each other.

As the regeneration control is advanced, the FET drain current is reduced markedly. This is because the regeneration acts to feed much stronger signals to the gate of the FET which consequently develops more negative gate bias.

If the regeneration control is adjusted too far, the circuit will go into oscillation. This will be evident in a number of ways. First, if the circuit goes into oscillation while it is being tuned to a station, the oscillation will be evident as a violent squeal. This is the heterodyne or difference frequency between the broadcast station's carrier frequency and the resonant frequency of the tuned circuit.

The reader may query this statement: If the circuit is tuned to the broadcast station, why should there be a difference between the incoming carrier frequency and the resonant frequency of the tuned circuit? This can be answered in a number of ways. First, the tuning gang does not have to be tuned exactly to the station's carrier frequency in order to hear signals — it can be moved away from it, if the signal is strong enough.

Another partial answer is the detuning effect of the regeneration control. You can easily show this, once the circuit has begun squealing, although you'll need to take the headphones off. Try varying the regeneration control — note its effect on the pitch of the heterodyne whistle.

If the circuit goes into oscillation between stations it will be noticeable as a single click in the headphones. The click is caused by the abrupt drop in FET drain current as the circuit goes into oscillation. This can be checked with a multimeter switched to a

PARTS LIST FOR COMPLETE RECEIVER

1 Chassis and panel to suit components.
1 12V battery and connections to suit.
1 Reinartz coil, Aegis type M12 or similar.
1 Tuning capacitor, 415pF (see text).
1 100pF variable capacitor for regeneration (see text).
1 2N5459 field effect transistor (FET).
1 BC108, 2N3565 or similar silicon transistor.
1 phone jack.
1 pair of headphones; minimum impedance 400 ohms.
1 on / off switch.
1 12-lug length of miniature tagboard.
1 2-way insulated terminal block.

RESISTORS
(½ or ¼ watt rating).
1 x 390K, 1 x 68K, 1 x 27K, 1 x 1K, 1 x 470 ohms.

CAPACITORS
1 x 10uF / 6VW electrolytic,
1 x 0.1uF / 25VW ceramic, polyester or paper,
1 x .001uF / 100VW ceramic, polyester or polystyrene,
1 x 100pF / 100VW ceramic or polystyrene.

MISCELLANEOUS
2 Knobs, screws, nuts, wire, solder, battery clamp.


PLESSEY

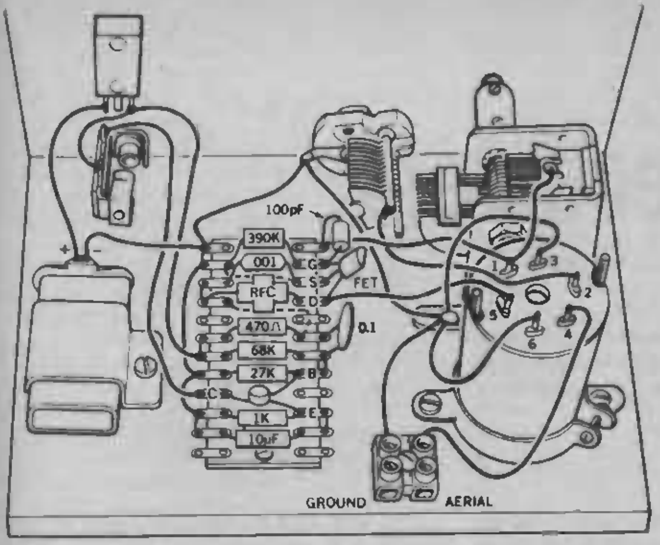
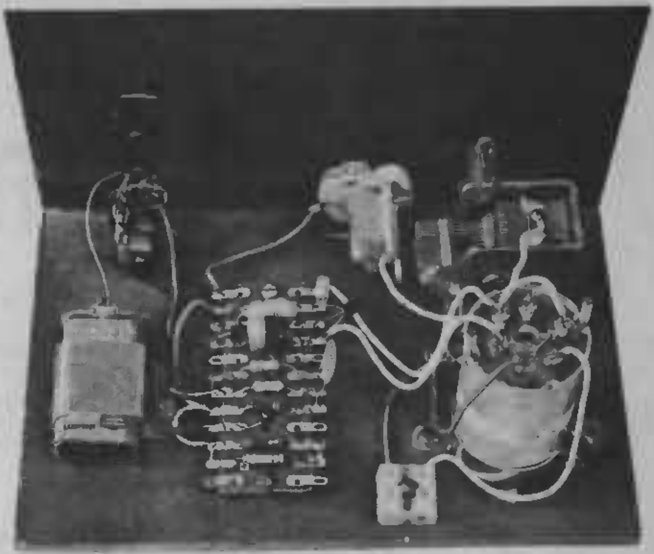


Fig 6: The wiring diagram for the final receiver, complete with regeneration and audio amplifier stage. With a good aerial and earth it should perform very well.



Our prototype receiver, built on a wooden baseboard and with a composition board front panel. Note that the receiver needs an earth as well as an aerial.

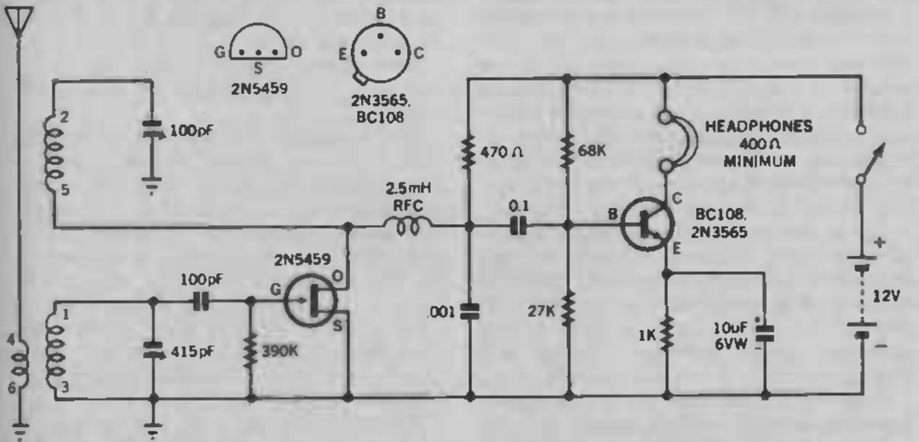


Fig 5: The circuit diagram of the complete receiver, as shown in the wiring diagram and the photograph above. The FET and transistor connections are depicted as seen, when looking from the lead end.

low current range. It can also be verified with the aid of a transistor radio, or any other radio for that matter.

You will be able to tune the radio to the resonant frequency of the circuit as it is radiated from the aerial. As you do so you will hear a slight hissing noise similar to that from a broadcast station during those brief periods when there is a break in the voice or music program. When in a state of oscillation, a regenerative receiver is virtually acting as a small transmitter.

If the receiver is allowed to oscillate while tuned to a broadcast station, a whistle may be heard in other receivers tuned to that station. The whistle is due to a beat or a heterodyne between the respective signals from the station and the receiver. In other words, it may cause interference in neighbouring receivers.

But this is hardly the purpose of building the receiver. We mention it as part of the discussion of its behaviour and to point out that an incorrectly used regenerative receiver can be a source of interference to radio communications.

The essential purpose of building these circuits is for the reader to discover the high degree of performance that can be obtained, with patience, from what is basically a very simple receiver.

The performance of the receiver of Fig 3 can be further increased by the addition of an audio stage to drive the headphones. The additional gain renders the receiver somewhat easier to use — the regeneration control does not have to be advanced so far to make stations listenable and consequently there is less likelihood of the FET detector breaking into oscillation.

Fig 5 shows the additional components for the audio stage. This is the unit shown in the photographs. In this circuit, the 470 ohm resistor becomes the audio frequency load for the FET and the signals developed across it are fed to the audio stage via a 0.1uF capacitor. No volume control is fitted to the circuit. One can use the regeneration control for this purpose, although admittedly, it is not ideal. We have omitted a volume control because it introduces a loss in gain.

The circuit is built along the same lines as the other two, with most of the smaller components mounted on a 12-lug length of miniature tagboard (see Fig 6). This should be completely wired as an assembly and then mounted on the board. As before, components salvaged from old radios may be used here but they should be checked before installation.

Resistors can be checked with a multimeter switched to the ohms range for correctness of value, and capacitors can be similarly checked for insulation resistance. It would be wise, though, to use a new electrolytic capacitor for the bypassing of the emitter resistor of the audio stage, because electrolytic capacitors deteriorate quite markedly with age.

The additional transistor is a general purpose silicon NPN bipolar type. Do not substitute other transistors unless you are sure they are directly equivalent.

The prototype receiver was tried out in the Western suburbs of Sydney, which is a strong signal area. With just a few feet of aerial, all the local broadcast stations romped in, plus a few country stations and Radio VL2UV, the University of NSW broadcast station on 1750KHz. With a longer aerial, it should do equally well in rural areas. But don't forget that, with a simple receiver like this, a good earth is almost as important as a suitable aerial.

CHAPTER 14

More Complex Receivers

Limitations of simple receivers — the RF amplifier stage, its operation and use — neutralisation — gain control of TRF receivers — the superheterodyne principle, its operation and advantages — autodyne converters — transistor receivers — more elaborate receivers.

While a beginner may concentrate initially on building small regenerative radio receivers of the type discussed in the last chapter, he must inevitably wonder about the design of larger receivers. In this chapter, we explain the basic idea behind two well-known types of receiver, the "TRF" and the "superheterodyne."

In an explanation of reasonable length, it is not possible to discuss individual circuits in detail — the how and why of every resistor and capacitor. The reader will have achieved something, however, if he can understand the general idea behind these circuits, particularly the superheterodyne.

Having grasped the basic idea, it should be possible to enlarge upon it later by studying the circuit and design information on actual receivers.

The TRF and, later, the modern superheterodyne receiver, came as a natural development from the desire to produce receivers which were more sensitive, more selective and more suitable for use by non-technical members of an ordinary household than regenerative receivers.

The story of the development of these types of receiver is really a continuation of the story told in chapter 12 about small receivers, and it must inevitably read like a piece of radio history.

As we pointed out in the previous chapter, a receiver having a regenerative detector followed by two audio stages is capable of receiving a great many stations, both on the broadcast and short-wave bands.

By using a power transistor or valve in the final stage, such a set can operate a loudspeaker at good volume on the stronger stations and, in the early days of radio broadcasting, many domestic receivers were of this general type.

For domestic use, however, such receivers have certain basic limitations.

In the first place, performance depends very largely on proper use of the reaction (regeneration) control. If it is too far advanced, the set oscillates, producing whistles in its own loudspeaker and in neighbouring receivers tuned to the same station. If the reaction control is not sufficiently advanced, perhaps to limit volume, selectivity is likely to suffer to the point where two or more signals are heard together.

While this is no special problem to anyone who understands what the controls are for and how they are supposed to be adjusted, it did prove an embarrassment in the early days for non-technical members of the household. Some less critical arrangement was obviously desirable for general use.

Another difficulty which was experienced with early simple regenerative sets lay in the fact that there was a practical limit to the amount of amplification one could provide following a detector. Thus, while one or two audio stages could be used to usefully increase gain and even selectivity (the latter by roundabout means), anything more than this tended to lead to difficulty.

Slight vibration in the detector valve, causing slight changes in plate current, could be amplified by subsequent stages to produce what are known as microphonic effects. Tapping the valve, or even normal vibration, would produce thumps and ringing noises from the loudspeaker.

Then again, noise, due to electron flow in the detector itself, could be amplified to the point where it produced a continuous background hiss. And in mains-operated receivers, very slight 50Hz or 100Hz voltages, coupled into the detector circuit from the heater and power supply tended to produce an audible hum from the loudspeaker.

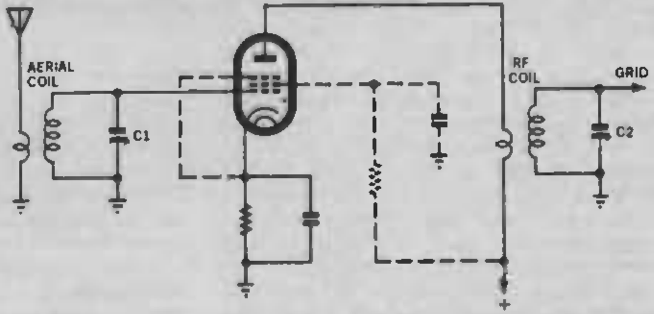


Fig 1: A valve-type RF amplifier stage, as used in early TRF receivers. The additional connections for a pentode are shown dashed.

Last, but not least, slight variations in high-tension supply voltage, caused by plate current variations in the output valve, could be fed back as a spurious signal to the plate of the detector. If regenerative, this feedback was able to cause an effect called motor-boating, evident as a regular pop-pop-pop noise from the speaker.

While all the problems could be minimised by careful design, they did set a limit beyond which the detector-plus-audio idea became rather impractical.

This limit was reached very early in the history of broadcasting, and designers had to find other means of improving the performance of their receivers. Since additional stages could not be added after the detector, the only alternative was to add stages ahead of the detector and to amplify the incoming signal at its own frequency.

Such stages were known as radio-frequency amplifier stages or simply RF stages.

Now an ordinary valve would not amplify radio frequency signals very effectively if merely coupled to the following stage by a resistance-capacitance network or by some kind of RF transformer. It would give a great deal more amplification if coupled to the following stage by means of a circuit tuned to the incoming signal frequency.

The essential circuit details of a valve-type tuned RF amplifier are shown in Fig. 1. This is the type of RF amplifier stage used in early receivers, and still used occasionally.

The incoming signal is fed through a tuned circuit to the grid of the RF amplifier valve. The coil of this tuned circuit is normally referred to as the "aerial coil". It is connected directly to the grid of the valve, without any capacitor or resistor, because the valve is intended to operate as an amplifier rather than as a detector. For the same reason the valve is provided with grid / cathode bias — here by means of a bypassed cathode resistor — to ensure operation as a class-A amplifier.

Amplifier signal output current from the plate circuit flows through the primary winding of a second coil assembly, and is coupled into a tuned secondary winding, the two windings forming what is commonly referred to as an "RF coil" or RF transformer.

The secondary windings of both the input and output tuned circuits, with their associated capacitors, must be capable of tuning over the entire broadcast band. To receive any given station, both tuned circuits should be set to the frequency of that station.

Under these conditions, the signal from the desired station is selected and passed to the RF amplifier grid, in preference to other signals which may be present. It is amplified by the RF amplifier valve and passed through the second tuned circuit, which also favours the desired signal and tends to reject signals on other frequencies.

In other words, the use of a tuned RF stage not only provides amplification, but also increases selectivity. This was — and is — a most important point.

Since an RF amplifier stage feeds into a circuit tuned to the signal frequency, this is the only frequency which it can amplify properly. Because there is no load resistor in the plate or collector circuit and no audio transformer, it cannot amplify significantly or pass on signals within the audio range. Therefore, it is not nearly as susceptible as an audio stage to hum, hiss, microphony or motor-boating. This, too, is important.

Many early receivers used one triode RF amplifier stage, a regenerative detector and two audio stages. They were the first "TRF" receivers, the letters indicating the use of a tuned radio frequency amplifier stage.

Such receivers were generally better than earlier types without the RF stage. They had better gain and selectivity and therefore relied to a lesser extent on critical setting of the reaction control. And because there was an amplifier stage between the detector and the aerial, the reaction setting was not affected so much by the type of aerial in use.

For all that, the basic problem remained that there was still a reaction control to set, and attempts were made to produce receivers with two RF amplifier stages ahead of the detector, but with no reaction.

Here designers came up against the full measure of a problem which was mentioned in chapter 7. They found that, because plate and grid in a triode were side by side, there was considerable capacitance between them and energy was being fed back from plate to grid as a result.

In detector or audio service it did not matter a great deal, but in RF stages, with both grid and plate circuits tuned to the one frequency, the feedback tended to cause oscillation. One low-gain triode RF stage was practicable (even if barely so) but two such stages were almost unmanageable.

A temporary answer to the problem was found in the so-called "Neutrodyne" principle, which enjoyed some popularity in the late 1920s. The primary winding of the RF coil was centre-tapped so that a signal voltage appeared at the lower end similar to but out-of-phase with the signal voltage at the plate end. A small variable capacitor was connected from the lower end of the primary to grid and adjusted to have the same value as the grid-plate capacitance of the valve.

Being connected in this fashion, this so-called neutralising capacitor fed back to the grid a signal equal to and out-of-phase with that fed back from the plate, so that the two cancelled out. As a result, the tendency to oscillation was overcome and two triode RF amplifier stages became practical.

The early "Neutrodyne" receivers were, therefore, a special type of TRF receiver, employing the principle of neutralisation.

In point of fact, neutralised TRF receivers did not enjoy a lengthy period of popularity because valve designers came to light with the screen-grid principle. Applied in RF tetrodes and pentode valves, it almost eliminated grid-plate capacitance, and therefore, eliminated the major source of instability in RF stages.

As a result, it became possible to achieve high figures of stage gain, and, further, to use two high-gain stages in sequence. Nor was there any great trouble with instability. By shielding the valves and coils and adopting a layout which kept input and output leads reasonably apart, such a set could remain completely stable, even under full gain conditions.

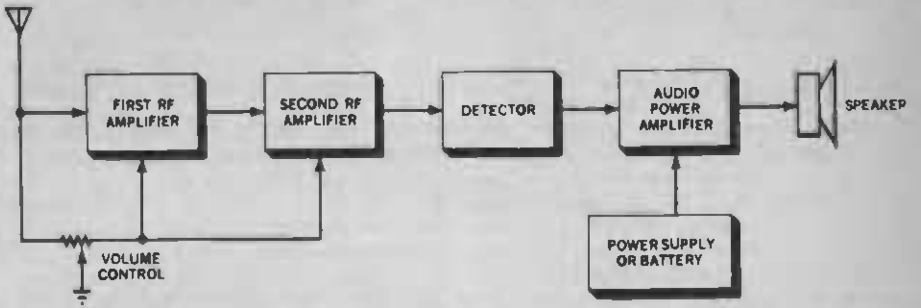


Fig 2: The basic arrangement used in early TRF radio receivers. Such receivers had their heyday in the 1930s.

With such gain available and the selectivity afforded by three tuned circuits, reaction became unnecessary, and the reaction control therefore largely disappeared from sets of the day.

By carefully matching tuning coils and adding small variable trimmer capacitors across each tuned circuit, designers were able to gang together the three tuning capacitors and operate them from a single tuning dial. This done, domestic receivers became really simple to operate for the first time — one dial to select the station and one knob to control the volume.

TRF receivers reached their heyday about 1930 and their general design followed the pattern shown in the block schematic diagram of Fig. 2.

The incoming signal was fed by the aerial coil to the first RF amplifier stage, then passed to the second RF amplifier stage and thence to the detector. This was followed by a single audio stage, but, using a sensitive pentode valve, which gave high amplification as well as ample power output to operate a loudspeaker.

Power to operate all these stages came, in mains receivers, anyway, from a power supply built on to the same chassis. This included a power transformer, a rectifier, a filter choke of some description and two or more filter capacitors. (More will be said about the operation of this section of the circuit in a later chapter.)

The volume control in such receivers usually took the form of a potentiometer connected at one end to the two RF amplifier cathodes, and at the other to the aerial terminal. The adjustable tapping went to earth. With the moving arm towards the cathode end, the RF amplifier valves operated with minimum bias and maximum gain, while the amount of resistance between aerial and earth was too high to make any real difference to its efficiency. Adjusting the control the other way applied high bias to the RF amplifier cathode and reduced the stage gain; at the same time it shunted the aerial to ground and therefore reduced the signal input.

It might be thought that the evolution of the TRF receiver would have largely halted receiver development in that it provided good gain and selectivity with plenty of acoustic output and simplicity of operation.

But it didn't.

About the same time, many new stations were coming on the air, crowding the broadcast band and ever increasing the demand for selectivity. The limitations of the simple TRF soon became apparent, particularly for the more difficult reception areas.

To add yet another RF stage or yet another tuning circuit gave only limited improvement at the cost of much greater complexity and with the attendant risk of instability. What was more there didn't seem to be any obvious way of making tuned circuits much more efficient. To return to reaction as an aid to selectivity was unthinkable to a commercial designer.

As a result, they began to look for other basic methods of receiver design and the one which seemed to hold the greatest promise was the superheterodyne principle.

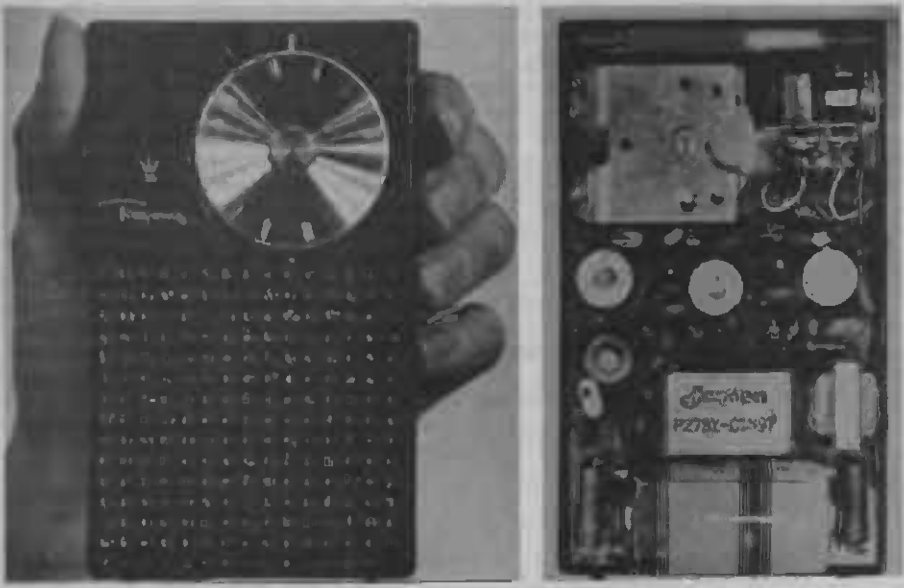
This was not new, having been developed by Edwin H. Armstrong of Columbia University in New York as early as 1921. It had been used in a limited way for many years, mainly in professional and military shortwave receivers. Could it be adapted for use in domestic radios? Designers soon found that it could.

Designed around the better valves available, and using modern circuit techniques, the superhet receiver quickly established itself in popular favour, and has remained undisputed leader ever since.

But how does the superhet work? At this point, we can drop the semi-historical sort of discussion and settle down to some straight theory. This is appropriate, because the superhet principle does not yet belong to history. Practically every modern broadcast, communications and television receiver uses the principle.

As the name suggests the superheterodyne receiver utilises a method of heterodyning or beating two signals together. Let's explain this.

It has been found that, when two signals are fed into a non-linear circuit, they combine to produce signal voltages at



Exterior and interior views of a "personal portable" transistor receiver. Receivers of this type use the superheterodyne principle, and have from six to eight transistors.

frequencies additional to and distinct from either of the original input frequencies. Further, that these new frequencies are equal to the sum and the difference of the original frequencies.

Consider, for example, two frequencies which we shall designate as f1 and f2. If fed into a non-linear amplifying stage, it would be possible to detect output voltages, as expected, having the original frequencies f1 and f2. But, in addition, we would find that output voltages were present at frequencies equal to f1-plus-f2, and f1-minus-f2 (assuming f1 to be the higher numerical value).

Taking an actual case, we may feed signal voltages at, say, 2,000 and 1,500 kilohertz into a non-linear stage. Both original signal frequencies would be present in the output, plus additional frequencies of 3,500 kHz (2,000 plus 1,500) and 500 (2,000 minus 1,500).

In actual fact, there may be other frequency components in the output, due to the presence or generation of harmonics, but we can afford to neglect these as being incidental to the main effect.

As we already know, stations on the broadcast band transmit on allotted frequencies, between the limits of 550 and 1600 kHz. To tune and amplify them on a TRF receiver involves the use of a ganged capacitor tuning two or three matched coils.

There are difficulties in the way of tuning more than about three coils in this way, so that the performance of a TRF receiver is largely limited by the selectivity and gain which can be achieved with three variable tuned circuits.

But, in the superheterodyne, the designers utilise the heterodyne principle to change the frequency of any and every desired incoming signal to a new, prearranged and fixed frequency. This is passed to and amplified in a section of the receiver, which can employ any desired number of fixed tuned circuits.

The new frequency is usually lower than the signal frequency but still well above the audio spectrum, which is perhaps the reason why it is commonly referred to as the "Intermediate" frequency — usually shortened to "IF."

The particular intermediate frequency is selected by the designer to suit his requirements. If high gain and extreme selectivity are the objectives, he may choose an intermediate frequency of about 200 kHz. But, with such a low frequency, great care has to be exercised to avoid receiving the same signal at two points on the dial — called "two-spotting" — owing to unwanted heterodyne effects.

An intermediate frequency of more like 2,000 kHz minimises double-spotting, but requires greater attention to the design of the tuned circuits, if gain and selectivity are not to be sacrificed.

A compromise figure, which is widely employed in this country, is an intermediate frequency of 455 kHz or thereabout.

Assume that a desired signal is on 1,000 kHz. The first obvious requirement then, is for the tuned aerial input circuit to be resonated to this figure. This is accomplished by tuning the aerial input coil with a variable capacitor, exactly as in an ordinary TRF receiver.

The desired 1,000 kHz signal is then fed into the "mixer" or "frequency changer." In the output, remember one desires a frequency equal to the selected intermediate frequency, which one may assume to be 455 kHz.

Essential for the frequency change is an oscillator, which delivers a locally generated signal voltage to the mixer stage. To obtain the desired result, the oscillator would be tuned to 1,455 kHz, above the incoming signal frequency by just 455 kHz.

It could alternatively be tuned to (1000-455) or 545 kHz, but the higher oscillator frequency is normally used.

At the output of the mixer stage, one would expect frequency components of 1,000, 1,455, 2,455 and 455 kHz. But the mixer invariably feeds directly into a tuned circuit, which would be resonated permanently at 455 kHz. This one frequency is, therefore, selected and passed on, while the first three mentioned above, together with all other incidental harmonic frequencies, are not amplified.

If the desired signal were on 1,020 kHz, then it would be necessary to increase the local oscillator frequency by another 20 kHz to ensure that the IF output remained at 455 kHz.

Thus, in a simple superhet, there are two variable tuned circuits. One gives initial selection at the signal frequency, and the other adjusts the local oscillator frequency to a figure which differs from the signal frequency by the selected intermediate frequency.

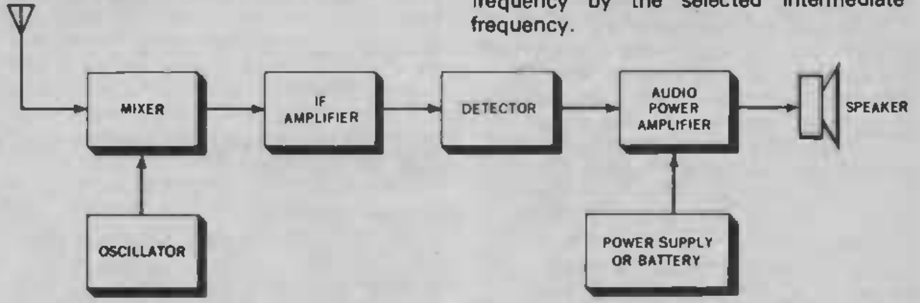


Fig 3: The basic arrangement of a single-conversion superheterodyne receiver. Most modern receivers are of this type.

In the earliest "superhets" the aerial and oscillator circuits were controlled by separate capacitors and tuning dials. But, in all modern sets, the coils are accurately adjusted and the oscillator tuned circuit arranged so that it maintains the required frequency difference automatically. This is called "tracking."

In most modern receivers this is arranged by means of a so-called "padderless gang", a two-gang variable capacitor in which one of the sections has smaller and specially-shaped plates. An alternative technique, generally used for short-wave receivers, is to use a conventional tuning capacitor with a so-called "padder" capacitor connected in series with the section used to tune the oscillator. The padder may be made variable to ensure that the oscillator frequency may be adjusted for optimum tracking.

The intermediate frequency generated from the received RF carrier retains the original modulation, so that it can be amplified and passed on to the detector in the usual way.

It is here that the advantage of the superheterodyne principle becomes evident. Since each selected signal is automatically transformed to a constant frequency (which we have assumed to be 455 kHz), the IF amplifier channel may be provided with any desired number of circuits, permanently tuned to the selected intermediate frequency.

Coupling coils between valves may have

both primary and secondary tuned, instead of the secondaries only, as in ordinary TRF practice. No variable tuning gang is necessary for this purpose, and the coils may be designed for compactness and efficiency, and thoroughly shielded for stability.

Intermediate frequency (IF) tuning circuits were frequently resonated in the past by means of small compression type mica trimmers, adjusted with a screwdriver. Alternative and common practice nowadays is to have a fixed mica or ceramic tuning capacitor and to vary the inductance of the coil by a small adjustable iron core.

Most ordinary superhets employ one stage of IF amplification, involving two IF transformers. Larger sets may use two IF amplifier valves or transistors with three IF transformers.

These tuned circuits in the IF channel are fully effective in discriminating against unwanted signals.

For example the desired signal may be on 1,000 kHz with an adjacent and interfering signal of 1,010. The single tuned circuit ahead of the mixer valve could not discriminate effectively against a signal only 10 kHz removed from the desired one, so that a substantial 1,010 kHz signal may reach the input of the mixer valve.

In the output of the latter, there would, therefore, be the desired heterodyne frequency of 455 kHz, plus another heterodyne produced by the unwanted carrier at 445 kHz. But, with four or more circuits to negotiate, all tuned to 455 kHz, the signal on 445 kHz would have little chance of reaching the detector at troublesome level.

Thus, even though the average domestic superhet uses only a two-gang tuning capacitor, there are generally something like four or five tuned circuits to discriminate against unwanted signals — as against two tuned circuits provided by a two-gang capacitor in the TRF arrangement.

Another reason for improved selectivity in the superhet is the basic fact that the frequency is changed to a lower value. As shown in the example above, the difference in frequency between the wanted and unwanted station — 10 kHz in this case — is retained when the frequency is changed. However, relative to 455 kHz, 10 kHz is a greater change, in terms of percentage, than the same change relative to 1000 kHz.

Thus, assuming tuned circuits of equivalent "Q", the one at 455 kHz will be better able to reject the unwanted signal than the one at 1000 kHz. It will also be understood why some circuits use even lower frequencies, 175 kHz and even 50 kHz being employed where very high selectivity is required.

The output from the IF amplifier stage ultimately feeds into a detector, which may be any one of several varieties. The output and power supply arrangements are exactly as for a TRF receiver.

Fig. 3 shows the sequence of stages in a typical superhet in block schematic form. The aerial input signal is fed to the mixer or frequency changing stage, where it is mixed with a signal generated by the in-built oscillator.

The resultant or intermediate frequency is then amplified in the IF stage and passed on to the detector, where the audio component is extracted. This is amplified in the audio stage and applied to the loudspeaker.

At first, the functions of mixer and oscillator were entirely separate, as indicated.

The mixer was normally operated under very high bias conditions, as employed for a detector. Hence, the mixer valve in these early superhets was commonly referred to as the "first detector." The normal detector for demodulation naturally gained the title of "second detector."

In the inevitable trend to simplification, it was found possible to obviate the separate oscillator valve, and the first detector was made simultaneously to fulfil the function of oscillator by connecting it to the oscillator tuned circuit.

This arrangement, employing generally the 57 or 6C6 pentode, was widely used around 1932. Known as the "autodyne" circuit, it proved quite efficient and adequate until the demand for dual-wave sets emphasised its non-suitability for such receivers.

Ultimately, the trend to superhet circuits, the popularity of dual-wave receivers and adoption of automatic volume control, led to the evolution of special valves for use as frequency changers. These varied a good deal in structure from one type to the next, but normally had a triode oscillator and a screen-grid mixer section within the one envelope.

With the advent of transistors, the autodyne principle has been revived with the first transistor acting as both local oscillator and mixer.

In fact, a modern "pocket portable" transistor radio illustrates very well how the superhet has been developed and simplified. Typical receivers of this type use only six transistors: one as an autodyne mixer, one as an IF amplifier, and the remaining four in the audio amplifier. The "second detector" function is generally performed by a germanium diode.

Receivers intended for specialised communications work invariably use the superhet principle, because of the high gain and selectivity which it offers. In many of these receivers more than one intermediate frequency is used, with a number of mixers used to change the frequency of the signals from one to the other. Thus one can have a "double-conversion" receiver, a "triple-conversion" receiver, and so on.

Many such receivers use special filter units in their IF amplifiers, to achieve either a very sharp selectivity response or a carefully adjusted wider response. There are a variety of such filter units available with names such as "crystal filter", "mechanical filter", and "ceramic bandpass filter".

The provision of these and other facilities at the fixed intermediate frequency is something which could not reasonably be duplicated in any TRF design.

At the same time, most high-performance superhet receivers do use at least one RF amplifier stage ahead of the frequency changer. An RF stage ahead of a superhet circuit makes a minor contribution to gain and selectivity and also helps exclude from the frequency changer strong signals at frequencies remote from the desired station. In special circumstances, such signals may cause spurious beats with harmonics of the local oscillator and penetrate the IF channel.

An RF stage also tends to have a lower inherent noise level than a mixer. By amplifying the incoming signal somewhat before its frequency is changed, a more favourable signal-to-noise ratio can be obtained.

Modern television receivers also use the superhet principle. The problem in this case is not to get extreme selectivity but a specific amount of selectivity — no more and no less. To meet this requirement in the variable tuned circuits of a TRF would be very difficult but in the IF channel of a superhet it can be provided without any special difficulty. Five or six tuned circuits are often used for this purpose.

Examples of more elaborate receivers. These are generally double or triple conversion superhets, with a variety of additional features.

CHAPTER 15

Power Supplies

Valve equipment power supplies — problems in using AC to heat valves — the indirectly heated valve — the full-wave rectifier, using valve diodes — smoothing and filtering with inductance and capacitance — the use of semiconductor diodes — the voltage-doubler rectifier — the half-wave rectifier — transistor equipment power supplies — the bridge rectifier — dynamic filtering and regulation.

For the sake of simplicity, most of our circuit discussions to date have assumed the provision of suitable DC supply voltages, without much emphasis on how such voltages are obtained. In this chapter, we explain how supply voltages are derived from the AC power mains.

In the early days of radio, receivers were invariably supplied from batteries. It was commonplace to use either an accumulator for the filament supply or a number of heavy-duty dry cells capable of supplying the requisite and often considerable filament current.

The grid bias voltages were taken from a special bias battery, not intended to deliver significant current, but with tappings at each cell junction to give voltages in 1½-volt steps to 4½ volts or 9 volts — to quote what were common figures.

For the plate supply, so-called radio "B-batteries" were used. These were large and rather expensive banks of dry cells, usually made up in 45-volt blocks and tapped at 22½ volts. Two such B-batteries in series could supply 90 volts, while three in series were commonly employed to give 135 volts. How cumbersome and expensive these batteries were tends to be forgotten in these days of transistor receivers.

While the early sets were simple enough from the designers' point of view, the need to provide, attach, and conserve batteries was a constant worry to radio set users and it was natural that efforts should be made to cut the operating costs, at least. As a result, various gadgets appeared aimed at supplementing or replacing the expensive batteries.

Numerous chargers or "trickle chargers" were put on the market for recharging the filament accumulators. The chargers might deliver currents up to 3-odd amperes and would top up a discharged battery in a day or so. Trickle chargers were designed to be left on more or less continuously, keeping the battery full at all times and saving the hitherto regular trip to the local garage for a battery re-charge.

So-called "B-Battery Eliminators" were released, to replace the high-tension batteries altogether. These incorporated a transformer, rectifier and filter system, rather like a modern AC power supply. Various resistors and tapping points were included so that they could supply the requisite intermediate voltages at the order of current drain commonly encountered in battery sets of the day.

Some B-battery eliminators also included auxiliary circuits to provide negative bias voltages, although the cost of a bias battery was never a major item.

These various units enjoyed a limited degree of popularity, but the obvious objection of having gadgets and accumulators attached to the family radio provided strong incentive to produce self-contained receivers which could simply be plugged into the power point and operated therefrom just like any other electrical appliance.

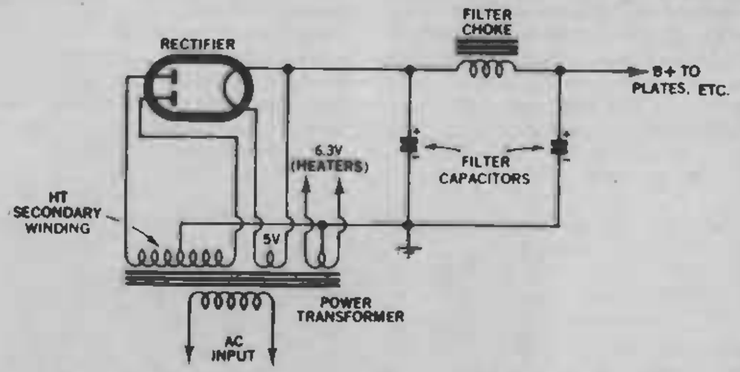


Fig 1: A typical power supply circuit using a valve rectifier, as used in older receivers and amplifiers. Supplies using semiconductor rectifiers are more common in modern receivers, being more compact and efficient.

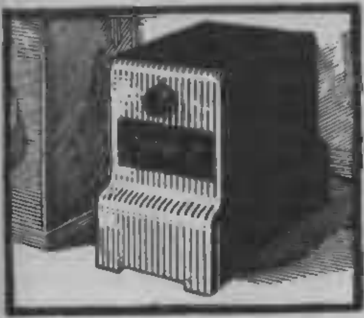
Initially, the main difficulty was that of providing filament supply. For reasons we shall see a little later, AC from the power mains could not readily be changed to DC at the voltage and current needed to operate a number of parallel-connected filaments. And there were — and still are — two basic objections to applying AC to the filament of a directly-heated valve.

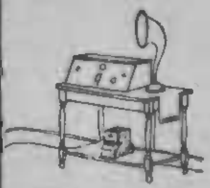
The first and perhaps most obvious objection is that because the filament has a certain voltage drop, the effective bias between filament and grid varies over its length. As a result, if the filament is heated by the application of AC, an alternating voltage component tends to be superimposed upon the desired DC bias to modulate the plate current and cause hum. It is possible to cancel out most of this superimposed component by accurately centre-tapping the filament AC supply, and this was done with early receivers designed to be operated directly from the mains. But unfortunately this technique does not avoid the second problem.

Because alternating current falls to zero twice in every cycle, the temperature of the filament tends to vary cyclically when AC is used to heat it. This causes a corresponding variation in the number of electrons produced, and thus still tends to modulate the plate current to produce an undesirable hum (in this case at twice the AC supply frequency). In an attempt to overcome this problem the directly heated valves used in early AC receivers had special low-voltage high-current filaments made from thick wire and therefore thermally sluggish. However, this was only partly successful.

A satisfactory solution to the problem only came with the introduction of valves having "indirectly heated" cathodes. Such valves were described in an earlier chapter.

With the development and release of valves having indirectly heated cathodes, the major problem with all-mains operation disappeared and numerous receivers were released using them. It was still necessary to produce from the mains a pure DC supply for the valve plates and screens, but, as we





1925 VERSION

This unit was advertised in 1925 under the name "Super-Ducon . . . Eliminates all B batteries". Produced by the Dubilier Condenser and Radio Corporation It sold for $47.50!

shall see, this was not — and is not — a major problem.

In most radio receivers, amplifiers and other equipment using valves, the DC supply for the plates and screens is provided by a power supply circuit using a transformer and rectifier. In early equipment the rectifier used was a valve, usually a double diode. In more modern equipment silicon diodes are used.

Fig. 1 shows a typical valve rectifier power supply circuit, whose operation we can proceed to discuss.

The heart of the supply is the power transformer, which is shown diagramatically as a number of windings adjacent to an iron core. The transformer is used to provide low voltage AC for the valve heaters as well as the plate supply.

The incoming power lead is connected across the primary winding, which will normally be rated to receive an input of 240 volts AC. It must be AC. A power transformer must not be connected across DC mains. If it is, it is almost certain to blow the fuses or burn itself up, or do both!

The reason for this is not hard to discover in that a transformer relies for its operation on a constantly changing magnetic field. As the alternating current from the power mains flows to and fro through the primary winding, it causes a strong magnetic field in the iron core to build up and collapse in cyclic fashion. The moving lines of force thus created induce current and voltage in the various secondary windings, obeying the laws of magnetism explained in an earlier chapter.

The alternating voltage developed across each secondary winding is almost exactly proportional to the ratio of turns between the primary and the secondary winding in question. Thus, if there are 1,200 turns on the primary winding, a secondary winding also having 1,200 turns would deliver the same 240 volts as fed into the primary — because the turns ratio would be 1:1. On the other hand, if a 6.3 volt winding is required to operate a number of valves with 6.3 volt heaters, then this heater winding would need to have 1,200 times 6.3 / 240, or approximately 32 turns.

In the above illustration we suggested 1,200 turns for a 240-volt winding on the assumption that the transformer might be wound on the basis of five turns for every volt of input or output. This is a likely enough figure, but, in practical transformers, the turns-per-volt figure may vary considerably from one type to another, according to the size of the core, the grade of the iron used and the ideas of the designer.

The thickness of the wire used on each winding depends on the current which it has to handle or deliver. In the case of a heater winding, which may be required to deliver several amperes, relatively thick wire has to be used and it is commonplace to see heater windings using 16-gauge enamelled wire or thicker.

It is important to realise that the gauge of wire used in a transformer winding determines only the amount of load current it can handle, without over-heating, if required to do so. Thus a winding rated to deliver, say, three amperes, can deliver up to three amperes without tending to overheat, according to the number of valves which may be connected to it. If only one valve were connected to the particular winding, the current drawn from it would probably be less than one amp.

Typical low power transformers designed for use in the power supply of a radio set, television receiver or similar equipment may have one, two or even three heater windings, to give the voltages and currents likely to be required. If designed in recent years for valve equipment, most heater windings are likely to be designed to produce a voltage of 6.3 volts RMS, to suit most modern valves.

In the circuit of Fig. 1 we have shown two heater (or filament) windings, one to supply the rectifier and the other to supply the heaters of all other valves in the receiver. The latter is shown as having a centre-tap connection, earthed to the chassis.

A typical modern transistorised regulated power supply suitable for use with experimental circuits. It provides an adjustable output voltage from 0 to 30V at a maximum current of 1A. The meters are included to indicate the output current and voltage.

Heater wiring is usually earthed for two reasons:

Firstly, the heater winding is very close, inside the transformer, to other windings producing high alternating voltages. Because there is some capacitance between them, some of the high voltage energy can be coupled capacitively to the heater winding and to the wiring connected to it.

This doesn't interfere in any way with the basic operation of the heater circuit but the high ripple voltage present on the heater wiring throughout the chassis can couple into grid circuits and produce an objectionable hum or buzz in the output.

A second reason is that wiring running from one stage to another throughout a high-gain receiver can transfer signals by stray coupling and produce troublesome regeneration.

Earthing the heater wiring largely obviates both effects. Although we have shown a centre-tap earth return, this is not strictly necessary except, perhaps, in equipment having very high audio gain. In many cases it is sufficient to earth one side only of the heater wiring.

For the plates and screens, AC from the power mains must be rectified and filtered till it becomes virtually pure DC. This involves, normally, a high tension secondary winding on the power transformer, a rectifier, a filter choke and two or more filter capacitors.

As might be expected, the high tension winding involves many turns of fairly fine wire, so that a considerable voltage is developed between its outer ends. Since the voltage across it is alternating, each end swings alternately positive and negative with respect to the other.

In valve rectifier circuits such as that shown, the high tension secondary winding has a centre-tap which is returned to chassis (shown as earth) so that half the total secondary voltage appears between earth and the respective ends. When one end of the winding swings positive with respect to earth the other end simultaneously swings negative by an equal amount.

As with the heater windings, the rating of the high tension secondary, in terms of voltage and current, varies with the size of transformer and the receiver which it is to supply. A small transformer, to supply a small mantel radio receiver, might typically have a HT secondary rating of 150 volts either side of the centre tapping, at a nominal current rating of 30 milliamps — this figure referring to the permissible DC load current.

A large transformer, intended to supply a television receiver, or amplifier, might have a voltage rating per side of up to 400 and a nominal DC load current of up to 250 or even 300 milliamps.

The two ends of the HT secondary winding are connected to the two plates of the rectifier valve, as depicted. This valve is virtually two diode elements in the one envelope, the plate and filament structure being expressly designed to carry a considerable amount of current.

A valve of this type, intended for use in a power supply and having two separate anodes or plates, was commonly referred to as a full-wave rectifier.

The filament of the rectifier is fed from a separate winding on the transformer, which is typically rated to deliver five volts at two or three amperes. It is quite usual for rectifier valves to consume considerable heater or filament power, the cathode or filament being designed to provide copious electron emission and thus allow the valve to pass heavy current without danger of early failure in service.

To follow the action of the rectifier, consider the instant when a positive voltage has appeared on the left-hand half of the HT secondary and therefore on the upper rectifier plate, as drawn.

Since the plate is positive, electrons will tend to flow to it from the heated filament. We can consider the result in a couple of ways, both of which lead to the same conclusion:

(1) In losing electrons, which are essentially negative charges, the filament of the rectifier must itself become positive.

(2) When conduction takes place through the rectifier, the impedance of the filament-to-plate path in the valve must decrease. The filament must, therefore, approach the plate potential, and, since this is temporarily positive, the filament must tend also to become positive.

Whichever way one cares to look at it, the result is the same — a positive potential on the plate and conduction through the valve produces a positive voltage at the filament.

When the same plate swings negative, during the next half-cycle, there is no conduction through the valve and, therefore, no tendency for the filament to develop a simultaneous negative potential.

On the contrary, as the first plate swings negative, the second plate simultaneously becomes positive and conduction takes place between the filament and this second plate. Once again, therefore, the filament tends to be carried positive.

In other words, during successive half cycles, when each plate in turn swings positive, current flow through one half of the rectifier or the other tends to carry the filament positive also. Since there are 100 half cycles per second with 50Hz power mains, 100 positive pulses are apparent at the rectifier filament per second.

The 50Hz alternating voltage at the rectifier plates is thus changed to pulsating DC at the rectifier filament, positive with respect to chassis and having a heavy ripple content of 100Hz.

This positive voltage is generated at the rectifier filament quite independently of the five-volts AC coming from the transformer winding to heat the filament. This latter voltage, applied across the rectifier filament, raises it to operating temperature. When the positive voltage is generated, it carries the filament as well as the transformer winding feeding it, to a high positive potential in respect to chassis.

Obviously enough, since the rectifier filament winding is expected to be at a positive potential with respect to chassis, it must not be earthed.

Instead a wire connected to one wire of the filament or its supply winding becomes the source of the positive potential which must ultimately be fed to the plates and screens of the remaining valves in the equipment.

A miniature mains supply for portable receivers and tape recorders.

However, the plates and screens must be fed with substantially pure DC, not a voltage which has a very high ripple content. To get rid of the ripple, it is necessary to use what is known as a filter system, which as shown in Fig. 1 may involve a filter choke or inductor and a number of filter capacitors.

The inductor consists normally of a large number of turns of wire wound within a laminated iron core, much like that used for small power or output transformers. It must be capable of carrying the amount of current involved in the particular supply and, with this current flowing, must have an inductance usually of several Henries.

In some older-type radio receivers a filter choke, as such, was not used. Instead the current from the power supply was passed through a winding around the pole piece of the dynamic speaker. This gave the requisite inductive effect for filtering, and the magnetic field created by the current served at the same time to energise the speaker's magnet system. The so-called "field" winding on the loud-speaker therefore served a double purpose.

It may be remembered that in an earlier chapter we learned that an inductor tends to resist any change in the amount of current flowing through it. If the current increases above average, part of the energy involved is diverted into creating a stronger magnetic field around the winding. If the current decreases, the magnetic field is reduced and returns some of its energy to the winding as current flow.

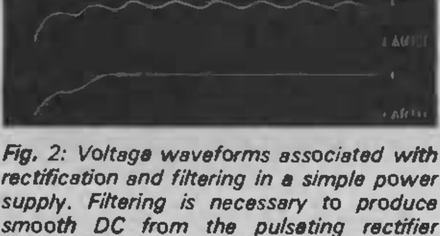


Fig. 2: Voltage waveforms associated with rectification and filtering in a simple power supply. Filtering is necessary to produce smooth DC from the pulsating rectifier output. The waveforms shown are: (a) the transformer secondary voltage, (b) the rectifier output, (c) effect of the choke, (d) effect of the capacitor, and (e) the output DC.

As a result of this action, current flowing through a filter choke loses a good deal of its ripple content and becomes more nearly pure DC. This is illustrated in figure 2, where (a) represents the transformer secondary AC voltage, (b) the basic rectifier output, and (c) the effect of the inductor.

As indicated earlier, capacitance is also involved in a filter system, its effect being more or less complementary to that of inductance. A capacitor tends to oppose any change in the potential or voltage across it. If the voltage rises above an average value, some of the energy involved is diverted into the capacitor as an extra charge. If the voltage then subsequently falls, the charge is released, tending to maintain the original potential.

When one or more capacitors is connected between the B-plus supply line and earth, as in Fig. 1, they naturally tend to oppose or absorb the change in potential due to ripple from the rectifier. They charge on "peaks" and release energy subsequently to fill the "troughs." Diagrammatically, the effect is as illustrated in Fig. 2d.

If properly designed, the combined effect of the choke and capacitors is to completely eliminate the ripple content for all practical purposes, and the output from the supply becomes virtually pure DC. (see Fig. 2e.)

The rectifier circuit of Fig. 1 is known as a condenser-input or capacitor-input filter, because the rectifier feeds directly into a capacitor. In the less common arrangement, where the rectifier feeds directly into an inductor, the filter is described as a choke input filter.

Filter capacitors normally need to have a large value of capacitance, certainly not less than eight microfarads each. To obtain this capacitance in small space and with adequate working voltage, not forgetting price either, they are invariably electrolytic types, as described in an earlier chapter.

The main point to remember about electrolytics is that they must be connected the right way round, with their positive terminal connected to the positive side of the circuit.

In recent years, much higher values of filter capacitance have become practical and, as a result, chokes having much lower inductances will suffice for the same degree of filtering. In point of fact, many small power supplies these days do not use a choke at all, relying only on large capacitors to give an adequate storage and filtering effect.

Readers may recall from the earlier chapter on semiconductors that a semiconductor diode behaves almost identically with a valve or thermionic diode. In view of this, it should not be very surprising to learn that semiconductor diodes can be used in rectifier circuits in place of diode valves.

In point of fact, they are somewhat better suited to this task than valves, as they require no heating power and also tend to conduct more easily during the part of the cycle when they are called upon to do so. They also have a longer life, and are more reliable in service. At present, their only disadvantage is that they tend to be vulnerable to damage from transient over-voltage "spikes" which are at times present on AC mains.

Two semiconductor diodes can be used in a full-wave rectifier circuit similar to that shown using a valve in Fig. 1, the only difference being that the diodes do not require a filament wiring on the transformer. They are simply wired with their cathode connections tied together as the output connection leading to the filter circuit and the load circuit, and each anode connecting to one end of the transformer HT secondary winding.

This type of rectifier circuit is not often used where semiconductor diodes are employed, however, because it requires the diodes to have a high peak inverse voltage rating. The peak inverse voltage is the reverse-bias voltage which appears across each diode when it is "off" and the other is conducting.

With the full-wave rectifier circuit, the reverse-bias impressed upon the diodes when they are non-conducting is actually 2.828 times the half-secondary RMS alternating voltage, and this can require the use of costly diodes having a very high peak inverse voltage rating.

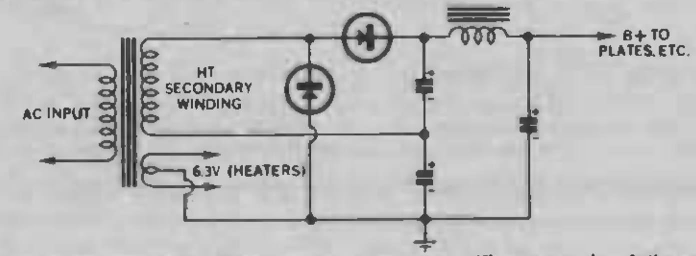


Fig. 3: Most power supplies using semiconductor rectifiers use the full-wave voltage-doubling circuit as shown here. It suits the rectifiers better, and is more compact and economical than the older full-wave system.

Because of this, it is often more desirable to employ what is called the full-wave voltage doubler rectifier circuit whenever moderate to high voltages and currents must be rectified by semiconductor diodes. Fig. 3 shows a circuit of this type.

A single untapped secondary HT winding is used on the power transformer, and the winding is arranged to produce an alternating voltage of only half (approx) the required DC output voltage. It should be noted, in passing, that this makes the power transformer somewhat simpler than in the full-wave circuit, and consequently somewhat less bulky and costly to produce.

Two semiconductor diodes are used as before, but this time they are connected in a different fashion. The first filter capacitor also undergoes a change, becoming two separate units which fill a more complex role than did the single unit of Fig. 1.

Neither end of the transformer HT secondary winding is earthed. Instead, one end goes to the junction of the two series-connected filter capacitors, while the other end goes to the two diodes. One diode has its cathode connecting to the winding and its anode earthed, while the other has its anode connecting to the winding and its cathode connecting to the top of the uppermost filter capacitor and the DC output circuit.

The operation is as follows: For the half-cycles when the top of the transformer winding is negative and the bottom positive, the "series" (upper) diode is reverse-biased and non-conductive. The "shunt" (lower) diode is forward-biased, however, being connected to the winding via the lower filter capacitor.

It therefore conducts, and in doing so it charges the lower capacitor to the peak value of the alternating voltage appearing across the winding. The capacitor voltage is as shown, with its earthed end negative with respect to the top end.

During the other half-cycle of the AC wave, when the top of the transformer winding is positive with respect to the bottom, the "shunt" diode is reverse-biased and non-conductive, while the "series" diode conducts. This time the upper capacitor is charged to the peak value of the AC secondary voltage, as it completes the circuit back to the lower end of the winding. The voltage across it has a polarity as shown.

As may be seen, the two capacitor voltages add together, and the total pulsating DC voltage available for filtering is twice the peak value of the transformer winding RMS voltage. Under load this voltage drops toward twice the RMS voltage.

The most important thing to realise about the voltage doubler circuit is that for a given and required DC output voltage, each diode has to deal with only half the voltage it would meet in a conventional full-wave circuit. Thus the doubler circuit allows the use of relatively inexpensive semiconductor diodes having but a moderate peak inverse voltage rating.

The voltage doubler circuit found almost universal acceptance in valve television receiver power supplies and in many other places where high current is required at a fairly high voltage. Silicon diodes are used almost universally in this circuit, as they are most easily arranged to have the required peak inverse voltage and forward conduction current ratings.

In power supplies where the voltage and current demands are very slight, it is possible to use a single diode valve or a single semiconductor diode in what is called a half-wave rectifier circuit. Such a circuit using a semiconductor diode is illustrated in Fig. 4.

A single untapped transformer secondary winding is used as with the doubler, but this time it needs to provide an RMS voltage approximately equal to the required DC output voltage. One end of the winding is earthed and the other connects to the first filter capacitor via the diode.

The circuit is in effect half the full-wave circuit, and the diode only conducts on every alternate half-cycle when the top of the HT secondary is positive. The half-cycles when the winding voltage is reversed are not used.

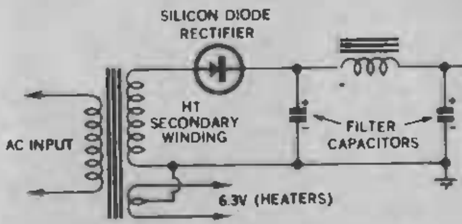


Fig. 4: When only very low current drain is involved, a half-wave rectifier system may be employed. A valve rectifier could be substituted for the semiconductor.

The half-wave circuit is thus rather inefficient, as it only uses half the energy available from the transformer. It is as a result only suitable for low current rectification and, as the diode has to have a peak inverse voltage rating of approximately 2.828 times the DC output (which is approx equal to the RMS voltage of the HT secondary) it is really only practical for low voltages as well.

The half-wave rectifier circuit delivers only one pulse of DC for each AC input cycle, so that its DC output ripple frequency is 50Hz. This makes filtering somewhat more difficult compared to the 100Hz ripple produced by the full-wave and doubler circuits.

So far in this chapter we have thought mainly in terms of power supplies required for the operation of valve receivers and equipment from the mains. Let us now look at the type of power supply required to operate transistor equipment from the mains.

As we saw in an earlier chapter, transistors are relatively low-voltage devices compared with valves. They typically operate with supply voltages of from 3 to about 80 volts, whereas valves normally use somewhat higher voltages.

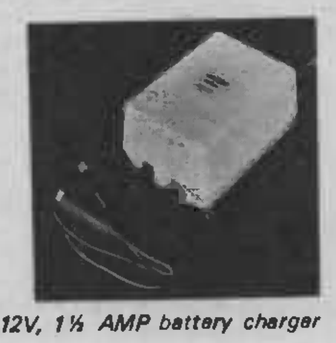
Where transistor circuits are required to deliver appreciable amounts of power — for example, in the case of transistorised audio amplifiers — they must accordingly be supplied with higher currents than valve circuits of equivalent performance. This is simply because to deliver power, they must be supplied with power, and power is effectively the voltage multiplied by the current.

Fig. 5 shows a fairly typical type of transistor power supply. The power transformer has only one secondary winding, an untapped low voltage winding. This is connected to a so-called bridge rectifier circuit, using four silicon diodes or a

Step-down transformer

FERGUSON Transformers Pty Ltd

MANUFACTURERS OF:– *TRANSFORMERS *FLUORESCENT LIGHTING BALLASTS *MERCURY VAPOUR CONTROL EQUIPMENT *SODIUM VAPOUR CONTROL EQUIPMENT *NEON TRANSFORMERS *OUTPUT TRANSFORMERS *TRANSISTOR OUTPUT *TRANSISTOR DRIVER *MATCHING TRANSFORMERS *GENERAL PURPOSE *FILAMENT TRANSFORMERS *LOW VOLTAGE TRANSFORMERS *SPECIAL PURPOSE *FILTER CHOKES *STEP-DOWN TRANSFORMERS



12V, 1½ AMP battery charger

A.C. / D.C. variable power supply used by N.S.W. Dept of Education in school science laboratores

Ferguson products are available from electrical and radio wholesalers and HEAD OFFICE — 331 High St., Chatswood. Phone: 40 0261

AGENTS

TAS.
Associated
Agencies Pty Ltd
25 Barrack St.
Hobart, 7000.
Phone: 2 1843

A.C.T.
Errol Nazer
Pty Ltd
84 Barrier St.
Fyshwick, 2600
Phone: 9 0325

QLD.
Keith Perty & Co
Pty Ltd
Waterloo St.
Newstead, 4006.
Phone: 51 5461

W.A.
H. J. McQuillan
Pty Ltd
107 Wellington St.
West Perth, 6005.
Phone: 21 4821

S.A.
K. Farmer Sales
Pty Ltd
49a Whitmore Sq.
Adelaide, 5000.
Phone: 51 4488-9

VIC.
Ferguson
Transformers
Pty Ltd
181-183 Hawke St.
West Melbourne

electronic kits for beginners

Popular imported electronic kits, no soldering, easy to assemble, battery operated, safe, suit all ages — children & adults, board type construction with easy to follow instructions that make them ideal gifts.

CRYSTAL RADIO KIT No. 28207, tunes AM broadcast band, simple 1 hour construction, no batteries, ideal for beginner, $4.25, post and pack 50c
AM TUNER AMPLIFIER KIT No. 28241, build your own 3 transistor tuner and amplifier, all parts transistors, tuning gang, transformers, speaker etc. $12.50, p.p. 75c
10 PROJECT ELECTRONIC KIT, No. 28202, 10 working projects, **SOLAR BATTERY,** builds radios, oscillators, signal generators, all solid state. $7.50, p.p. 75c
15 PROJECT ELECTRONIC KIT No. 1544, learn electronics with each project. Build these, morse code oscillator, radios, alarms, sirens etc. $9.90, p.p. 85c
IC-20 20 PROJECT ELECTRONIC KIT, learn about integrated circuits with this educational kit, 20 working projects including integrated circuit. $11.90, p.p. 95c
50 PROJECT KIT No. 28201 DELUXE MODEL. 50 working projects, educational entertaining, all solid state, includes everything, nothing to buy, constructed in hardwood case, panel meter, radios, amplifiers, burglar alarms, tachometer, test equipment, good value — $19.50, p.p. $1.20
DELUXE 150 ELECTRONIC PROJECT KIT using integrated circuits. Contains all parts for 150 different working projects including I.C. diode & transistor radio, electronic switches, relays, alarms, test equipment, etc, etc. Very good value, Prices $30.95, p.p. 95c
MODEL C-100 POCKET MULTIMETER, 1000 ohms / per Volt. AC Volts: 0-10 / 50 / 250 / 1000 (1000 opv). OC Volts: 0-10 / - 50 / 250 / 1000 (1000 opv). DC Current: 0-100mA. Resistance: 0-150K ohms (3K centre), 2 colour scale. Range Selector Switch. Dimens.: 3½ x 2¼ x 1 in. — $6.75 post free.
MODEL 200H MULTIMETER, 20,000 opv. DC Volts: 0-5 / 25 / 50 / 250 / 500 / 2500V (20,000 opv). AC Volts: 0-15 / 50 / - 100 / 500 / 1000V (10,000 opv) DC / Amps: 50uA / 2; 5mA / 250mA. Resistance: 0-60K / 6M ohm (scale centre 300 ohm — 30K ohm. Capacitance: 10uf to .001uF — .1uF, D3 scale 20 db to plus 22 db. Size 4½ x 3¼ x 1 1-8. —$11.75, P.P. 40c.
MODEL CT330 MULTIMETER, 20,000 / OPV, DC Volts. 0-6 / 6 / 30 / 120 / 600 / 1.2K / 3K / 6K Volts. AC Volts: 0 / 6 / 30 / - 120 / 600 / 1.2K Volts (10K / OPV). DC / Amps: (0.06mA / 60mA / 600mA. **RESISTANCE:** 0-6K / 600K / 6M / 60M / 600 Megohm. (30 / 3K / 30K / 300K ohms) centre scale: Capacitance: 50 uf to .01 uf .001 to 0.2 uf. Decibels: - 20 to plus 63 db size, approx. 5½ x 3⅝ x 1⅜ —$18.75, P.P. 40c.

HAM RADIO SUPPLIERS

323 ELIZABETH ST., MELBOURNE, VIC., 3000. Phone: 67-4286
2 DOORS FROM LITTLE LONSDALE STREET

ESTABLISHED 1947

selenium "stack" (as used in battery charger rectifiers), and thence to a very high value filter capacitor C1 and a further regulator and filter circuit using a transistor.

The bridge rectifier is full-wave, in that it operates on both half-cycles of the AC wave. It differs from the full-wave circuit of Fig 1 in that it does not require a tapped supply winding, and it differs from the doubler in that it does not supply a DC output voltage twice that of the RMS input voltage.

In the bridge circuit, two diodes conduct during each half-cycle. When the top end of the transformer winding is positive, diodes D1 and D3 conduct, and when the lower end of the winding is positive diodes D2 and D4 conduct.

The peak reverse voltage across the diodes when they are non-conductive is 1.414 times the RMS supply voltage and (approx) the DC output voltage, so that the bridge circuit is midway between the full-wave and doubler circuits in its demands upon the diodes regarding their peak inverse voltage rating.

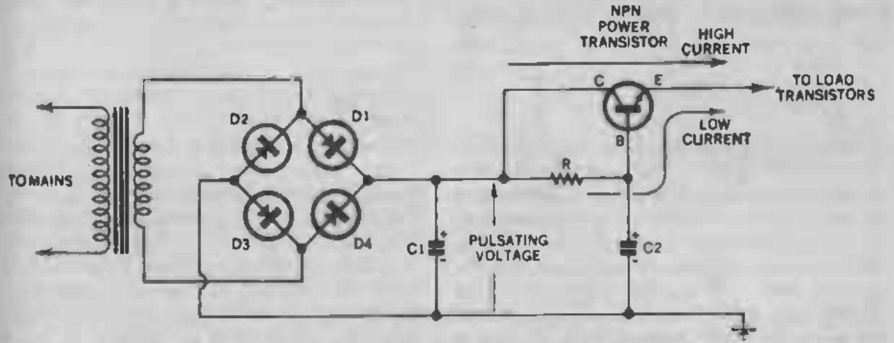


Fig 5: Transistorised equipment normally needs a much lower supply voltage than valves but at a higher current. This typical transistor supply shows a bridge rectifier system, a high value filter capacitor, and a transistor dynamic filter circuit.

The low-voltage, high current requirement of transistor power supplies makes filtering the AC ripple from the DC output a difficult task. A very large first filter capacitor is required (some supplies use 10,000 uF or higher), and as we have shown a transistor filter circuit must often be used for additional filtering. To maintain the output voltage constant under load it may also be necessary to add further circuitry for regulating the output.

The transistor is used to give what we might think of as "amplified" smoothing of the power supply output. Its operation relies upon the fact that the bipolar transistor is a current amplifier. It is capable of controlling large currents when supplied with small input or "bias" currents.

The general principle of transistor filtering and regulation is that the transistor is made to control the relatively large current drawn by the load circuit by supplying its control electrode — the base — with a smoothed and / or regulated source of bias current. As this reference source is required to supply only the small control current of the transistor, it is a relatively easy matter to provide it with filtering and regulation.

As may be seen from Fig. 5, the transistor (here an NPN type) has its collector connected to the pulsating DC output of the rectifier. Its base is supplied with smoothed bias current by means of the filter circuit formed by resistor R and capacitor C2, while the emitter becomes the output electrode and connects to the load transistors which must be supplied with power.

The simple resistor-capacitor filter circuit used to supply the base bias for the transistor is sufficient to provide adequate smoothing, because the base current required is relatively small. However the fact that the transistor is fed with well-smoothed base current means that its collector-emitter current — which is an amplified version of the base current — also tends to be well smoothed. Hence the relatively high current fed to the load transistors is smoothed, and the effective output voltage produced at the emitter of the filter transistor is also smoothed.

This type of transistor filter circuit is often called a "dynamic" filter, because the filtering is achieved by the transistor effectively varying its instantaneous resistance to compensate for the pulsations at the rectifier output. Because the load voltage tends to duplicate the reference voltage at the base of the transistor, a transistor connected in this way is also said to be connected as an "emitter follower".

The feature of the emitter follower mode of connection which is of particular importance from the viewpoint of dynamic filtering and regulation is that the load voltage is more or less independent of the transistor collector voltage. As long as there is sufficient collector supply voltage to supply the requirements of the transistor and load, any pulsations or variations present in the collector supply voltage tend to have little if any effect upon the load current and voltage.

Often the action of a dynamic filter is pictured by considering the transistor to have "amplified" the filter capacitor C2 to a value given by the product of C2 and the transistor current gain. Thus, it is said to act as a "filter capacitance multiplier."

For example, if C2 has a value of 500uF and the transistor has a gain of 100, the effective filtering is considered to be equivalent to a capacitor of 50,000uF shunted directly across the load.

While this comparison is fairly accurate as far as the filtering is concerned, it is not accurate as far as the source impedance seen by the load is concerned. This point is a little too involved for our purposes at present, but it should be remembered that the concept of "capacitance multiplication" is rather limited in its application.

As mentioned earlier, a transistor connected like that in Fig. 5 (as an emitter follower, in other words) can also be used to "regulate" the output of a power supply. This means that it can be arranged to keep the supply voltage substantially constant at the correct value, despite changes in the current drawn.

As you might have already guessed, this is done by holding the voltage supplied to the base of the transistor constant, so that the transistor has no choice but to maintain substantially the same voltage at its emitter. Usually the base voltage of the transistor is held constant by using a circuit with a so-called "Zener diode", which is a special sort of semiconductor diode made to be operated in the reverse breakdown condition. The operation of the circuit depends upon the fact that the voltage drop of such a zener diode remains effectively constant for a wide range in currents.

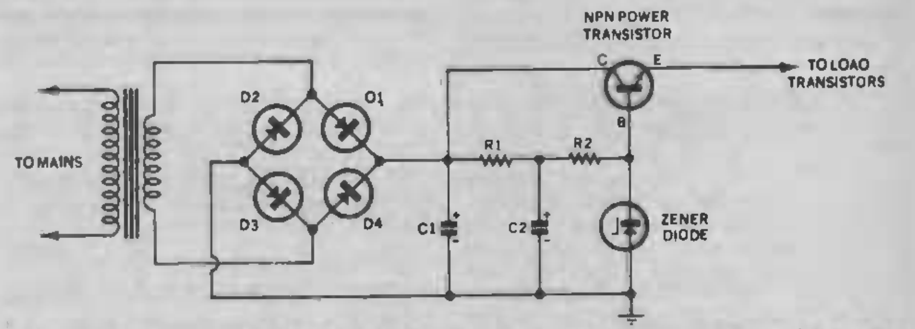


Fig 6: This transistor supply is similar to that of Fig. 5, but the transistor circuit performs voltage regulation as well as dynamic filtering.

Fig. 6 shows the circuit of a very simple regulated power supply using a zener diode. Basically the supply is identical to that of Fig. 5, but the resistor in the base circuit is now divided into two, with capacitor C2 now connected between their junction and the negative line (which is earthed). The zener diode is connected between the transistor base and negative, holding the base above ground by the diode breakdown voltage.

In closing the discussion of power supplies, filtering and voltage regulation, it should be mentioned that, although the principles of dynamic filtering and regulation have been explained by reference to transistors, the same principles apply to valves. Dynamic filtering is not often employed in valve circuits — principally because it is fairly easy to achieve adequate filtering using normal inductor-capacitor filters — but valve-type voltage regulators are quite often used in test equipment and other equipment requiring well-regulated supply voltages.

CHAPTER 16

More Basic Concepts

Basic circuit configurations for active devices — common cathode, emitter, source — common plate, collector, drain — common grid, base, gate — biasing — thermionic valve biasing — fixed and back bias — cathode or self bias — bipolar transistor biasing — simple biasing — divider biasing — divider biasing with an emitter resistor — FET biasing — self bias — modified bias — power supply decoupling — feedback, positive and negative — the effects of feedback.

In this chapter we will look a little more closely at some of the general circuit concepts which have been touched upon in earlier chapters. This should enable the reader to approach unfamiliar electronic circuits with greater confidence, and provide a more solid foundation for further study.

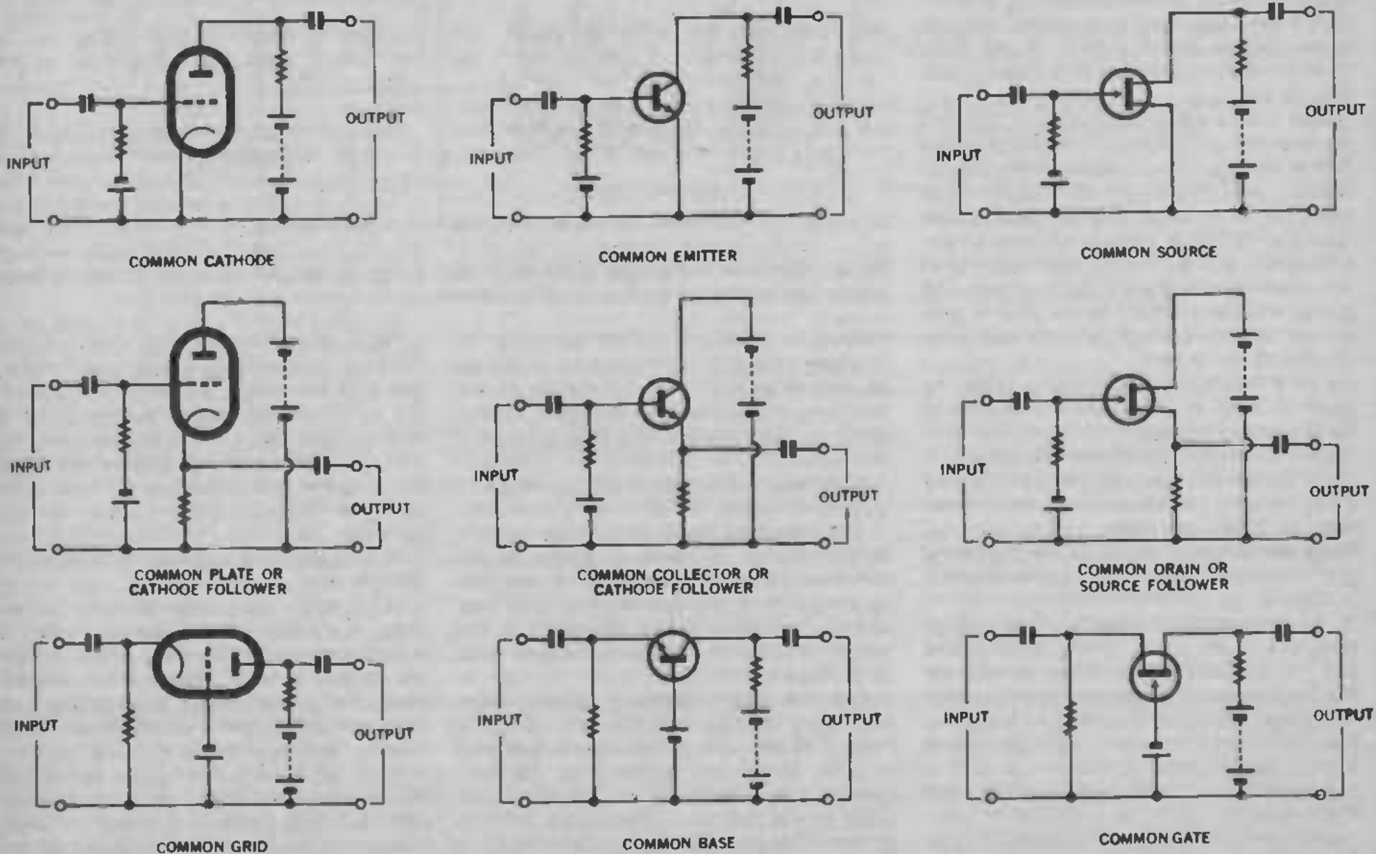
Let us look first at the various ways in which amplifying or "active" circuit devices such as the thermionic valve, the bipolar transistor and the FET are connected into circuits.

As we have seen in earlier chapters, active devices generally have three main terminals where AC signals are concerned: the cathode, grid and plate for a valve; the emitter, base and collector for a transistor; and the source, gate and drain for a FET. It may have already become apparent that the devices are not always connected into circuit with these terminals playing exactly the same roles with respect to the input and output of signals.

In fact, there are three basic ways or "configurations" in which each of the active devices may be connected into circuit, and each configuration tends to make the device concerned behave in a different way.

The three basic circuit configurations used for the valve, the bipolar transistor and the FET are shown in elementary form in the diagrams of Fig. 1. Note that although for simplicity the diagrams show a triode valve, an NPN transistor and an N-channel FET, virtually the same configurations are used for other devices.

The three similar configurations shown at the top of the figure are those which are most commonly used for general-purpose amplification. Here the cathode, emitter or source is used at the terminal of the device



COMMON CATHODE

COMMON EMITTER

COMMON SOURCE

COMMON PLATE OR CATHODE FOLLOWER

COMMON COLLECTOR OR CATHODE FOLLOWER

COMMON DRAIN OR SOURCE FOLLOWER

COMMON GRID

COMMON BASE

COMMON GATE

Fig. 1. The three basic configurations in which a valve, bipolar transistor and a FET can be connected into a circuit. Those in the top row are the most common, but the configuration in the centre row would normally be preferred where the stage is required to exhibit a low output impedance.

common to both the input and output circuits. Hence the terms "common cathode", "common emitter" and "common source".

Broadly speaking, when used in these configurations the devices give the greatest voltage and current gain, and hence the highest power gain. However, the gain may tend to fall away significantly at higher signal frequencies, because it is possible for some of the output to be fed back to the input through the internal grid-plate, base-collector or gate-drain capacitances of the devices to cause partial cancellation. Generally, the input resistance of these configurations tends to be moderately high, and the output resistance also fairly high.

Look now at the centre three configurations of Fig. 1. Here the devices are connected with their plate, collector or drain as the terminal common to the input and output circuits (note that the supply battery or power supply connected in series with these terminals is effectively a short-circuit for AC signals). Hence the terms "common plate", "common collector" or "common drain".

In these configurations, the devices still provide current gain, but in contrast with the previous configurations the voltage gain is slightly less than one. In other words, there is a slight loss, the output signal voltage usually being about 0.9 times the input. This occurs because in these configurations the output signal is effectively in series with the input signal as far as the real input terminals of the device are concerned. The output signal therefore effectively tends to cancel out the input, so that it must always be slightly less than the input for the device to operate.

In effect, the output signal from these configurations tends to almost duplicate or "follow" the input in terms of voltage. For this reason the configurations are often given the names "cathode follower", "emitter follower" and "source follower".

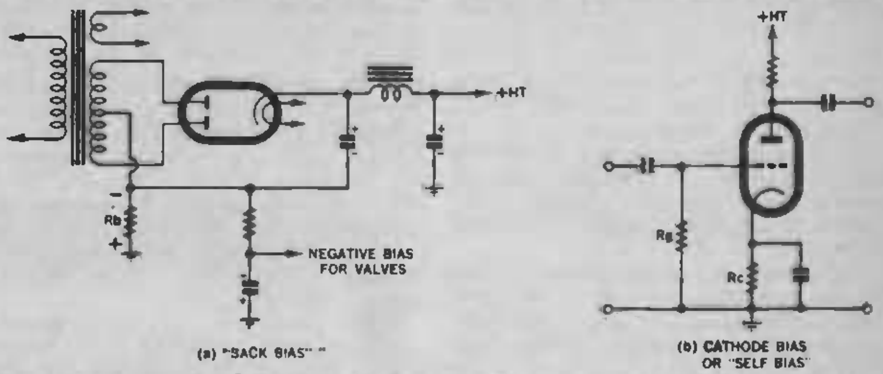
It may seem that because the output voltage only "follows" the input, these configurations would be of little value. However, they still provide current amplification, so that they can be useful wherever it is necessary to feed a signal into a low resistance load. Because they also tend to exhibit the highest input resistance possible with the device used, they are also useful in places where a signal must be taken from a circuit without significantly loading that circuit.

The last three configurations shown in Fig. 1 are those where the grid, base or gate of the device is used as the terminal common to both input and output circuits. These are naturally known as the "common grid", "common base" and "common gate" configurations.

The configurations again provide voltage gain, in fact slightly more than that of the first three configurations. This is because the input signal is effectively in series with the output, and tends to add to it. However, the current gain is now quite low, being approximately unity. In other words, the output current is virtually identical with the input current.

The main advantage of these last configurations is that because the grid, base or gate is effectively connected to the "cold" side of both input and output, it is able to act as a shield between the input and output terminals. This prevents output energy being fed back through the internal capacitance of the device, and hence enables the device to amplify efficiently at high frequencies.

Unfortunately, the configurations also tend to give the devices a very low effective input resistance, so that they are really only suitable for circuits where the input signal can be provided from a very low resistance source. In general, this means that they are used mainly for power amplification at radio frequencies.



(a) "BACK BIAS"

(b) CATHODE BIAS OR "SELF BIAS"

Fig 2. Producing the bias to make the grid of a valve negative with respect to the cathode. Self bias (b) is currently the more favoured method.

Having looked at the general ways in which active devices may be connected into circuits, let us now briefly consider the subject of biasing. As you have no doubt gathered from earlier chapters, a valve or FET must be provided with a suitable direct voltage between its grid and cathode or gate and source, in order to establish the correct operating current and conditions for amplification. Similarly, a bipolar transistor must be provided with the appropriate direct base-emitter current. The circuit components or network used to provide this voltage or current is known as the "biasing" circuit.

In order to bias a thermionic valve correctly, its grid must be made negative with respect to the cathode. With early battery-operated equipment, this negative bias was provided by means of a small auxiliary battery called the "C-battery". However, when equipment began to be operated from the mains, it naturally became desirable to derive the necessary bias voltage from the power supply used to produce the heater and high tension voltages.

One way of producing the required negative voltage is to have a separate winding on the power transformer, with a rectifier valve or semiconductor diode together with a filter circuit. This was occasionally done, but it tended to be rather costly, and designers understandably sought to achieve the same results in a simpler way.

One of the methods which they developed is shown in Fig. 2(a). Here a resistor Rb was connected in series with the negative output of the basic transformer-rectifier circuit used for the high tension plate supply. In this position, all of the current flowing from the rectifier circuit through the cathode-plate circuits and other wiring connected to the high tension supply must ultimately flow back through the resistor, and in so doing the current sets up a voltage drop across it. This voltage drop has the polarity shown, ie, the top of the resistor becomes negative with respect to chassis.

The voltage produced across the resistor tends to vary due to variations in the plate current of the various valves, and it also tends to have a certain amount of superimposed AC ripple. However, by using a suitable filter circuit such as the series resistor-shunt capacitor type shown, it could be smoothed and used for grid biasing.

This type of biasing was known as "back bias" for fairly obvious reasons, and the resistor Rb was known as the back bias resistor.

Note that the negative voltage produced across the back bias resistor is really part of the output of the high tension rectifier circuit. In other words, back biasing produced a small negative voltage at the expense of a corresponding drop in the effective HT voltage available for the valve plates.

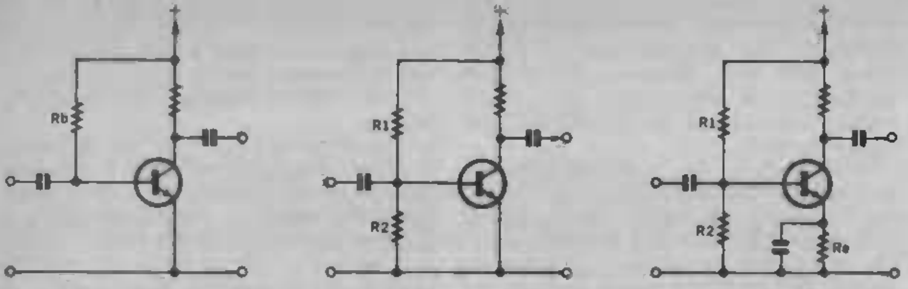
Another method used for valve biasing is shown in Fig. 2(b). Here each valve is used to generate its own bias, by means of a small resistor Rc connected between the cathode and chassis. The valve's own plate current flowing through the resistor sets up a voltage drop as shown, with the cathode positive with respect to chassis.

By connecting the grid to chassis via a high value resistor Rg, this voltage is effectively used to provide grid bias. It does not really matter that the grid is at "ground" or zero voltage, while the cathode is positive; as far as the valve is concerned, the voltage drop set up across Rc is effectively connected between cathode and grid, with the grid negative with respect to the cathode.

This method is known as "cathode bias" or "self-bias", and Rc is usually referred to as the cathode bias resistor. In designing the stage, the value of Rc is chosen so that it will produce the right value of bias voltage when the valve is drawing the desired amount of plate current (and screen grid current, in the case of a tetrode or pentode).

Note that just as with back biasing, cathode biasing effectively provides the grid bias for the valve at the expense of plate-cathode voltage. However, in this case each biasing circuit affects only the plate-cathode voltage supplied to its own stage.

A capacitor is normally connected across the cathode bias resistor, as indicated in Fig.



(a) SIMPLE BIAS (b) VOLTAGE DIVIDER BIAS (c) VOLTAGE DIVIDER WITH EMITTER RESISTOR

Fig. 3. Biasing a bipolar transistor with a small forward base-emitter current. As explained in the text, the simple system (a) has definite shortcomings and there are good reasons for preferring the more complex systems (b) and (c).

2(b). This is to ensure that the cathode is still connected to the chassis or "common" side of the circuit as far as AC signals are concerned. If this is not done, a varying AC voltage drop is also produced across Rc, which effectively tends to cancel out the input AC signals as far as the valve is concerned. The gain of the stage as an amplifier is accordingly much lower than if the resistor is shunted or "bypassed" with a capacitor.

Cathode biasing is generally the method used in modern valve circuitry, as it is usually more convenient. However, back biasing and other forms of fixed biasing are still occasionally used.

With bipolar transistors, biasing involves feeding a small current through the base-emitter junction in the forward direction. This can be done in a variety of ways, one of which is to use a small separate battery or supply as shown in the diagrams of Fig. 1. However, this is often not very convenient, so that, as with valves, the designer generally tries to provide the bias from the same source used to supply the collector-emitter voltage.

Happily, the collector-emitter supply voltage is of the correct polarity to supply forward bias current for the base, and no special arrangements need be made. Thus in the simplest case, bias can be provided simply by connecting a suitable high value resistor between the base and the collector supply, as shown by Rb in Fig. 3(a).

While simple, this method of biasing has two distinct disadvantages. The first arises because there is considerable variation in current gain among transistors, even for those of a given type. This means that the fixed base current produced by a simple series resistor tends to overbias some devices, and underbias others. The second problem is that the high resistance in series between the transistor base and the rest of the external circuit (particularly the emitter) tends to make the transistor's behaviour very dependent upon temperature. This is because the temperature-dependent collector-base leakage current is able to significantly affect the biasing.

For this reason it is generally preferable to supply the base with current via a resistive voltage divider connected across the collector-emitter supply, as shown in Fig. 3(b). The lower resistance source of current provided by the voltage divider tends to make the bias current into the transistor base adjust itself to suit the requirements of transistors having different current gains, and also tends to stabilise the transistor behaviour against temperature changes.

Even better results may be obtained by adding a small resistor in series with the emitter, as shown in Fig. 3(c). The voltage divider must be altered when this is done, as the voltage drop across the emitter resistor due to the collector-emitter current tends to oppose the forward bias on the base. Resistors R1 and R2 must therefore be adjusted to increase the effective base supply voltage.

Because the voltage drop across the emitter resistor is directly proportional to the collector-emitter current, the effect of the resistor is to adjust the bias automatically to suit the requirements of transistors with different gains. With a high gain transistor, the increased voltage drop across Re tends to reduce the bias to a suitably small value, while with a low gain transistor the low voltage drop tends to increase the bias to the higher value needed for correct operation. In the same way, the effect of the resistor is to give improved stabilisation against temperature changes.

As before, the emitter resistor is normally bypassed with a high value capacitor to ensure that the emitter is effectively connected to the chassis for AC signals. If this is not done, the gain of the stage as an amplifier is significantly reduced due to partial cancellation of the input, as with the valve circuit of Fig. 2(b).

Like the grid of a thermionic valve, the gate of a FET must normally be reverse-biased with respect to the source, for correct operation. As before, this reverse bias voltage may be provided by a separate battery or supply; it could even be provided using a back-bias scheme like that shown in Fig. 2(a). However, these methods are not very convenient, and are in fact rarely used these days.

The method most generally used is self-bias, as shown in Fig. 4(a). This works in exactly the same way as the cathode bias scheme of Fig. 2(b), but in this case the voltage drop set up across the series source resistor Rs is due to the flow of drain-source current. As before, the resistor is bypassed to ensure that the source is tied to the chassis common as far as AC signals are concerned.

Even more so than with bipolar transistors, the gain of FETs tends to vary quite significantly among those of nominally the same type. Because of this, it is sometimes necessary to increase the value of the series source resistor Rs beyond the value required for the correct bias voltage, in order to increase the ability of the bias circuit to adjust itself automatically for different devices. When this is done a forward bias voltage must be applied to the gate to restore the bias to its correct value. This is done by using a voltage divider across the drain-source supply, as shown in Fig. 4(b).

Although the output load resistors and biasing resistors in Figs. 2(b), 3 and 4 are simply shown as connecting to a common power supply, there are two points which should be explained about these connections.

Almost invariably, modern high fidelity amplifiers make extensive use of negative feedback to achieve wide, level frequency response and low distortion.

As you can see from the diagrams of Fig. 1, the power supply used to provide the plate-cathode, collector-emitter or drain-source voltage actually completes the output circuit of each device. Because of this it is generally essential that the supply represents an effective short-circuit as far as AC signals are concerned.

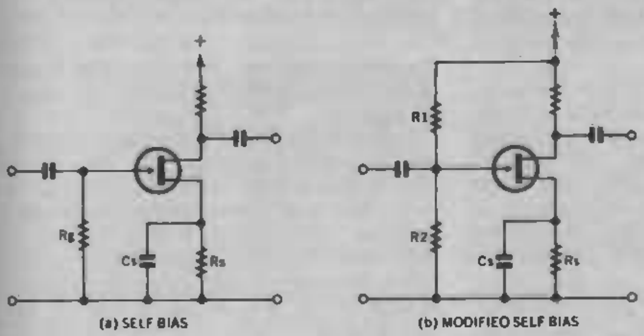
This is particularly true in the common situation where a number of devices are fed from a common power supply. If the supply does not represent an effective short-circuit for AC signals, the supply voltage for each device tends to vary with the current drawn by the others, so that the signals from the various stages tend to become jumbled together.

The second point to realise about the power supply circuitry is that with equipment operating from the AC mains, there is generally a significant amount of AC ripple superimposed on the DC output voltage of the power supply. Generally, this ripple is not sufficient to disturb the operation of devices in parts of the circuit dealing with large signals, like output stages, but it can seriously upset the operation of devices dealing with very small signals. Thus the supply voltage fed to such stages must generally be given additional filtering.

The usual way in which these two basic power supply requirements are met is shown in Fig. 5. A large capacitor (C1) is connected across the main output of the power supply circuit. This is known as the "reservoir" capacitor, and besides serving to smooth out some of the AC ripple superimposed on the DC output of the supply, its purpose is also to serve as an effective short-circuit between the active supply line and the chassis common.

Because the value of C1 can never be made large enough either to reduce the ripple to zero, or to provide a completely effective short-circuit for AC signals, the voltage across C1 is used only to supply the devices in the circuit which deal with large signals. The other stages are fed through so-called "decoupling" circuits, which usually consist of series resistor and shunt capacitor combinations such as that represented by R and C2 in Fig. 5. In some cases a number of such decoupling circuits may be used, one after the other, to feed devices dealing with smaller and smaller signals.

The series resistance-parallel capacitor arrangement of the decoupling circuits acts to provide further filtering of the supply voltage, reducing the AC ripple. In addition, it acts to ensure that the signals from the devices dealing with large signals do not find their way into those dealing with small signals. The shunt capacitors effectively provide "individualised" AC short-circuits across the supply fed to the devices handling small signals, while the series resistors give additional isolation.



(a) SELF BIAS

(b) MODIFIED SELF BIAS

Fig. 4. Biasing the gate of a FET negative with respect to the source.

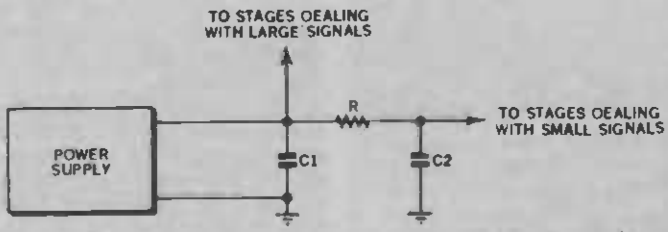


Fig. 5. A power supply should present an effective AC short-circuit and include sufficient decoupling to prevent interaction between stages in a circuit. It should also minimise AC ripple.

The last general circuit concept which we should look at in this chapter is that which is usually given the name "feedback". As the name itself suggests, this is basically nothing more than the technique of feeding back part of the output signal of a circuit to its input.

It is very important to distinguish between two different forms of feedback — positive feedback and negative feedback. Positive feedback occurs when the signal fed back from the output of a circuit has the same effective polarity as the input signal, so that it reinforces the input. Negative feedback occurs when the signal fed back from the output has the opposite effective polarity to the input, so that it opposes and tends to cancel the input.

Broadly speaking, positive and negative feedback tend to have opposite effects on the behaviour of a circuit.

Positive feedback tends to increase the gain or amplification, but also tends to cause instability and oscillation. In fact, an oscillator may be considered as an amplifier with sufficient positive feedback to provide all of the input signal required. An example of positive feedback may be seen by turning back to Figs. 6 and 7 of chapter 10.

Other effects of positive feedback are that it tends to exaggerate any distortion, nonlinerarity or changes in gain at different frequencies.

In contrast with positive feedback, negative feedback tends to reduce the effective gain of a circuit. However, it also tends to stabilise the gain, making it less dependent upon variations in the behaviour of the active devices in the circuit. Negative feedback also tends to reduce distortion, improve linearity, and smooth out any changes in gain at different frequencies. Similarly it tends to extend the range of frequencies over which the circuit will operate — ie, broaden the "frequency response".

An example of negative feedback is the effect on AC signals when the bypass capacitor is not connected across the cathode bias resistor of Fig. 2(b) in this chapter. Similarly the effect when the bypass capacitor is omitted from across Re in Fig. 3(c), and that across Rs in Fig. 4.

Note that in the same examples, the worthwhile effect which the resistors have on biasing stability is due to the fact that they actually produce negative DC feedback. Negative DC feedback is quite often used in semiconductor circuits to improve the biasing stability.

Positive and negative feedback also have various effects on the input and output resistances of a circuit, but this is a little beyond the scope of our present study course. The main points to remember are those in the foregoing.

CHAPTER 17

Receiver Alignment

The concept of alignment — why it is necessary — general techniques — tools required — the effect of AGC — instability — aligning IF stages — aerial and oscillator circuits — short wave alignment — alignment using instruments — television receiver alignment.

The inclusion of a chapter on alignment may seem rather out of keeping with earlier and more elementary material. However, it is reasonable to assume that by this stage many readers will have tried their hand at building one of the receivers described in "Electronics Australia" or elsewhere, and may by now be facing up to this very subject.

Such being the case, it seems appropriate to do three things:

(a) Explain what is behind this matter of alignment and why it is necessary at all;

(b) Discuss some of the side-issues as, for example, dial setting, alignment tools and so on.

(c) Give an alignment procedure for a typical modern superhet receiver.

Even if the reader does not happen to have a receiver on the table, awaiting attention, the information should be useful and will be available against the day when it may be required.

In the early days, the word alignment was virtually unknown. For the most part, each tuning circuit in receivers of the day was brought out to a separate large dial and tuning involved adjusting each dial to the appropriate setting for the particular station. In fact, a standard accessory at the time was a card listing local stations, with two or three spaces alongside each, in which the set owner could insert the dial readings for optimum reception.

As time went by, designers sought to simplify matters by making all the tuning coils and capacitors as nearly identical as possible, so that the dial settings would correspond fairly closely. Thus, the owner could remember if he wished that a certain station came in with all the dials set to about 70, another station with the dials about 55 and so on.

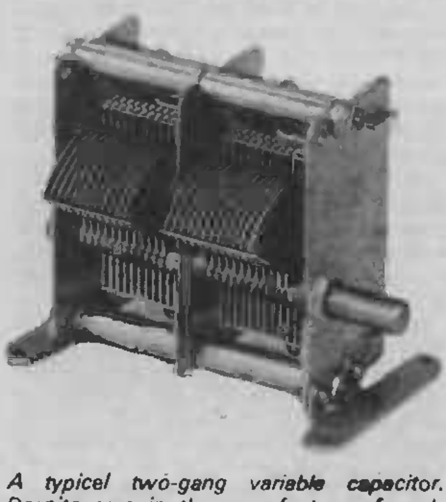
From here, it was an obvious step to take even greater care with the tuning circuits and arrange them so that they could all be adjusted simultaneously by rotating a single tuning knob. At first, the individual tuning capacitors were linked behind the panel with gears or belts. Later they were combined into the one assembly as a two, or three or even four-gang capacitor. Thus "single-dial tuning" became the vogue.

Single-dial tuning has been with us ever since and is obviously a very convenient feature. However, several tuning circuits will not remain exactly in step — or "track" — accurately just because they are superficially alike. Special provision has to be made to ensure that the ganged tuning circuits all resonate to the correct frequency at each and every setting of the tuning dial.

By very careful control of the number of turns on the coils during manufacture and by matching the tuning gang sections, fairly good tracking can be obtained, as a matter of course, near the low frequency end of the range: that is, with the tuning capacitor plates well into mesh.

However, at the other end of the band, with the capacitor plates well out of mesh, the exact tuning of the circuits is affected, as much as anything, by "stray" capacitance — that to do with the position of the components, connecting leads and so on. Variation in this stray capacitance between one tuning circuit and another leads to tracking error at the higher frequencies, and therefore loss of efficiency.

To overcome this problem, it has become accepted practice to connect small trimmer capacitors in parallel with each tuned circuit. They are commonly adjustable between about 10 and 30pF. Sometimes they are separate components; sometimes included as part of the ganged capacitor.



A typical two-gang variable capacitor. Despite care in the manufacture of such units and the associated coils, the tolerances of these components plus stray capacitance effects make it necessary to align or "peak" tuned circuits operating from a common control.

Normal intention is for the trimmers to be set at about half capacitance, the tuning coils and the gang itself then being designed to cover the requisite band of frequencies — from about 1,700 to 535kHz for the ordinary medium-wave radio broadcast band.

If the stray capacitance across one or more of the tuned circuits happens to be a little high, then the relevant trimmer or trimmers are unscrewed a little. Conversely, if the strays happen to be low, the trimmer can be screwed in a fraction to increase the amount of capacitance to the anticipated figure.

Provided that the coils and gang sections are accurately manufactured, the tuning circuits thus aligned at the high frequency end of the band, remain reasonably in step over the whole tuning range.

In more recent years, a further technique has been evolved which leaves even less to chance at the low frequency end of the tuning range. This involves the provision of an adjustable iron dust slug inside each tuning coil. Moving the core in or out of the winding changes the inductance by quite a large percentage.

Thus, in modern receivers, very accurate alignment can be achieved by adjusting the coil cores for exact inductance balance at the low frequency end of the band, and the trimmers for capacitance balance at the **high frequency end** of the band.

Note that the cores and trimmers perform different functions and they should always be adjusted to fulfil those basic functions. The fact that tuning circuits include variable iron cores therefore does not obviate the need for trimmers.

There is another facet to this matter of alignment which must be mentioned.

Originally, receivers used tuning dials simply numbered 0-100, leaving the set owner to memorise the tuning position for each station. Provided the receiver tuned over the necessary frequency band, the exact position of the stations within that band did not matter a great deal.

Thus trimmers could be set for proper tracking, without special reference to the dial reading. The only point which needed to be watched was that the trimmers were not all screwed in so far that they restricted coverage of the receiver at the extreme high frequency end of the band.

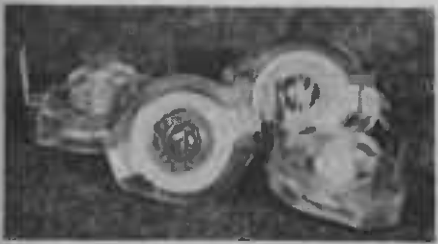
inappropriate adjustment of the cores could similarly affect the low frequency coverage. The remark is rather academic, however, because 0-100 dials had largely disappeared by the time iron cores were introduced.

More recently and, in fact, for many years, frequencies or station names have been marked directly on the dial face. This means that the trimmer and core settings and the position of the pointer relative to the tuning gang shaft must be determined for accurate indication of the incoming station as well as for accurate alignment.

One important point should be mentioned: There is some variation from one type of ganged capacitor to another in the maximum and minimum capacitance figures and the shape of the moving plates. This affects the distribution of stations across the dial scale.

In an existing receiver, it can usually be assumed that the dial has been calibrated to match the particular ganged capacitor and correct alignment should therefore bring the stations in on their calibrated positions.

In a home-constructed set, however, there is a chance that a dial might be used having a scale calibrated for some gang capacitor other than the one used. In this case, no amount of manipulation of the cores and trimmers may succeed in getting all the stations to come in at the right positions on the dial.



Trimmer capacitors come in various shapes and sizes. These illustrated have a clear polystyrene base and measure about ¾in diameter. As the screw is turned clockwise, it forces the top springy plate down, increasing the capacitance between it and the fixed plate beneath.

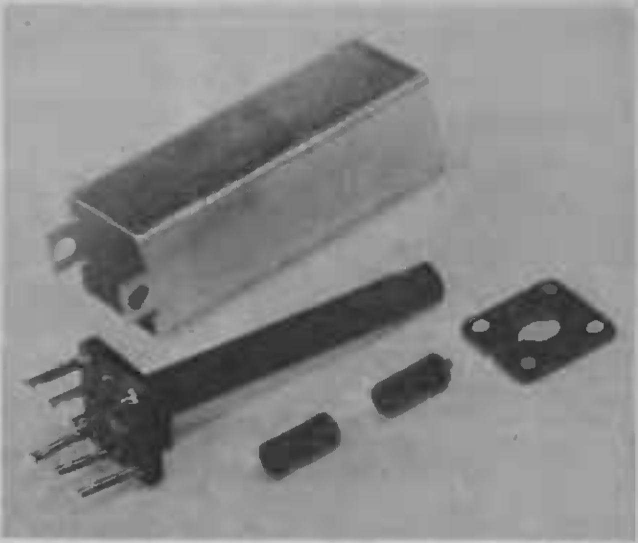
In assembling components, constructors should therefore see to it that the dial scale is for the type of gang capacitor selected.

By now, the reader should have a fairly clear idea as to what alignment is all about. It should be equally clear that a receiver which has not been aligned cannot operate efficiently, because its various tuning circuits will be a long way out of step.

In a factory, receivers are aligned with the aid of a signal generator and an output meter and there is no doubt that the most accurate job can be done in the shortest time with their aid.

Since this is essentially an article for beginners, however, we do not plan to say overmuch about alignment procedures using such instruments.

The average home constructor is likely to have to rely on broadcast stations for a source of signal and to rely on his ears to indicate whether an adjustment has brought about an increase or a decrease in sound volume. Nevertheless, providing care is taken, quite good alignment is possible by these means.



The mechanical components of a typical IF transformer. The alignment slugs are threaded to screw directly into the central former. The hexagonal hole in the ends of the slugs is to take a special alignment tool. In some older types the slugs were moulded on to the ends of threaded brass studs.

If the alignment is to proceed smoothly, a few preliminary points must be checked. Firstly, there is the action of the dial mechanism and the setting of the pointer relative to the dial scale and tuning gang plates.

Make sure that the dial will rotate the capacitor between the limits of its travel without obstruction, or straining the cord in a cord-drive type. If there is any such trouble, it should be corrected before spending time over the alignment.

In cord-drive dials, the tension of the cord is also important. If it is too loose, the cord will slip. If too tight, it may bind and ultimately break. In general, it is best to have the cord twisted twice around the control knob shaft, with the tension no higher than necessary to ensure positive drive.

If the receiver has a simple 0-100 dial scale with no station call sign marked, it is only necessary to see that the dial drives the tuning gang smoothly between the full-in and full-out positions and that the pointer travels over the scale with any overlap about equal at the two ends.

The pointer travel can be corrected in some dials by loosening a screw or slipping the drive cord through a loop. In others, it involves loosening the grub-screws locking the dial to the gang shaft and retightening the screws with the two in different relative positions.

With dials having the stations marked on them, the position is rather more confused. Because the capacitor plates are specially shaped, precise tracking can only be expected if the pointer and its travel is locked to the gang shaft in one specific position — that for which the dial was originally calibrated.

With the two in the wrong relative positions, stations may be brought to correct calibration at the two ends of the bands by manipulation of the cores and trimmers, but those near the centre may be displaced slightly one way or the other.

Some dials — but not many — have a "dial set" line marked just beyond one end of the scale. The intention is that the dial shall be locked to the gang shaft with the pointer set to this reference line and the gang plates either full in or full out, depending on which end of the scale is involved.

Where there is no "dial set" line, the pointer can only be locked in a likely position and the alignment procedure followed out. If the stations can duly be made to fall in the calibrated positions, the pointer can be left set; if not, the pointer may have to be reset slightly one way or the other in relation to the gang shaft and the alignment procedure repeated, noting whether the new position has improved matters or otherwise.

Depending on the receiver, a special alignment tool may or may not be required.

Many trimmers are adjustable with an ordinary small metal screwdriver. As a rule, they are connected into circuit so that the screwhead makes contact with that trimmer plate which returns to earth or to the "earthy" side of the circuit. If touching the trimmer with a metal screwdriver alters the signal level, the chances are that it has been installed the wrong way round.

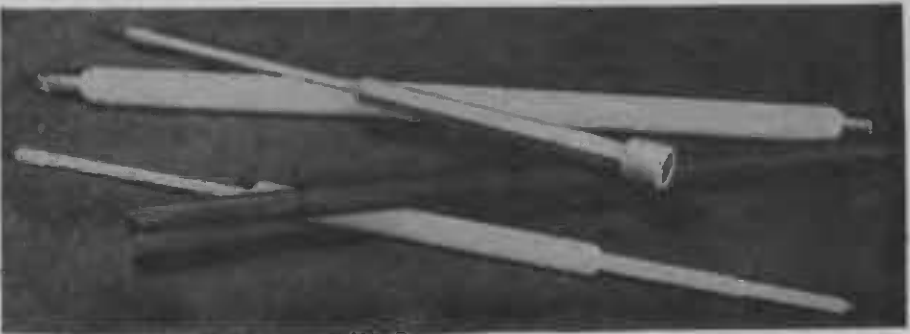
The trimmers in many early type IF transformers are likewise adjustable with a small metal screwdriver. Fingers should be kept off the shaft, however, and the blade kept clear of the metal can, because the trimmers on the plate side often connect internally to the HT circuit.

Iron cores are often adjustable also with a plain metal screwdriver, notably cores which are attached to threaded brass rods protruding from top and / or bottom of the shield cans. In some cases, where the threaded rods are not earthed, touching them with a metal object will affect the behaviour of the coil, making it difficult to pick the proper peak position.

To overcome this difficulty, a type of alignment tool has been available for many years having a very small metal screwdriver tip embedded in a moulded handle.

Alternatively, if the cores are not too stiff in their action, a suitable non-metallic screwdriver can be made by filing a piece of ebonite rod or an ordinary plastic knitting needle to the appropriate shape.

In recent years alignment "screwdrivers" have been made available of nylon or similar materials and these are also very handy for adjusting cores which have a screwdriver slot moulded directly into their ends. A still



A selection of alignment tools used for adjusting radio and television receivers. They are normally non-metallic, or use a minimum of metal, to minimise any effect on the circuit being adjusted by the proximity of the tool.

further type of alignment tool has a hexagonal end designed to engage a similar hole in moulded cores of a somewhat different type.

Irrespective of the core style and tool, however, it is wise to keep in mind that core adjustment systems are seldom very robust. If a slug has become jammed, don't try to force it. It may shatter or become detached from its brass shank (if any) or the whole coil former may be twisted from its mountings inside the can.

Not long after the general adoption of the superhet, designers began to adopt also the technique of Automatic Volume Control, abbreviated commonly to "AVC." The general principle involved has since been extended to television receivers and other equipment concerned with the reception of signals other than those from the "entertainment" radio stations, and a more appropriate modern name for the technique is Automatic Gain Control — "AGC" for short.

In an ordinary broadcast receiver the AGC circuit is usually fairly simple. A DC control voltage, developed by the detector across its load resistor, is fed back as a bias voltage to one or more of the amplifier devices in the tuning section — RF amplifier, converter or IF amplifier.

On weak input signals, very little bias is developed by the detector or fed back to the controlled stages, so that these stages operate at almost full gain.

With stronger input signals, however, the detector develops a large control voltage and this, fed back to the controlled stages, reduces quite drastically the gain or amplification.

As a result of this quite automatic action, the receiver operates at full gain for weak signals but at much reduced gain for stronger signals. "Blasting" and overload effects are largely eliminated, together with at least some of the fading experienced when listening to distant signals.

A volume control still needs to be provided, of course, but it normally operates in the audio system. It allows listening volume to be set to the required level and adjustment to be made for any residual difference between weak and strong signals, not fully compensated by the AGC system.

AGC has much to recommend it from the user's point of view. However, it does complicate alignment somewhat and for a fairly obvious reason. With the receiver tuned to any given input signal, the AGC voltage attains a level depending on the strength of that signal and the gain of the receiver.

Now, if adjustment of a trimmer should increase the effectiveness of a tuned circuit, the resulting increased signal at the detector will produce more AGC voltage. This will decrease the gain, making the effect of the trimmer adjustment much less apparent than it would otherwise be.

Conversely, an adjustment which reduces the effectiveness of a tuned circuit will also reduce the AGC voltage and allow the gain to rise in consequence.

In other words, an AGC circuit in a receiver tends to mask the effect of any adjustments, and quite substantial changes in the efficiency of a tuned circuit through peaking might make only a slight audible difference in the loudspeaker output.

The best way to counter this masking effect is to align such a receiver on very weak signals, as from distant transmitters. Unfortunately, during evening hours, when most homebuilders would want to work on a receiver, there is often a hopeless confusion of weak signals between the strong locals, most of them subject to fading effects which can be most misleading.

If you have to rely on stations for alignment, the best plan is to carry out a rough alignment procedure on local stations to make the dial track correctly, leaving fine adjustment of the trimmer, etc, to some daylight period when it should be possible to pick up a couple of weak but steady signals at the respective ends of the band.

Out of all this comes one of the golden rules of alignment. Peak all receiver circuits as far as possible on weak input signals, advancing the volume control, if necessary, to make them audible.

So much for AGC action and its complicating effect on alignment. Now for a few words on instability.

In a well-designed receiver, it is possible to peak all adjustments for optimum performance, without any complication arising. Each adjustment merely increases the gain and sensitivity till it reaches the maximum of which the set is capable.

In some cases, however, poor design, wrong choice of components, wrong placement or long active leads may allow an excessive amount of signal from one or more stages to couple into an earlier point in the circuit, producing a positive feedback or regenerative effect.

As the gain is increased by progressive adjustments, the receiver may suddenly become unstable or "burst into oscillation," to use another very common phrase.

Oscillation in a large receiver sounds much the same as oscillation in a small regenerative receiver, except that it is produced deliberately in the latter case and controlled by the "reaction" knob.

Instead of the station signals being heard clearly, each one is accompanied or even blotted out by a loud whistle which varies in pitch as the set is tuned across the station carrier.

Many superhets produce faint whistles on odd stations, particularly when operating near powerful transmitters. They are fairly distinct, however, from the strong whistles on every station produced by instability. And while a set is unstable, complete alignment is impossible.

The cure for instability in most cases involves elimination of the cause — improved design if the circuit is of doubtful origin, use of the proper components or rearrangement of the wiring and layout.

Needless to say, this is good argument for care in the first place, when building a receiver, whether it be a TRF or superheterodyne.

So much then for instability or oscillation.

In the case of a superhet receiver, there is more to alignment than merely getting two or three tuned circuits to track, one with the other. This much will be evident from our explanation in chapter 14, of the superheterodyne principle.

Alignment involves getting all the IF transformer windings resonated to the appropriate frequency, usually 455kHz. It then involves getting the oscillator circuit to track with the aerial and possibly RF circuit, the requirement, in this case, being that the oscillator tunes at all times 455kHz higher than the signal frequency.

Fortunately, the task of aligning a superhet receiver is not as difficult as it might appear at first encounter and a home constructor can do a passable job of alignment without instruments and without help, provided a certain routine is followed.

First connect the set to an aerial and earth, preferably the ones with which it is to be used; connect the speaker and switch the power on. Tuning across the band, you will probably be able to hear quite a few stations, if the set is in working order.

Try to find a weak but steady station near the low frequency end of the band.

Tune the receiver as accurately as possible to the station you choose. The best way to do this, or in fact to make any of the adjustments about to be described, is to rock the setting backwards and forwards over the correct point, gradually converging on it.

You can now adjust the IF transformer cores or trimmers as the case may be (most modern IF transformers have core adjustments) for maximum sound output from the speaker.

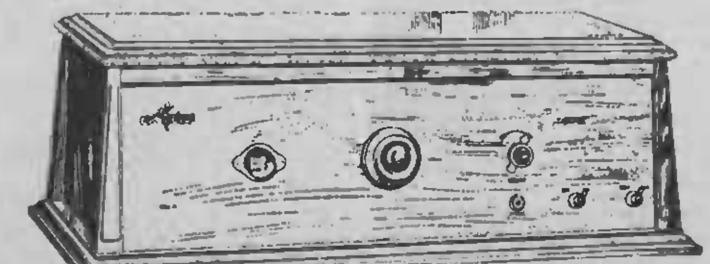
It is possible that what was previously a weak signal now becomes a strong signal as the sensitivity of the receiver rises. This being the case, do the best you can on the original station, and then tune accurately to a weaker adjacent station and go over the procedure again.

The IF windings do not usually need to be peaked in any special order but be sure not to miss any. There may be a slight amount of interaction between the adjustments, so it is a good idea to go over the IF adjustments a second time.

When you have finished with the IF transformers they should be fairly close to the nominal frequency. This is 455kHz for most receivers nowadays, but occasionally you will strike a receiver with a different arrangement.

For this method of alignment it doesn't

A spot of history: In 1925, a prestige product from the Formica Insulation Company (Ohio, USA) was this front panel, finely engraved for three tuning dials and two supplementary controls. The same issue of "Radio News" carried an advertisement for the Pfanstiehl "Overtone" receiver, one of the very few current models offering single dial control. The scale, visible through the window, was calibrated directly in wavelength.



matter and, further, even if you do finish up a few kHz off the specified frequency, it will not be serious. The main thing is that all IF circuits be accurately aligned to the one frequency.

Having satisfactorily completed the alignment of the IF amplifier it remains to adjust the aerial and oscillator circuits. If your receiver has an RF amplifier stage, the adjustments for the RF coil are exactly the same as for the aerial coil.

With most modern receivers the adjustments consist of trimmers and variable slugs associated with both aerial and oscillator coils, and the padder or tracking capacitor is a fixed mica type.

In the case of the variable slugs, tune to a station which you can easily identify toward the low frequency end of the band and adjust the oscillator coil slug until the station coincides with its position as marked on the dial.

Then tune to a station toward the high frequency end of the band and adjust the trimmer associated with the oscillator section to bring this station to its correct position on the dial. Repeat the above adjustments a couple of times because each adjustment does have some effect on the other.

If everything is in order, ie, the dial and tuning capacitor correctly matched, all stations should now coincide with their marked positions on the dial.

Should the AGC system tend to make the strong local stations appear broad it would be in order to remove the aerial or operate the set with a short length of wire while making the adjustments to the station positions.

Finally, the aerial coil slug and the aerial trimmer should be adjusted to obtain the strongest signals. The slug adjustment should be made with the receiver tuned toward the low frequency end of the band and the trimmer adjustment toward the high frequency end.

The aerial circuit should preferably be adjusted with the receiver's actual aerial. It is also a good idea to go over it several times to make sure that you get the best results.

In passing, it should be noted that many modern portable and "mantel" receivers do not have an aerial or RF coil of the type used in older sets and more elaborate modern receivers. Instead they employ a "ferrite rod aerial", or "loopstick", which performs the function of an aerial as well as that of an input tuned circuit. By their very nature, the inductance of these aerials cannot be varied over more than a small range, but this also means that they tend to require little if any adjustment. If needed, however, the adjustment is performed by sliding the coil along the ferrite rod.

By this time the set will probably be very sensitive, and you may not be able to find exactly the sort of signal you require. In this case it is quite in order to tune the set off a station and adjust for the greater noise output.

Modern aerial coils are designed so that the alignment is little affected by the aerial be it long or short, but some of the earlier aerial coils may not be above reproach in this respect. In any case, we suggest that you do the final adjustment of the aerial trimmer with the aerial connected.

Typical RF signal generators. At top is a commercial unit, while immediately above is a do-it-yourself design from "Electronics Australia".

It may be worth mentioning in passing, that some broadcast band superhet receivers have no padder capacitor, either fixed or variable. The necessary tracking between the aerial and oscillator tuned circuits is ensured by having dissimilar sections in the tuning capacitor. The aerial (and possibly RF) tuning section is normal but the oscillator section plates are smaller and differently contoured.

The alignment procedure is exactly as set out for the fixed padder type of receiver.

To this point we have spoken only of the broadcast band. The procedure for the shortwave band or shortwave bands is essentially the same, but there is the added difficulty that it is often hard to find and identify a suitable station for alignment.

Conditions vary a great deal, and sometimes change within a matter of minutes. Therefore, do not be discouraged if first results are not very satisfactory. Of course, the aerial is more important than in the case of the broadcast band, because you frequently wish to listen to very weak signals. However, shortwave stations are heard strongly in Australia, and even a poor aerial will often receive them at considerable strength.

Another problem is that most conventional dual-wave receivers will "double spot" on strong signals. Double spotting is due to the fact that the local oscillator can produce the required 455kHz beat when it is in either of two conditions — 455kHz higher or lower than the wanted frequency. Thus, a second spot is always twice the IF (910kHz in this case) away from the correct dial setting.

While the simple aerial tuned circuit will easily reject the second spot on the broadcast band, it is inadequate on the short wave bands, and a strong signal will inevitably be found at two points on the dial. For alignment purposes, and assuming normal circuitry, the setting having the higher frequency of the two is almost always the correct one.

An excellent source of signals for shortwave alignment and frequency calibration of receivers are those radiated by standards stations such as the American WWV and WWVH, and the Australian station VNG in Lyndhurst, Victoria. Generally WWV and WWVH are heard in Australia at good strength on 5MHz, 10MHz and 15MHz, while VNG may be found on 4.5MHz, 7.5MHz and 12MHz.

The signals from these stations can easily be identified by the audio tones which are superimposed on the RF carrier and the fact that the tones are interrupted by a one-second pulse. There are interruptions at regular intervals for call sign announcements and other purposes.

If suitable test instruments are available, of course, a more precise job can be made of alignment. As we said earlier, we are not so much concerned in this chapter with readers who are sufficiently advanced to own or have access to test instruments, but a brief explanation may help the beginner understand what it is all about.

The best instrument for alignment is a modulated oscillator or the more elaborate instrument which usually goes under the name of a signal generator. These instruments can produce radio frequency signals anywhere in the spectrum required for alignment and the strength of the signals can be controlled by turning a knob on the front panel.

The signal is modulated usually by a 400Hz tone, so that when reproduced by the receiver a single whistling note is heard from the loudspeaker. With a signal thus available at any desired frequency, at any desired strength and producing a constant output tone, alignment is much simplified.

The alignment of television receivers involves special techniques and special equipment and in no circumstances should beginners tamper with TV tuners or IF systems.

Unlike the tuned circuits in broadcast receivers, those in a television receiver are not peaked for maximum gain. They have to be adjusted to pass a band of frequencies, from 5 to 6 megahertz wide, over which the audio and video signals from a television station are distributed. If the tuned circuits in a TV receiver were simply peaked in the normal way, the receiver might well become unstable. In any case it would produce only a poor picture, with no sound, or sound with little or no picture.

Even the sound channel in a television receiver is "special," involving a frequency modulated signal, as distinct from amplitude modulation used by ordinary broadcast stations.

Alignment of picture IF channels in a television receiver is normally performed with the aid of a "sweep and marker" signal generator and a cathode-ray oscilloscope.

The alignment problems inside a television tuner are even more complicated than for the IF systems, such that tuner alignment is rarely attempted other than at the factory or at special tuner service depots.

CHAPTER 18

Building a Superhet Receiver

Here is an ideal project to build up in order to reinforce the knowledge you have gained in the previous chapters. It is a simple superhet radio receiver of modern design, battery operated and with its own loudspeaker. The circuit uses six bipolar transistors and a FET, and is built on a printed wiring board to give you experience with this widely-used construction technique.

Having studied various aspects of electronics as applied to radio receivers in previous chapters, together with the construction of simple receivers in Chapter 13, you are now in a position to tackle a more involved type of receiver using the superheterodyne principle. Perhaps it may be wise to refresh your theory of operation of the superheterodyne before proceeding, by referring back to chapter 14.

Although superheterodyne receivers can often be very complex affairs, nevertheless it is possible to design quite a simple version. With this in mind, we will discuss details of a simple superhet (for short) and show how it may be built quite easily. It must be pointed out here that although the receiver to be described is simple, it embodies all the basic superhet principles.

As you may well have gathered by now, the superhet principle applies only to the sections or stages of a receiver up to and including the detector. The signal emerging from the detector is the wanted audio in the form of speech or music at low level. So that the audio signal may be used to drive a loudspeaker, it is necessary to amplify the signal and this is done in the audio amplifier, which is another essential part of a complete receiver. A modern audio amplifier is incorporated into this receiver and it is sufficient to drive a loudspeaker to a level adequate for normal listening purposes.

The other essential part to complete the receiver is a power supply. In its simplest form, this can be a suitable battery and this is the type of supply which we have used here. By taking this approach, we are able to avoid the use of a mains power supply, which has the advantages of reduced initial cost and the avoidance of any hazards which may attend the use of mains power by a novice. Hum problems are reduced to vanishing point, and the receiver may also be operated where no mains supply is available.

Having dispensed with the preliminaries, we shall now take a look at the complete circuit and discuss it in some detail. It will be seen that there are seven transistors in all, with only three of them in the "tuner" and the other four in the audio amplifier system.

The first stage which the signals meet after entering from the aerial is an RF amplifier and it uses an N-channel junction FET, either a Motorola 2N5485 or a Fairchild FE5485. A FET was chosen in preference to a bipolar transistor for this stage, on grounds of circuit simplicity.

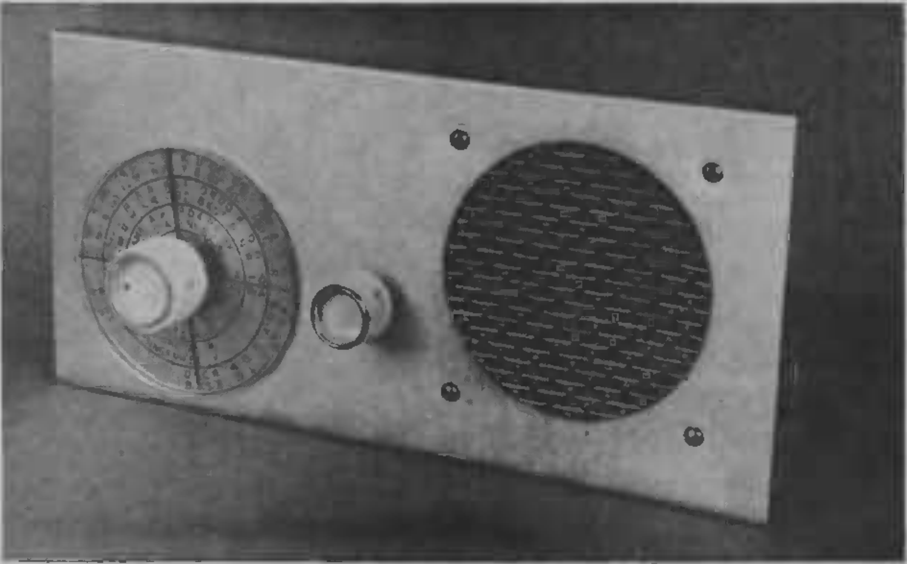
In the gate circuit is a conventional type of miniature aerial coil. This coil is type No S203, made by Aegis Pty Ltd. It may also be seen that there is a 220pF capacitor from the top of the primary winding of the coil, to the tap on the secondary winding. This has been added to increase the coupling, more particularly at the high frequency end of the RF band.

In many receivers which have an RF amplifier, the output as well as the input of this stage are tuned. However, for the sake of simplicity, we have not tuned the output. Instead, in the drain circuit of the FET is a 2.5mH RF choke. The output of the RF amplifier, at the drain of the FET, is a high impedance and the input to the next stage is a relatively low impedance. To meet this situation, we must devise some sort of matching arrangement.

This problem has been solved by connecting two capacitors in series, across the RF choke. By arranging the correct ratio between the two capacitor values, an impedance transformation is achieved at the junction of the two capacitors. From this point, the signal is fed directly to the base of the next stage.

In order to provide a standing negative bias for the gate of the FET, with respect to the source, a resistor is placed between the source and earth and the voltage drop across this resistor is effectively the bias for the stage. Generally, there is a capacitor across such a resistor, to provide a low impedance path for the signal. However, in this case, the capacitor has been omitted so that negative feedback or degeneration is introduced. Also, the value of the resistor is larger than the optimum value for maximum gain. Again, this has been done to increase the amount of degeneration.

At this stage, it may well be asked, why all the degeneration in the RF amplifier, which reduces the gain? The answer is that a compromise must be reached between



A front view of the finished receiver showing the location of the controls. Note that the callsigns are calibrated directly on the scale. In some cases the scale may be blank.

YOU WILL NEED THESE PARTS

1 Chassis-panel, 9in long x 4 ¼in high x 5in deep
2 Terminals, 1-red, 1-black
1 Dial, Roblan 6 / 1, with knob
1 Knob to match, for volume control
6 Spacers, ½in long x ¼in dia, tapped ⅛in Whitworth
1 Speaker, 3in, 8 ohms
1 Battery, 9V type 2364 or similar
1 Battery plug
1 Printed board, 72 / R9, 7in x 3in
1 Jabel flexible coupling, ¼in x ¼in
1 Aerial coil, Aegis S203
1 Oscillator coil, Aegis S201
2 IF transformers, Aegis ST45C
1 Murata filter, SFD455B
1 2.5mH RF choke
1 Transistor, 2N5485, FE 5485
2 Transistors, BF115, TT1002, SE1002
1 Transistor, BC108, TT108, AY1121
1 Transistor, BC178, TT608, 2N3638A
1 Transistor with heat sink, TT801, AY8139
1 Transistor with heat sink, TT800, AY9139
3 Diodes, OA91
Screws, nuts, hookup wire, solder, etc.

RESISTORS (½ watt)
1 22 ohms
1 33 ohms
1 220 ohms
2 390 ohms
3 470 ohms
1 2.2k
3 4.7k
1 18k
1 33k
1 180k
1 330k
1 Potentiometer, 500k log with switch
1 1M
1 4.7M
1 5.6M

CAPACITORS
1 220pF 630V polystyrene
1 3-gang, 200pF-90pF-200pF, Roblan
2 470pF 630V polystyrene
2 .001uF 630V polystyrene
1 .0012uF 400V polyester
1 .0022uF 400V polyester
2 .0039uF 400V polyester
1 .0047uF 25V ceramic
3 .01uF 25V ceramic
1 .01uF 160V polyester
1 .047uF 25V ceramic
1 0.1uF 25V ceramic
1 25uF 25VW electrolytic single ended
1 47uF 25VW electrolytic single ended
1 220uF 25VW electrolytic single ended
1 1000uF 10VW electrolytic single ended

Note: resistor wattage ratings and capacitor voltage ratings are those used with our prototype. Components with higher ratings may generally be used providing they are physically compatible. Components with lower ratings may also be used in some cases, if ratings are not exceeded.

---

conflicting requirements. While maximum gain is desirable, there is another consideration which must also be met.

It will be explained later that automatic gain control (AGC)voltage is applied to the gate of the RF amplifier and the RF amplifier is followed by a self-oscillating mixer. Now the varying AGC voltage causes a change in capacitance at the output of the RF amplifier and this change is seen by the following mixer stage. The capacitance change tends to cause corresponding mixer oscillator frequency changes, and this cannot be tolerated as it upsets the tuning of the receiver. By introducing the degeneration previously referred to, the effect is reduced to a minimum.

As mentioned earlier, the next stage is a self-oscillating mixer, also called an "autodyne." One transistor, which is a bipolar type BF115, TT1002, SE1002, etc, performs the dual task of oscillator and mixer. In order to achieve this, the circuit is rather interesting, as even a casual glance will indicate.

The transistor is biased in the usual way, with a voltage divider in the base circuit and a bypassed resistor in the emitter circuit. A special oscillator coil has been designed for this arrangement. There is the usual winding tuned by one section of the gang, together with two small windings to provide feedback between collector and emitter. It will be seen that these windings are connected in series with the collector and emitter, respectively.

To complete the oscillator circuit, the base of the transistor should be earthed to AC, although not to DC. This can be met by placing a bypass capacitor from base to earth. However, if we use a large value of capacitance, with a very small reactance, for this purpose, the oscillator requirement will be met but the wanted signal from the RF stage would be short-circuited, preventing the stage from acting as a mixer.

Once again, we must compromise but fortunately this can be done quite easily, without detriment to performance. In order to get the oscillator to function satisfactorily, it is not necessary to bypass the base with a very large capacitor. It is possible to arrive at a capacitance value which will bypass the base sufficiently to allow proper oscillator operation and still permit the signal to be fed into the base.

If we take a careful look at the circuit, it will be seen that the base of the autodyne is bypassed by the .0022uF capacitor at the RF choke, and the 0.1uF capacitor from the +9V line to earth. These two capacitors are effectively in series for this function. It may be remembered that the .0022uF capacitor is part of the voltage divider across the RF choke, mentioned earlier. This same capacitor also performs a third function of DC blocking between the drain of the RF amplifier and the base of the autodyne.

The circuit requires a 2-gang variable capacitor, with one section designed for an oscillator incorporating the padderless technique. This section has a lower maximum capacitance and consequently smaller plates. It will be noted that we have actually used a 3-gang unit, 200pF-90pF-200pF, made by Roblan. This was chosen because it was most physically suited to our purpose and the extra cost was only slight. The extra 200pF section is left unused.

So far, we have considered the function of the mixer as an oscillator, and the wanted signal at its base. You have read in an earlier chapter about the process of frequency conversion of the incoming signal, to an "intermediate frequency" (IF). The IF we are using is nominally 455KHz.

This means that as well as the oscillator and RF signals at the collector of the autodyne, there will also be a 455KHz component. So that the latter may be separated out and fed into the succeeding stages, the first IF tranformer primary is connected in series with the oscillator coil winding in the collector circuit. As the IF transformer winding is tuned to the wanted 455KHz frequency, this component is accepted and the oscillator and signal components rejected.

This has brought us to a very important section of the receiver, and where most of the selectivity is realised. Two IF transformers, intercoupled with a two section ceramic filter, form a complete "filter block." We have already explained that the tuned winding of the first IF transformer is connected into the collector circuit of the autodyne stage. For reasons which should become apparent, the low impedance winding of this first IF transformer is left unused.

Now the impedance of the IF transformer is high and the input of the ceramic filter into which the transformer feeds is a relatively low impedance. We have a very similar situation here that we had when we had to couple between the RF amplifier and the input of the oscillator-mixer, and although different values of capacitance are used, the matching system used is the same. In this case the total capacitance across the winding has a value which will allow the transformer to be tuned to 455KHz and the ratio of the two capacitances is chosen to give the right impedance transformation.

The ceramic filter consists of two sections, and it is necessary to provide coupling between the two sections. This is done with a 68pF capacitor. The value of this capacitor, within definite limits, determines the bandwidth of the ceramic filter assembly. Increasing the capacitance up to 150pF increases the bandwidth and reducing it to 27pF reduces the bandwidth. The value we have chosen is about right for our purpose.

Coupling from the output of the ceramic filter into the tuned winding of the second IF transformer is done in the same way as before. This brings us to the point where we have to couple from the filter block into the base of the IF amplifier, which again is a relatively low impedance. As mentioned before, there is a low impedance winding provided on the IF transformer for this purpose. A look at the circuit will show just how this is done.

The two IF transformers (type ST45C) and the oscillator coil (type S201), like the aerial coil mentioned earlier, are made by Aegis Pty Ltd. The ceramic filter is type SFD455B, made by the Japanese firm Murata and imported by IRH Components Pty Ltd.

The IF amplifier is a conventional bipolar transistor amplifier and there are only a couple of points of interest as far as we are concerned. Although it is used as an IF amplifier, the collector circuit is not tuned and consists simply of a 4.7k resistive load. It will also be noted that the collector circuit is decoupled with a 220 ohm resistor and a 0.1uF capacitor. This is done to prevent any possibility of IF signals being fed back into earlier stages via the power supply line, which would cause instability.

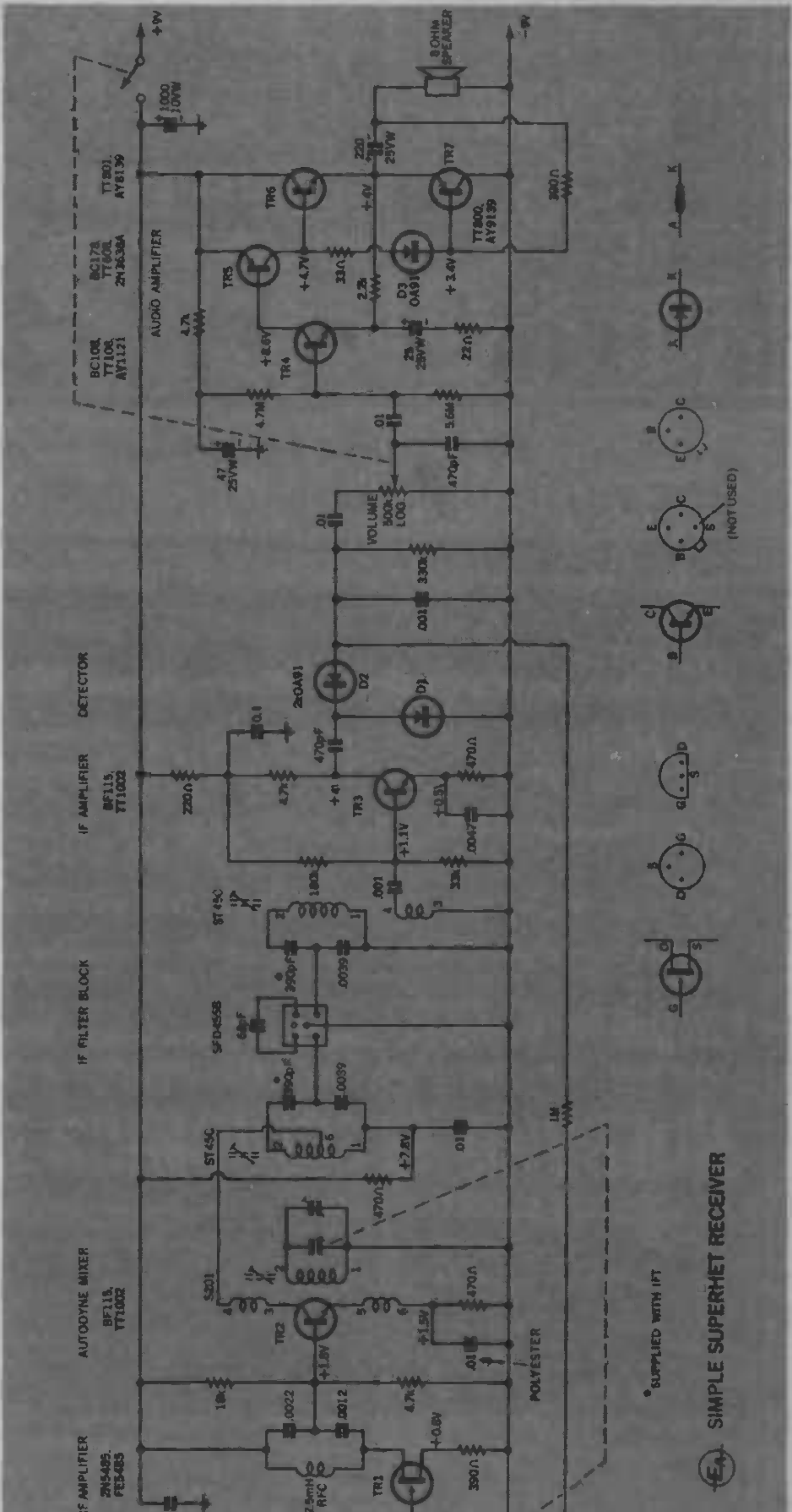
The detector follows, and is quite an interesting circuit. Two germanium diodes are used and they are connected in a half-wave voltage-doubler rectifier circuit. The IF amplifier transistor takes the place of the usual transformer and the 470pF and .001uF capacitors perform the functions of doubling and filtering. The 330k resistor forms a load for the circuit and the rectified 455KHz signal appears across the 330k load. This takes the form of a DC voltage, together with the audio signal recovered from the modulation. The advantage of using the voltage doubler type of detector is that we get almost twice the recovered audio compared with a single diode detector.

The DC voltage which accompanies the audio is also greater than that from a single diode detector. As this voltage varies in level according to the signal strength, it is possible to use it for automatic gain control (AGC). The gain of the FET in the RF amplifier may be controlled by feeding a negative voltage to its gate. By connecting the two diodes in the detector the right way around, we can make them produce a negative DC voltage across the 330k resistor.

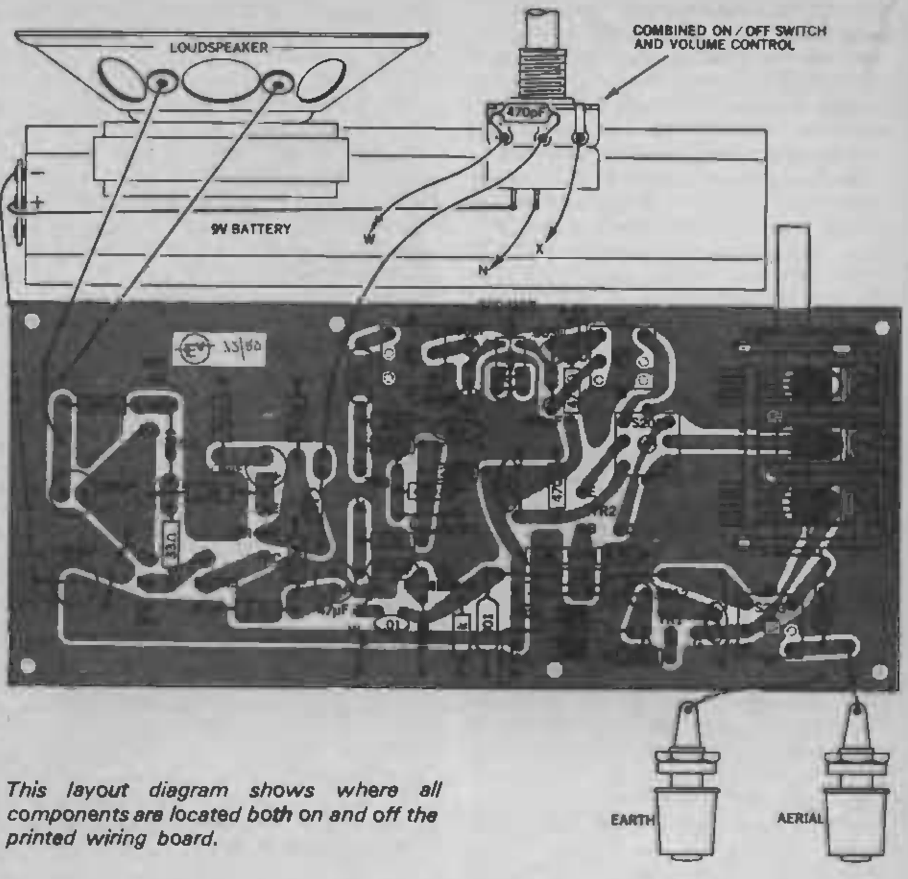
As mentioned before, this DC voltage has the audio superimposed on it. Before the DC voltage can be used for AGC purposes, the audio component must be removed and this is done with a simple filter consisting of a series 1M resistor and a .047uF capacitor. This capacitor performs the dual role of filter and a low impedance path to earth for the secondary winding of the aerial coil. Thus, the DC control voltage for AGC action is fed from the detector output to the gate of the FET, via the 1M resistor and the secondary of the aerial coil.

As the audio output from the detector is not sufficient to drive a loudspeaker, it must be amplified in an audio amplifier system. So the audio from the detector is fed through a volume control into an audio amplifier consisting of three stages. The battery On / Off switch is mounted on the back of the volume control.

The audio amplifier circuit is quite straightforward and in line with modern design. It is transformerless and uses a complementary-symmetry PNP-NPN pair of silicon transistors in the output stage. Two more silicon transistors, NPN and PNP, are direct coupled and form a high gain system to drive the output pair. DC feedback of 100 per cent is achieved through the 2.2k resistor, to stabilise the operating current. AC negative feedback is also applied via the same resistor, but can be controlled over a wide range by varying the value of the resistor in series with the 25uF capacitor. The input impedance of the stage following



The circuit diagram of the Superhet receiver. The components are marked and identified to allow you to follow the description in the text. Note the transistor lead identification details.



This layout diagram shows where all components are located both on and off the printed wiring board.

the volume control is high, in the order of a megohm or so, and is largely dependent upon the degree of negative feedback adopted.

The output transistors are operated in class B, with a consequent low value of quiescent current. In its simplest form, the bases of the transistors of this type of output stage are tied together. This can lead to some "crossover" distortion, particularly at low levels. To avoid this, we have introduced a 33 ohm resistor and an OA91 germanium diode in series, between the transistor bases. This sets up the bias such that crossover distortion is avoided and is well worth the slight increase in cost.

The voltage at the emitters of the output transistors should be as close as possible to half the supply voltage, to ensure that the amplifier can deliver full output without premature clipping. As the three stages in the main amplifier are DC coupled, this voltage is determined mainly by the voltage divider at the base of the stage following the volume control. It may be necessary to adjust the value of the resistor, shown as 4.7M, if the voltage at the output emitters is not approximately +4.5V when the set is first switched on.

The loudspeaker is fed from the output transistor emitters via a DC blocking capacitor — a 220uF electrolytic. For our requirements, this value is sufficient as we are only driving a small speaker. If we had to consider a much larger speaker in a system where good bass response would be expected, then the capacitor would have to be larger, up to 1000uF or so.

The speaker we have chosen is a type 3U made by Magnavox, with an 8 ohm voice coil. A similar speaker with a 15 ohm voice coil may be used but somewhat less volume will be available. Other brands of speaker with similar characteristics to the one we used would be quite suitable.

The 9V battery is an Eveready type 2364. This is the most suitable type for our purpose, being compact in size and still an economical proposition. However, it should be pointed out that if you wish to use a larger 9V battery and either make the metalwork larger to take it, or leave it external to the receiver, a longer life would be possible for only a little extra outlay. Just as a broad example, supposing that the type 2364 battery lasted for 60 hours of operation, a type 276P may be expected to last for about 130 hours. The cost of the larger battery is only marginally greater than the small one.

The above figures are only given as a guide and a very important point to watch with regard to battery life is the volume at which the set is used. The class B output stage draws current from the battery to a degree roughly proportional to the volume. Clearly, if the set is operated at a modest volume, the battery will last correspondingly longer. The choice is up to you.

The dial is quite a simple unit but adequate for the job. Many of the Australian broadcast stations are marked in groups for the six States, around the circular scale. A planetary drive is incorporated and the movement is very smooth, with a reduction of about 6:1. This makes tuning pleasant and easy. The complete dial assembly is marketed under the name of Jabel by Watkin Wynne Pty Ltd, and should be available from most radio components suppliers.

Construction of this little set has been made about as easy as it could be, as most of the components are contained on one printed board 7in x 3in. The only components external to the board are the loudspeaker, battery and volume control. The latter has a 470pF capacitor mounted between the earthed and rotor lugs.

Mounting and soldering of the components on the board may best be done by following the coded picture which we have provided. Alternatively, some board manufacturers, such as RCS Radio Pty Ltd, actually print the code directly on the board. This makes assembly doubly easy.

Before proceeding with the assembly, it must again be emphasised that care should be taken not to overheat the components. This applies particularly to transistors and the point has been covered in more detail in Chapter 13.

A systematic approach to construction will make the job easier and here are some suggestions as to how it may be done.

Start with the smaller components and more specifically, all resistors may be mounted and soldered in. To make a neat job, the resistors should have their pigtails bent at right angles and at points equidistant from the body of the resistor. The pigtails are then dropped through the appropriate holes. With the body of the resistor against the surface of the board, turn the board over and bend the two pigtails outwards by about 45 degrees. This stops the resistor from falling out during the soldering process. Cut the two leads off, leaving about 1/16in protruding and then carefully solder.

Having done all the resistors, the small capacitors are next. Treat them in a similar way to the resistors. It will be seen that some of the holes on the board are too close together to allow capacitors to be mounted against the board. There are six capacitors so-affected and these are the ones across the two IF transformers and the .0022uF and the .0012uF across the RF choke. These should be stood on end, with one end hard against the board and the other pigtail bent over and down, so that it goes through the other hole.

Perhaps it should be noted that we have specified 630V rating types for many of the small value polystyrene capacitors. This has been done simply because in these capacitors, most suppliers stock only the 630V types. Needless to say lower voltage capacitors may certainly be used if your supplier has them.

The four electrolytic capacitors are next and should present no problems. However, care should be taken to make sure that they are mounted with due regard to polarity. If you should have any difficulty in getting "single ended" electrolytics, as we did with the 1000uF unit, then an ordinary one may be used and mounted in the same way as we have described for the other capacitors which were mounted on end. In some cases, it may be necessary to lengthen one pigtail by soldering another pigtail offcut to it.

The RF choke is mounted in a similar way to the resistors and with the coil resting against the board. The little ceramic filter must be given careful attention, as it is quite small and perhaps a little delicate. Make very sure that it is mounted with the dot on the case facing the 68pF capacitor. Press the filter hard against the board and the four corner tags may be bent over so that they lie against the copper to which each has to be soldered. Start with the unbent centre tag and carefully solder it, then follow with the other four. Be very careful not to allow any solder to run across and so bridge the gaps between the lugs and adjacent copper on the board.

The four coils should be tackled next. Although it is not possible to put the coils into the board the wrong way, care must be taken to see that each coil is placed in its correct position on the board. Push the lugs of the coil through the holes on the board and press the can hard against the board. Do not bend the lugs over, but first solder the two lugs of the can. Then solder all other lugs which connect to copper conductors. It will be noted that in some cases, lugs are left insolated and no attempt should be made to solder them.

There are still three OA91 diodes and these may be fitted now in the same way as the resistors were done. There are seven transistors and it may be seen that there are heat sinks on the two output transistors. Before soldering these in, it may be easier to fit the heat sinks. These are stretched over the metal case of the transistors. When fitting the transistors to the board, make sure that each one is in its right place and make sure also that the connections are correct. When the leads are pushed through the holes in the board, about 1 / 16in should protrude for soldering.

When the transistors are fitted, the driver transistor may be touching the heat sink of the TT801 or equivalent. This is simply remedied by bending the driver transistor slightly away from the other one.

The three-gang capacitor is the largest and last component to be mounted. There are three mounting lugs, two rotor earthing lugs and three stator lugs to be soldered. Push them all through the appropriate holes and push the body of the gang hard against the board. Solder the three mounting lugs first.

It will be necessary to use much more solder than for other joints and to approach the lug from all directions, until it is soldered all around. The rotor lugs are treated in a similar manner, although they are not so heavy. Bend the stator lugs over so that they lie along the copper pad provided and then solder each one. In common with the other lugs, these also require a fair amount of solder.

To complete the board, nine leeds of hook-up wire have to be soldered in. There are three for the volume control, two for the battery, two for the speaker and one each for the aerial and earth terminals. A variety of colours may be used to advantage and each lead should be long enough to reach the terminating point when the board is finally mounted on its chassis.

The completed printed board assembly, along with the other items, is mounted on a one-piece chassis-panel in the form of an "L" made from a piece of aluminium or other metal. The unit is 9in long, 4¼in high and 5in deep. Along the top of the front panel is a ½in fold, for added strength and to improve its appearance. Similarly, the base has a 1in fold up along the back, again for strength as well as to provide a mounting position for the aerial and earth terminals.

For readers who wish to make up their own metalwork, we have provided a drawing giving all the necessary details. One alternative is that readymade metalwork may be available commercially. Also, may we hasten to point out that it is not necessary to have a metal chassis and panel. Readers who feel so inclined may make up their own base and front panel from plywood, hardboard, etc.

Before attempting to mount any components on the panel and chassis, it would be wise to make a thorough check of the printed board assembly, to make sure that there are no errors or omissions. It is also a good idea to see that all joints have been properly soldered and that none has been missed.

Satisfied that all is well, we may now proceed with the final assembly. The two terminals on the back should be screwed in place, with the insulating washer provided for the aerial terminal, but without the insulating washer for the earth terminal. In each case, a solder lug is placed under the nut. When mounting the speaker on the front panel, a piece of grille cloth or perforated metal should be placed behind the panel to protect the cone of the speaker.

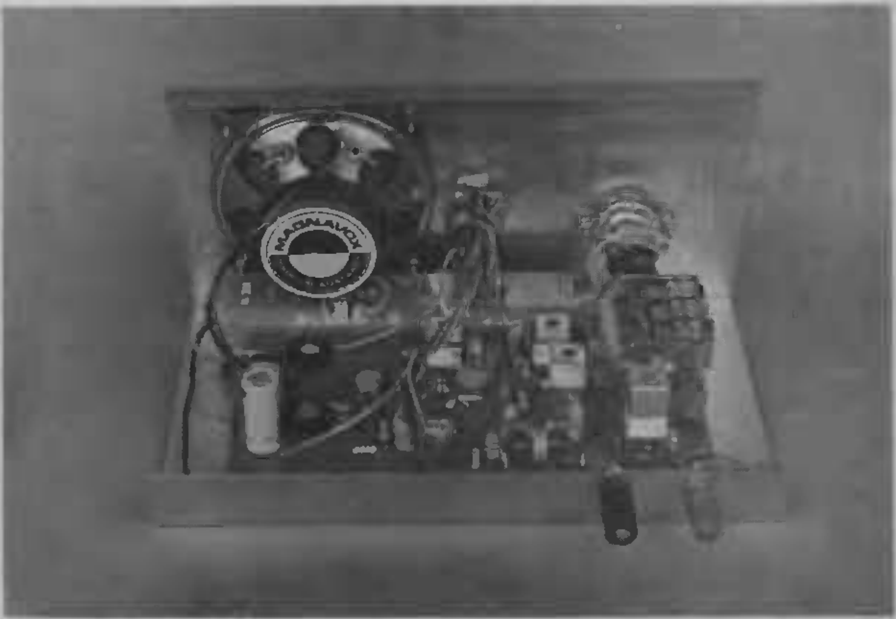
Before attempting to mount the dial assembly, undo the two screws and remove the circular cursor from the reduction drive and put the cursor aside for the moment. The dial movement is held in place with two countersunk head screws, each with a nut behind the panel. Another nut is run on to each screw and so placed that it provides a means of spacing the dial movement lugs from the back of the panel by the required amount. The correct spacing will be determined a little later on. With the dial movement in place, run another nut on to each screw and tighten lightly.

The dial scale is fixed to the front panel by the adhesive on the back, after the backing paper has been peeled off. Fixing the scale should be done with great care, so that it is central with the drive spindle and the line on the scale truly horizontal. When satisfied, the scale is gently rubbed in place. It is possible to lift it to change the position, by gently peeling it off and starting again.

The cursor may now be screwed back onto the drive assembly. The cursor must clear the dial scale by about ⅛in and this is adjusted by means of the spacing nuts on the screws which fix the lugs of the drive assembly to the front panel. When adjusted correctly, all nuts are finally tightened. The knob may also be fitted to the dial.

The printed board assembly is fixed to the base plate with six ½in spacers. It does not matter whether you fix the spaces to the base plate first, or the board. During the process, the dial drive is connected to the gang spindle, via a flexible coupling. It will be necessary to provide a piece of ¼in shaft, about ⅝in long. This may be obtained from the volume control shaft, which may be cut to the wanted length at this time.

The battery and volume control are the only items left. Slide the battery in just behind the front panel, from the speaker side and locate it laterally so that it does not foul the flexible coupling. When mounting the volume control, fix the crescent-shaped clamp at the same time, so that it presses down on the battery and holds it in place.



Rear view of the receiver showing the assembly of the major components. Note how the battery is fitted and the leads left long enough for easy removal. Tight leads on any of the components may strain the printed board with resultant cracks in the copper connections, causing open circuits. The loudspeaker terminals are placed at the top to clear the battery.

All the wires from the board are now cut to length and terminated at the appropriate points. The two battery leads will be terminated in the battery plug and care must be taken to ensure that the correct polarity is observed. The positive lead will go via the switch on the back of the volume control.

Before attempting to align the receiver, it is important to have a good aerial and earth system. The earth wire should be run to a clamp on a water pipe, or to a metal plate buried in moist ground. For best results, the aerial wire should be as long and as high above ground level as possible. However, many readers in metropolitan areas may find that they will obtain adequate performance with just a few feet of wire.

As a preliminary to alignment, set the gang rotor plates fully in and see that the dial cursor is in line with the horizontal line on the dial scale. All grub screws in the drive between the dial and gang should be firmly tightened in this position.

With aerial and earth connected, switch on the receiver and advance the volume control. Then tune in the first station nearest the low frequency end of the dial. This will vary according to the location and will be 2FC in Sydney, 3AR in Melbourne, etc. With a proper aligning tool, proceed to adjust the slug in the second IF transformer for maximum response. Then, similarly, adjust the slug in the first IF transformer. Rock the dial about the station to make sure that the station is correctly tuned and proceed to touch up the previous adjustments.

The station already tuned in may not coincide with its marked position on the dial. If it does, then all is well. If not, then turn the dial pointer to the correct position and proceed to adjust the slug in the oscillator coil until the station reappears. Then adjust the slug in the aerial coil for maximum response. If by now the station is so strong that it is difficult to determine the correct adjustments, remove the aerial and substitute a shorter one, at least for the time being.

Now tune in the local station which is nearest the high frequency end of the dial. Check its position on the dial scale and more than likely, it will not be in its correct place. Reset the pointer to the correct position and proceed to retune the station as before, using this time the trimmer on the oscillator section of the gang. The trimmer on the aerial coil section of the gang should then be adjusted for maximum response.

In common with all superhet receiver alignment, it is now necessary to go back to the station near the low frequency end of the dial, as it is almost certain that the pointer will no longer correspond exactly with the station. This must be corrected with the slug in the oscillator coil, after which the slug in the aerial coil is readjusted. Then return to the other station once again and go through a similar procedure, using the two trimmers. This procedure must be repeated as often as required to bring the stations to their correct position on the dial scale, always remembering to use the coil slugs at the low end and the trimmers at the high end. With this, alignment is complete.

If you live in a metropolitan area and consider that you are in a fairly strong signal location, then you may wish to consider the use of a ferrite rod aerial coil (Aegis type S211) in place of the aerial coil shown in the circuit. Only the two end leads of the rod coil are used, being soldered to the points or the printed board corresponding to pins 1 and 2 of the aerial coil. The lead from the tap on the coil may be neatly taped to the body of the coil, out of the way.

If you intend to try both coils, then to avoid the problem of removing the aerial coil after it has been firmly soldered to the board it is a good idea to just push the coil pins through the board, so that there is even less than 1 / 32in for soldering. This will make unsoldering easy when you wish to remove this coil and try the rod coil.

With your new receiver complete, you may expect many hours of enjoyment from it. It may be left as it is, or you may even make a cabinet of wood or other material for it. The cabinet may be so made that the receiver slides in from the front. Two cleats may be provided, one at each end of the cabinet so that a screw at each end of the panel will hold the whole assembly together.

HOW THE SUPERHET BEGAN

The concept of the Superheterodyne principle, it is reported, originated in the mind of Major Edwin Armstrong, US Army, during an air raid in France, in the latter part of World War I. Searching for a method to detect the approach of enemy aircraft, he thougnt of listening for the faint "radio" signals emitted from the ignition systems in the form of electrical interference. This called for a receiver having far more than the usual sensitivity, and the principle of the superheterodyne was evolved. It showed promise but it was not until May 1918 that it reached a practical form. By then it was too late for the immediate purpose.

The principles of the circuit were rather complicated for that era and it enjoyed only limited popularity as a result. It had its problems too: heterodyne whistles, radiation and the general complexity of adjustment. Even so, several companies made prestige superheterodyne models during the next few years and in fact AWA manufactured a Radiola, "Super Model C26" in the year 1927. The deteils of this receiver were featured in our forerunner magazine of that time, "Wireless Weekly." The description emphasised the excellent finish of the cabinet, simple controls and a station-calibrated tuning control as a major feature. A rotatable frame aerial was fitted, being all that was necessary for interstate reception. It was also noted that at no time were there any "howls and squeals" to spoil reception of "music, sweet music." It added, "The volume control was a most useful refinement."

Around 1930 the Superhet became the standard type of circuit for domestic use because it offered the selectivity necessary to cope with the rapidly increasing number of stations. By this time many of the difficulties had been ironed out and superhets became virtually as easy to produce as the TRF type.

CHAPTER 19

Test & Measuring Instruments

Testing and measuring instruments — the moving coil meter and its use in measuring voltage, current and resistance — the multimeter or VOM — the electronic voltmeter, VTVM or SSVM — the cathode-ray oscilloscope, CRO or "scope" — oscillators and signal generators — measuring bridges — other less common instruments.

In many of the preceding chapters we have had occasion to mention a number of test and measuring instruments, such as the multimeter, the oscilloscope and the signal generator. It would therefore seem wise at this stage to spend a little time looking at the various measuring instruments used in electronics, to give the reader some idea of their operation and use.

A basic component in many measuring instruments is the moving-coil meter. This is a device which has an indicating pointer so arranged that, when a current is passed through the meter, the pointer moves along a scale by an amount which is directly proportional to the amount of current flowing.

Fig 1 gives the general idea, and may be used to explain how the moving-coil meter works. The heart of the meter is a rectangular coil of fine insulated wire, pivoted at opposite ends on jewelled bearings so that it can rotate. Current is fed through the coil via two delicate spiral springs, one at each pivot. The springs also supply what is known as the "restoring torque," which will be explained in a moment.

Attached to the coil is a long but extremely light and delicate pointer, along with a set of three short arms and small weights. These are used to counterbalance the pointer so that the whole rotating assembly remains balanced, irrespective of the position in which the meter is mounted. The pointer moves over a dial plate having a measuring scale printed upon it.

Behind the dial plate is a strong "horseshoe" — shaped permanent magnet; this is fitted with two pole-pieces which concentrate its magnetic field through the movable coil. There is also a cylindrical soft iron core arranged to be within the coil, but not to rotate. It is fixed, and serves to ensure that, no matter where the coil is, it always moves perpendicular to the magnetic field. This must be done if an evenly-spaced or "linear" scale is required.

In an early chapter, you may recall, we saw that a current flowing in a coil of wire produces a magnetic field; this is precisely what happens when current flows through the meter coil. Two magnetic fields are thus present around the coil — that due to the permanent magnet, and that produced by the current being passed through the coil.

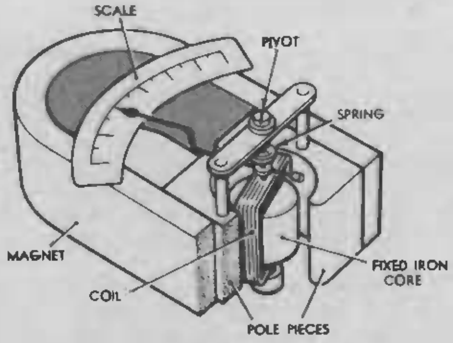
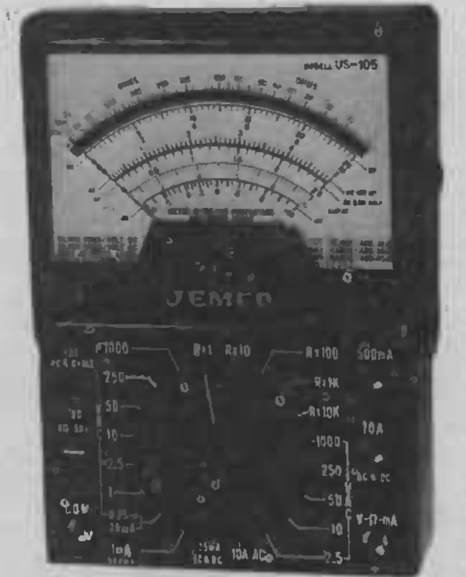


Fig. 1: A basic moving-coil meter known as a "d'Arsonval" movement. The pole pieces are cut away to show the moving coil over the fixed iron core. The coil and core are suspended in the magnetic field of the pole pieces, being free to rotate in the jewel pivots positioned in the front and rear supports.



A typical modern multimeter or VOM. It features more than 20 ranges. The meter movement is generally 0-100uA for the 10k ohms per volt type and 50uA for the more sensitive 20k OPV meters. Earlier types used an 0-1mA meter movement as described in the text.

The two fields interact, and the coil experiences a turning force or torque; it therefore tends to rotate and, in so doing, to move the pointer needle along the indicating scale. When it does this, the spiral springs begin to have an effect; coil rotation compresses one spring and expands the other, so that both springs tend to resist such movement. The further the coil rotates, the greater the force of "resistance" produced by the springs, which are all the time trying to restore the coil to its original position. Hence we say that the springs provide "restoring torque" as well as providing flexible connections to the coil.

The springs are thus fundamental to the operation of the meter. Because of their "proportional-to-angle-turned" restoring force, the coil and pointer are brought to rest at a point where the pointer needle indicates a value directly proportional to the actual current flowing.

If a strong current flows through the coil, its forward torque will be large and the coil will be able to rotate through quite a large angle before the restoring torque of the springs is able to counteract the current torque and bring it to a halt. On the other hand, a small current will only produce a small forward torque and the coil will only be able to move through a small angle before the springs are able to stop it.

For every value of coil current there will thus be a corresponding coil and pointer position, providing that the meter is not "overloaded" by passing through it a current greater than it is designed to handle.

The moving-coil meter can be used to measure very small orders of current. It can also be used to measure larger currents, as well as voltage and resistance, as we shall explain.

In passing, however, it should be noted that there are other types of basic meter movements besides the moving-coil type. These types are not as common as the moving-coil meter and, for this reason, they need not concern us here. It is sufficient merely to mention that they exist.

When we discussed Ohm's Law in an earlier chapter, we saw that resistances in parallel share any current which may be flowing. In fact, the proportion of the total current which flows through each of a number of resistances in parallel is inversely

THE OSCILLOSCOPE

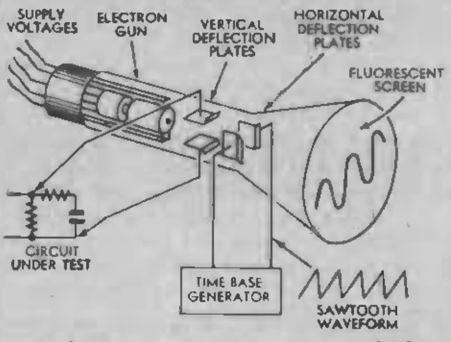


Fig 2: (above) Basic operation of the CRO or "scope". The horizontal plates are shown connected to the time base or sweep circuit and the vertical plates are connected to the circuit under test. Shown at left is a typical small instrument of this type. Note that the vertical amplifier control is calibrated in volts per centimetre sensitivity and the time base in milliseconds and microseconds.

— and exactly — proportional to their resistance. The lowest resistance will take most of the current, while the highest resistance will take least current, and so on.

Because of this fact, a moving-coil meter may be arranged to measure large currents. Its coil has a certain value of resistance and, by placing in parallel with it a smaller resistance called a "shunt," the meter coil will receive only a known minor part of the current.

A typical meter movement giving full-scale deflection of the pointer needle with only 1 milliamp through its coil can thus be arranged to read as a 1 amp meter, by wiring it in parallel with a shunt which takes 999 / 1000ths of the current flowing through the two. The shunt would simply have a resistance of 1 / 999 that of the meter coil so that, when 1 amp flows through the two, the shunt takes 999 milliamps and the meter receives its correct 1 milliamp.

If the total current should be less than 1 amp, the meter will read the same proportion of 1 milliamp. Thus ½-amp of total current would read half-scale on the meter.

When we discussed Ohm's Law we also saw that, when a voltage is applied to a resistance, a current flows which is proportional to the applied voltage and inversely proportional to the resistance. It is this fact which permits us to use the moving-coil meter to measure voltage.

All we have to do is connect the meter coil in series with a resistor, which is called the "multiplier." The multiplier is made to have a resistance which, when added to that of the meter coil, will draw the full-scale meter current when the intended full-scale voltage is applied to the two.

An example should again make this clear: Suppose we have a 1-milliamp meter and we want to use it to read from 0 to 100 volts. All we need do is work out from Ohm's Law the resistance which draws 1 milliamp when 100 volts is applied, which works out to be 100k (100 divided by 1 / 1000, giving 100,000). To find the multiplier resistor value, we simply subtract the resistance of the meter coil from this figure. Thus, if the meter coil has a resistance of 100 ohms (a typical figure) we will need a multiplier of 99,900 ohms in series with the meter to convert it into a 0-100 voltmeter.

In practice, if subtracting the meter resistance only affects the multiplier resistor value by the small amount shown above, it would be neglected. An accurate 100k resistor would be quite close enough as the multiplier.

In these circumstances, 100 volts applied to the combination would make the meter read full scale. With 50 volts applied it would read half-scale and with 25 volts, quarter-scale. By marking the meter scale in volts to 100, any voltage up to 100 can be read directly.

Different voltage ranges may be provided simply by selecting different values of multiplier resistance.

There are a number of different arrangements whereby the moving-coil meter may be used to measure resistance. The resistance-measuring circuit most often used is basically little more than a battery wired in series with the meter.

When an unknown resistance is connected into the circuit between the appropriate terminals, it completes the circuit and a current flows through the meter. The amount of current (and hence the meter reading) will depend upon the value of the resistance. Small resistors will produce a large current, and large resistors a small current; thus the scale of the meter can be marked in terms of resistance.

The moving-coil meter may also be used to measure AC, with the aid of a small rectifier circuit to change the AC into DC. Multiplier resistors may be added to the meter-rectifier combination to measure alternating voltage, while a small transformer is used to allow the combination to be used to measure heavy alternating currents.

The common "multimeter" or "VOM" (short for Volt-Ohm-Milliammeter) is simply a meter movement fitted into a case along with a variety of current shunts, multiplier resistors and resistance measuring circuits, to enable it to be used to perform a wide variety of measuring tasks.

It is usually provided with switches to select the various shunts or multipliers, etc. Alternatively, it may have a series of pinjacks or terminals to which the test leads may be connected by the user, depending on the measuring job to hand.

The multimeter is one of the most useful instruments in electronics. A modern instrument of the type shown in the photograph may provide thirty or more different measuring ranges, covering voltages (AC and DC) from a fraction of a volt to many thousands of volts; currents from microamps to amps; resistance from a fraction of an ohm to many megohms. It may also provide built-in protection for the meter, to guard against damage due to improper setting of the controls.

Because the moving-coil meter is essentially a current reading device, the multimeter always draws a small current when being used to measure voltage. With modern, sensitive meter movements this metering current may be as low as 10 or 20 microamps for full-scale deflection, but even this current can load some circuits unduly and produce reading errors.

It was principally because of the loading imposed by the moving-coil multimeter that electronic voltmeters and multimeters were developed. These use either valves, bipolar transistors or FETs to increase the sensitivity of the basic moving-coil meter, at the same time giving it a very high effective input resistance so that it does not seriously load even high resistance circuits being tested. Typically the input resistance of an electronic voltmeter is around 10 megohms, while special types may go as high as several hundred megohms.

The first type of electronic voltmeter to be developed was the vacuum-tube voltmeter or "VTVM", also called the valve voltmeter. This generally used two triode valves, or a twin triode, in a balanced circuit. Electronic voltmeters and multimeters using valves are still available, but most modern instruments of this type use either bipolar transistors or FETs, and thus go under the name of solid-state voltmeter or "SSVM".

The principal disadvantage of VTVMs and other electronic meters using valves is that considerable power is needed to run the valves, so that they must generally be operated from the AC mains. This limits their versatility compared with a simple moving-coil multimeter. Solid state voltmeters are less subject to this limitation, however, as they can operate for quite long periods on a small battery supply. Modern SSVMs are in fact fast becoming used for many of the jobs formerly handled by the multimeter, because of their higher input resistance.

Multimeters and electronic voltmeters permit the measurement of voltages, currents and resistances in circuits, but they do not allow us to see the ways in which currents or voltages may be changing — unless the changes are taking place very slowly. The cathode-ray oscilloscope or CRO, also called the "scope", is an instrument which allows both fast and slow changes to be seen. It may also be arranged to measure voltages and frequencies.

The heart of the CRO is the cathode-ray tube, which is a small-scale version of the picture tube used in television receivers. In addition to the size difference, the cathode-ray tube used in most oscilloscopes uses what is termed electrostatic deflection, rather than the magnetic deflection used by television picture tubes, and has no "yoke" mounted on its neck.

Fig 2 should help in understanding how the cathode-ray oscilloscope works. The tube consists of three main parts — a group of electrodes called the "electron gun," two pairs of flat electodes called the deflection plates, and a fluorescent screen.

It is the screen which is visible at the front of the oscilloscope, and it is on the screen that a "picture" of circuit voltage or current changes appears. Repeating changes, like those of alternating voltages and currents, produce fixed wave-line patterns on the screen — hence the use of the term "waveform" to describe the screen patterns and the circuit variations which they represent.

The purpose of the electron gun is to produce a fine stream of electrons aimed at the fluorescent screen. The gun has a heated cathode similar to that in a normal valve, and a system of cylindrical and disc electrodes used to control and guide the electrons into a narrow beam.

If the deflection plates were not there, or if they were not connected to anything, the electron beam from the gun would strike the centre of the fluorescent screen and cause the phosphor powder at that spot to glow. All that would be visible would be a small bright dot at the centre of the screen.

Consider what happens to the beam when one of the pairs of plates — say the pair marked "vertical deflection" — is connected into a circuit so that a changing circuit voltage appears across the two plates.

The two deflection plates are in effect a parallel-plate capacitor, and the voltage difference between them will set up an electric field in the space between them — through which the electron beam is travelling. As the voltage of the circuit varies, the electron beam will therefore find itself in a varying electric field.

Electrons, it will be recalled, are negatively charged, and the electrons of the beam will thus tend to be moved or "deflected" by the electric field toward the more positive plate. The beam is travelling quite fast, so the electrons will not actually be able to reach the more positive plate, but the beam will be bent in the direction of the plate and the electrons will hit the fluorescent screen at a new spot somewhat removed from the centre of the screen. The exact distance moved will depend upon the voltage applied to the two plates.

As the voltage of the circuit varies, the bending of the beam will also very and the glowing spot on the screen will move up and down in sympathy. Whenever the top plate is more positive than the lower plate, the beam will bend upward and the spot will move up. Conversely, when the lower plate is more positive than the upper plate, the beam will be deflected downward and the spot will move below the centre of the screen.

The amount of beam deflection produced at any instant will be proportional to the circuit voltage present at that instant, and so the distance moved by the spot on the screen will be directly proportional to the circuit voltage at all times. If too much voltage is applied to the plates, the beam will be deflected right into the glass neck of the tube, and the spot will disappear off the top or bottom of the screen.

So far, the cathode-ray tube is simply acting like a meter with an electron-beam indicator "needle". But here is where the second set of plates come in — those marked "horizontal deflection." These are very similar to the first set, but are closer to the screen (for mainly physical reasons) and are turned sideways so that any voltages applied to them will tend to move the beam and spot horizontally.

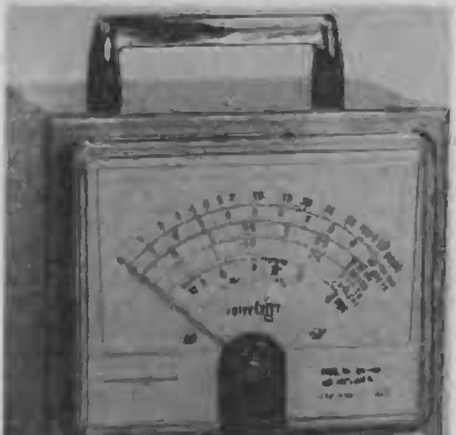
A circuit, called the "timebase generator", applies, to this second set of plates, a voltage which changes linearly (smoothly) for a certain period then drops back to its inital value, then changes linearly again, and so on. The waveform of this voltage is thus shaped like the teeth of a rip-saw, and is accordingly called a sawtooth sweeping voltage.

The effect of this sweeping voltage is to move the beam and spot smoothly across the screen from one side to the other, then quickly back again, then smoothly across again, and so on. The speed at which this occurs can be adjusted over a wide range by controls in the timebase generator circuit.

By this horizontal movement of the beam and spot, the timebase waveform "spreads out" the up-and-down spot motion produced by the signal so that it can be seen.

In effect, the cathode-ray tube plots a graph of the test voltage compared with time; the time represented by one sweep of the time base waveform can be worked out or measured.

The cathode-ray oscilloscope thus allows us not only to see the circuit voltages changing, but to measure by how much they change — given by the height of the pattern — and how long they take in changing — which may be deduced from a knowledge of the period of time represented by one sweep of the timebase waveform. It is thus an extremely useful instrument.



A modern solid-state voltmeter or *SSVM, using FET devices.*

Most modern oscilloscopes include an amplifier (called the "Y" or "vertical" amplifier) to enable very small circuit voltages to be made large enough to produce a visible deflection of the spot. The amplifier has a switch to allow the selection of various amounts of amplification, and the switch is marked directly in terms of the amount of input voltage which is represented by 1 centimetre of vertical spot deflection. Voltages can thus be measured quite easily.

The speed of the timebase generator is adjustable by means of other switches, and these again are marked directly in terms of the period of time (in seconds, milliseconds or microseconds) represented by 1 centimetre of screen width. Time duration and frequency can also be measured quite easily.

In passing, it should be noted that some oscilloscopes, notably the older types and the simpler modern types, do not have such "calibrated" vertical amplifier and timebase controls. They often have just a variable gain control on the vertical amplifier and have timebase controls either unmarked or marked in terms of the approximate timebase sweeping frequency (in Hertz). Such instruments are intended mainly for "looking" at circuit goings-on, and are not really suitable for making measurements.

Most oscilloscopes have knobs to control the brightness and focus of the spot, and to set the spot to the centre of the screen when voltages having a large steady component are being measured. These latter are called the "shift" or "centering" controls.

Oscilloscopes are also fitted with circuits to enable the timebase to be locked or "synchronised" with the voltage under inspection so that the screen pattern is held steady. Depending upon the exact type of circuit used, the controls associated with this feature may be marked "synch." or "triggering" or "locking" adjustments.

In some of the preceding chapters, we have referred to sources of RF alternating voltages called variously modulated RF oscillators or RF signal generators depending upon their degree of refinement. It was explained that such devices are basically valve or transistor type oscillators, provided with a tuning capacitor and various coils to set the desired frequency band.

They usually include provision for modulating the RF voltage with a fixed audio tone (usually 400 or 1,000 Hz), and with a reliable control over the amount of output voltage delivered — the so-called "attenuator."

In laboratory parlance, the name "Signal Generator" is usually reserved for instruments which have a meter to monitor the output voltage and the degree of modulation present, and an output voltage attenuator capable of setting the output to a known level between a fraction of a microvolt and a few volts.

Just as there is a need for instruments able to supply RF voltage, there is a similar need for instruments able to supply low and audio-frequency alternating voltages. Such instruments are known as audio oscillators or audio generators, the latter being the counterpart of the RF signal generator in terms of accuracy and refinement.

One early type of audio oscillator was the so-called "beat frequency oscillator" or "BFO", so named because it generated the desired audio signal by beating together two RF signals. This type is now rarely used because it tended to have poor frequency stability and a rather high distortion level. Most modern instruments are based on oscillators which generate the audio signal directly, a common type being based on the so-called "Wien bridge" network. This consists of a network of resistors and capacitors connected in a feedback loop around a high-gain audio amplifier.

Modern instruments of this type can typically deliver audio signals from as low as 1Hz to above 200kHz, at levels from a fraction of a millivolt to many tens of volts. The output waveform may be either a sinewave, with perhaps a distortion level as low as .01%, or other waveforms such as square wave or triangular wave.

There are many occasions when resistors, capacitors and inductors must be accurately measured to determine their value of resistance or reactance. While a reasonably accurate measurement can be made using various circuits provided on a multimeter or electronic multimeter, it is usually necessary to employ what is known as a measuring bridge in order to make really accurate measurements.

Basically, a measuring bridge is a device which balances the unknown resistor, capacitor or inductor against a known or standard unit, to give an indication of the relative value of the unknown component. This type of measurement is accurate because the bridge simply performs a comparison between the unknown unit and a highly accurate standard unit; it does not rely upon the voltage of an internal battery or oscillator, or the accuracy of a meter. If a meter is used in the bridge, it is simply used as a balance indicator and not used as a measuring device.

For the measurement of inductance a rather elaborate bridge is required, whereas quite a simple bridge can measure resistance and capacitance fairly accurately. For this reason most of the simpler measuring bridges are called R-C bridges to signify that they are really only suitable for measuring resistance and capacitance.

The R-C bridge shown in the photograph has inbuilt standard resistors and capacitors for six comparison ranges, with a pair of terminals provided so that additional standard resistors or capacitors may be connected if desired. The comparison is carried out at a frequency of 50Hz and a small meter used to indicate when a balance is achieved. On older bridges this function was served by a special valve known as a "magic eye" or electron-ray indicator which indicated balance by means of two overlapping fan-shaped glowing segments on a fluorescent screen.

The bridge is adjusted for a balance by means of the large dial knob, which effectively "tries out" various ratios between the unknown and standard components. When it finds the ratio which produces a balance, the dial reading gives the value of the unknown component relative to the standard. If it reads "0.5" when the standard is 100pF, for instance, the unknown capacitance is 50pF. If it reads "3.4" with the same standard, the unknown would be 340pF.

The instruments which we have looked at so far are perhaps the most common types which are met in electronics. Before we leave this topic, however, it might be worthwhile briefly mentioning a few of the many not-so-common instruments.

The sweep and marker generator is an instrument used in the alignment of TV receivers and similar applications. It is in effect two RF oscillators or signal generators in one. The sweep section generates an RF signal which is swept back and forth in frequency, and may be used in conjunction with a CRO to show how tuned circuits and amplifiers behave over the band of frequencies being swept. The marker is a fairly normal RF oscillator used to identify or "mark" the various frequencies being swept through.

Valve testers and transistor testers are instruments used to check valves and transistors for proper operation. Simple types may only indicate the difference between useless or very poor units and those which should operate more or less normally. The more elaborate instruments place the valve or transistor being tested under its correct operating conditions and measure just how well it performs.

Distortion meters are audio testing instruments which may measure one or more of a number of different types of signal distortions. They are often combined in the one case with an instrument called a millivoltmeter — which, as the name suggests, measures very small alternating voltages. The millivoltmeter is very useful for measuring the performance of microphones and gramophone pick-ups, as it can measure their output voltage directly.

A typical small R-C measuring bridge. The value of an unknown component is read on the large dial.

A modern digital multimeter made by Hewlett-Packard. Readout is by light-emitting diodes or LEDs.

Dip oscillators are small RF oscillators which have externally-mounted resonance coils and a meter which indicates their strength of oscillation. They are used to determine the tuning frequency of resonant circuits in receivers and other equipment. When the coil of the dip oscillator is brought near the coil of the unknown tuned circuit, the frequency is varied until the meter indicates that the test circuit is absorbing some of the oscillation energy — this shows as a "dip" in the meter reading. The dip oscillator dial then indicates the resonant frequency of the unknown circuit.

The first dip oscillators used valves, and were called "grid-dip oscillators" or "GDOs" because the strength of oscillation was monitored by measuring the oscillator valve's grid current. Nowadays, most dip oscillators use either a bipolar transistor or a FET in the oscillator, resulting in a smaller and more flexible instrument.

The signal tracer is effectively a sensitive radio receiver fitted with a switch which permits signals to be fed into it at any of the various points along the signal path. It is used to follow the path of radio signals through a receiver under test, in order to find out speedily the section of the receiver which is faulty.

Digital counters or digital frequency meters (DFMs) are instruments used to measure frequencies and time periods, and to count pulses. They operate by converting the input signals into a stream of rectangular pulses, and counting the number of these pulses occurring in a known period of time. The counting is performed by groups of circuit modules known as "flip-flops", which are bi-stable circuits capable of flipping from one stable state to the other when triggered by a pulse. Readout of the count is in the form of glowing numbers, usually from either special gas discharge tubes or from arrays of light-emitting semiconductor diodes or "LEDs".

Digital voltmeters or "DVMs" are electronic voltmeters which, like digital counters, operate by counting rectangular pulses and displaying their reading in the form of glowing numbers. However, in a DVM an input circuit is used to produce a group of pulses whose number is directly proportional to input voltage, so that when they are counted the reading may be arranged to indicate the input voltage. The voltage-to-pulses conversion is generally performed by a circuit which generates a voltage "ramp" which is linearly rising or falling from zero or a set value, and opens a digital "gate" to pass the pulses until the ramp voltage just equals the voltage to be measured.

Although we may seem to have looked at quite a number of test and measuring instruments in this chapter, there are a great many more that we have not been able to mention. All sorts of test and measuring instruments have been developed in order to make the job of the electronics worker a littler easier, faster and surer.

In a later chapter we will discuss instruments in relation to fault finding.

CHAPTER 20

The Electronics Serviceman

Electronics servicing — A servicing career — Training to be a servicemen — "Trades" courses, private study, private colleges. Avoiding the "dead end" job. Practical experience and putting training to work. Servicing in practice. Diagnosis. Typical aquipment. The multimeter. Sensitivity and loading problems. The electronic voltmeter. The service oscillator. The sweep generator. The cathode ray oscilloscope. The audio generator. Transistor testars. Service data.

Servicing is a vitally important part of the whole electronics industry and, today, this embraces a much wider field than that of radio and television sets. Electronics, in one form or another, has invaded almost every phase of our modern life and, wherever there is a piece of electronic equipment, there is a need for someone with skill and knowledge to maintain it in proper working order.

This fact holds a promise not only of a continuity of service work, but of increasing scope.

But what is a radio — or should we say an electronics — serviceman? How does he acquire this status? What are his qualifications? How does he go about his job? What equipment does he use? These, and similar questions, are invariably asked by members of the rising generation who are seriously considering a career in electronics, and to whom servicing may have an appeal.

It is the purpose of this chapter to try to answer these questions, at least in a general way. Fairly obviously, we cannot tell you how - to - be - a - serviceman - in - one - easy - lesson; no one can hope to do that. But we can provide some useful background, which will better enable the reader to appreciate what is involved if he chooses such a calling.

For those who are seeking something other than the routine of factory mass production but who feel that, for various reasons, they cannot advance to engineering level, servicing could be a wise choice.

In spite of all the attempts to make it otherwise, a good serviceman still has to think and, if he has any liking for his job at all, there is little chance that he will stagnate technically. On the contrary, the rapid growth of electronics provides a continual challenge to move beyond the routine to more specialised — and more remunerative — fields.

How does one become a serviceman?

At present, there are few, if any, hard and fast rules but, these days, anyone who shows an aptitude for the job should have little difficulty in finding employment which will provide the necessary practical training. At the same time, the would-be serviceman must be prepared to study diligently if he really wants to get somewhere.

What are his qualifications?

Again, there are few hard and fast rules, a fact which many people deplore, believing that a certain minimum standard of proficiency should be demanded before a person is allowed to offer his services to the public. Be that as it may, the truth is that, in this country, there is very little to stop almost anyone calling himself a radio / TV serviceman.

Whether he can maintain such a front is quite another matter, involving honesty, personality, business ability — and technical knowledge.

However, for his own benefit and that of future customers, the aspiring serviceman should consider nothing less than a thorough basic training at around trade-course level, plus as much practical experience as he can get. There is little doubt that a good serviceman is a balanced combination of basic training and practical experience, with each aspect as important as the other.

In line with this, the would-be serviceman should aim, as a minimum requirement, at passing the examination at the end of three or four years of secondary school, which normally qualifies for a tertiary "trades" course, apprenticeship and so on.

The difficulty then arises as to where to go from there. For most other trades, there is an established pattern of apprenticeship and training but, because of its relatively short history, the "tradesman" pattern in the electronics industry is still very much in the formative stage.

Nevertheless, at many centres where tertiary education is available, students should be able to find courses offering a substantial content of electrical and electronics fundamentals. They may not be aimed at producing certificated electronics servicemen and may, in fact, produce a trade qualification of a quite different type. But such qualification will be no load to carry and it will provide a very useful practical and theoretical basis from which a career in servicing can be developed.

Vocational guidance officers are available to assist in choosing the best course at the various centres.

By and large, "trade" courses are not so onerous as to prevent the willing student from supplementing the official syllabus with parallel reading and some kind of parallel practical work.

The latter is most important if the student is to develop any tie-up between theory and practice.

Do-it-yourself projects, such as described in "Electronics Australia," are one way of gaining additional experience. Linking up with a local radio club or an amateur operator are other possibilities. Occasionally, local radio dealers will offer sparetime bench work to trainees, working under supervision.

If circumstances are such as to make it impossible for a trainee to attend a Government-sponsored college, training at a private institution can be considered, with a preference to those which offer practical sessions in addition to lectures and / or correspondence.

In fact, there is something to be said for doing a private servicing course, more or less in parallel with a trades course, where the latter is only obliquely related to electronics servicing.

The vital point is for the intending serviceman to be committed to some definite course of study, involving fundamentals, which will impose upon him obligation and discipline. "Private study" can lapse all too easily, added to which the student has nothing tangible to display for his efforts, when later applying for a position.

Strangely enough, and despite all this, there are still people who believe that basic principles, or "theory" as they are inclined to disparagingly call it, is unimportant to a "practical" serviceman. In fact, the "valve jockey" so often encountered in the TV servicing field, relies for his existence on this kind of thinking.

While service organisations can make a case for "valve jockeys" on economic grounds, this approach could pose a serious hazard for the younger would-be serviceman. While there is no valid objection to anyone starting as a "valve jockey," such a position can easily turn out to be a "dead end" job, unless the individual is encouraged to educate himself out of it.

In five years time he will still be a "valve jockey" — a somewhat more experienced one perhaps — but still working by rote to a monotonous formula, with all his advancement (if any) behind him, and in a job which, for him, holds no future.

In fact, with the advent of solid state techniques, and the gradual phasing out of valves, his future looks bleak indeed.

Summing all this up, we might say that training should involve all the elementary basic theory — the nature of resistance, inductance and capacitance; the behaviour of tuned circuits, RC networks, etc; the operation of valves and transistors and so on. On top of this should come the more practical training; how the basic theory applies to the practical circuits of radio and TV receivers or, in short, how they work.

Thus a serviceman should be able to explain, at a fair technical level, just how any section of a receiver works. And, arising out of this, he should be able to make a reasonable estimation of what would happen if a particular component failed in a particular manner. In short, he should begin to anticipate faults as his training progresses.

The other aspect of this training is practical field experience; an excellent teacher for those equipped — by study — to learn. This is where the student learns the things not found in the ordinary text book: That Blank and Co.'s model X123 TV set used a batch of vertical oscillator transformers that were unusually prone to breakdown; that poor sensitivity in a certain radio receiver is almost always due to faulty resonating capacitors in the IF transformers; or that crystal pickups, where still in use, often fail in hot, humid weather.

Similarly, it does not take the astute person long to realise that most sets have weaknesses which are peculiar to them by reason either of the circuit they employ or the components they use. Knowledge of this kind helps enormously in rapid diagnosis and repair of the more routine faults.

This last statement may suggest that such an approach to diagnosis is no more than an extension of the valve jockey idea; that one simply learns what are the most likely faults in each make of set, and equips oneself to treat the set on this basis.

But there is a vital difference. The valve jockey type is "through," once he has exhausted his fund of experience. Unless he can fluke a cure on a costly trial-and-error basis, he has nothing to fall back on. He must either waste still more time or seek assistance.

The well-trained man, on the other hand, expects that he might have to search for the fault. To be sure, he will first investigate the most likely possibility, based on experience; to do otherwise would be foolish. But when he draws a blank, in terms of expected faults, that is when he really starts to work.

Knowing how this part of the set should function, he combines his basic training, his experience with voltages or waveforms normally found in this part of the circuit, and his ability to make and interpret measurements, to go over the section systematically until, ultimately, he is rewarded.

This, then, becomes experience on which he can draw, next time, to do a quicker job. In short, the trained man constantly adds to his own experience. The valve jockey can do so only on a very limited basis, with the additional hazard that his "knowledge" will be tainted with a good many wrong ideas.

A serviceman checks the voltages in an old-style valve TV set. Ready to hand is a book of circuits which carry information on the voltages and waveforms to be expected.

It is that practical experience that helps to consolidate basic training, so that the student not only learns but understands. And when he understands he is well on the way to becoming a good serviceman and a good technician, with a latent ability to tackle other tasks.

How does a serviceman go about his job?

Whether he realises it or not, every serviceman attacks each job in three stages: First he notes the symptoms; from these he diagnoses the fault; then he effects the necessary repair.

Even the valve jockey works this way. In his case, however, the symptoms noted will normally be confined to the most obvious ones, from the diagnosis to nominating the section of the set at fault, and the repair to replacing the appropriate valve or valves.

This brings us to the **first requirement of good servicing — the ability quickly and accurately, to diagnose a fault.**

Perhaps this may appear to be stating the obvious, but it is not always fully appreciated that diagnosis is usually the most difficult part of the job. Almost anyone who is handy with pliers, screwdriver and soldering iron can replace a faulty component once it is nominated, but it often takes real skill to nominate it.

The need to do the job in a reasonable time is fairly obvious; time is money in business and anyone who consistently takes longer than necessary to locate a fault will soon price himself out of the market — or starve!

Accuracy is equally important. A wrong diagnosis means wasted time even the most experienced will go off on a false trail sometimes. The person who consistently needs several bites at the one problem will prove as costly as anyone who is too slow for any other reason.

But even more important is the question of customer satisfaction. Wrong diagnosis can involve the customer in unnecessary expense or, possibly worse, result in a fault seeming to be fixed when it is not, with consequent customer inconvenience and dissatisfaction. The latter occurs most frequently when the fault is of an intermittent or nebulous nature.

What makes a good diagnostician?

There are many things. Fairly obviously he needs adequate training, adequate experience, and adequate testing facilities. But most of all he needs an inquiring mind and an ability to reason things out for himself; an ability — and a willingness — to collate all the signs, symptoms, and history of a fault and then, from basic principles and experience, to nominate its most likely cause.

This implies, in turn, that he has the ability to collect all these data, and this is also part of his training. Keen observation plays a large part, but he must also learn how to ask questions in a pleasant and tactful manner, and without risk of antagonising the customer.

Most "signs" can be taken in at a glance, almost automatically. They are all the things we observe about a set, other than its actual performance; whether it has suffered physical damage, whether it appears to have been tampered with, the type of aerial it uses, the presence of appliances which might cause interference, and so on.

Symptoms refer to the set's actual performance. The more accurately the serviceman observes them, the better his chances of a good diagnosis. More important, however, is the desirability of having the customer demonstrate the symptom about which he is complaining. If every serviceman made this a rule, there would be fewer cases of customers being charged for something about which they didn't complain, while the "important" fault remains uncorrected.

In practice, history is all that you can encourage a customer to tell you about the set in general, and the fault in particular. Much of it will be valueless prattle, but it is surprising how often they let drop a vital clue. And if it is one that draws attention to an obscure or intermittent fault, then it may well mean the difference between holding and losing a customer.

What equipment does a serviceman need?

As with the other aspects of servicing, there are no hard and fast rules as to what constitutes the ideal combination of test instruments. At the one extreme we have the type who used to boast — mostly before the days of TV — that he could service any set with a screwdriver. To him, the need to use even a multimeter was a sure sign of incompetence!

At the other extreme we have the type who buys — or wishes he could buy — every piece of test equipment in the catalogue. He firmly believes that if he can only acquire enough pieces of equipment he will become a top-notch serviceman, lack of training and / or experience notwithstanding.

Fairly obviously, the ideal lies somewhere in between. Too little equipment makes for inefficiency and probably leads to secondary faults, and faults in the making, being overlooked.

For example: It may not be difficult for an experienced person, working on a conventional radio set, to diagnose that the reason the set failed is because the output valve has packed up. So, all that is necessary to get the set going again is to replace the output valve.

But how is he to know, without the aid of some equipment, that the HT voltage is significantly below par, probably because the rectifier is on the way out too? Detection of such a condition is something the customer has a right to expect, otherwise he is likely to find himself involved in the inconvenience and expense of another service call in only a few months' time. Such a situation does not enhance the serviceman's image in the customer's eyes.

At the other end of the scale, the "buy everything" type ignores two vitally important fundamental facts: that equipment alone never made a serviceman and that, in the hard world of business, every piece of equipment in the workshop will have to pay for itself in a reasonable time.

In short, what the individual finally buys in the way of test equipment will be a compromise between what he would like and what his business can afford — the latter based not so much on available cash, but on what the equipment will earn.

All this simply means that some pieces of equipment are essential, some are highly desirable but not essential, and some virtual luxuries. But just where the line of demarcation occurs is a matter for individual decision. About all this article can do is discuss the more usual items in the approximate order that a serviceman might consider them.

In so doing, we would remind readers that in a previous chapter we discussed test and measuring instruments in their own right. There it was possible to cover many of the finer points of design which we cannot hope to cover in this chapter. It is suggested that the reader refer to this chapter for detailed explanations of the operation of measuring instruments.

One piece of equipment, the worth of which no one is likely to question, is the multimeter. This is an essential item, used for tracking down a large percentage of routine faults. About the only decision needing to be made is the type to purchase.

A popular version in bygone years was based on a 0-1mA meter movement and provided 12 basic ranges; four DC voltage, four DC current, and four resistance. In addition, a rectifier provided AC voltage ranges, and an "output" version of the latter was normally provided.

Such a meter was — and still is in some cases — the favourite instrument of the serviceman. It is compact, relatively rugged, and will cope adequately with the majority of voltage measurements in a radio and TV set. Even when something better is acquired, it still comes in for a lot of routine work.

Nevertheless, it does have limitations. When used as a voltmeter it has a sensitivity of only 1,000 ohms per volt and this can lead to serious measuring errors in certain parts of a circuit.

The sensitivity of a meter is a measure of the degree to which it loads the circuit under test or, looking at it another way, the amount of power it must take from the circuit in order to move the pointer. Whether the circuit can supply this power, without seriously altering its operating conditions, is the point to be determined when considering the suitability, or otherwise, of the meter.

The "ohms per volt" refers to the resistance of the meter — as seen by the circuit under test — relative to the full scale value of the voltage range selected. For example: In the case of a 1,000 ohms per volt meter the 10V range would have a resistance of 10,000 ohms (10k), the 50V range 50k, and so on.

This is a very practical way to express sensitivity, because it is usually fairly easy to assess the resistance of the voltage supply being measured. Provided this is significantly lower than the selected meter range, the error introduced by the meter's presence will not be serious. However, when it approaches or exceeds the meter resistance, the error will be gross.

Typical examples are the screen and plate circuits of an ordinary resistance coupled audio stage, say a pentode with a 0.25M plate load resistor and a 1.5M screen dropping resistor. Here, the plate voltage supply has a resistance of at least 0.25M and the screen voltage supply at least 1.5M.

If we attempt to measure the plate voltage, using, say, the 250V range of the meter, we find that the supply resistance of 0.25M is the same as that of the meter, and the latter will create considereble error. If we drop to the 50V range, which is not unreasonable, the meter resistance is only one fifth that of the supply resistance, and the meter reading becomes quite meaningless. It need hardly be stressed that the position is just so many times worse in the case of the screen voltage.

In practice, most servicemen overcome this problem, to a degree, by familiarising themselves with typical readings made in these circumstances, error notwithstanding. These then become the voltages one expects to read at these points if all was well. The fact that the readings themselves are wrong matters little.

The AGC voltage is another having a very high resistance source, but it is also one which can reveal a lot about a set's behaviour if we can measure it accurately. Unfortunately, it is not possible to read this with even an allowable error, using the conventional 1,000 ohm per volt meter.

This limitation of the simple multimeter has been emphasised for two reasons: to provide a logical background against which the advantages of more elaborate instruments can be considered, and to introduce the idea that the presence of any test instrument — multimeter, CRO, generator, tracer etc. — can, in certain circumstances, upset the behaviour of the circuit under test. An important lesson a serviceman has to learn is to recognise such a possibility and take precautions against it. This includes selecting test instruments which are least likely to create such problems.

But to return to the multimeter. There are two broad approaches to improved sensitivity, (1) substitute a more sensitive meter movement or (2) interpose some form of amplifier between the meter and the input to the instrument and use the amplification so provided to improve the sensitivity.

In the first case we may typically substitute a 100uA (0.1mA) or 50uA (.05mA) for the previous 1mA movement, giving an increase in sensitivity of 10 or 20 times respectively. (10,000 or 20,000 ohms per volt.) Either represents a very significant improvement and will cope with most measuring problems without significant error.

A further worthwhile improvement is the extension of the ohms ranges by a factor of

This service department, operated by a distributor of high fidelity sound equipment provides an object lesson in tidiness and organisation. Test equipment is ranged on shelves above the workbench, while the lighting is bright but even. Note the drawers at right for holding spare parts and drawers under the bench for cables and manuals. Not only does an orderly situation save time in the long run, but it encourages a better standard of workmanship.

10 (or 20) times. Thus, where a 1mA meter will normally be limited to reading about 1M maximum resistance with a convenient size battery, a 100uA movement will extend this to 10M.

While such instruments were relatively expensive in the past, the position is a good deal better today. As a result, they are a popular choice, either in place of or, better still, addition to the 1,000 ohms per volt variety.

The advent of transistors, incidentally, has introduced another factor into the formulation of an "ideal" multimeter. Whereas valves involve voltages widely spread up to 300 or more, transistor circuits seldom operate above about 50 volts. The voltages at various points in the circuit may be quite low and, for easy testing of transistor equipment, it is an advantage if a multimeter has ranges of the order of 0-1.5V and / or 0-3V, instead of earlier practice of 0-10V as the most sensitive range.

A valuable addition to any multimeter, high sensitivity or not, is a system of protective diodes. Properly fitted, as a part of the overall instrument design, they make a meter virtually "bash proof."

Approach No 2, the use of amplification, gives us the electronic voltmeter; the solid state version of what was called the VTVM (vacuum tube voltmeter) a few years ago. In either version, these instruments normally have a constant input resistance of around 11M, regardless of the voltage range selected. This means that its sensitivity is about on a par with a 10,000 ohms per volt meter on the upper ranges (1,000V) but is many times better in the lower ranges. This makes it particularly suitable for checking the low voltage, high resistance circuits.

The electronic voltmeter is a naturally protected meter, by reason of the amplifier operating conditions and it is also capable of measuring a very wide range of resistance values.

A minor disadvantage of the majority of electronic voltmeters is that they include only voltage and resistance ranges, with no provision for measuring current.

On the other hand, a major disadvantage of the older VTVM — the need to operate them from a power point — no longer exists, making them much more suitable for field work than was formally the case.

While the ability to measure voltage, resistance and, occasionally, current, will enable a serviceman to track down a good many faults, not all faults manifest themselves as a significant change in one of these characteristics. Many are a good deal more subtle and need other equipment to help track them down.

In radio servicing, the service oscillator or signal generator is one of the most used items, next to the multimeter. Although primarily intended as a means of alignment, it plays a much wider role in practice.

A good generator will deliver a signal suitable for feeding into any part of the receiver, from the speaker to the aerial terminal, at audio, intermediate, or radio frequencies, and at a known level. Using these facilities, an experienced serviceman can quickly isolate a faulty stage, whether it be completely dead or simply below a reasonable standard of performance.

Many radio servicemen use the generator as the first line of attack, to pinpoint the faulty stage. If a valve set, they then test the appropriate valve — often by substitution and using the generator signal as a basis for comparison — and, if this does not completely cure the trouble, attack the stage with the multimeter.

Thus the signal generator provides a quick means of isolating a faulty stage, to show up poor performance due to other causes, and to put the set into top working order.

The technique of signal injection, using a signal generator, has become even more popular with the advent of solid state receivers which, due to their compact construction, printed wiring, and general unfamiliarity, often appear more difficult to service than the older, valve types. On a value for money basis the signal generator is well worthwhile.

In the TV field, the equivalent is the sweep generator, used in conjunction with a CRO. However, it is seldom used as an aid to diagnosis — except, perchance, when the fault lies in the tuner or IF strip — and is usually reserved for its intended role, that of alignment. This is something that the average TV serviceman avoids as far as possible; it normally involves returning the set to the service bench, and can be tricky and time consuming.

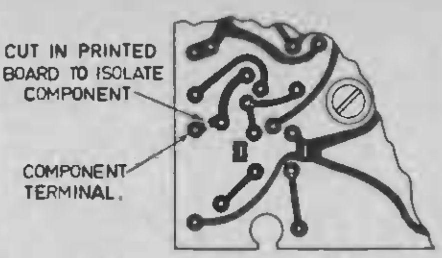
As a rule, the faulty section of a TV set is usually fairly evident from the set's behaviour, enabling the serviceman to start with the faulty area already well defined. However, a sweep signal source may sometimes be useful when tracking down some of the more obscure faults.

Mention of the sweep generator leads naturally to its companion instrument the CRO — or cathode-ray oscilloscope. For TV servicing this is an essential item, not only as a companion to the sweep generator, but also as a test instrument in its own right. A glance at the waveforms on any manufacturer's service diagram will explain this latter point. These represent correct operating conditions and any significant departure from them generally indicates a fault.

Ideally, a CRO should be capable of displaying accurate waveforms up to about 5MHz, the video bandwidth of our TV system. However, if such an instrument is too costly some compromise is necessary. A good minimum figure would be about 1MHz, with 2MHz a better one if it can be managed.

At one time the CRO was regarded as being of little value for radio servicing, and it is true that the average radio set seldom needs its help. However, the rapid growth of high-quality audio systems has changed the picture markedly, and anyone who hopes to service this kind of equipment will find it essential. Remember, however, that it needs some form of audio generator as a companion instrument, and one suitable for the job might cost nearly as much as the CRO.

A service instrument which had to be considered in the past was the valve tester, but very few servicemen regarded it as essential. In practice most prefer, whenever possible, to test a valve by substituting a known good one in the receiver socket, and judging the results on performance. Performed intelligently, such a test is normally quicker and more accurate. Furthermore, since servicemen normally carry a stock of likely valves when on outside service, such valves provide a means of testing without the need to carry a bulky instrument.



A practical tip from the service data of an Australian-made receiver. To isolate a component for a particular test, the manufacturer suggests cutting the copper track on the board with a sharp knife. After the test, the gap can easily be bridged with solder.

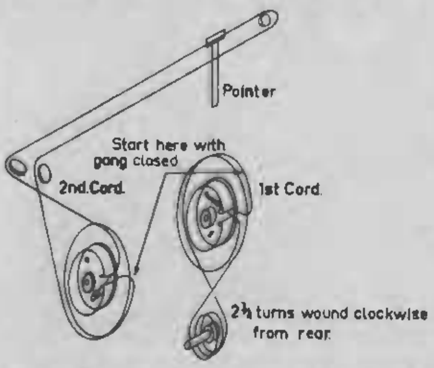
These days, with solid state devices progressively taking over from valves, the budding serviceman might well think twice about investing in such an instrument, particularly as they are fairly expensive.

The natural equivalent is the transistor tester, although the situations are not exactly parallel. Valves are often the first thing tested, partly because they are easy to remove, and partly because their life is known to be limited. Transistors on the other hand, tend to be the last things tested, partly because they are wired directly into the circuit and partly because they do not wear out, in the same sense as valves do.

As a result, most servicemen tend to test around the transistors, check their operating voltages, and remove them for test only when most other possibilities have been discarded. Once again, most servicemen seem to favour direct replacement as the most satisfactory and convenient test. On the other hand, a tester can be a handy item when a replacement is not immediately available, or when trying to determine the characteristics of foreign type number for which no data are available.

For these reasons, and because simple transistor testers are available for quite modest prices, there is probably a good case for one in every serviceman's kit.

No discussion on a serviceman's needs would be complete without mention of data sheets and service manuals. Service data, usually in booklet form, are normally made available in the first instance by the set manufacturer. However, they may be reprinted by independent organisations (by arrangement with the manufacturers) and offered as bound books or in loose-leaf form in suitable binders. In the latter case there may be an arrangement whereby the purchaser is entitled to additional or corrected pages as they are printed, to keep his file up to date.



A typical dial drive assembly drawing showing the method to be used when replacing the dial cord.

In the days before TV there did not appear to be a great need for such data. Sets had become pretty well standardised around a four-or five-valve chassis, using circuits which differed only in minor details. Most servicemen soon found their way around a new model, and knew pretty well what to expect in the way of voltages, etc, from past experience.

Nevertheless, those who studied these booklets generally picked up a point or two which saved valuable time later on. Often, this was nothing more than the best way to remove the chassis, how to string the dial, or details about the mechanics of the record changer, if one was used.

Similarly, some car radios are a good deal easier to handle if one is forewarned about some of the mechanical problems. And, where time is money, one should not be too proud to note what the maker has to say about a new model. Once the lesson has been learned, it will seldom be necessary to refer to the manual again. Nevertheless, it is useful to have on hand, particularly if it is a set which is not encountered very often.

With the advent of TV the need for such data became intensified. The simplest TV set is many times more complex than the most complex radio set, and it is almost impossible to keep all the data, for all the makes of set, in one's head. And for the beginner, the circuits, waveforms, voltage tables, mechanical details, etc, supply a very real need.

Much the same applies to present day radios In contrast to the older valve models, modern solid state sets, with their compact construction, call for all the assistance one can summon. The printed wiring boards, in particular, present a major problem, and much valuable time can be saved if a coded diagram is available.

A good serviceman would be wise to make sure that he acquires such a manual for every new set which appears, and keeps it handy the first couple of times he handles such a set. He may seldom need it for routine faults after that, but it can be mighty handy when a real "stinker" comes along.

Those servicemen who followed this rule when TV commenced, have a very good reason to be glad they did so. Many of the firms which mushroomed in the initial boom have long since passed from the scene, leaving a legacy of receivers for which no one is responsible except the serviceman. When substitute parts have to be provided, the data in the service manual can be very useful.

Another point worth remembering is that very few new sets are released on the market without some modifications becoming necessary after they are in the field. These may be minor, but they can often waste a lot of time if the serviceman is unfamiliar with them. The serviceman should therefore make sure that he is in a position to be advised of all such matters as soon as they are released.

Amateur Radio Stations

Early history of amateur radio. Amateur radio since World War II. Amateur radio defined. Typical activities and interests. The Wireless Institute of Australia. Licensing requirements in Australia. The AOCP examination. Sitting for the examination. Suggestions about the Morse Code.

There are, in the world today, some people to whom amateur radio represents the most satisfying, most exciting of all hobbies. To these people this scientific hobby provides a means of gaining personal skill in the art of electronics and, at the same time, an opportunity to communicate with their fellow enthusiasts over their own private short-wave — or high frequency — radio installation.

In this chapter we will endeavour to give the reader some idea of just what amateur radio is and does, and to discuss ways and means of obtaining the licence to operate an amateur station.

The history of amateur radio goes back a long way. In fact, Guglielmo Marconi, who astounded the world with his first experiments showing that messages could be sent between distant points without wires, liked to style himself as the first radio amateur.

Amateur radio really began, however, when private citizens saw in this new marvel a means of personal communication with others and set about learning enough of the art to construct a home-made station.

The equipment of these early stations consisted of spark transmitters and simple crystal or coherer detector receivers. In those days the range of the highest powered transmitters, under the most favourable conditions, was very small, and stations were so few and far between that they frequently took, as a call sign, the name of the town or area in which they were located.

Interest in the new art grew rapidly, however, and soon there were large numbers of private and commercial stations competing for "air" space. At this stage the need for some form of control became obvious and around 1904-5 various "Wireless Telegraphy" acts were passed in various countries giving the respective Governments licensing control over stations and the right to allocate call signs and frequencies of operation.

Australian amateurs were given three letter call-signs prefixed by the letter X and quite a number were active in these early days. In 1910 the Wireless Institute of NSW (now the Wireless Institute of Australia) was formed and it is interesting to note that this was the world's first institute devoted to amateur radio. Great Britain followed with the London Wireless Institute in 1913 (changed to the Radio Society of Great Britain in 1922) and American amateurs formed the American Radio Relay League in 1914.

Most amateur station operators are male, but there are exceptions as, for example, Mrs Ruth Sinclair of St Annes-on-sea, Lancashire, England. Using her call-sign G3TNN, she has had numerous contacts with amateurs in Australia, as well as in many other countries around the world. As well as operating on phone, she is a competent Morse operator.

The war of 1914-1918 brought a temporary halt to amateur activities as such, but a majority of amateurs joined the armed forces and gave valuable service as instructors and operators. Radio was, in fact, regarded as a "secret weapon" in this war in much the same manner as radar was in the early years of World War II.

At the conclusion of World War I, radio amateurs were quite naturally eager to return to their hobby but Governments, notably those in the American continent, were loth to relinquish the supreme control over all communications which they exercised in wartime. It was only the determined fight put up by amateur institutes established before the war which eventually caused the Governments to relent and allow the resumption of amateur radio activities.

From the very start, however, post-war amateur radio took on a quite different aspect. The pressures of wartime had stimulated technical developments and there were new types of equipment available. Licensing regulations restricted amateurs to wavelengths of 200 metres and below (above 1500kC) in the mistaken belief that they would "never get out of their own backyard" on these frequencies.

Amateurs, however, adapted the newly discovered "valve" to these frequencies and they were soon covering many times the distances covered with their pre-war, lower frequency spark equipment. Contacts over 1,000, 1,500 and then 2,000 miles were made, and amateurs began to dream of trans-Atlantic and trans-Pacific contacts.

In December, 1921, the ARRL sent a prominent amateur to Europe with the best receiving equipment then available. Tests were run on the 200 metre band and 30 American amateur stations were heard. The news greatly excited the amateur world and the following year further trans-Atlantic tests were carried out. This time no less than 315 American calls were logged by

European amateurs and, what was more, one French and two British stations were heard in America.

None of these tests, unfortunately, was heard in Australia and, furthermore, they were all one-way contacts.

Technically it might have been possible to span the Atlantic with a two-way contact at this time and on 200 metres but, for one reason or another, this was not achieved. In consequence, amateurs began to look for ways of achieving reliable two-way contact and extending the range of operation.

Further increases in power were out of the question since most were already at the usual legal limit of 1,000 watts. Better receivers didn't seem to be the answer since they were already using superheterodyne circuits and it did not seem possible to make any great advances in that direction. It was at this stage that they began to wonder about the wavelengths *below* 200 metres. Engineers had said these were useless but maybe they were wrong.

Experiments were begun on 130 and then 90 metres and it was noted that, as the wavelength decreased,, the results were better. On April 22, 1923, the first signals of an American amateur station were heard in Australia. The Pacific had been spanned, one way to be sure, but it was a good start.

Finally, in November, 1923, the first two-way amateur communication across the Atlantic became a reality when two stations in America worked one in France for several hours, all three stations using a wavelength of 110 metres. Other stations dropped to 100 metres (3.00MHz) and found that they, too, could easily work two-way across the Atlantic.

The exciting possibilities of this discovery did not go unnoticed and soon hundreds of commercial companies were rushing stations into the 100-metre band. Chaos threatened for a while but finally the first of a series of radio conferences partitioned off specific bands of frequencies for the various services clamouring for space.

Although thought, at this time, was mainly centred around 100 metres the various amateur organisations at these conferences came to the conclusion that the surface had probably only been scratched and, quite wisely, petitioned for frequencies not only at 80 metres but also at 40, 20, 10 and even down to 5 metres. (These bands are now commonly referred to as 3.5, 7, 14, 28 and 52MHz.)

Many amateurs promptly moved to 40 metres and the band responded by enabling two-way trans-Pacific communications between America and Australia, New Zealand and South Africa. A second jump to 20 meters brought about another dream of amateur radio, daylight two-way communication over long distances.

In many countries, and particularly in Australia, amateurs were the first to set up stations broadcasting program material of a type similar to that transmitted by the "broadcast band" stations we know today. The interest aroused by this amateur program material encouraged many people either to purchase or to build suitable receivers and, eventually, commercial broadcasting with paid advertising became a reality.

In this, as in many other countries, the earliest commercial broadcasting set-ups developed from amateur stations but, upon doing so, they became legally divorced from the "amateur service" and came under separate licensing.

British radio enthusiast Peter Blair, aided by his wife Eileen, was the first amateur operator to bounce a signal off the moon — in 1964. The signal was transmitted from the Blair's aerial shown, directed towards the moon by hand, and costing about £5 for materials. The signal was received on a US Government radio telescope in Puerto Rico which, in the same currency, would have cost £3,000,000.

Even so, amateur station operators, in Australia at least, continued to broadcast musical programs until just before World War II, but they did so purely for experimental purposes and without pecuniary reward.

In the 20s and 30s, amateur radio expanded rapidly, both in technical advances and the numbers licensed to operate stations, until World War II once again brought a temporary halt to activities.

In this war some 25,000 amateurs used their skill and knowledge in the armed forces and many, on both sides, paid the supreme sacrifice in the defence of their countries.

After the war amateur radio took up where it had left off, at least initially. However, wartime pressures had produced dramatic technical advances in communications equipment and much of this equipment, now declared surplus by the various armed forces, was made available to amateur operators. Then, too, the ranks of amateur radio enthusiasts were swelled by many who had received training in communications during the war and decided to continue their interest in peacetime.

That phase of amateur radio has largely passed. Advanced as it was at the time, war surplus equipment has been worn out or simply discarded and its former owners have become the veterans in the amateur ranks.

A whole new generation of amateurs has emerged, who are far more familiar with solid-state devices and technology than with valves. The bulky equipment racks and the huge power supplies of the previous era have given place to table-top equipment, which is nevertheless far more advanced technically.

Amateurs are also exploring the most modern communication systems. In the VHF bands, notably the 2 metre segment, there is world wide interest in repeaters: unattended relay stations strategically sited for best VHF coverage, which accept an incoming signal on one frequency and re-transmit it on an adjacent frequency.

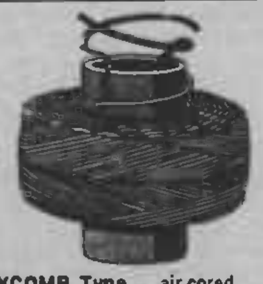
The result is to greatly extend the range of 2 metre equipment, particularly mobile and portable equipment having limited power and modest aerials. In Australia the Wireless Institute of Australia and affiliated clubs have established a number of such repeaters in capital cities and large provincial centres. They are available for use by any amateur.

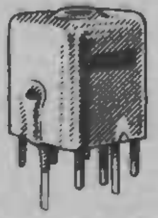
Amateurs have extended this theme to satellite repeaters. Through their own organisation, AMSAT, based in the USA but open to amateur organisations world wide, they have built and launched a number of satellites. The first was OSCAR 1 (Orbital Satellite Carrying Amateur Radio) launched in 1961. The latest, at the time of writing, is OSCAR 6 which is providing international amateur communication using the 2 metre band (in) and the 10 metre band (out). Plans are in hand for a geostationary amateur satellite called SYNCART (SYNChronous Amateur Radio Translator) to be launched around 1975.

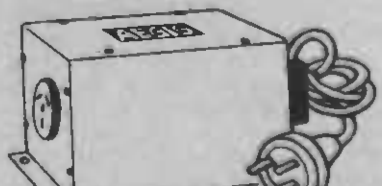
So far we have given a very brief history of amateur growth and in so doing may have given some idea of what an amateur is and does. We will now elaborate a little on this point.

The "Handbook for Operators of Radio Stations in the Amateur Service" published by the Australian Postmaster-General's Department provides the following definition:

> **"Amateur Service" means a service of self-training, intercommunication and technical investigation carried on by amateurs, that is by duly authorised persons interested in radio technique**






AEGIS

> **solely with a personal aim and without pecuniary interest.**

On the operation of amateur stations, the Handbook has the following:

> **The use of radio stations in the amateur service shall, as a general rule, be confined to technical investigations, research into or instruction in radio communication techniques, without pecuniary interest.**

From this we can see that an amateur is a person who is (a) interested in electronics with an emphasis on the transmission and reception of signals by "wireless" and, (b) owns and operates a private station for the purpose of communicating with fellow amateurs and, in doing so, furthering his knowledge.

This is not to say that all amateurs are cast in the same mould, for indeed they obtain pleasure from their chosen hobby in many and diverse ways.

Some are mainly interested in the technical aspects of their equipment and spend many hours constructing various pieces of "gear" and trying out new ideas. These amateurs might spend very little time on the air, only coming on when they have a new piece of gear to try.

Other amateurs find that, once having set up a station, their greatest interest is in operating it. They may be of a gregarious nature and enjoy long conversations (called "rag chews" in amateur parlance) with fellow amateurs on subjects of mutual interest.

Still other amateurs find pleasure in "DX'ing" (long distance working) with a view to contacting as many amateurs in as many countries as possible. Special awards are available for this and, to obtain these awards and as a confirmation of their contacts, these amateurs like to exchange "QSL" cards. An illustration of a typical QSL card sent by an overseas amateur to the author is included in this article.

Then, too, these amateurs variously interested in experimenting, operating and DX'ing can often be subdivided according to favourite bands of operation and the different modes of transmission they employ.

Some operate on one or all of the HF bands (3.5 to 28MHz) while others prefer to confine their operating to the VHF bands, above 52MHz. A very limited band of enthusiasts confine their activities to operation on "microwaves" — or frequencies above 1000MHz.

Amateurs may variously employ most of the available modes of transmission but the most popular are AM (amplitude modulation), SSB (single sideband), CW (continuous wave), FM (frequency modulation), RTTY (radio teletype) and TV (television), probably in that order. SSB is rapidly growing in popularity and may soon become the method used by most.

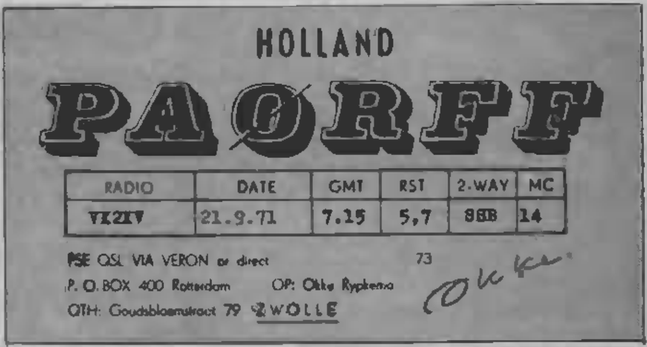
Some amateurs have elaborate stations with maximum power transmitters on several bands employing several different modes of transmission and feeding into complex aerial arrays. Others are satisfied with the bare minimum of equipment which, quite typically, might be stowed in a bedroom cupboard when not in use. Fortunately, for the latter amateur, the pleasure to be derived from the hobby is NOT directly proportional to the quantity and quality of the equipment owned.

Amateurs are a fraternal lot; their common interest makes them "brothers under the skin". When visiting strange towns an amateur often looks up friends with whom he has become acquainted over the air and even if he knows no amateurs in a given vicinity, his amateur call makes him more than welcome.

Amateur radio clubs have been formed all over the world and most feature meetings with both elementary and advanced technical talks, study lessons and code classes, social contacts, and "eats".

In this country the Wireless Institute of Australia has branches in each State and is affiliated with most of the country radio clubs and the school radio clubs. The WIA keeps its members in touch with current affairs by the publication of a monthly journal and some divisions and clubs publish their own bulletins.

Each State has its own official station and these broadcast news to members, usually on Sunday mornings. Some of these stations are also available to intending members wishing to gain practical experience in the operation of transmitters, and some broadcast slow code transmissions on certain nights of the week for those studying for the amateur licence.


HOLLAND

PA0RFF

RADIO	DATE	GMT	RST	2-WAY	MC
VK2XV	21.9.71	7.15	5,7	SSB	14

PSE QSL VIA VERON or direct

73

P. O. BOX 400 Rotterdam

OP: Okke Rypkema

Okke

QTH: Goudsbloemstraat 79 ZWOLLE

A typical QSL card from an amateur operator in Holland confirming contact with an Australian amateur on the 14MHz band, using the SSB mode of transmission.

Amateur radio recognises no barriers of race or creed. These two young men are pursuing a common interest in the hobby despite another barrier: both are blind but they are smiling because they had just learned of a pass in a Youth Radio Scheme Course examination conducted by the Wireless Institute of Australia.

The aims and objects of the WIA are to encourage and assist all persons interested in any or all aspects of amateur radio and allied techniques. Membership is divided into two grades:

Grade "A" ("Full Members") — Those who have obtained their Amateur Operators' Certificate of Proficiency or the Limited Amateur Operators' Certificate of Proficiency, and

Grade "B" ("Associate Members") — Those studying for their Certificates or who are interested in some aspect of radio science.

Honorary life membership has been bestowed on some who have rendered valuable assistance to a division or to radio science.

Apart from being a fine hobby, there is always the possibility that amateur radio will lead to a lifetime career in electronics. Operators with amateur experience are several steps ahead when it comes to the operation of commercial stations because they have developed a "feel" for radio equipment.

Many of today's research scientists and engineers, men with full professional standing, can trace their careers back to an early interest in amateur radio.

Some of these men, while radio amateurs, made major contributions to the art, and there is no doubt that they, and the rising generation of amateurs, will continue to make such contributions.

Most would-be amateurs know that they are required to sit for an examination before obtaining a transmitting licence, but not all of them realise just why this is necessary and how they can best prepare themselves for their test.

Radio communication is now essential to the navigation of thousands of ships and aircraft, not to mention international message handling and mobile communications circuits.

Co-ordination between nations on the use of the radio spectrum is therefore necessary so that there will be a minimum of mutual interference. Ships and aircraft moving between countries should also be able to communicate with control stations, weather stations, etc, with as little trouble as possible.

The way in which the frequency spectrum is split up between all these essential services, plus the amateur requirements, is decided at an international telecommunications convention at which the governments of all the major countries of

A TYPICAL AOCP THEORY PAPER

TIME ALLOWED — 2½ HOURS

NOTE: Seven questions only to be attempted. Credit will not be given for more than seven answers. All questions carry equal marks.

1 (a) Draw a circuit diagram of the plate-modulated radio-frequency amplifier and modulator stages of a 150 watt D.C. input amateur band transmitter.
(b) Describe fully how 100 per cent modulation is obtained.

2 (a) What is meant by the following terms when used in reference to an iron cored transformer:—
(i) turns ratio, and (ii) impedance ratio
(b) List the losses associated with the operation of a power transformer and state how these may be minimised.
(c) A transformer has a primary winding of 600 turns to which 240 volts AC is applied. Neglecting losses, calculate the number of secondary turns required to give a voltage of:—
(i) 16 volts, and (ii) 6.8 volts

3 (a) Discuss factors you consider desirable in a microphone used in mobile operation.
(b) With the aid of a sketch describe the construction and theory of operation of a microphone which you consider meets these requirements.

4 (a) Assisted by a circuit diagram describe the operation of a Grid Dip Oscillator or a Transistorised Dip Oscillator.
(b) Give a practical example of the use of such an instrument.

5 (a) With the aid of sketches explain how frequency-modulation differs from amplitude-modulation.
(b) Describe a detector suitable for resolving the audio component of a frequency modulated signal.

6 Describe an instrument which, when inserted into the transmission line between the transmitter and aerial, will indicate whether the aerial is correctly matched to the impedance of the transmission line and the output stage of the transmitter. A circuit diagram will assist.

7 (a) With the aid of a circuit diagram explain the operation of the radio frequency stages of a receiver suitable for reception of UHF signals in the 420-450 MHz amateur band.
(b) Indicate and give reasons for any techniques used in the UHF receiver which would not apply to a receiver operating on 7 MHz.

8 (a) Describe a quarter-wave vertical groundplane aerial suitable for use in the 14 MHz amateur band. Show dimensions and approximate feeder line impedance.
(b) Discuss the advantages and disadvantages of this type of aerial when used for both transmitting and receiving.

9 Calculate:— (i) the frequency at which a capacitor of 3.3 pico-farads (mmf) has a capacitive reactance of 159.2 ohms. (ii) the power dissipated in a resistor of 180 ohms when a potential of 3 volts exists across the resistor.

the world are represented. The major amateur bodies are also invited to send representatives to these conferences with an aim to protecting their interests.

The authority which carries out the decisions of these conferences in Australia is the Postmaster-General's Department, through its various Radio Branches. The department carries out in detail the local rules and administration of all radio stations in this country.

Matters such as the frequency or range of frequencies in which individual stations may operate, the allowable power level and, more to the point of this article, the minimum knowledge and skill required of the operator in order that he may operate his station efficiencly and without interference to others are decided by this authority, which acts within the framework of international agreement.

The required minimum standard is established by the Amateur Operators' Certificate of Proficiency (AOCP) which is issued after the applicant has passed the appropriate PMG examination.

Having obtained the certificate, it is largely a matter of routine for the individual to obtain a station licence and be allotted a call-sign. The latter will be prefaced "VK" to indicate an Australian station, a following number to indicate the State and a final combination of two or three letters for individual identification. If you live in New South Wales you could be VK2XYZ or if in Queensland VK4XYZ.

The qualifications required to obtain the AOCP are set out in publications available from the PMG radio branches.

The qualifications may be summarised as follows:

(1) An elementary knowledge of wireless telegraphy and wireless telephony and electrical principles.

(2) A knowledge of such of the radio communications regulations for the time being in force under the telecommunications convention and related to the operation of amateur stations.

(3) Ability to send correctly and to receive correctly by ear in Morse code a test in plain language, including figures, at a speed of 10 words a minute.

The examination to prove these qualifications is in three sections as follows:

SECTION K (REGULATIONS). A half-hour written examination consisting of one paper containing questions based on the "Handbook for the Operation of Radio Stations in the Amateur Service" issued by the PMG's Department.

SECTION L (TELEGRAPHY). (a) One Morse test covering the correct transmissions of text in plain language — including figures — at a speed of 10 words a minute. Sending test, 2½ minutes. (b) One Morse test covering the correct reception by ear of text in plain language — including figures — at a speed of 10 words per minute. Receiving test, 5 minutes.

In both the above tests the text averages five characters to the word, each figure counting as two characters.

SECTION M (THEORY). A two-and-a-half-hour written examination of one paper containing questions based on the theory of wireless telegraphy and wireless telephony as applied to amateur transmitting and receiving systems, and the elementary theory and practical applications of the principles of electricity and magnetism.

In order to gain a pass the candidate must obtain at least 70 per cent of the total marks allotted in each section.

There is also a Limited AOCP certificate which does not require the candidate to have passed the telegraphy tests (Section L). The limitation is that transmissions must be confined to the amateur bands at 52MHz and higher.

This certificate will be of interest to many who can reach the necessary standard in radio theory but who may not have the time to reach the necessary standard in Morse. An unrestricted licence can always be obtained at a later date simply by passing the Morse tests.

Two examinations are held each year, on the third Tuesday of February and August, in all capital cities and in cities and towns in which a District Radio Inspector is stationed. A list of these places is contained in the PMG handbook for amateurs. Examinations in telegraphy only are also held in the same places on the third Tuesday of May and November.

There is also provision for people living in areas away from the cities. A candidate living more than 50 miles from the nearest Radio Branch office can, provided the facilities are available, arrange to take the examination under the supervision of the Postmaster at the official post office nearest to his home.

Applications to sit for the examination must be made on a printed form obtainable from the PMG Radio Branches, located in capital cities, also, in NSW, in Newcastle, Wollongong, Armidale and Wagga; in Victoria, in Bendigo, Ballarat and Sale; in Queensland, in Rockhampton and Townsville. The completed form must be returned not later than two weeks before the examination date.

Applications to sit for the examinations will be accepted only from people 14 years of age and over. An official copy of the applicant's birth certificate, and an entrance fee of $2, must be forwarded with the completed form.

After receiving examination results from the PMG, successful candidates may apply for the AOCP (or AOLCP). This will be issued without further charge, when the following particulars have been supplied:

Height, colour of eyes and hair, complexion and any special physical peculiarities. A recent head and shoulders photograph, autographed on the front, is also required, for attachment to the certificate.

Those who pass the examinations before their fifteenth birthday will be required to wait until that date before the certificate is issued.

The bare outline of the requirements does not provide a very good idea of how to set about studying for the examination. One method adopted by many is to obtain copies of previous examination papers, available free of cost from Radio Branch offices. From the same source, at a cost of 30c, may be obtained the "Handbook for Operators of Radio Stations in the Amateur Service," which is essential for Section K (Regulations) of the examination.

To give some idea of a typical paper we publish, in this article, the theory section of a recent examination.

For the theory section of the examination the PMG recommend the following textbooks: The "Radio Amateurs Handbook"(ARRL), "Radio Handbook" by Editors and Engineers Ltd, "The Amateur Radio Handbook" (RSGB).

In fact, these books contain far more detailed information than is necessary to pass the AOCP examination and the student needs only to concentrate on those sections which deal with fundamentals and the basic theory which is called for by the Department. As already mentioned, copies of earlier theory papers should provide a good guide and, if assistance can be obtained from a club group, so much the better.

Keep your eye open also for manuals, even if dated, published by valve and transistor manufacturers. They often contain useful background material.

The program of study you set yourself will depend on how much previous knowledge you have of the subject.

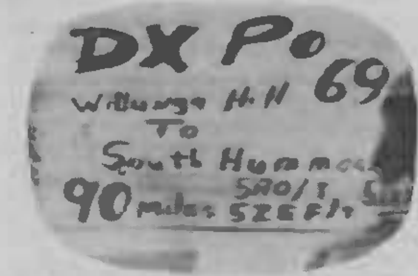
It may be worth while to inquire in your area in respect to radio colleges, evening classes for adult training, or radio clubs. It is just possible that you may discover a course that could assist in obtaining your amateur certificate.

The WIA also runs a complete course lasting 12 months and this may be taken either by correspondence or with practical lectures given twice weekly. Morse code practice is included.

By following a recognised course, the student can be sure that nothing is left out and it should ensure that he does progress from week to week. There is a tendency when studying privately to imagine a thing is understood when, in fact, he has only a sketchy outline. Some students become disheartened when working alone for long periods, and it does help to have sympathetic assistance from a qualified tutor or an interested amateur.

You can become proficient at Morse by working on your own and listening to the slow Morse transmissions from the WIA stations. Those who own a tape recorder can get extra practice by using the tape hiring service operated by NSW Division of the WIA. There is every advantage, however, in having assistance from someone already proficient, particularly in the matter of sending, since it is easy to develop bad sending habits which may be difficult to correct on your own. Wait until you have had a few weeks receiving practice before attempting to send.

A Drake model R-4B receiver, especially designed for use by amateur radio operators. It employs valves, transistors and IC's.



An amateur television transmission involving Australian stations 5AO/T and 5ZEF/T. Transmission distance was 90 miles.

Always try to practise receiving Morse at a speed a little too fast for perfect copy. This is essential if you are to pick up speed. You should aim at a standard several words a minute ahead of the required 10 words per minute, for under the stress of examination even the best do not show their full capabilities.

If you can locate and make friends with a nearby amateur, he will almost certainly help you by sending to you on the key. Use a standard PMG key, if possible, because you are sure to be examined on this type.

When sitting for the theory paper, take time out to read each question carefully and then make sure that your answer gives ALL the information the paper requests. As in the typical paper we have published, many questions consist of either 2 or 3 parts and many of these parts could again be subdivided into 2 parts.

For example, a question may ask the applicant to draw the circuit diagram of the plate modulated radio frequency amplifier and modulator stage of an amateur band transmitter AND explain how 100 per cent modulation is obtained. If you were to draw a perfect circuit diagram for such a transmitter, but fail to explain how 100 per cent modulation is achieved, you could easily fail in the question.

At the same time do NOT offer information which is not specifically asked for in the question. To do so is not only time wasting but if, in offering additional information, you should make several bad errors your examiner may take it into account when marking the paper.

Keep your answers brief and to the point — the examiner is not interested in your abilities as a marathon writer — only in your knowledge of the questions. Remember that the examiner does not have to be convinced that you are a genius at radio, he just wants to be sure that you can build a transmitter and place it on the air without interference to others.

When you feel that you are ready to sit for the examination, do not hesitate to do so. Even if you should not obtain a pass, you can sit again for theory in six months time (or in three months time for the Morse test). In fact, you can sit as many times as are necessary to obtain a pass. When you do pass, you will find that the pleasures of amateur radio as a hobby will make all the effort and study worthwhile.

Amateur radio stations have provided emergency communication on numerous occasions in Australia and elsewhere, when natural disasters have cut normal means and normal access. Working from a schoolroom in Corning, NY, USA, this team of amateurs helped co-ordinate medical and rescue services during the 1972 floods in the area and were responsible for saving many lives.

CHAPTER 22

Audio Equipment & Techniques

The importance of audio techniques — Microphones — Distortion in audio reproduction — Magnetic recording, equalisation, fidelity, noise, wow and flutter — Disc recording and playback — Pickups — Audio amplifiers — High fidelity amplifiers — Dynamic loudspeakers — Impedance and output transformers — The need for loudspeaker baffling.

As the heading indicates, this chapter is devoted to a discussion of audio frequency amplification and sound reproduction. It leads naturally into an explanation of terms such as "Hi-Fi," "Stereo" and "Quadraphonic."

The chapter is much longer than most others in the course, mainly because the subject is of such wide general interest and application. In the developed countries of the world, there would be very few individuals who are not exposed in a number of ways to audio equipment of one type or another. To some, the exposure or involvement is quite casual; to others, notably Hi-Fi enthusiasts, it is the basis of an intense personal interest.

As you have probably gathered, the word "audio" comes from the same root as "audible," "audition" and so on. In electronics, it is used primarily to describe phenomena, techniques, circuits and equipment related to frequencies in the audible range. In broad terms, this means frequencies within the range 20Hz to 20kHz. These are not strict limits, however, and engineers in the audio field are likely to regard as significant sub-sonic or supersonic phenomena which may have some bearing on sound reproduction.

The most prolific use of audio components is found in ordinary radio receivers.

Basically, a radio receiver makes available at its detector a small audio signal, which is a close replica of the speech and music signal being radiated by the broadcasting station. Apart from "crystal" sets and such like, all radio receivers use one or more audio amplifier stages to build up the signal to a level where it can adequately operate a loudspeaker.

Much the same thing happens in a television set, the two applications adding up to literally millions of audio amplifiers in Australia alone.

Added to these are all the record playing units and tape recorders currently in use. Every one involves an audio amplifier — in fact two, where a stereo system is involved.

Then there are public address amplifiers for reinforcing speech; amplifiers for theatre sound; amplifiers to do with electronic organs and other musical instruments; amplifiers in paging and intercommunication systems; and so on into a variety of less obvious applications.

Little wonder that engineers can specialise at length in audio techniques and components, without any fear of exhausting the subject.

A discussion of audio techniques, at least for our present purpose, can logically begin with a look at the sources from which signals for amplifying equipment may be derived. These would include, most commonly, radio and TV receiver detectors, as already mentioned, magnetic tape and tape heads, and phonograph records and pickups.

However, microphones must qualify as the most fundamental of signal sources, because they perform the primary task of changing sound from a pattern of pressure waves in the air to an equivalent pattern of electrical signals. The original sound may be voice, music, noise of one kind or another, or any combination of the three; anything, in fact, which we can hear and which we may wish to amplify or record.

An ideal microphone would be capable of changing any audible sound pressure wave into an exactly equivalent electrical signal without **distortion**. This is a collective term indicating any departure from perfect reproduction.

Fig 1: A typical dynamic microphone from the Philips range especially suitable for use by vocalists in stage or studio situations.

It would exhibit no **frequency distortion**; in other words, it would handle all sound with equal efficiency, irrespective of the frequency or pitch of the sound. If its signal output voltage were plotted against frequency, the resulting graph or curve would be flat. An audio engineer might simply say that it should have a flat response.

An ideal microphone would not exhibit any **harmonic distortion**; in other words, it would not add to the electrical output signal extra frequencies or harmonics which were not present in the original sound.

Again, an ideal microphone would not introduce any **intermodulation** effects. This means that it would not produce extra and spurious output signals as a result of heterodyning or mixing of original signals. To quote a simple example, there would be no tendency for two tones (say 500Hz and 700Hz) fed to the microphone to heterodyne (or intermodulate) to produce extra spurious signals at 200Hz (700 — 500) or at 1,200 Hz (700 + 500).

Perhaps it would be a good plan to re-read these last three paragraphs and to look again at the terms which have been put in black letters. While we have chosen to introduce them in connection with microphones, they have an application to all audio components. Frequency distortion (or frequency response), harmonic distortion and intermodulation effects are terms which you will strike over and over again in audio literature — and which you will certainly strike again in this chapter.

Not surprisingly, there is no such thing as an ideal or perfect microphone, although good quality studio type microphones do not suffer to an obvious degree from these failings. However, the same cannot be said for the class of microphones more commonly handled by audio hobbyists, or those used for utility public address systems.

A hobbyist or small-scale P.A. operator might class a microphone in the $15 to $30 group as "good," and it may be so to the extent that it can give acceptable results on speech, singing and non-critical instrumental groups. However, to varying degrees, microphones in this group do evidence some distortion — using the word collectively — which is responsible for the reproduced sound being not quite natural.

In the still cheaper grades of microphone – and this includes many of those supplied in the past with domestic tape recorders – distortion is plainly evident, particularly frequency distortion in the way of a strong prominence in the middle of the range (say about 3,000Hz), with diminished response in the upper register and at the bass end. That is why speech and music recorded through such microphones commonly has an unnatural or "metallic" quality.

Microphones use a variety of basic principles to convert sound pressure waves to electrical signals.

Studio quality microphones commonly employ one of three basic principles, being described respectively as "ribbon," "dynamic" or "condenser" types.

Among medium quality microphones, one finds an occasional ribbon type, a preponderance of dynamic types and a few "crystal" (or piezo-electric) units.

In the inexpensive class one finds again a preponderance of dynamic types and a diminishing proportion of crystal types.

Fig 2: A modern reel-reel two-channel tape recorder manufactured by the Sony Corporation. It can be switched to operate on any one of three different speeds and can impose on quarter-inch tape either four mono tracks or two stereo pairs. It has six heads, allowing the tape to be recorded or played back when moving in either direction.

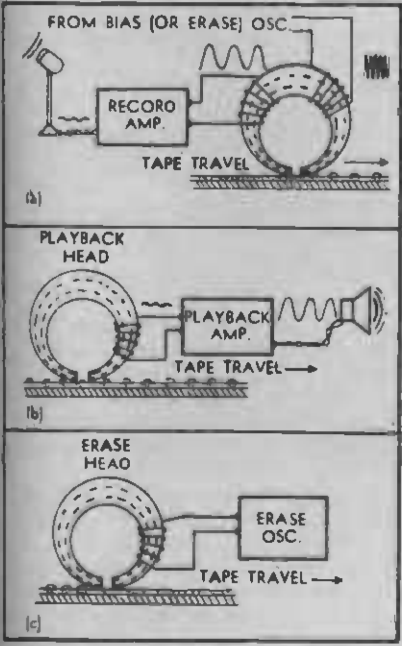


Fig 3: These diagrams illustrate, in elementary fashion (a) the tape recording process (b) tape playback and (c) signal erasure when the recording is no longer required or when a new recording is to be made on the same track.

At this stage, these must remain mere terms, although some clue as to their meaning may be picked from later references.

The electrical signal available from most microphones is quite small, being commonly very much smaller than that available from the detector of a radio or television receiver. This being so, an amplifier used with a microphone must have a higher amplification or gain than one intended to operate only from a radio detector.

But enough about microphones. Let us move on.

If sound is to be recorded for future replay, it is most commonly done, these days, first of all on magnetic tape equipment. The tape itself, on which the recording is impressed, is a ribbon of thin plastic, about 6.5mm (0.25") wide, coated on one side with a thin layer of finely divided and magnetically sensitive particles.

A tape recording system provides the necessary spooling and traverse mechanism to pull the tape at an even speed past a magnetic head assembly. Most readers will be sufficiently familiar with tape recorders to picture the mechanism.

Internally, a tape recorder contains an erase **oscillator**. This is a normal oscillator circuit but arranged to provide output in the supersonic region, usually between about 50kHz and 120kHz.

During the recording process, a large proportion of this supersonic energy is fed to an erase head, being the first head over which the tape rides as it is drawn from the feed spool to the takeup spool. The signal in the erase head creates a powerful, supersonic magnetic field which is intended to erase or cancel any signal that may be on the track on which the new recording is to be made.

A small proportion of the erase signal is taken also to the record head, where it is combined with the audio signal to be recorded. Most likely this latter will have come from a microphone and will have passed through a number of amplifier stages to bring it to a level where it can produce adequate signal current through the windings in the record head.

The reason for feeding some of the erase signal energy to the record head involves some rather involved theory and is beyond the scope of this present article.

We will have to be content with the statement that magnetic recording relies heavily on the presence of a fairly critical amount of supersonic signal in the record head, along with the signal to be recorded. Though it is usually taken from the erase oscillator, it is referred to, in relation to the record head, as the high frequency **bias** voltage.

The audio signal and bias voltage fed to the record head produce a magnetic field evident across a tiny gap in the exposed face of the head. As the tape is drawn across this gap, the particles are magnetised according to the field which is present at the particular instant. This they retain so that, in effect, the tape can be regarded as being coated with an extremely large number of separate magnets of particle size, each one magnetised in a particular way, according to the original signal which produced it.

To play back the tape, the erase (or bias) oscillator is switched off, the tape is rewound and then drawn across the head assembly again, at the same speed as it was recorded. The tiny magnetised particles, drawn across the heads, now induce a signal current in the windings of the heads, more or less equivalent to the signal originally recorded. When the induced current is amplified and fed through a loudspeaker, the original sound is recreated.

In simple recorders, the same head is commonly used for record and playback, being switched from one function to the other by operation of the controls and, in particular, the "record-playback" switch. In more complex equipment, separate record and playback heads are used, allowing the designer to suit the characteristics of the individual heads more specifically to the job they have to do.

Either way, the signal recovered from a magnetic replay head is quite small, being comparable to that produced by a microphone. As with a microphone, a high order of amplification is required from the amplifier itself, to bring the signal to a level adequate to operate a loudspeaker.

And here we must introduce another consideration, which will have implications wider than in tape recorders alone . . . that of **equalisation**.

The natural characteristics of a magnetic recording system is that it tends to record and replay most efficiently frequencies in the middle of the audio range – say between about 1000Hz and 3000Hz – lower frequencies and higher frequencies being very much reduced or attenuated by comparison.

The amount of high frequency attenuation varies, in turn, with the speed of travel of the tape. At slower tape speeds, say 3¾ inches per second and below, the attenuation is much more serious than at higher tape speeds, say 7 ½ inches per second and above.

To counteract this peculiarity, steps must be taken to boost the bass response and the treble response during recording and / or playback so that their level, as finally heard, is equivalent to that of the middle frequencies. The process of equalising the performance of the system at all frequencies is commonly referred to as **equalisation.**

Here again, having mentioned the word and the concept, it is necessary to leave the reader to follow it up independently, as opportunity occurs. At the very least, however, the reader should take note of magnetic tape and tape heads as a signal source and to register that they involve the need for (a) high amplification and (b) suitable equalisation.

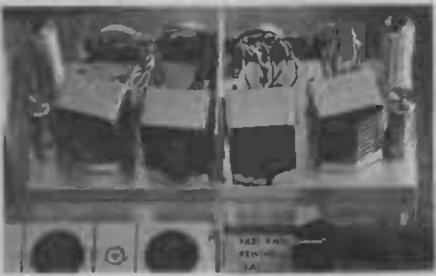


Fig 4: Close-up view of a row of tape heads on a professional quality Ampex tape recorder. This particular machine, used mainly to make multi-track studio master recordings, can impose eight tracks simultaneously on half-inch tape.

In the matter of quality of reproduction, much the same observations apply to tape equipment as have been made already about microphones.

The best studio and commercial tape recording equipment is very good in terms of equality or **fidelity**, exhibiting a minimum amount of **intermodulation distortion, harmonic distortion and frequency distortion**. In view of what we have just said about equalisation, this last remark implies that the equalisation in the internal circuitry of a good quality tape system exactly balances its inherent losses at bass and treble, so that the final response curve is exactly equalised, or is flat.

With less costly and elaborate tape equipment, and because of resultant design compromises, the overall fidelity may not be as good, the output containing a higher content of intermodulation, harmonic and frequency distortion.

Tape equipment can also suffer from certain other important limitations on its performance.

One of these is **background noise**, usually evident as a soft hiss heard behind the recorded sound. Some of this hiss can come from the recording amplifier, and some from the playback amplifier, being an indirect result of having to amplify signals of such a small original magnitude.

Noise can also be introduced as a result of random passage of the tiny tape particles past the gap in the replay head. A tape can be "noisy" if the magnetic particles are too coarse or too irregular in their distribution, or if the tape has not been properly erased prior to recording.

Noise will also be more evident if the laminations in the replay head are partially magnetised rather than being magnetically neutral.

Unfortunately, while background noise is most common in inexpensive equipment, it is noted all too frequently in tapes made at professional level on professional machines.

Other important limitations on the performance of tape equipment take the form of **wow** and **flutter**. Both are mechanical in origin.

Tape equipment is said to exhibit "wow" if there is a slow, periodic variation in tape speed (a repetition rate of say not more than once per second) causing a variation in the pitch of the recorded sound. Such wow is usually caused by some eccentricity in the drive system, a bearing which tends to bind or a spool clutch mechanism which causes a varying load on the drive motor(s).

Flutter is a much more rapid speed variation caused notably by imbalance or eccentricity in the motor and spindle which actually pulls the tape past the head.

In most modern recorders, domestic as well as professional, the wow and flutter content are held to within acceptable limits but it was a problem, particularly in early model, inexpensive battery powered units.

Fig 5: This picture, reproduced by courtesy of Festival Australia, shows an LP disc being removed from the press. Provided the process is carefully controlled, the grooves in the pressing are a very close replica of those in the original master. The fact that records can be reproduced in this way, relatively cheaply, gives them a tremendous commercial advantage over tape for pure playback purposes.

Perhaps brief mention is warranted, just here, of the varying track arrangements used in tape equipment. Normal recording tape is 6.5mm (¼in) wide overall and earliest practice was to record right across the full width of the tape. This is still done in professional recordings.

Subsequently, it became the practice to record two tracks on the tape, side by side, each somewhat less than half the width of the tape. **Half-track** recording allows twice the playing time from a given reel of tape at any given speed.

Nowadays, **quarter-track** recording is quite commonplace.

In equipment using tape **cartridges**, as many as eight tracks are used, while in **cassette** players, either two or four tracks are accommodated on tape approximately 3.5mm (1 / 7in) wide.

The use of narrower tracks does involve some loss in overall performance, but quarter-track recording can still give entirely acceptable results for domestic applications.

While tape recording and replay equipment is currently in very wide use, the most popular medium for recorded sound entertainment remains the disc; the descriptive words, "phonograph" or "gramophone" have become almost redundant in modern usage except perhaps for the contractions "gram" or "gramo."

At a manufacturing level, discs have the basic appeal of ease and economy of mass production, to the extent that there is no embarrassment at all if a "hit" disc sells a million copies in a relatively short period. At the user level, discs are easy to handle, easy to code in terms of artists and selections, lend themselves to attractive packaging and can produce sound quality which is adequate for present-day requirements.

On a disc, the recording is retained as tiny deviations in a spiral track. In a sense, the track can be regarded as a mechanical graph or plot or signal waveform against time.

An original or master disc recording is made on a **recording lathe** which carries a recording head slowly across the surface of the disc, from outside to inside, as the disc rotates at predetermined speed beneath it. The resulting spiral **track** is cut by a specially shaped diamond stylus protruding beneath the recording head.

A coil and magnet system within the head, fed from an external amplifier, drives the stylus, so that it vibrates laterally, and sometimes vertically as well, in sympathy with the audio signal which is to be recorded. As a result, the spiral track deviates in accordance with the actual pattern of audio waveforms which has been fed to the stylus.

By a process involving spraying and electroplating, the original recorded track can be transferred to metal **stampers** or **dies** and used to stamp out any required number of copies of the master disc for distribution to the public. For all practical purposes, the disc which you buy carries a track identical to the one cut in the original master.

To reproduce the recorded sound, the disc is replayed — nowadays — with a **pickup**. This has a fine **stylus** which traces the spiral groove as the disc rotates. The tiny waveform serrations in the groove impart a corresponding motion to the stylus which actuates a mechanism inside the pickup head capable of generating a corresponding electrical signal.

This is built up in an amplifier to a level where it can operate a loudspeaker in the usual way.

The disc technique has undergone drastic revision within the living memory of many readers.

Originally, disc recording and playback was an entirely mechanical process, involving no electrical amplification whatever. The sound to be recorded was collected in a large horn and conducted through tubing to a diaphragm attached to the recording stylus, so that the sound waves actuated the stylus directly.

For playback, the disc was tracked by a steel needle attached directly to a diaphragm. This vibrated and created air waves which were conducted into the listening room by tubing and a flared horn.

Because the system relied on the direct transfer of mechanical energy throughout, its specifications needed to be "robust" so that heavy diaphragms, steel needles, relatively large grooves and coarse-grained record materials were accepted without too much question. This was the era of the early 78rpm discs.

Later, the techniques of using electrical cutting equipment and then electrical reproducing equipment removed many of the limitations of the mechanical system. A greater variety of sound could be sucessfully recorded, the frequency response of the system could be widened and flattened, harmonic distortion reduced, with a reduction also in the noise content contributed by the coarse grained disc materials.

However, while enormous advances were made along these lines, the traditional standards of groove dimension and record speed were retained up till the end of World War II.

At this juncture sound engineers began to accept the fact that the old standards were needlessly cumbersome in a situation where disc recording no longer relied on the direct transfer of mechanical energy from cutting to reproducing stylus. In all technically advanced countries, electrical amplification was universal in record playing equipment.

Groove dimensions could be drastically reduced, record speed likewise, non-abrasive disc materials and styli — ground from gems — could minimise noise and wear problems, and playing time could be dramatically increased.

Vital to the latter point was the fact that, meanwhile, magnetic recording had been developed to the point where original recordings could be assembled on a tape, mistakes could be edited out and the whole thing checked and timed. Then the material could be transferred to a disc without fear of the effort being ruined by an inadvertent error on the part of a performer.

So after the war came the era of "fine groove" or microgroove discs, intended to be played with sapphire or diamond styli, contoured to a spherical shaped tip of radius .001in — about one-third the earlier dimension.

Two major standards emerged — 7-inch diameter discs turning at 45rpm and 12-inch diameter discs turning at 33rpm.

Nowadays, 7-inch discs are issued as "singles" (single play) or "EP's" (extended play). The "singles" carry one item per side, while "EP's" carry two items per side.

The 12-inch discs are used for major musical works, with a playing time of 20-30 minutes per side, or four popular recitals with about five or six tracks per side.

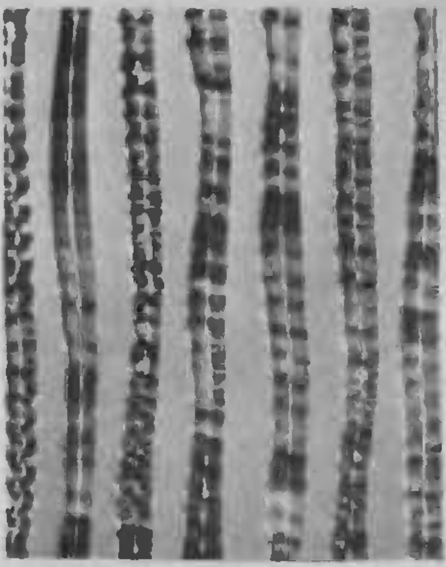


Fig 6: A greatly magnified photograph of the grooves in en LP stereo disc. (Picture by Cecil E. Watts)

The quality of reproduction from current fine groove (or long playing or LP) discs varies widely.

In the popular "hit parade" field, quality is made secondary to immediacy of issue and, often, the overall distortion is tolerated only because they are played on nondescript equipment by a non-critical audience.

For the average long playing disc much higher standards of reproduction are sought and maintained, while there is a high proportion of modern discs which qualify — by current standards — in the "excellent" to "superb" class. As often as not, the ultimate quality of such discs is limited, not by their own inherent characteristics, but by the original recording situation and the original tape equipment.

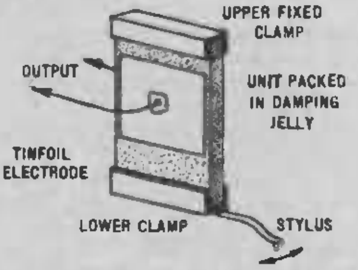


Fig 7: This greatly simplified diagram illustrates the principles of a piezo electric pickup. Movement of the stylus stresses the crystal or ceramic element, producing a corresponding signal voltage between its opposite faces.

Fig 8: A typical turnover crystal pickup cartridge, pictured about normal size and in the position it occupies when fitted into an arm. Note the stylus protruding from the bottom.

As with magnetic recording, problems of equalisation occur also with disc recording, though for different reasons.

In the case of disc recording, there is a tendency for low frequency audio waveforms to produce excessive deviation, leading to a danger of waveforms on one section of the track breaking through into waveforms on adjacent tracks.

At the other extreme, waveforms corresponding to high audio frequencies tend to diminish in physical dimensions to scarcely perceptible serrations in the sides of the groove.

To minimise these problems, and by common agreement, recording companies now manipulate the frequency response of their recording system to restrict (a) the natural increase in waveform magnitude with decreasing frequency and (b) the natural decrease in magnitude with increasing frequency.

The amount of such restriction is stated in the so-called RIAA recording standard and its effect is to produce something much closer to a **constant amplitude** recording characteristic.

If one were to feed into such a system a tone of constant loudness but gliding in frequency from very deep bass to very high treble, the resulting waves recorded on the disc, while varying in length along the groove, would not exhibit too great a variation in side-to-side deviation.

Putting this another way, it can be said that present-day practice is to attenuate the bass and boost the treble, while recording a disc, (in order to achieve a more convenient recording characteristic).

This has to be taken into consideration for playback, where bass needs to be boosted and treble needs to be attenuated for the final sound, as heard, to be level, or equalised.

For replaying records a so-called *pickup* is required. Most early pickups were designed as integral units, with the stylus assembly and generating mechanism assembled into an enlarged portion of the tracking arm. The more common approach, these days, is to regard the arm and base as one item and to fit into it a *cartridge* which combines the stylus and signal generating components. A complete phono pickup therefore comprises an arm (and base), a cartridge, and the necessary connecting leads.

Two classes of cartridge are widely used. By far the most numerous is the so-called "crystal" type or, to give it its proper name the "piezo-electric" cartridge.

In this type the stylus is attached, directly or through a tiny lever system, to a slab of material exhibiting the piezo-electric effect. Such materials produce an electrical potential between certain faces when they are subjected to physical stress.

In a piezo-electric cartridge, movement of the stylus stresses the material, producing a signal voltage between two of its faces, directly related to the movements of the stylus and the deviations of the recorded groove. The signal voltage developed by the piezo-electric element is picked up by metal foils cemented to the relevant faces and made accessible by attached wires.

When this signal is passed through an amplifier to a loudspeaker, it produces a sound similar to the sound originally recorded.

Two important characteristics of piezo-electric cartridges account for their very wide use.

The first is that they tend inherently to boost the bass response and to attenuate the treble, in playing a record. By a certain amount of manipulation of their mechanical design, it is possible to make this inherent characteristic just about balance out the reverse characteristic used for recording. As a result, they largely obviate the need for equalisation in the amplifier system.

The second factor is that their signal output voltage is fairly high, so that the signal can be fed into the same general class of amplifier as might be used, for example, to follow the detector of a radio or television receiver. This makes it very simple to produce combination radio-receivers and record players, because the same amplifier and loudspeaker system can be switched to serve either function.

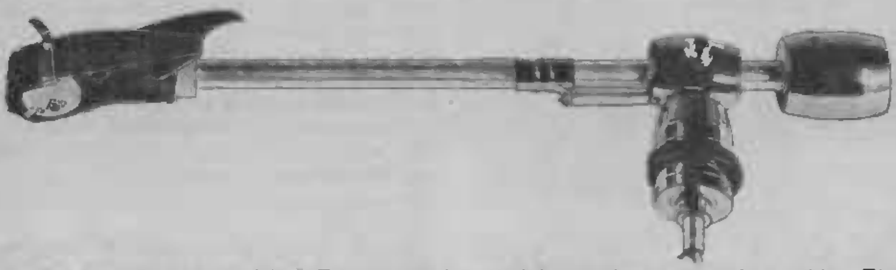


Fig 9: A complete phono pickup. The headshell at the left contains a magnetic cartridge. The pivot system is towards the right, with the signal leads protruding from the bottom. The counterweight (extreme right) helps to balance the weight of the headshell.

Two varieties of piezo-electric cartridge are in common use. In the original and still popular *crystal* type the piezo-electric element is a slab of rochelle salt. In practice, this material suffers the major disadvantage of being liable to deterioration under conditions of high temperature and high humidity, a certain number of failures occurring on this account.

The second type is the so-called *ceramic* cartridge where the piezo-electric element is a piece of specially manufactured ceramic. Cartridges of this type are not affected by temperature and humidity and have gained favour on that account. They deliver a somewhat lower signal output voltage than a typical crystal type, but it is usually sufficient still to avoid the necessity for special amplifier circuitry.

Crystal and ceramic cartridges commonly have a "turnover" facility, allowing them to be used alternatively with the now normal microgroove recordings, or with the old coarse-groove records intended to spin at 78rpm. In some cases the entire cartridge turns over inside the pickup headshell, presenting one of two quite distinct stylus assemblies in the playing position. In other types, only the stylus lever turns over, exposing either of two jewel tips.

As distinct from piezo-electric cartridges, the other major class is the *magnetic* type. These involve coils of wire and a magnetic field provided by a tiny, in-built permanent magnet. Their mechanisms vary widely in detail, but they all rely on the stylus effecting a change in the relative position of turns of wire and magnetic lines of force. As a result of this relative movement, signal voltage is produced across the coil (or coils) and is brought out through leads for subsequent amplification.

A great deal of effort has been directed to refining the design of magnetic type pickups to achieve the highest possible order of fidelity, along with low-tracking weight, the interests of minimum record wear. The refinement has involved not only the signal generating system — or cartridge — but the tracking arm and the bearing system on which it moves.

This effort has resulted in a range of magnetic pickups being available throughout the world, with extremely good performance characteristics — but quite highly priced in comparison with the more common piezo-electric types.

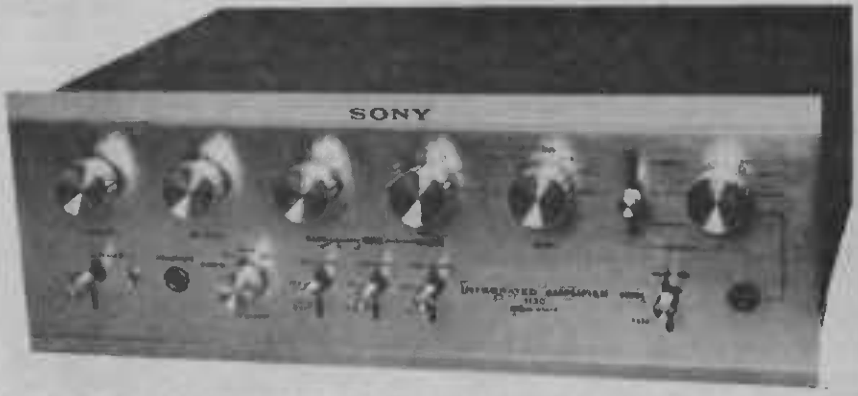


Fig 10: A modern high fidelity amplifier. The multiplicity of knobs and switches gives the operator close control over the general balance of the reproduction and allows the amplifier to be used easily with radio tuner, phono pickup, tape deck, & c.

Magnetic pickups, as a class, deliver a relatively low level of output signal, approximating that from a microphone or magnetic replay head. Unlike piezo-electric pickups, they provide no inherent correction for the original recording characteristics so that the amplifier into which they operate must provide high gain and full frequency compensation as well. The need for a special amplifier involves extra complication and cost. As a result, the use of magnetic pickups tends to be confined to the more expensive and elaborate class of record playing equipment.

Magnetic cartridges (therefore pickups) are rarely, if ever, manufactured with a "turnover" facility; complication of the stylus assembly would almost certainly prejudice their ultimate performance and this would be unacceptable in a prestige component.

Now for a brief look at the subject of amplifiers.

Because the subject is large and the space available is restricted, it may be helpful to cover some of the necessary ground, glossary style, listing and explaining terms which occur very frequently in discussions about audio amplifiers.

POWER OUTPUT: This is the amount of audio power which an amplifier, without producing serious distortion, can deliver to a loudspeaker. To be meaningful, it is best based on the RMS voltage which the amplifier will deliver across its intended load, when fed with a continuous or steady tone input signal. The resulting figure is referred to as the "RMS", or "continous tone", or "steady tone" power output — terms which have a fairly basic significance, not likely to be misunderstood.

Unfortunately, a practice has grown up of quoting the power output of amplifiers on the basis of a pulsed input signal and / or the peak rather than the RMS output voltage. The resulting output figures in terms of "music power", "peak power" or "peak music power" look much more impressive in specification but that is their main virtue (or should it be their main vice?). When comparing amplifiers, it is best to do so on the basis of the RMS figure, unless the significance of other figures is clearly understood.

In transistor portable receivers, the available audio power lies in the range 0.1 to 0.5 watt RMS. In mains powered domestic radio and television receivers, from 2 to about 10 watts is usually available. The power output from specialised high quality amplifiers is more commonly in the range 10 to 50 watts.

GAIN or AMPLIFICATION: This refers to the amount of amplification which is available in an audio amplifier system and is closely related to the magnitude of signal which has to be fed into it to obtain its maximum power output. From what has been said in earlier pages, it will be apparent that an amplifier which has to work from a microphone or magnetic pickup will need more gain or amplification than one which has to work from a crystal pickup or radio tuner.

FIDELITY or QUALITY: These are broad terms which describe the correspondence between the reproduced and the original signal. A good amplifier is one which is characterised by high-quality reproduction or *High Fidelity.* Putting it another way, it could be said that such an amplifier will introduce a minimum of distortion.

DISTORTION: This term has already been explained. As with microphones and pickups, an amplifier can introduce frequency distortion, harmonic distortion and intermodulation distortion.

INSTABILITY: An amplifier can oscillate, either due to an internal design fault or because of incorrect external connections. The oscillation may occur at an audible frequency and be heard as a "plopping" noise, a growl or a whistle, depending on its pitch. Alternatively, it may occur at a supersonic frequency and, while it will not then be directly audible, it will generally affect the performance of the amplifier in other respects, often causing noticeable distortion and a severe reduction in power output.

HUM: Hum is a fairly familiar problem in amplifiers operating from the power mains, being evident as a consistent low-pitched humming sound. It is due to energy from the AC power mains reaching the amplifier circuitry through lack of filtering in the power supply, through poor design in other respects, a faulty component or by unsuitable arrangement of the signal input leads.

NOISE: As distinct from hum, an amplifier may produce a certain amount of noise, most evident when the volume control is turned towards maximum. Most amplifiers produce some noise, evident as a gentle hissing or rustling sound, but it can reach objectionable proportions as a result of poor design or a faulty component.

By far the largest number of audio amplifiers in common use are those which serve as the audio systems of ordinary radio and television receivers. Their power output, as already indicated, is modest, being no greater than necessary — or convenient — for the particular class of use.

Their gain, or amplification, is also modest, since they need operate only from the comparatively large signal available from the detector.

The order of fidelity required might typically warrant such descriptions as "fair only" for small transistor receivers or "acceptable" in the case of domestic radio and television receivers. Listeners usually accept the sound from radio and TV receivers without much question, provided it is not positively "unpleasant" or "harsh" – which is as good as saying that the distortion level may be significant but not intolerable.

In terms of facilities, the minimum requirement is a *volume control,* being normally a potentiometer which affords control over the amount of signal passing through the amplifier and therefore the loudness of the sound, as ultimately heard from the loudspeaker.

A large number of receivers also have a *tone control* which usually takes the form of a potentiometer, so arranged in the circuit, that its operation allows the treble response to be reduced at will.

It might be considered that use of tone control in this way is a form of deliberate frequency distortion and this may well be true. However, many listeners seem to like the "mellow" kind of sound which results from attenuating the higher frequencies and provision of a control for this purpose is, at one time, a concession and a sales feature.

Fig 11: Rear view of a typical loudspeaker cone, with the voice coil cemented to its apex. Special flexible leads connect the voice coil to lugs or terminals mounted on the cone housing. Weight and texture of the cone has an important bearing on the performance characteristics.

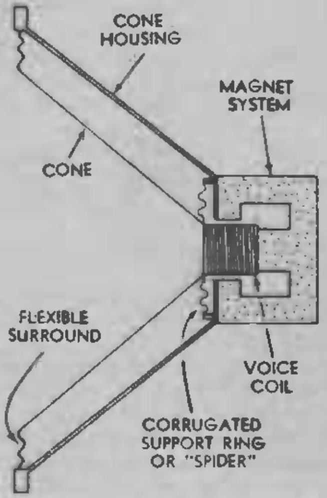


Fig 12: A dynamic (ie moving coil) loudspeaker, sketched in cross-section. Note the flexible support around the edge and the apex of the cone. The magnetic field surrounding the voice coil can be provided either by an electromagnet or a permanent magnet, the latter being by far the most common.

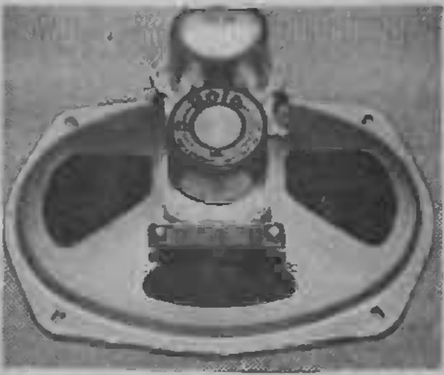


Fig 13: A 6 / 9-inch oval permanent magnet dynamic loudspeaker — a size commonly used in car radio installations. The magnet is attached centrally to the cone housing. Above it is the output transformer — an item that has largely disappeared as valves have given place to transistors.

Where a receiver is to be used as part of a "radiogram," with the combined facility of playing records, there follows the implication that the user may be rather more critical than usual, and more demanding in terms of fidelity.

There is a tendency, therefore, for the amplifiers in radiograms to be designed to marginally higher standards and, in some cases, to provide separate knobs for controlling independently the level of bass and treble.

While these controls can conceivably be used to improve the sound from a record which exhibits questionable bass-to-treble balance, they are more frequently used to modify the sound to the way the listener happens to prefer it.

Over and above the large number of people who own and use radiograms on a fairly routine or casual basis, there are many with an intense interest in reproduced sound and these fill the ranks of the so-called high fidelity enthusiasts. "Hi-Fi" enthusiasts, as a group, tend to purchase or build amplifiers with facilities and performance figures well ahead of the amplifiers used in ordinary radiograms.

Almost invariably, such amplifiers include an extra stage, referred to as a *pre-amplifier,* to enable them to operate from the low level signal produced by a magnetic pickup. This provides the extra amplification necessary, as well as the requisite frequency compensation or equalisation.

Again, almost without exception, high fidelity amplifiers incorporate a special tone control stage providing facilities for treble boost or treble cut, and bass boost and bass cut. These are often supplemented by rumble and noise filters, loudness compensation circuits and so on. With such facilities, an enthusiast can vary the balance of sound to suit his own tastes, or cope better with recordings which are technically suspect.

As already mentioned, the power output available from special quality amplifiers is usually quite high, in the range of say 10 to 50 watts.

Precautions are taken also to ensure a low order of harmonic and intermodulation distortion from the amplifier. This is secured automatically, in part, by providing for high power output so that, in practice, the amplifier is always operated well below its maximum capacity. In addition, however, the circuit and components are selected with greater emphasis on performance and less on cost and simplicity.

Considerable attention is paid also to filtering and to other points in the design which might affect the inherent hum level. Because the loudspeaker systems used in a high fidelity installation are likely to be very efficient at 50 and 100Hz, the two main hum frequencies with 50Hz power mains, an amplifier hum level which might be unnoticed in a fairly ordinary receiver or record player, may be quite objectionable in a premium quality installation.

As far as actual design is concerned,

quality amplifiers vary widely, both in physical appearance and electrical circuitry. However, while many claims and counterclaims are made, it is not difficult to produce amplifiers, these days, which have excellent characteristics and which add very little to the total distortion of a reproducing system.

LOUDSPEAKERS

From amplifiers, we pass to a brief discussion of loudspeakers. Since at least 1930, the vast majority of speakers used for sound reproduction have been of the so-called *dynamic* type, using the moving coil principle of operation.

In this type of loudspeaker, a small solenoid coil, wound usually on a paper former, is attached to the apex of a cone. The cone is suspended within the frame of the loudspeaker by a corrugated ring or a "spider" around the apex, and a corrugated cloth or ring around the periphery. So suspended, it is able to move back and forth, although it always tends to return to an intermediate or rest position.

The moving coil or *voice coil* attached to the apex of the cone is arranged to be inside a magnet structure which envelopes it in a strong magnetic field (see diagram). Initially, the coil structure is not affected by this field because it is made of copper wire and wound on a paper former — both non-magnetic substances.

However, if audio current from the amplifier is passed through the voice coil, the current sets up a magnetic field which interacts with the fixed magnetic field, causing the coil to move rapidly back and forth with the alterations of the audio current. In fact, it behaves as a simple type of motor.

As the coil vibrates back and forth, it imparts the same motion to the cone, setting up air waves which recreate the original sound.

During the 1930s the vast majority of dynamic (i.e. moving coil) loudspeakers used an electromagnet system to provide the fixed magnetic field, because this was the most convenient and effective way at the time. Surrounding the voice coil was a structure of soft iron enclosing a quite large coil containing, usually, many hundreds or even many thousands of turns of copper wire.

Current was passed through this so-called *field* coil, sometimes from a separate DC power supply and sometimes by wiring the field coil directly across the high tension supply in the associated receiver or amplifier. A third and popular arrangement was to wire the field coil in series with the power supply, so that it functioned also as a filter choke.

The DC resistance was an important characteristic of a field coil, since it had to suit the voltage and current available from any particular power supply, receiver or amplifier.

The DC resistance of loudspeaker field coils ranged, commonly, from about 700 ohms to about 7000 ohms. Values from 700 to 2500 ohms were most frequently involved in the "filter" connection, while those above 2500 ohms were frequently connected directly across HT supply circuits.

The actual wattage dissipated in the fields, being also a measure of the strength of the magnetic field produced, ranged from 3 or 4 watts in very small loudspeakers to about 25 watts in very large ones.

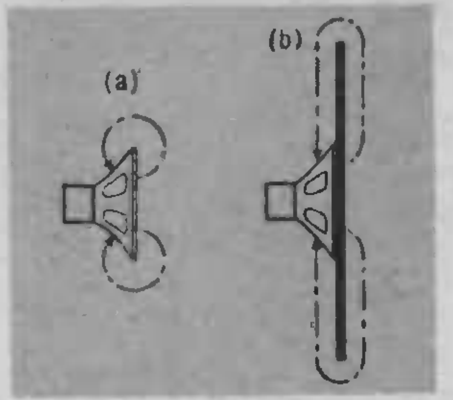


Fig 14: Fitting a baffle (b) increases the path length from front to rear of the cone and improves propagation of low frequency sound. Enclosures are a development from this basic concept.

Since that era, it has become possible to produce permanent magnets for loudspeakers which can provide the same field density — or greater — much more easily and cheaply than from electromagnets.

As a result, field coil type or *electrodynamic* loudspeakers have given place to loudspeakers using permanent magnet fields, described in full as "permanent magnet dynamic loudspeakers." In fact, this type of loudspeaker is now so universal that a permanent magnet field is taken for granted, electrodynamic units being found only in very old equipment.

A second major factor to do with dynamic loudspeakers is the resistance or impedance of the voice coil. This coil has to be relatively small and light, to be able to vibrate at high audio frequencies. With rare exceptions, it has to be limited to a few dozen turns of fine copper or aluminium wire wound on the cylindrical former. The impedance at audio frequencies lies usually within the range 2 to 35 ohms, with 4, 8 and 16 ohms being the values most commonly encountered. The DC resistance of a voice coil, as measured on an ohmeter, is usually about 0.8 of the nominal impedance.

An impedance of this general order is much too low for direct connection to valve type audio amplifiers. These as a rule require to work into a load impedance typically within the range 2000 to 10,000 ohms.

As a result, it has been standard practice to provide a coupling transformer — or *output* transformer — between a valve amplifier and the loudspeaker. The transformer is given a turns ratio such that it can *match* the output load requirements of the amplifier to the resistance or impedance of the loudspeaker voice coil. The output transformer is sometimes fitted into the amplifier, sometimes attached to the loudspeaker frame.

While a huge number of valve type amplifiers are still in service (in record and tape players, radio receivers and TV sets) most new audio systems use transistors. As distinct from valves, audio power transistors can operate directly into the order of impedance which can be presented by a loudspeaker voice coil. As a result, the need for an output transformer has largely disappeared.

The small output transistors used in battery-powered portable radio receivers and record players commonly require to operate into a voice coil impedance in the range 20 to 50 ohms. Mains powered radio and television receivers use loudspeakers with a nominal impedance in the range 10 to 20 ohms, while high fidelity systems commonly require loudspeakers with an impedance in the range 4 to 16 ohms.

With transistor amplifiers it is particularly important to observe their load requirements when attaching or substituting loudspeakers. A transistor amplifier should never be operated into a loudspeaker impedance lower than its specifications allow; the higher peak currents which can flow in these circumstances may ruin the output transistors. Operating into a higher than optimum load will not usually cause damage, though it will normally reduce the available power output.

As most readers will probably know already, loudspeakers vary in size from the very tiny units used in transistor portable receivers to very large units, 12 to 18 inches in diameter, used in home hi-fi systems, electronic musical instruments, public address, theatre systems and so on.

Very tiny loudspeakers are broadly capable of doing no more than producing

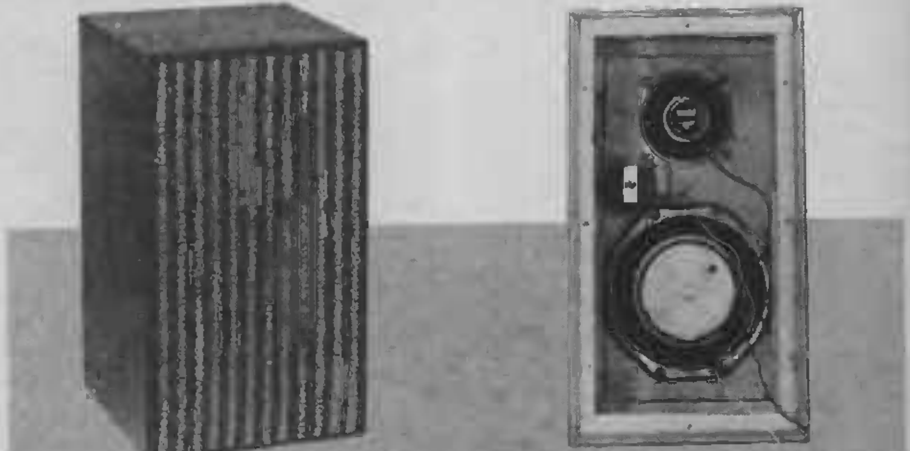


Figure 15: A typical compact speaker system measuring 14 x 8 x 8½ inches. The rear view shows the two speakers and the frequency divider network.

the limited kind of sound heard from baby transistor receivers. The potential quality rises with cone size but, from nominal 6-inch loudspeakers upwards, other characteristics, as well as mere size influence the ultimate quality of sound.

If advertisements are to be believed, there are no poor loudspeakers in this class, only good through to superb! In fact, however, quality of reproduction varies enormously and the newcomer, surveying the scene, must rely partly on independent recommendations and the general principle that quality follows price, at least in a measure.

Perhaps it should be mentioned that the majority of loudspeakers are designed for and used as full-range units; that is to say, they are fed with the full range of frequencies available from the amplifier and reproduce those frequencies to the extent of their ability to do so. Not unexpectedly, some loudspeakers are much better than others in this respect.

Where extra cost can be tolerated, there is some advantage in splitting up the frequency range so that all the lower frequencies are handled by one loudspeaker (say below 3,000cps) while the higher frequencies are handled by another. Or, again, three loudspeakers may be used as a group with one handling frequencies below say 400Hz, another frequencies between 400Hz and 5,000Hz and a third handling frequencies above 5,000Hz.

Where this is done, it is usual to select loudspeakers designed to handle the appropriate range of frequencies.

Those designed to handle only the lower frequencies are commonly referred to as *"woofers,"* while high frequency loudspeakers have acquired the name *"tweeters."* Where a *mid-range* loudspeaker is used, it is commonly a general-purpose type, but fed only with the middle frequencies.

To divide up the signal in terms of frequency, so that it can be diverted to this loudspeaker or to that, *frequency divider networks* are commonly employed, involving one or more capacitors and, frequently, one or more inductors.

BAFFLE SYSTEMS

For proper results, a moving coil loudspeaker must be provided with some form of *baffle* system.

When the loudspeaker cone vibrates at an audio rate, it alternately compresses and rarifies the air adjacent to its surface, creating sound waves. These move away from the cone at the normal speed of sound – approximately 1,100 feet per second.

At frequencies above a few hundred Hertz the sound waves radiate fairly efficiently into the surrounding area, more or less as a broad beam of sound diverging from the surface of the cone. Sound is actually produced from both surfaces of the cone — front and rear — but that from the rear is liable to be obstructed by the body of the loudspeaker and whatever supports it.

However, toward the lower end of the frequency spectrum, there is an increasing tendency for air, compressed as the cone moves in either direction, simply to flow around to the other side of the cone, where there is a zone of rarefication. Thus, instead of the loudspeaker radiating low frequency sound into the surrounding area, it tends simply to "pump" air back and forth around the periphery of the cone.

HIGH FIDELITY

Traditionally, the term high fidelity (often abbreviated to "hi-fi") has been used within the audio industry and by audio enthusiasts, to describe components and systems which could provide a higher quality of reproduction than they currently accepted as "standard." In purchasing a "high fidelity" pickup, amplifier or loudspeaker, one could reasonably expect to receive a better unit than the ordinary and cheaper item not so classified.

Infringements of this general idea certainly occurred from time to time but the industry and enthusiasts generally have found it convenient to sustain it.

The term, being a relative one, has reflected the rising standards of reproduction through the years. Products that might reasonably have been regarded as "high fidelity" at a particular time would be no better than was common-place a few years later.

In the 1950 era, following the introduction of magnetic recording and microgroove discs, standards of musical reproduction in the home rose sharply and this seemed to provide the signal for all and sundry to attach the words "high fidelity" to their products, relative quality notwithstanding.

To be sure, quite ordinary equipment of the post-1950 era could sound as well as earlier high fidelity equipment but the term rapidly lost its meaning to the public. It now appears on a large proportion of tape and record players offered for sale, and even on records and tapes of rather dubious quality.

However, while the term has been rendered valueless, as far as the public is concerned, it is still used within the industry and amongst audio enthusiasts in its original sense — to describe equipment capable of better than average performance by reason of extended frequency response, minimal distortion and adequate power handling capacity. It will probably continue to be so used.

As far as the listener is concerned, the loudspeaker becomes less and less efficient as the frequency is lowered and the reproduction therefore appears to be lacking in bass.

To overcome the effect, it is necessary to limit the flow of air around the periphery of the cone and thus ensure that the low frequency pressure waves are propagated into the listening area.

The most elementary baffling method is to mount the loudspeaker over a hole cut in a flat board. The larger the board the better, since it then provides more complete isolation between the front and rear of the cone. If the loudspeaker could be mounted in a dividing wall between two rooms, the isolation would be virtually complete.

Because neither flat baffle boards nor holes cut in a wall are a very practical solution to the problem, it is much more usual, for domestic situations, to have the loudspeaker of a radio / TV receiver or radiogram mounted in the front face of an ordinary radio or TV cabinet.

A large cabinet provides a fair degree of baffling and ensures reasonable bass response but, as the cabinet is diminished towards "mantel" or "portable" size, bass response tends to diminish with it. The extreme is seen in personal portable transistor sets where the combination of a tiny loudspeaker and very little effective baffling results in reproduction which is completely lacking in bass.

While the vast majority of loudspeakers are mounted in radio and TV cabinets of one kind or another, those interested in better quality reproduction have put an enormous amount of effort into devising baffle systems — some to give optimum bass response, irrespective of size and price, right through to the other extreme, where the prime objective is small size with bass response preserved as far as possible.

Baffle systems other than mere radio cabinets are commonly referred to as "enclosures". Loudspeaker enclosures are designed and built to a variety of patterns and given a variety of names, some of which are basic technical terms, others being names which zealous proponents have dreamed up!

The discussion has been confined to ordinary dynamic types and their baffle requirements for the reason that they are, far and away, the most widely used. It must suffice merely to mention other types which the reader may encounter:

P.A. HORN LOUDSPEAKERS: Commonly used for outdoor public address work, these have a compact moving coil driver unit feeding into the small end of a metal horn. Speakers of this type are substantially weatherproof, highly directional and efficient over the range of "speech" frequencies, being therefore well suited to public address applications. However, restricted frequency range severely limits their usefulness for reproducing music.

ELECTROSTATIC LOUDSPEAKERS: These have a very thin metallic membrane adjacent to, but separated from, a metallic back plate. When audio voltages are fed between membrane and back plate, the membrane vibrates and creates sound waves. The device may be likened to a capacitor having one plate flexible enough to vibrate in the presence of applied signal voltage. Small electrostatic loudspeakers have been used frequently as "tweeters" but full-range electrostatic loudspeakers are available. The latter are highly regarded for their performance but have not, as yet, posed any significant challenge to the conventional dynamic type.

RIBBON LOUDSPEAKERS: These are the converse of the ribbon microphone. In the loudspeaker, a thin non-magnetic metal ribbon, usually about 3in by ¼in, is suspended in an intense magnetic field. When signal current is passed through the ribbon, it vibrates and sets up sound waves. Ribbon loudspeakers are suitable only as tweeters and, while capable of excellent results, are expensive and rather easily damaged.

CHAPTER 23

Stereo Sound Reproduction

Limitations of mono reproduction. The search for sound "dimension". Two-channel stereo discs. Two-channel stereo in the home. The latest concept: 4-channel sound. Simulated 4-channel reproduction. Matrix-type 4-channel disc records. The CD-4 or "discrete" 4-channel disc record.

In the discussion of radio receivers and audio amplifiers to this point, it has been assumed that the objective has been to transmit, or record, or amplify, or reproduce one distinct audio signal — be it music, speech, sound effect or a mixture of all three. In short, we have assumed a single channel or **monophonic** signal — a term that is commonly abbreviated to **mono**.

Up to about 1950, it was scarcely necessary to consider anything else, because most familiar sound reproducing systems employed only a single channel — medium-wave broadcast receivers, television receivers, sound motion pictures, and tape and disc recordings.

Apart from the cost and complexity of providing more than a single sound channel, there appeared to be no strong demand for anything more elaborate. The most obvious requirement was for clean, enjoyable sound at an adequate power level and free from distortion. Even today, most ordinary broadcast, short-wave and television stations provide only single-channel (or mono) sound, as do most film theatres for most of the time.

Audio signals can be extremely complex containing, as they do, "information" about the frequency, phase and amplitude of each individual component of the original sound, However, a mono signal contains no information about the direction from which the original sounds came relative to the microphone (or microphones).

Thus, while an orchestra may be heard and enjoyed as a whole, it is not possible to nominate from the reproduced sound the relative positions of the original instruments or instrumental groups. At best, one might guess that the violins were at the left, percussion at the rear and so on, but there is no way of being sure.

This remains true, whether the original sound is picked up by one microphone or by several. Information relative to direction is lost immediately the total signal is combined into a single amplifier or a single recording or a single transmission channel.

Once "directional" information has been sacrificed there is no way in which it can genuinely be recovered. At best, it can only be subsequently contrived or faked.

Over the years, generations of audio enthusiasts have tried, in a variety of ways, to counter the lack of directional information in single-channel sound and to achieve a more natural sense of spread, or dimension. They have sought to get away from the impression of all the sound issuing from one localised source, as if it were coming through a small round hole in the studio wall! In fact, monophonic reproduction is sometimes referred to, derisively, as "hole-in-the-wall" sound.

In an effort to overcome this effect, some have exploited the use of multiple loudspeakers, woofer / tweeter combinations, &c, with the individual units widely separated. Others have mounted loudspeakers at oblique angles, to bounce the sound from adjacent walls and thus render the actual sound source less apparent.

At best, however, such measures can only be a palliative and no real substitute for the directive quality missing from the reproduced sound.

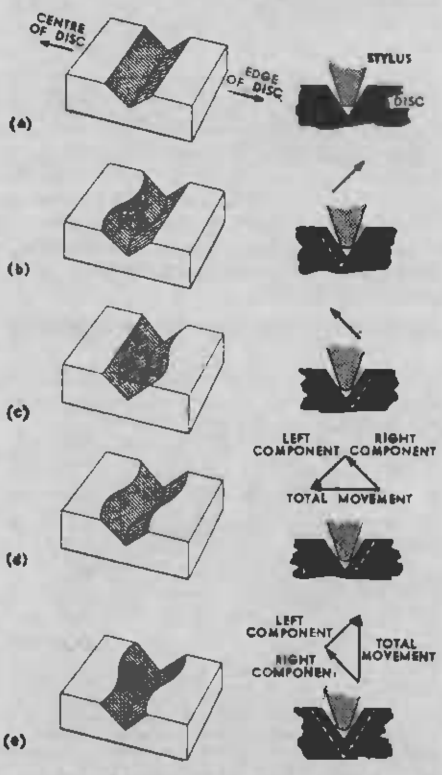


Fig. 1: The basis of a 2-channel stereo recording. (a) plain groove; (b) left-hand channel modulated; (c) right-hand channel modulated; (d) both channels modulated in phase; (e) both channels modulated out of phase.

Audio engineers have known for the best part of 50 years that directional qualities could be imparted to sound reproduction by using multiple channels. For example, a number of microphones might be ranged in front of an orchestra and their signals conveyed directly, or by way of a multi-channel recording, to as many loudspeakers similarly placed in front of a remote audience.

The loudspeakers will then virtually reconstitute the "spread" of the original sound, allowing the audience to listen selectively to individual sections of the orchestra.

The real problem was to take advantage of this basic knowledge, without running into unacceptable complexity and cost.

Its first important practical application was to the sound tracks of large-budget feature films. Signals from multiple microphones and other sources were recorded optically (later magnetically) onto the release prints. After amplification, the signals were fed to loudspeakers ranged behind the screen and even, in some cases, around the auditorium itself. By such means, it was possible to have sounds coming from left, right or centre, appropriate to the visual action, and to reproduce orchestral music with a real sense of spread.

However, it was one thing to provide multi-channel sound in a theatre; it was quite another to envisage multi-channel sound for ordinary listening in the ordinary home situation — no matter how desirable it might seem from the viewpoint of realism, therefore of ultimate fidelity.

In terms of radio broadcasting, there appeared to be no easy way to superimpose multiple signals on the one station carrier, while it was equally impractical to allot more than one transmitter to each program — to say nothing of the need for more than one receiver in every listening room.

Similarly, there seemed no easy way to record multiple signals on the one track of a disc record — or at least not without introducing serious attendant problems.

Until the mid-fifties it seemed that audio enthusiasts would be denied dimensional sound in the home, having to content themselves instead with winning the best possible sound from a mono system.

At the time, magnetic tape players offered the most obvious promise of a breakthrough, because there was no particular difficulty about recording and playing back two or more parallel tracks on tape, each track representing a discrete microphone / loudspeaker link. However, the domestic mass market was not ready for tape and tape players in any form, let alone in one which would demand multiple amplifiers and loudspeakers!

Nevertheless, a potential demand was there and disc recording engineers began to consider seriously the possibility of adapting to microgroove records an idea that had been put forward by the brilliant British engineer Blumlein in the early thirties. This offered a means of recording two substantially independent audio signals in the one groove. By using a suitably designed pickup cartridge, the two signals could be recovered and fed via two amplifiers to separate loudspeakers, one to the front-right and the other to the front-left of the listening position.

A typical stereo record player system intended to be accomodated on shelves in a domestic living room. The amplifier offers full tone control facilities and a power output of 8 watts RMS per channel. (Picture by courtesy of EMI)

A fortunate aspect was that a twin-channel recording, as proposed, could use the existing microgroove standards for disc size and speed, with only a minor variation in the exact groove dimensions. The way was open, therefore, for the industry to introduce a complete new generation of twin-channel records and record playing equipment, on which existing mono records could still be played — without the stereo effect, of course.

While there was some debate as to whether two channels were really sufficient, it soon became apparent that well-recorded two-channel sound could create the sense of direction and dimension that was so obviously missing from mono reproduction. Public acceptance was immediate and, with the disc showing the way, twin-channel tape systems also "took off." Nowadays, a large proportion of reel-to-reel machines, tape cartridge and tape cassette players provide for twin-channel reproduction.

Attention turned also to radio broadcasting and means were devised of modulating two channels on to FM broadcasting stations, at least. A large proportion of FM broadcasting stations around the world now transmit two-channel program material. However, partly for technical reasons, medium-wave broadcasting stations, short-wave stations and television stations still transmit in mono only.

Twin-channel recordings, designed to feed loudspeakers placed one to the left and one to the right of the listener, came to be known as **stereophonic recordings**, commonly abbreviated to the single word **stereo**. Because of the association, some have mistakenly assumed that the term "stereo" signifies "two." In fact, it has the connotation of "solid," or "having dimension" and needs to be broadened in its usage. We should strictly be talking here about "two-channel stereo," bearing in mind that certain other stereo sound systems (as in large theatres) use more than two channels.

Stereo disc records and stereo disc record players — or **stereograms** — are now virtually standard in Australian homes.

Conventional disc recordings, carrying only the one signal, are known as **monophonic** or simply **mono** records. (The term **monaural** should be avoided in this context, since it really means "one ear.")

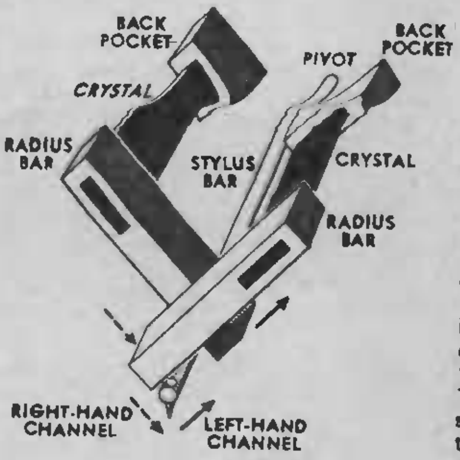


Fig 2: A simplified illustration of one particular type of stereo crystal cartridge. Stylus movement, transmitted by radius bars, twists one or both crystals producing an equivalent output.

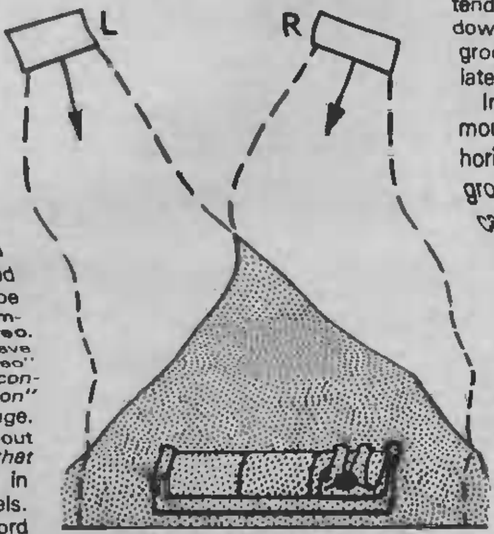


Fig 3: A typical listening room situation for 2-channel stereo. A virtual sound image can be created at least as broad as the space including and between the two loudspeakers. They would normally be placed between 10ft and 12ft apart.

Just as there are mono and stereo disc recordings and playing equipment, so there are mono and stereo tape recordings and players.

To all appearances, a stereo disc record looks the same as a mono record but there is a vital difference in the way in which the groove is inscribed.

In a conventional mono record, the groove deviates laterally and causes the pick-up to generate a signal when its stylus is vibrated from side to side.

For a stereo record, a special cutting head is used in which the stylus tip is driven obliquely and independently by "left" and "right" signals, fed to different driving coils. The left channel signal tends to drive the stylus tip in one direction, at 45 degrees to the surface of the record, while the right channel tends to drive the stylus at 45 degrees in the other direction. (See Fig 1)

For all practical purposes one may consider that the left and right-hand signals appear as separate indentations on opposite walls of the V-shaped groove.

However, because the recorded signals tend to drive the cutting stylus up and down, as well as from side to side, the groove in a stereo recording varies both laterally and in depth.

In a stereo pickup, the stylus is so mounted that it can move vertically, horizontally and obliquely to follow the groove. The generating system inside the cartridge, be it magnetic or piezo electric, is arranged to deliver two distinct output signals, corresponding to respective 45-degree movements of the stylus tip and to the original "left" and "right" channel information. (See Fig 2.)

The two signals so produced are taken through two amplifiers to separate loudspeakers placed, preferably, several feet apart. To listeners seated in front of the loudspeakers, the sound appears to be spread out in a much more natural fashion than is possible from a mono system.

It is generally agreed that, for ordinary domestic listening, the loudspeakers should be about ten feet apart, while the optimum listening position is towards the opposite end of the room, as indicated in Fig 3.

Perhaps it should be mentioned that producers of stereo recordings follow various techniques to secure the effect they want.

A typical stereo record player system intended to be accomodated on shelves in a domestic living room. The amplifier offers full tone control facilities and a power output of 8 watts RMS per channel. (Picture by courtesy of EMI)

the most obvious promise of a breakthrough, because there was no particular difficulty about recording and playing back two or more parallel tracks on tape, each track representing a discrete microphone / loudspeaker link. However, the domestic mass market was not ready for tape and tape players in any form, let alone in one which would demand multiple amplifiers and loudspeakers!

Nevertheless, a potential demand was there and disc recording engineers began to consider seriously the possibility of adapting to microgroove records an idea that had been put forward by the brilliant British engineer Blumlein in the early thirties. This offered a means of recording two substantially independent audio signals in the one groove. By using a suitably designed pickup cartridge, the two signals could be recovered and fed via two amplifiers to separate loudspeakers, one to the front-right and the other to the front-left of the listening position.

A fortunate aspect was that a twin-channel recording, as proposed, could use the existing microgroove standards for disc size and speed, with only a minor variation in the exact groove dimensions. The way was open, therefore, for the industry to introduce a complete new generation of twin-channel records and record playing equipment, on which existing mono records could still be played — without the stereo effect, of course.

While there was some debate as to whether two channels were really sufficient, it soon became apparent that well-recorded two-channel sound could create the sense of direction and dimension that was so obviously missing from mono reproduction. Public acceptance was immediate and, with the disc showing the way, twin-channel tape systems also "took off." Nowadays, a large proportion of reel-to-reel machines, tape cartridge and tape cassette players provide for twin-channel reproduction.

Attention turned also to radio broadcasting and means were devised of modulating two channels on to FM broadcasting stations, at least. A large proportion of FM broadcasting stations around the world now transmit two-channel program material. However, partly for technical reasons, medium-wave broadcasting stations, short-wave stations and television stations still transmit in mono only.

Twin-channel recordings, designed to feed loudspeakers placed one to the left and one to the right of the listener, came to be known as **stereophonic** recordings, commonly abbreviated to the single word **stereo**. Because of the association, some have mistakenly assumed that the term "stereo" signifies "two." In fact, it has the connotation of "solid," or "having dimension" and needs to be broadened in its usage. We should strictly be talking here about "two-channel stereo," bearing in mind that certain other stereo sound systems (as in large theatres) use more than two channels.

Stereo disc records and stereo disc record players — or **stereograms** — are now virtually standard in Australian homes.

Conventional disc recordings, carrying only the one signal, are known as **monophonic** or simply **mono** records. (The term **monaural** should be avoided in this context, since it really means "one ear.")

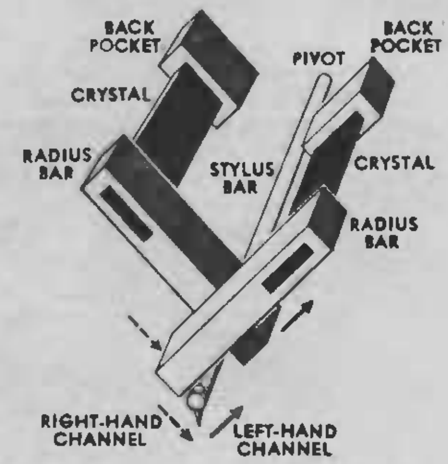


Fig 2: A simplified illustration of one particular type of stereo crystal cartridge. Stylus movement, transmitted by radius bars, twists one or both crystals producing an equivalent output.

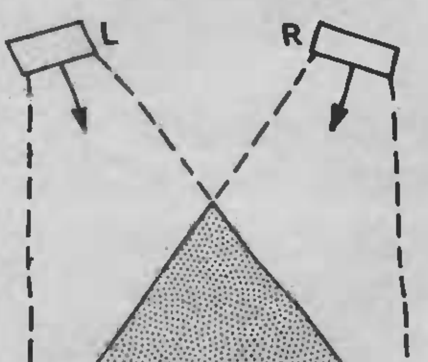


Fig 3: A typical listening room situation for 2-channel stereo. A virtual sound image can be created at least as broad as the space including and between the two loudspeakers. They would normally be placed between 10ft and 12ft apart.

Just as there are mono and stereo disc recordings and playing equipment, so there are mono and stereo tape recordings and players.

To all appearances, a stereo disc record looks the same as a mono record but there is a vital difference in the way in which the groove is inscribed.

In a conventional mono record, the groove deviates laterally and causes the pick-up to generate a signal when its stylus is vibrated from side to side.

For a stereo record, a special cutting head is used in which the stylus tip is driven obliquely and independently by "left" and "right" signals, fed to different driving coils. The left channel signal tends to drive the stylus tip in one direction, at 45 degrees to the surface of the record, while the right channel tends to drive the stylus at 45 degrees in the other direction. (See Fig 1)

For all practical purposes one may consider that the left and right-hand signals appear as separate indentations on opposite walls of the V-shaped groove.

However, because the recorded signals tend to drive the cutting stylus up and down, as well as from side to side, the groove in a stereo recording varies both laterally and in depth.

In a stereo pickup, the stylus is so mounted that it can move vertically, horizontally and obliquely to follow the groove. The generating system inside the cartridge, be it magnetic or piezo electric, is arranged to deliver two distinct output signals, corresponding to respective 45-degree movements of the stylus tip and to the original "left" and "right" channel information. (See Fig 2.)

The two signals so produced are taken through two amplifiers to separate loudspeakers placed, preferably, several feet apart. To listeners seated in front of the loudspeakers, the sound appears to be spread out in a much more natural fashion than is possible from a mono system.

It is generally agreed that, for ordinary domestic listening, the loudspeakers should be about ten feet apart, while the optimum listening position is towards the opposite end of the room, as indicated in Fig 3.

Perhaps it should be mentioned that producers of stereo recordings follow various techniques to secure the effect they want.

the most obvious promise of a breakthrough, because there was no particular difficulty about recording and playing back two or more parallel tracks on tape, each track representing a discrete microphone / loudspeaker link. However, the domestic mass market was not ready for tape and tape players in any form, let alone in one which would demand multiple amplifiers and loudspeakers!

Nevertheless, a potential demand was there and disc recording engineers began to consider seriously the possibility of adapting to microgroove records an idea that had been put forward by the brilliant British engineer Blumlein in the early thirties. This offered a means of recording two substantially independent audio signals in the one groove. By using a suitably designed pickup cartridge, the two signals could be recovered and fed via two amplifiers to separate loudspeakers, one to the front-right and the other to the front-left of the listening position.

A fortunate aspect was that a twin-channel recording, as proposed, could use the existing microgroove standards for disc size and speed, with only a minor variation in the exact groove dimensions. The way was open, therefore, for the industry to introduce a complete new generation of twin-channel records and record playing equipment, on which existing mono records could still be played — without the stereo effect, of course.

While there was some debate as to whether two channels were really sufficient, it soon became apparent that well-recorded two-channel sound could create the sense of direction and dimension that was so obviously missing from mono reproduction. Public acceptance was immediate and, with the disc showing the way, twin-channel tape systems also "took off." Nowadays, a large proportion of reel-to-reel machines, tape cartridge and tape cassette players provide for twin-channel reproduction.

Attention turned also to radio broadcasting and means were devised of modulating two channels on to FM broadcasting stations, at least. A large proportion of FM broadcasting stations around the world now transmit two-channel program material. However, partly for technical reasons, medium-wave broadcasting stations, short-wave stations and television stations still transmit in mono only.

Twin-channel recordings, designed to feed loudspeakers placed one to the left and one to the right of the listener, came to be known as **stereophonic** recordings, commonly abbreviated to the single word **stereo**. Because of the association, some have mistakenly assumed that the term "stereo" signifies "two." In fact, it has the connotation of "solid," or "having dimension" and needs to be broadened in its usage. We should strictly be talking here about "two-channel stereo," bearing in mind that certain other stereo sound systems (as in large theatres) use more than two channels.

Stereo disc records and stereo disc record players — or **stereograms** — are now virtually standard in Australian homes.

Conventional disc recordings, carrying only the one signal, are known as **monophonic** or simply **mono** records. (The term **monaural** should be avoided in this context, since it really means "one ear.")

A typical stereo record player system intended to be accomodated on shelves in a domestic living room. The amplifier offers full tone control facilities and a power output of 8 watts RMS per channel. (Picture by courtesy of EMI)

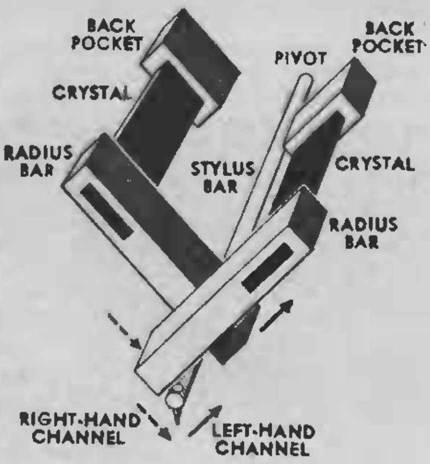


Fig 2: A simplified illustration of one particular type of stereo crystal cartridge. Stylus movement, transmitted by radius bars, twists one or both crystals producing an equivalent output.

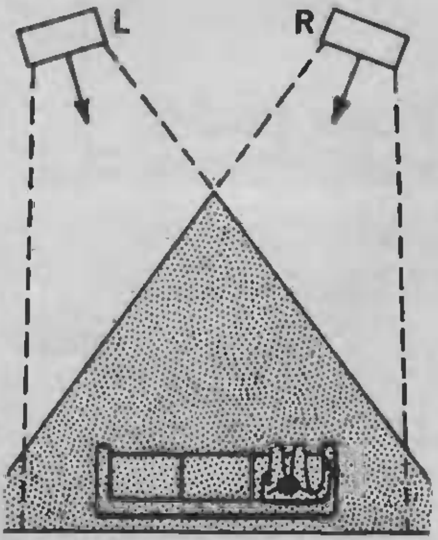


Fig 3: A typical listening room situation for 2-channel stereo. A virtual sound image can be created at least as broad as the space including and between the two loudspeakers. They would normally be placed between 10ft and 12ft apart.

Just as there are mono and stereo disc recordings and playing equipment, so there are mono and stereo tape recordings and players.

To all appearances, a stereo disc record looks the same as a mono record but there is a vital difference in the way in which the groove is inscribed.

In a conventional mono record, the groove deviates laterally and causes the pick-up to generate a signal when its stylus is vibrated from side to side.

For a stereo record, a special cutting head is used in which the stylus tip is driven obliquely and independently by "left" and "right" signals, fed to different driving coils. The left channel signal tends to drive the stylus tip in one direction, at 45 degrees to the surface of the record, while the right channel tends to drive the stylus at 45 degrees in the other direction. (See Fig 1)

For all practical purposes one may consider that the left and right-hand signals appear as separate indentations on opposite walls of the V-shaped groove.

However, because the recorded signals tend to drive the cutting stylus up and down, as well as from side to side, the groove in a stereo recording varies both laterally and in depth.

In a stereo pickup, the stylus is so mounted that it can move vertically, horizontally and obliquely to follow the groove. The generating system inside the cartridge, be it magnetic or piezo electric, is arranged to deliver two distinct output signals, corresponding to respective 45-degree movements of the stylus tip and to the original "left" and "right" channel information. (See Fig 2.)

The two signals so produced are taken through two amplifiers to separate loudspeakers placed, preferably, several feet apart. To listeners seated in front of the loudspeakers, the sound appears to be spread out in a much more natural fashion than is possible from a mono system.

It is generally agreed that, for ordinary domestic listening, the loudspeakers should be about ten feet apart, while the optimum listening position is towards the opposite end of the room, as indicated in Fig 3.

Perhaps it should be mentioned that producers of stereo recordings follow various techniques to secure the effect they want.

Classical orchestral records are generally recorded in such a way as to produce an even spread of the orchestra across the intervening space between the loudspeakers. Some popular music is recorded this way also.

However, plenty of popular music is arranged to emphasise channel separation, so that a lead sound may issue from one speaker, and the accompaniment from the other.

It is also possible to manipulate a recording so that a soloist appears to be located close-up, midway between the loudspeakers, while the accompaniment is widely spread. In a well balanced system, this "3-channel" effect can be quite startling.

Stereo records should only be played with stereo pickups, since these can be expected to have the necessary lateral and vertical compliance (ie the ability to move in a lateral or vertical direction) to follow the modulation in the stereo groove.

While a mono pickup may seem to play a stereo record and produce from it an acceptable monosignal, there is a chance that it will damage the groove in so doing, particularly if it is an old style cartridge having very little vertical compliance.

However, while a mono pickup should not be used with anything but mono records, a stereo pickup can be used freely with either stereo or mono records.

FOUR CHANNEL STEREO

While 2-channel stereo discs and tapes have dominated the home entertainment market for more than a decade, a new trend is now gaining momentum — 4-channel!

For many years, musicians have realised that a recording containing ONLY sound information obtained from the performers may be precise — but may also be lacking in "atmosphere." Recordings may benefit by the presence of some natural reverberation from the auditorium or recording studio. In actuality recordings, a certain amount of audience noise can increase the sense of listener involvement.

This kind of "ambience" is present in most mono and 2-channel stereo recordings, being achieved either by placing the microphone(s) sufficiently far away from the performer(s) or by deliberately mixing with the main channel the output from one or more remote microphones.

While this contributes to the enjoyment of present-day recordings, it has the disadvantage that the reverberation and audience noise comes through the main loudspeakers in front of the listener, whereas it should be coming from beside and behind the listening position.

Why not extra channels and extra loudspeakers, as indicated in Fig 4?

This leads naturally into further speculation: Musicians and producers who have hitherto exploited the interplay of instruments through left and right channels, would be able to project sounds at will from behind the listening position. From this can flow any amount of argument as to whether listeners really want to be made aware of echoes and audience noise, or be assaulted by sound from all directions.

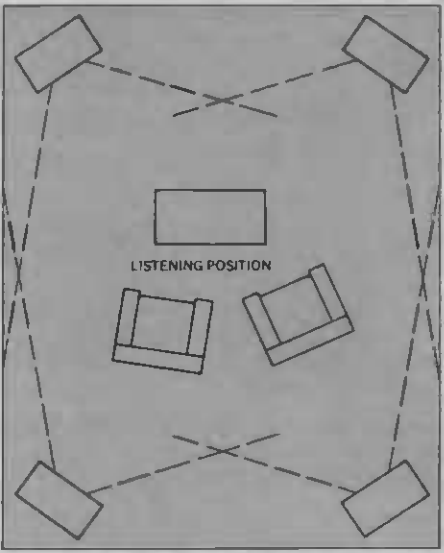


Fig. 4: Ideally the listening room for quadraphonic reproduction should be arranged like this.

From the domestic viewpoint, there can be obvious difficulties in having to place another two loudspeakers in the listening room and re-locate the seating. For many, the ideal of Fig 4 will be out of the question, leaving Fig 5 as the most likely compromise.

And, of course, there is a natural reluctance to spend money on extra equipment and replace existing stereo discs and tapes with new, four-channel releases. In fact, these observations are reminiscent of argument which went on during the mono / stereo transition and it seems likely that the transition from 2-channel to 4-channel will follow the same pattern.

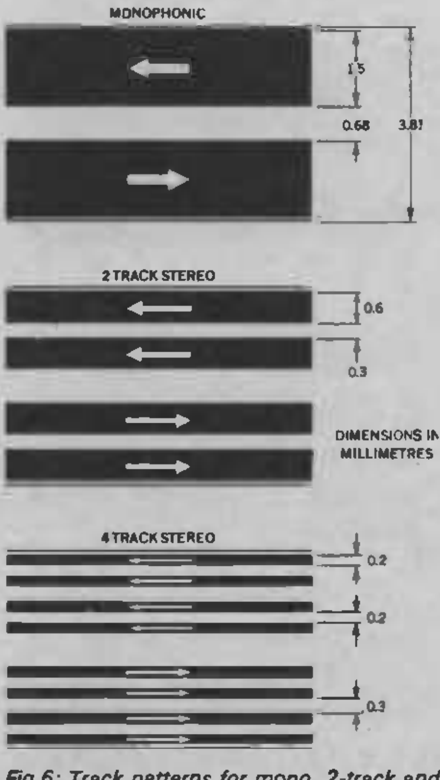


Fig. 6: Track patterns for mono, 2-track and 4-track stereo configurations on cassette tape. They are compatible but the 8-track configuration would demand critical mechanical accuracy.

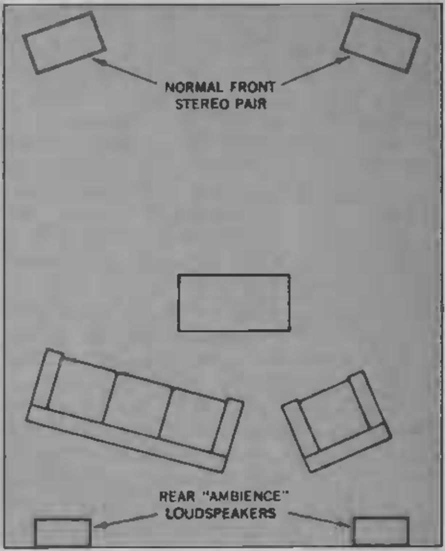


Fig. 5: To save space, small loudspeakers may be placed high up on the rear wall, as shown.

In one sense, history has already repeated itself. Initially, it was assumed that any demand for extra channels would have to be met by the tape medium. A logical approach seemed to be to modify the existing 2-channel stereo format so that four tracks would be laid down in place of two. Ideally, such a configuration could be made compatible by arranging that the four tracks could be scanned by a four-section head structure for four-track stereo, by a two-section head to give a logical two-channel stereo output, or by a single wide-gap head to replay in mono mode. (See Fig 6.)

The logic of this approach, plus the fact that most listening rooms have four corners, set the pattern that two-channel stereo should ultimately be replaced by four-channel stereo — not by three- or five-channel.

Out of this came a variety of coined terms: quadrasonic, quadrisonic, quadsonic, quadphonic and so on. The one which seems to have gained favour is quadraphonic, mainly because it has a syllabic relationship to monophonic and stereophonic. Either that or "four-channel stereo," having in mind that stereo basically means "solid," as we have already observed.

One other word should be mentioned at this point. Since the tracks on a quadraphonic tape can be entirely independent in terms of what they contain, each potentially as important and distinct as the other, such a recording has been termed four-channel "discrete."

Efforts to market 4-channel tapes and equipment met with only limited success and, while such equipment is represented in the catalogues, it is very much a minority interest. As happened in a previous decade, it became apparent that any transition to a new system of sound reproduction would have to be effected primarily by disc equipment and in a way which would cause the least possible dislocation. The demand for 4-channel tapes would then follow as a matter of course.

Out of this situation emerged, in the short term, various schemes for simulating quadraphonic reproduction using existing 2-

channel disc recordings. The basic idea was — and still is — to combine the "right" and "left" signals of a 2-channel recording out of phase, so that a "difference" signal is produced which has no counterpart in what is fed to the front loudspeakers. By feeding this difference signal to loudspeakers behind the listening position, another and distinct sound is produced which can readily be interpreted as one or more additional channels.

In orchestral recordings made originally with a single stereo microphone cluster, the difference signal tends to contain a minimum of the direct frontal information, because combining the two signals out of phase cancels it out. Signals from other directions, including echoes, are not cancelled to the same extent. As a result, the difference signal can provide a larger proportion of ambience information.

Again, where a record has been made with solo effects unique to one channel, they will appear without any significant cancellation in the difference signal, and therefore in the rear loudspeakers. The end result can be a rather accidental and sometimes startling distribution of the sound sources around the room, not very different from what might be done deliberately with a true 4-channel system.

In terms of circuitry, four channel simulation can be approached in a number of ways. The simplest is probably the system popularised by Hafler-Dynaco in America, where two additional loudspeakers, connected out of phase, are fed with a modified difference signal picked up directly from the loudspeaker outlets of an existing stereo amplifier. (See Fig 7.)

An elaboration of the idea, first introduced in Australia by AWA, introduces transformers into an add-on simulation unit, giving somewhat greater versatility in the way in which the rear loudspeakers can be interconnected.

A still further idea, popularised by "Electronics Australia" magazine, blends the right and left signals in an adaptor unit using four transistors and a few other components, producing signals which are intended to be fed to a second stereo amplifier and thence to a pair of rear loudspeakers. Though more elaborate, it gives complete control over the rear channels and can produce pleasant and convincing simulated quadraphonic reproduction.

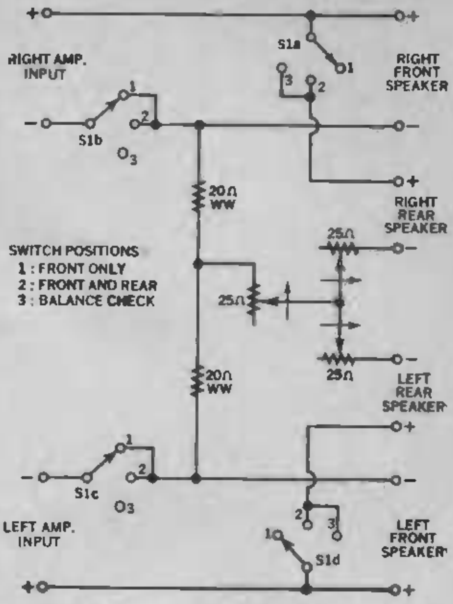


Fig 7: Circuit details of the QD-1 Quadaptor marketed by the Dynaco Company. It feeds a "difference" signal to two additional loudspeakers, without involving an additional amplifier.

Over and above these approaches, overseas manufacturers have produced a variety of very elaborate 4-channel synthesisers, some separate, some combined with complete amplifiers.

The concept of simulated 4-channel stereo is important in the marketing sense because it provides a vital bridge between existing 2-channel technology and equipment and what lies ahead. Audio enthusiasts can purchase supplementary equipment of one type or another, at any time they choose, and go right on playing their existing two-channel records or tapes — either in two-channel stereo or simulated quadraphonic.

But, so equipped, they are in a favourable position to play genuine 4-channel material, provided it is compatible. And from this remark flows what has been undoubtedly the greatest trauma in the hifi industry during the last decade — the problem of producing 4-channel discs which hopefully, will be compatible with 2-channel equipment and technology.

At first glance, the problem may seem incapable of solution. While a stylus may be driven in two virtually independent directions as per Fig 1, there is no obvious way for it to be driven in four independent directions, as would seemingly be necessary to produce four independent signals.

In fact, however, a partial solution has been evolved, based largely on the work of an American enthusiast-engineer-musician by the name of Peter Scheiber.

Scheiber used a matrix or combining circuit (Fig 8) to merge four independent audio signals on to two channels. He showed that, by reversing the procedure, it was possible to recover four signals from the two channels. With such a system, a large amount of intermixing or cross-talk is inevitable, with the result that the four recovered signals differ significantly from the four originals. However, Scheiber claimed that by careful design of the encode and decode matrices, an acceptable result would be obtained.

A vital point about Scheiber's method was that it made no new demands on existing recording and playback techniques, being equally adaptable to disc or tape recording and to stereo FM broadcasting. Encoded four-channel material could be recorded or broadcast, and recovered either as mono or stereo on existing equipment, or decoded and reproduced as a quadraphonic program.

After a period of uncertainty, audio engineers around the world addressed themselves to the matrix system and came up with quite an array of names, circuits and patents relating to their particular and favoured matrix configuration. Some of them have found common ground but there is still a great deal of rivalry between the "SQ" system favoured by Sony, CBS and EMI, the "QS" system favoured by Sansui, and "RM" (regular matrix) favoured by those who have been prepared to sink their differences.

For optimum decoding, playback equipment needs to include matrices designed for the specific systems and, in fact, many adaptors and amplifier systems now appearing on the market do have selector switches offering this facility.

How important such a facility is has yet to be proven in the marketplace. Most decoder circuits tend to produce a reasonably convincing "surround" effect from almost

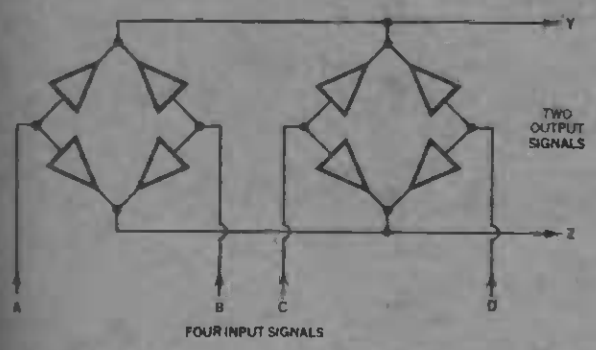


Fig 8: A basic 4:2 encoding matrix, from four channels to two channels. The proportions of A, B,C & D which appear on Y and Z depend on the gain, phase and frequency characteristics of the individual matrix amplifiers.

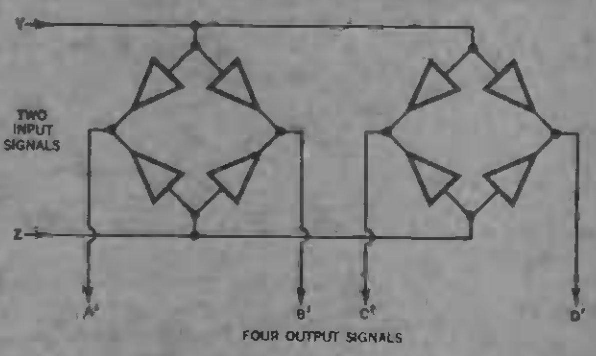
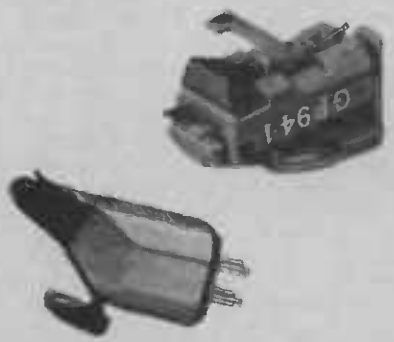


Fig 9: A 2:4 decoding matrix, from two channels to four channels. Note that the four output signals are branded as A', B', C' and D', implying that the four output signals are not identical with the original input signals A, B, C & D.

all encoded program material — and from a great many existing 2-channel recordings as well. The subtleties of apparent sound. source position may well be more important to engineers than they are to the listening public! Some of the fervour may therefore drain out of arguments about the respective matrix systems.

One important group of manufacturers, headed by JVC-Nivico in Japan and RCA in America have adopted the attitude that the matrix system in any form is inadequate, because of its inherent cross-talk problem. They have evolved a quite different system which aims to provide four distinct (or discrete) signals from a disc record.

During recording, all the right hand signals (front + rear) are fed to the stereo cutter in the normal way and appear on one wall of the standard stereo groove. Similarly, all the left hand signals (front + rear) are inscribed on the other wall. When played back on ordinary 2-channel equipment, the record is heard in normal stereo, with normal frontal separation. In this sense, the so-called CD-4 discrete disc is completely compatible.



Typical stereo cartridges able to cope with mono, 2-channel stereo or 4-channel matrix. At top a ceramic type, below a magnetic cartridge.

There is more to it, however. A difference signal is derived (front — back) for each side. These difference signals are modulated on to a high frequency carrier and the resulting band of frequencies from 20kHz to 45kHz are impressed on the groove wall, along with the main audio modulation. Ordinary 2-channel equipment simply ignores this extra information so that it does not compromise compatibility.

However, by playing a CD-4 recording with a suitably designed pickup and passing the signal through a demodulating and decoding unit, the four signals so derived can be combined to produce the four signals originally fed to the recording equipment.

The CD-4 system is technically more complex and more expensive and requires more specialised replay equipment in the home for optimum results. Its proponents claim that this is justified by results and that it avoids the compromises of the matrix method. Champions of the matrix approach take the reverse position — naturally!

Four-channel sound is undoubtedly here to stay but the consumer will be the final arbiter as to the relative popularity of program sources: tape reel, tape cartridge, tape cassette, matrix disc or discrete disc!

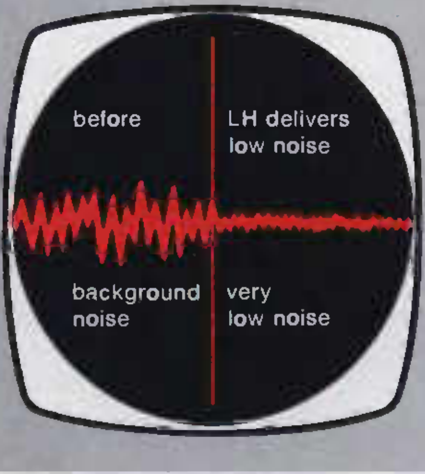
Advertising Directory

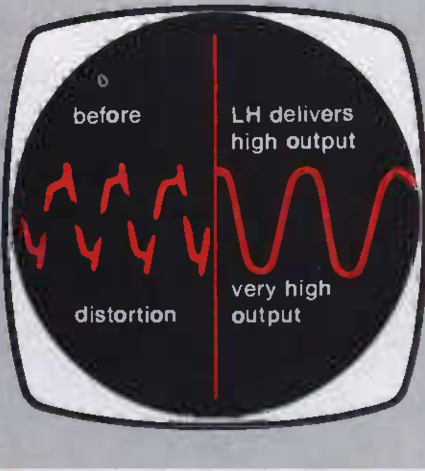
AEGIS PTY LTD, 347 Darebin Rd, Thornbury 3701	112
AMATEUR ASTRDNOMER, 118 Clarke St, Crows Nest 2065	55
AMPLION (A'SIA) PTY LTD, 29 Majors Bay Rd, Concord 2137	108
MSP, 554 Parramatta Rd, Ashfield 2131	60
AUSTRALIAN GENERAL ELECTRIC PTY LTD, 86-90 Bay St, Ultimo 2007	12
BWD ELECTRONICS PTY LTD, 329 Burke Rd, Gardiner 3146	4
DICK SMITH ELECTRONICS PTY LTD, Pacific Highway, Gore Hill 2065	89
EDGE ELECTRIX, 25A Burwood Rd, Burwood 2134	92
FAIRCHILD AUSTRALIA PTY LTD, 420 Mt Dandenong Rd, Croydon 3136	26
FERGUSON TRANSFORMERS PTY LTD, 331 High St, Chatswood 2067	58, 84
HAM RADIO SUPPLIERS, 323 Elizabeth St, Melbourne	84,112
H. ROWE & CO PTY LTD, 185 Hoddle St, Richmond 3121	11
I.C.S., 400 Pacific Highway, Crows Nest 2065	16, 35
IMPORTED COMPONENTS, PO Box 1683P, Melbourne 3001	68
JACOBY KEMPTHORNE PTY LTD, 469-475 Kent St, Sydney 2000	IBC, OBC
J. H. McGRATH, 208 Little Lonsdale St, Melbourne 3000	61
KITSETS AUSTRALIA PTY LTD, PO Box 176, Dee Why 2099	100
MAURICE CHAPMAN PTY LTD, 276 Castlereagh St, Sydney 2000	IFC, 58
McGILLS NEWSAGENCY PTY LTD, 187 Elizabeth St, Melbourne 3000	75
PHILIPS INDUSTRIES LTD, 95 York St, Sydney 2000	2
PLESSEY DUCON PTY LTO, PO Box 2, Villawood 2163	42
PLESSEY ROLA PTY LTD, The Boulevarde, Richmond 3121	74
R. H. CUNNINGHAM PTY LTD, 439-499 Victoria St, West Melbourne 3003	29
RADID DESPATCH SERVICE, 869 George St, Sydney	93
SIDEBAND ELECTRONICS, PO Box 23, Springwood 2777	71, 112
TECHNICAL BOOK & MAGAZINE CO, 295-299 Swanston St, Melbourne 3000	65
WATKIN WYNNE PTY LTD, 32 Falcon St, Crows Nest 2065	55
WILLIS TRADING CD, 445 Murray St, Perth 6001	108

Why is BASF Magnetic Tape LH hifi?

1. low noise and high output 2. faithful reproduction of music and speech 3. small distortion factor 4. large frequency range 5. wide dynamic range 6. no oxide rub off 7. easy on equipment 8. extremely supple 9. super-tensile strength 10. can be stored indefinitely — resistant to tropical conditions

That's why





BASF

Distributors: Sydney (Head Office): Maurice Chapman & Co. Pty. Ltd., 276 Castlereagh St., 2000. **Newcastle:** W. L. Redman Agencies, 11 Hall St., N.S.W., 2300. **Canberra:** Sonny Cohen & Sons, 20 Isa St., A.C.T., 2600. **Melbourne:** Maurice Chapman & Co. Pty. Ltd., 146-150 Burwood Rd., Hawthorn, Vic., 3122. **Brisbane:** Chandlers Pty. Ltd., 399 Montague Rd., West End, Qld., 4101. **Adelaide:** Neil Muller Pty. Ltd., 8 Arthur St., Unley, S.A., 5061. **Perth:** Anderson-Tedco, 11-13B Belmont Ave., Belmont, W.A., 6104. **Launceston:** P. & M. Distributors, 87A Brisbane St., Tas., 7250. **Darwin:** Pfitzners Music House, 2 Darwin Arcade, Smith St., N.T., 5790.


SONY
SONY
SONY
SONY