As you are now the owner of this document which should have come to you for free, please consider making a donation of £1 or more for the upkeep of the (Radar) website which holds this document. I give my time for free, but it costs me in excess of £300 a year to bring these documents to you. You can donate here https://blunham.com/Radar

Please do not upload this copyright pdf document to any other website. Breach of copyright may result in a criminal conviction.

This document was generated by me Colin Hinson from a document held at Henlow Signals Museum. It is presented here (for free) and this version of the document is my copyright (along with the Signals Museum) in much the same way as a photograph would be. Be aware that breach of copyright can result in a criminal record.

The document should have been downloaded from my website https://blunham.com/Radar, if you downloaded it from elsewhere, please let me know (particularly if you were charged for it). You can contact me via my Genuki email page:
https://www.genuki.org.uk/big/eng/YKS/various?recipient=colin

You may not copy the file for onward transmission of the data nor attempt to make monetary gain by the use of these files. If you want someone else to have a copy of the file, please point them at the website (https://blunham.com/Radar).

Please do not point them at the file itself as the file may move or be updated.

I put a lot of time into producing these files which is why you are met with this page when you open the file.

In order to generate this file, I need to scan the pages, split the double pages and remove any edge marks such as punch holes, clean up the pages, set the relevant pages to be all the same size and alignment. I then run Omnipage (OCR) to generate the searchable text and then generate the pdf file.

Hopefully after all that, I end up with a presentable file. If you find missing pages, pages in the wrong order, anything else wrong with the file or simply want to make a comment, please drop me a line (see above).

It is my hope that you find the file of use to you personally – I know that I would have liked to have found some of these files years ago – they would have saved me a lot of time !

Colin Hinson
In the village of Blunham, Bedfordshire.

AIR PUBLICATION 1762
Issued October, 1939

AIR MINISTRY

ELECTRICAL AND RADIO NOTES FOR WIRELESS OPERATORS

LONDON
HIS MAJESTY'S STATIONERY OFFICE
Price 3*s* 6*d* net

WIRELESS TELEGRAPHY *offers* a lucrative and successful career in the Merchant Navy. Positions as Radio Officers assured when qualified. We have had more than 40 years experience in training students for the above career.

Students accepted from 16 years of age. Moderate fees, payable by instalments, if desired.

School is equipped with the latest Marconi ship's wireless apparatus, identical to that installed at the General Post Office for the practical test in the examinations.

199 Certificates of Proficiency awarded to our students during the past three years.

We have in addition Postal Courses of instruction in the theory of Wireless Telegraphy. These are most useful to those who are proficient in sending and receiving in the Morse Code.

The School is recognised by the Ministry of Education as efficient for training students for the Postmaster General's examinations.

Prospectus free on application to

THE BRITISH SCHOOL OF TELEGRAPHY LTD.

179, CLAPHAM ROAD, LONDON, S.W.9. (*Est.* 1906)

CONTRACTORS TO H.M. GOVERNMENT, THE G.P.O.
& FOREIGN GOVERNMENTS

RADIO SERVICING EQUIPMENT

B.P.L. SUPER RANGERS
B.P.L. SIGNAL GENERATORS
B.P.L. UNIVERSAL BRIDGE
B.P.L. WOBBULATORS
B.P.L. VOLTASCOPE

Write for details to

BRITISH PHYSICAL LABORATORIES

HOUSEBOAT WORKS, RADLETT, HERTS.

Telephone: RADLETT 5674-5-6
Cables: BEEPEELLE, RADLETT

AIR PUBLICATION 1762
Issued October, 1939

AIR MINISTRY

ELECTRICAL AND RADIO NOTES FOR WIRELESS OPERATORS

LONDON: HIS MAJESTY'S STATIONERY OFFICE
1940
(Reprinted 1951)

This handbook is issued for the information and guidance of all concerned.

By Command of the Air Council.

CONTENTS

CHAPTER I .. .. *page* 1

THE ELECTRONIC STRUCTURE OF MATTER.

Structure of the atom. Free electrons. Conductors and insulators. Current, potential difference, resistance. Electrical units.

CHAPTER II .. .. *page* 6

CELLS AND BATTERIES.

The simple cell. Leclanché and dry cells. Cell grouping. Resistance, Ohm's law. Specific resistance. The series circuit. Volts drop. The parallel circuit. The mixed circuit. Cell on open and on closed circuit. E.M.F. and P.D.

CHAPTER III .. .. *page* 16

THE EFFECTS OF A CURRENT.

The heating effect. Energy considerations. The chemical effect; chemistry of the accumulator. Charging circuit. The magnetic effect; fields of straight wire and of solenoid. Iron cored solenoid. Hysteresis.

CHAPTER IV .. .. *page* 29

THE ELECTRIC MOTOR ; MEASURING INSTRUMENTS.

The motor principle. Simple motor; commutator. Field excitation. Shunt motor starter. Moving coil instruments; shunts and multipliers. Moving iron instruments. Hot wire instruments. Thermo-junction instrument. Metre bridge. Megger.

CHAPTER V .. .. *page* 40

ELECTROMAGNETIC INDUCTION.

Faraday's experiment. Faraday's law. Magnitude of induced E.M.F. Lenz's law. Self-induction, self-inductance. Mutual induction, mutual inductance. Time constant. The dynamo principle. Motor generator. Field excitation. Back voltage of motor. Eddy currents; laminations.

CHAPTER VI .. .. *page* 52

ELECTROSTATICS.

Electric fields around charged conductors. Potential in the electric field. Capacity of an isolated conductor. Simple condenser; capacity of a condenser. Fixed and variable condensers. Grouping of condensers. Field energy. Charge and discharge of a condenser. Exponential discharge; oscillatory discharge. Closed oscillatory circuit; damping. Open oscillatory circuit; radiation. Electrostatic voltmeter.

CHAPTER VII .. .. *page* 64

ALTERNATING CURRENTS (1).

Sinusoidal alternating quantities. Phase difference. Vector addition of alternating quantities. The series A.C. circuit. Series resonance. The parallel A.C. circuit; parallel resonance. Dynamic resistance. Coupled circuits. Transformer; matching.

CHAPTER VIII .. .. *page* 77

ALTERNATING CURRENTS (2).

High frequency circuits; irregular behaviour of components. Coil losses and dielectric losses; equivalent resistances. Selectivity. Radiation from H.F. circuits; electromagnetic waves. Simple receiving circuit. Sound; characteristics of audible notes. Microphone, telephone and loudspeaker. Filter circuits.

CHAPTER I

THE ELECTRONIC STRUCTURE OF MATTER

It has long been an accepted scientific theory that all forms of matter are built up of molecules, a molecule being the smallest portion of any piece of matter which can exist and still retain the characteristic properties of the substance.

Molecules are exceedingly small, so small that we can never hope to see them with the most powerful microscope because the light, to which our eyes are sensitive, is far too coarse in its structure for such a delicate task. But their minuteness is no hindrance to our knowing quite a lot about them, and the science of chemistry has taught us that molecules can be broken down into smaller particles still, called atoms ; and, moreover, having analysed all the varied types of molecules, we learn that only 92 different types of atoms exist. This is a tremendous simplification of the study of matter, and it was regarded for many years as the end of the story. In fact, the very word atom means indivisible. If we look in any chemistry text book we shall find a list of these atoms, arranged in ascending order of atomic weight, with the hydrogen atom as No. 1, and at the other end uranium as No. 92. The number of any atom in this series is called its atomic number, and its weight (taking the oxygen atom as having a weight of 16 units) is called its atomic weight.

Two or more similar atoms combining together to form a molecule produce substances we call *elements*, and the names of these elements have been adopted as the name of the corresponding atom. Hydrogen, carbon, oxygen, sulphur, iron, platinum and gold are elements, and their molecules consist entirely of atoms of the same name. Substances whose molecules contain different atoms are called *compounds*. Water (2 hydrogen atoms, 1 oxygen atom, formula H_2O), Chalk ($CaCO_3$), Copper Sulphate ($CuSO_4$), Sulphuric Acid ($H_2 SO_4$), are some common compounds.

Up to about the close of the 19th century, electricity had been regarded as something of a mystery. Two kinds were recognised, negative electricity, such as is produced when amber is rubbed with flannel, and positive electricity, produced by rubbing glass with silk. It was known that similar kinds of electricity exerted repulsive forces on one another, and that opposite kinds attracted one another. It was also known that by means of voltaic batteries electricity could be made to stream along wires and do useful things, like lighting lamps and ringing bells, but little was known about its actual physical nature.

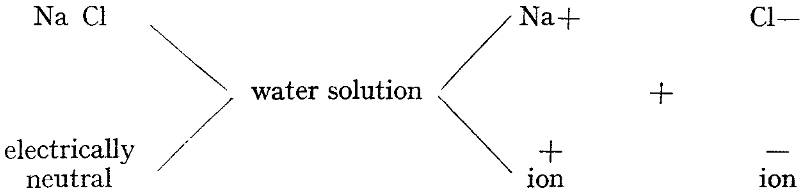
It was Sir J. J. Thomson who threw the first light on the mystery and opened up a new universe of science when he discovered that negative electricity appeared to consist of multitudes of minute particles, each having 1/1800th part of the mass of the hydrogen atom and identical electric charges. He found that these particles were present in all forms of matter and were always the same, whatever the material from which they came, and to summarise in a few words the outcome of his work, we now know that the so-called indivisibles, the atoms, are themselves constructed entirely from particles of positive electricity (PROTONS) and of negative electricity (ELECTRONS).

We believe for example that the hydrogen atom, which is the lightest and simplest of all atoms, consists of a central core or nucleus consisting of 1 proton with 1 electron spinning round it in an orbit like the earth around the sun, but instead of completing. one revolution in a year it completes many millions in a second.

The proton or elementary positive charge, is relatively heavy, having about the same mass as a hydrogen atom, but like the electron, it is incredibly small. If a single hydrogen atom [see fig. 1.1 (a)] could be magnified to the size of St. Paul's Cathedral, the proton and the electron would each be about the size of a pin's head : this means, of course, that the atom is nearly all empty space.

POTENTIAL DIFFERENCE (P.D.). We also speak of cells and dynamos as possessing an ELECTROMOTIVE FORCE (E.M.F.), for they can set electricity in motion if a complete circuit exists. The term E.M.F. refers to the electrical pressure generated in the interior of any source, which results in the production of a P.D. between its terminals. P.D., on the other hand, always implies a *comparison* between *two* points, e.g. we say " the P.D. between the ends of a wire ", but *not* " the E.M.F. across it ".

Solid conductors are not the only things which contain electric particles capable of being set in motion ; if common salt for example, is dissolved in water its molecule (1 sodium atom, 1 chlorine atom, Na Cl) is ionised according to this scheme :—



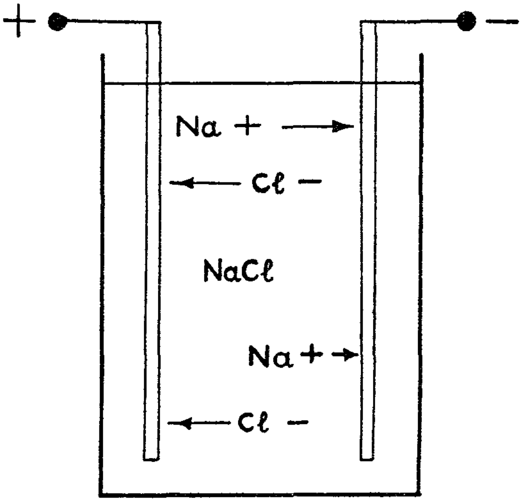


FIG. 3.1. Conduction in an electrolyte

Thus, if the terminals of a battery are connected to two plates immersed in a salt solution (fig. 3.1) these ions, by virtue of the electric charges they carry, are set in motion, the positive ion moving towards the low potential or negative end of the battery and the negative chlorine to the high potential or positive end.

Inside the conducting fluid or ELECTROLYTE there are two sets of moving particles, the positive ions moving down the potential slope, from high to low electric pressure, and the negative ions moving upwards towards the positive plate, from low to high electric pressure.

Similarly gases (under ordinary circumstances very good insulators), can be made to conduct quite well, and, as in liquid conductors, there are positive and negative carriers moving in opposite directions.

We are in something of a difficulty here, because when we tackle the problems of engineering we do not want to have to stop and enquire whether we are considering the motion of positive particles down the potential slope or of negative ones up the slope, so we resolve the difficulty by a convention, " the direction of a current in a conductor is taken as the direction in which the positive ions move, or tend to move " that is to say, in spite of our knowledge of electrons drifting along metallic conductors, we shall still think of electricity as a positive flow, down the potential slope, like water flowing downhill. When we want to speak of moving electrons specifically we shall speak of the " electron stream " : this is opposite in direction to that agreed upon as the direction of the electric current. This convention will be adhered to throughout this book.

Units

The common unit employed to measure water is the gallon, and we are so used to speaking about gallons that the idea presents no difficulty, but when we come to select units for measuring such a subtle thing as electricity it is not quite so easy.

Negative electricity consists of electrons, and positive electricity of protons ; they contain identical but opposite quantities, and therefore either would do as a unit of electricity, but they are so minute that even a small electric charge measured in such electric units would have to be expressed in enormous numbers : it would be like expressing a gallon of water as so many molecules. We agree therefore to lump 10^{19} electrons together as our unit of electricity and call it a COULOMB. A coulomb therefore is the unit of electricity. P.D. and E.M.F. are measured in VOLTS ; the P.D. between the terminals of an ordinary accumulator is about 2 volts : the P.D. between the commercial supply mains is about 200 volts.

Electric current and its measurement require a little consideration, because the word current in its common usage has not quite the same meaning as in its scientific sense and this difference is sufficient to cause confusion.

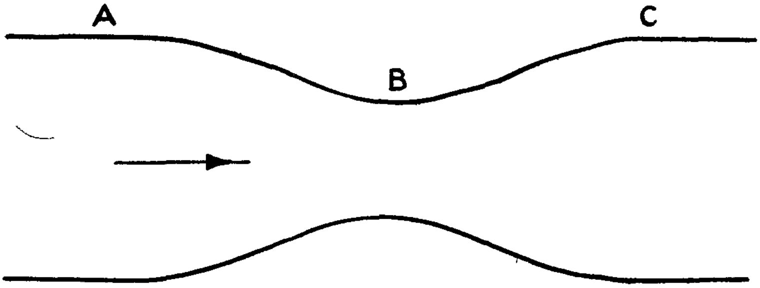


FIG. 4.1. To illustrate the meaning of current

Fig. 4.1 represents a water pipe with a constriction at B : water is driven along from A through B and C : in general conversation we might say that the current is higher at B than it it is A or C, and our meaning would be understood, but strictly speaking, while the *speed* of the water is higher at B than at A or at C, the current, or more properly the strength of the current, is the same at all three points. This is in consequence of two things (i) the definition of current strength, and (ii) the incompressibility of water.

The current strength at any section such as A is defined as the number of gallons of water passing that section per second. Since water is incompressible, the gallons passing A per second must be just the same as at B and C as well : that is, if there are no inlets or outlets, the current strength is the same at all points along a pipe. If, while the water is flowing, another constriction such as B is made in the pipe, the current strength falls, not only in the neighbourhood of the new constriction, but all the way along the pipe, but again at this new value the current has the same strength at every point.

Similarly when electricity is caused to stream through a conductor, the number of coulombs passing any section per second is taken as a measure of the current strength. For convenience we have chosen a single word AMPERE to stand for coulomb per second : thus a conductor in which there is a current of 10 amperes passes 10 coulombs of electricity per second.

Let us follow this water analogy a little further. The friction of the walls of the tube and the viscous drag between the water particles constitute a kind of opposition to the flow : the pipe, we say, offers resistance to the water flow. It is not difficult to imagine that electricity streaming through the inter-atomic spaces of a conductor will also experience opposition due to the collisions between the drifting electrons and the fixed atoms : every conductor in fact has electrical resistance, the value of which is expressed in OHMS. We shall discuss this point a little more fully in the sequel.

To summarise

(1) Electricity is the material out of which all the atoms (and therefore all forms of matter) are constructed. Positive electricity consists of protons having the same mass as a hydrogen atom, and negative electricity consists of electrons, having a mass 1/1800th part of that of the hydrogen atom.

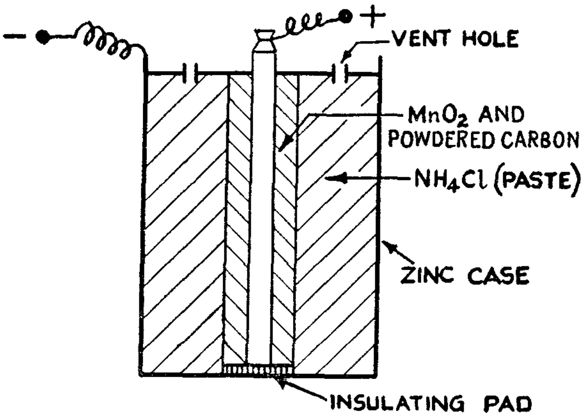


FIG. 3.2. Dry cell

Fig. 3.2 shows a dry cell in section. The container is a zinc cylinder, also serving as the negative plate. A carbon rod surrounded by manganese dioxide stands on an insulating pad in the middle of the cylinder, and the space between it and the zinc is filled with a paste (gelatine or plaster of Paris, or even sawdust) containing the excitant, sal-ammoniac.

The INERT cell is another special type which can be stored for long periods without deterioration, and to make it function it must be ACTIVATED. This is accomplished by introducing a little water into one of the vent holes, thereby rendering the excitant sufficiently moist to do its work. A dry cell has a voltage of about 1·5 volts.

Secondary cells or accumulators are used whenever possible in preference to primary cells such as we have just examined; the discussion of their action will be taken up in the next chapter, but it may be observed here that even accumulators are but another modification of the simple cell.

It should be noted carefully that the voltage of a cell depends only upon the materials of which it is constructed and not on its size.

Cell grouping

(1) *Series connection*

A, B and C (fig. 4.2) are three cells: the negative pole of A is connected to the positive of B, the negative of B to the positive of C. There are now two free terminals, the positive of A and the negative of C.

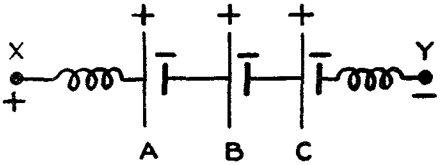


FIG. 4.2. Cells in series

These cells are said to be connected in series, and the combination is a " battery of three cells in series " whose positive terminal is X and whose negative terminal is Y.

The voltage of the battery is equal to the sum of the voltages of the component cells and for this reason series grouping is of great use when a high voltage is wanted and only low voltage cells are available. A wireless high tension battery is a case in point, in which a large number of small dry cells are connected in series to give a battery of anything up to 120 volts. Intermediate voltages are obtained by " tappings " which give access to the positive poles of the cells at suitable points as indicated in fig. 5.2. This battery consists of 100 cells of 1·4 volts each.

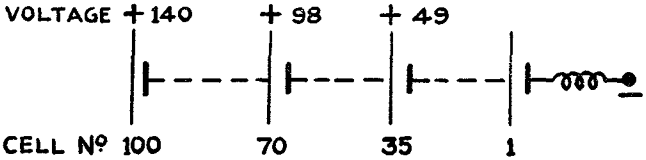


FIG. 5.2. H.T. Battery

(2) *Parallel connection*

The three cells A, B and C (fig. 6.2) have their positive terminals connected to a common lead terminating at X; the negative terminals are similarly joined by a lead terminating at Y.

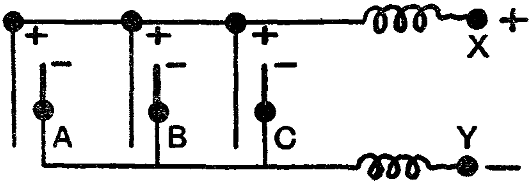


FIG. 6.2. Cells in parallel

Here, the cells are connected in parallel. The positive and negative terminals of the battery are X and Y respectively. The voltage between X and Y is the same as that between the terminals of any one component, so that parallel connection does not make for high voltage: all we have done in this form of connection is, in effect, to make a cell similar to any component cell but three times as big. Now while the size of a cell does not affect its voltage, it does have an effect on the working life of the cell: in the above example our battery of three cells in parallel will last three times as long as a simple cell, doing the same job. We express this by saying that the capacity of a bank of cells in parallel is equal to the sum of the capacities of the individual cells.

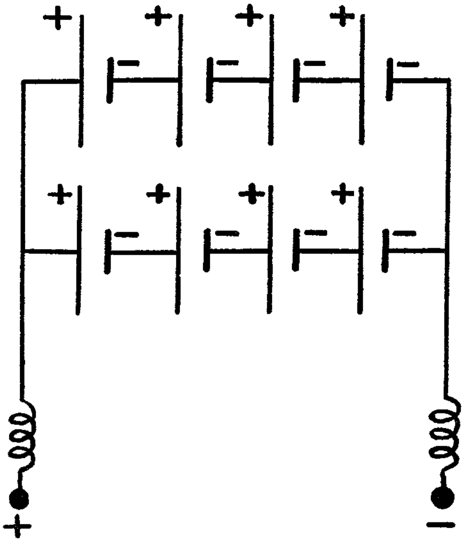


FIG. 7.2. Mixed grouping of cells

(3) *Mixed grouping*

A mixture of series and parallel grouping gives "mixed grouping" as illustrated in fig. 7.2.

By a careful choice of the number of cells in a row, and the number of rows in parallel we can build up a battery to have any desired voltage and capacity.

Resistance

It has already been noted that any path along which electricity is driven offers opposition to its passage; in other words every conductor possesses electrical resistance. It does not matter whether the conductor is a copper wire or the fluids of a cell, but we distinguish the latter from the former by speaking of the internal resistance of a battery, while by external resistance we mean the resistance of those parts of a circuit outside the source of P.D. as indicated in fig. 8.2.

At this stage the reader should examine for himself the common measuring instruments used in electrical work. A voltmeter should be examined and then used to find the voltage of cells and batteries of cells. An ammeter should be examined and then connected in a circuit to measure the current flowing. It might be advisable also to make a preliminary study of Chapter IV, in which some of the principal types of electrical instruments are described. Examine also a variable resistance or rheostat.

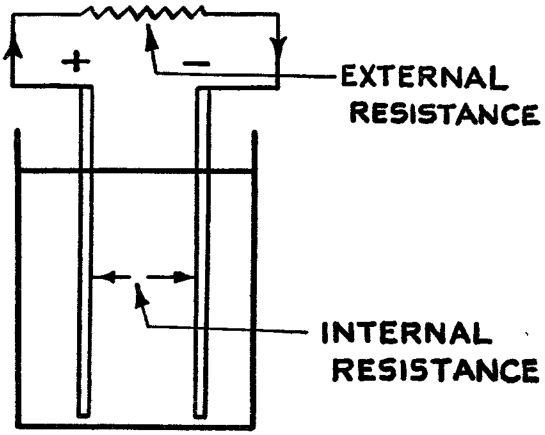


FIG. 8.2. Internal and external resistance

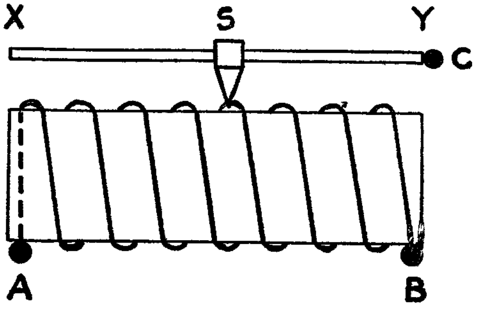


FIG. 9.2. Rheostat

Fig. 9.2 is a diagram of a rheostat. A coil of bare wire AB is wound on an insulating former, and provided with terminals at A and B; the bar XY is thick brass and practically resistanceless. A slider S makes good contact with the bar XY and also with coil AB. Current entering at A flows along the coil till it reaches the slider when it passes up to the bar (which offers no further resistance) and leaves the rheostat by the terminal C. The only resistance encountered is between A and the slider, so that as the latter is moved nearer to X, the resistance offered decreases, and as the slider is moved to Y the resistance increases.

We could ignore the bar and the terminal C, and use only terminals A and B: the system then behaves as a fixed resistance and the slider has no effect on it.

The next step in the study of resistance consists in performing an experiment. Assemble 10 or 12 accumulators and join them in series. From one terminal (say the positive) of the battery take a stout copper wire to the positive terminal of an ammeter, and from the negative of the instrument continue through a fixed resistance of 3 or 4 ohms (see fig. 10.2).

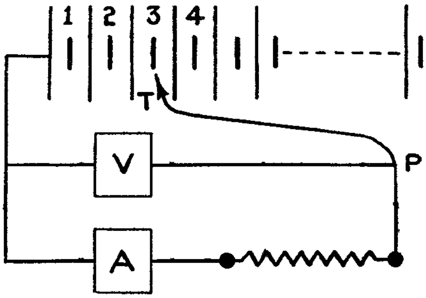


Fig. 10.2. Experiment on Ohm's law

A voltmeter is also connected to the positive pole of the battery on one side, and the other side is joined to the end of the fixed resistance. From P (or junction of voltmeter lead and fixed resistance) a stout flexible cable is taken, ending in a bare end T, which can be connected to the negative terminal of any cell in the battery. With T connected to point 1, we shall have the voltage of a single cell driving current (measured by the ammeter) through our fixed resistance: with T on point 2 we shall have the voltage of 2 cells in series (indicated by the voltmeter) and so on.

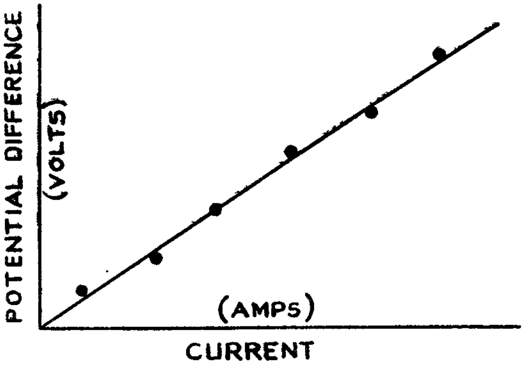


Fig. 11.2. Volts-amps graph

With this circuit we can take 12 pairs of readings of the voltage acting on our fixed resistance (voltmeter reading), and the corresponding current. These readings should be plotted on graph paper, voltage axis vertical, current axis horizontal, and they will give a result something like that shown in fig. 11.2.

The more accurately the readings are taken, the more nearly will the corresponding points lie on a straight line, and we are able to make a deduction of great importance:—

"The current in the conductor is directly proportional to the applied voltage"

which simply means that if we double the voltage, we shall double the current—other things being constant.

This experiment has been done so many times, for such a variety of conductors, always giving the same result that the conclusion has been raised to the status of a law, called *Ohm's Law* which in its simplest form is:—

"The current in a conductor is directly proportional to the applied voltage".

If we take, say, half a dozen pairs of readings (volts-amps) from our graph, and then divide the voltage by the current in every case we shall get the same answer every time.

This result is, of course, only another aspect of the proportionality mentioned above but it enables us to write :—

$$\frac{\text{Voltage (V)}}{\text{Current (I)}} = \text{constant, for a given conductor.}$$

and it is this constant, the ratio of V to I which we take as measuring the resistance of the conductor. Thus, if in any given case the ratio V/I worked out to 4, the resistance (R) of our conductor is 4 ohms, so that we can now write :—

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

or by transposition :—

$$I = \frac{V}{R} \text{ or } V = I \times R$$

This, of course, is the algebraical counterpart of Ohm's Law, and it enables us to work out one of the quantities, given the other two. Like most physical laws, Ohm's Law is not universal : some conductors do not obey it at all and these very conductors (valves, crystals, arcs, etc.) are of immense use in the technique of wireless. The law is true for all metallic conductors.

Non-compliance with the law would mean that the volts-amps graph was not a straight line, but a curve of some kind. Whereas the straight line graph leads at once to Ohm's Law and a simple algebraical expression, the curved volts-amps " characteristics " of the non-ohmic conductors defeat our attempts at either verbal or mathematical expression and we have to be content always to refer to their characteristics if we want to know their volts-amps relationships.

Again, most metallic conductors increase in resistance when they are heated, that is to say that the ratio V/I would not be constant unless the temperature were kept constant, and to meet this possibility we use an extended form of Ohm's Law :—

> " The current in a conductor is directly proportional to the voltage and inversely proportional to the resistance ".

The resistance of copper rises quite markedly with temperature. Certain alloys such as manganin and constantan preserve their resistance constant through a considerable range of temperature, and are used for the wire of which standard resistances are made. Unless otherwise stated, the resistance of any coil or instrument means the resistance at 15° Centigrade.

Specific Resistance

The resistance of a uniform conductor (at 15° C) will depend on three things :—

(i) Its length (l).
(ii) Its cross sectional area (A).
(iii) The material of which it is made.

When we wish to refer to the resistive properties of a material, we speak of the resistance from face to face of a unit cube of the material, and call it the " specific resistance ". Notice carefully the difference, in this connection, between a cubic centimetre (which may have any shape) and a centimetre cube (whose shape is entirely defined). The specific resistance (s) of a material is the resistance in ohms between the opposite faces of a unit cube (usually a centimetre cube) of the material.

All three factors are brought together in a formula :—

$$R = \frac{s \times l}{A}$$

R = resistance of uniform conductor in ohms.
s = specific resistance in ohms per cm. cube.
l = length in cm.
A = area of cross section sq. cm. The same formula holds true if all dimensions are expressed in inches.

Circuit theory

(1) *The series circuit*

FIG. 12.2. Conductors in series

The three conductors A, B and C in fig. 12.2 are said to be connected in series, and the whole circuit consisting of only one possible path for the current is a series circuit. Let us suppose that V, the voltage source is without internal resistance, for simplicity, when we have :—

$$R = r_1 + r_2 + r_3$$

R standing for the total resistance of the circuit.

This is an important rule " The overall resistance (R) of a number of conductors in series is equal to the sum of the resistances of the component conductors ".

A direct application of Ohm's Law gives us the current (I)

$$I = \frac{V}{R} = \frac{V}{r_1 + r_2 + r_3}$$

and it is important to observe that this is the value of the current at all points in the circuit. " In a series circuit the current has the same value in all the component conductors ".

Fix attention on conductor B for a moment : current is flowing in it, and it has resistance. Therefore there must be a potential difference across it. We can easily find what this P.D. is by applying Ohm's Law which, in one of its forms is :—

$$V = I \times R$$

In this particular case, the voltage across B will be

$$\left[\frac{V}{r_1 + r_2 + r_3} \times r_2\right] \text{ volts.}$$

or Ir_2 volts.

Similarly the P.D. across A will be Ir_1 volts, and across C it will be Ir_3 volts.

These voltages which are developed across resistive coils, due to current being driven through them are called VOLTS DROPS. Notice that the sum of the volts drops in the series circuit is :—

$$Ir_1 + Ir_2 + Ir_3 = I(r_1 + r_2 + r_3) = IR = V$$

that is to say " In a series circuit the sum of the volts drops is equal to the applied voltage " and also that this equation involves the statement $R = r_1 + r_2 + r_3$.

This may conveniently be depicted by means of a volts drop diagram.

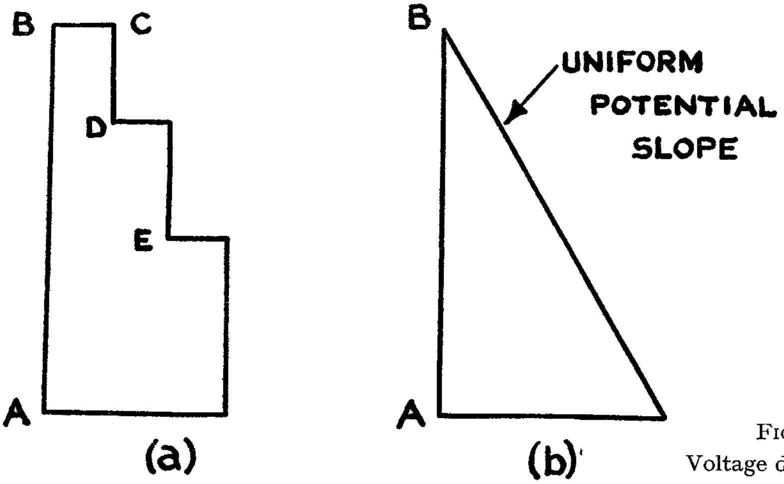


FIG. 13.2.
Voltage drop diagrams

The vertical line AB [fig. 13.2 (a)] represents the applied voltage: the level at A represents the potential of the negative pole of the battery, and the level of B that of the positive pole.

Passing along r_1 the potential falls to the level D, the volts drop being CD. Finally, at the end of r_3 the potential has reached the A level again.

If, instead of having resistance concentrated in three small regions in the circuit, we had a single wire with uniformly distributed resistance, the volts drop diagram would be as shown in fig. 13.2 (b), and instead of having potential steps we should have a potential slope. Any point of a circuit which is in conducting connection with the earth through a path of negligible resistance is regarded as being at a zero level of potential. If, for example, we had earthed the negative pole of the battery referred to above, the level A [fig. 13.2 (a)] would have been at zero potential or " earthy ". The potential at level B would have been numerically the same as the voltage of the battery, V. It is instructive to think of a battery as a lift, which raises the electricity to a certain level from which it falls, through the resistance in the external circuit to its original level at the low-potential terminal of the source of P.D.

The parallel circuit

Conductors are said to be in parallel when they provide alternative paths for the passage of electricity. Physically it means that the beginnings of all the conductors are joined together, and their ends are also joined. Fig. 14.2 shows two common ways of representing a set of conductors in parallel :—

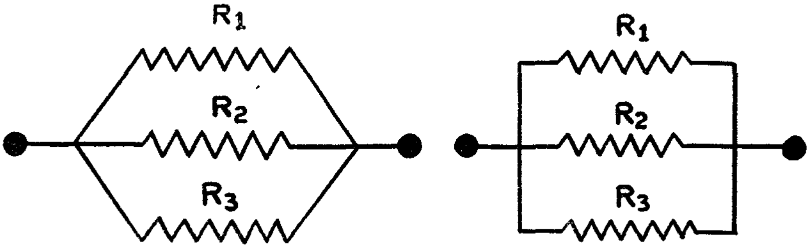


FIG. 14.2. Conductors in parallel

In this type of connection, it is a little more troublesome to find the overall resistance (R) : stated algebraically :—

$$\frac{1}{R} = \frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3} + \&c.$$

or " The reciprocal of the overall resistance is equal to the sum of the reciprocals of the component resistances ".

The reciprocal of a resistance is called the *conductance* of a conductor, measured in MHOS or SIEMENS, and usually denoted by the letter G : so that for a parallel system :—

$$G = g_1 + g_2 + g_3 + \&c.$$

Fig. 15.2 shows a group of three conductors in parallel (r_1, r_2 and r_3) connected to a source of voltage V volts. We can apply Ohm's Law directly and say that the main current I is given by :—

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

R being the overall resistance, but it is more instructive to start from first principles.

Notice that the same voltage is applied to each conductor—a fact which is always true of parallel systems. It follows therefore, that the currents in the separate conductors will be:—

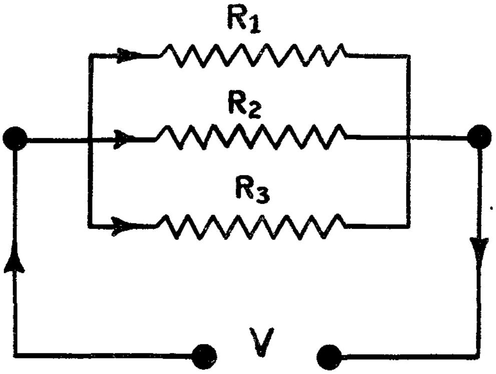


FIG. 15.2. Parallel circuit

$$\frac{V}{r_1}, \frac{V}{r_2} \text{ and } \frac{V}{r_3}$$

and the total current I will be equal to the sum of these sub-currents so that :—

$$I = \frac{V}{R} = \frac{V}{r_1} + \frac{V}{r_2} + \frac{V}{r_3}$$

The reader will observe that this last step brings us back to where we started, because cancelling the common factor V, we arrive at the formula stated at the beginning of this section for finding the overall resistance. The two important points to notice in this type of circuit are :—

(i) Every member of a parallel system has the same P.D. across it.
(ii) The sum of the sub-currents is equal to the main current entering the parallel system.

The mixed circuit

This is the most general type of circuit which contains both series and parallel elements. It does not involve any new principles.

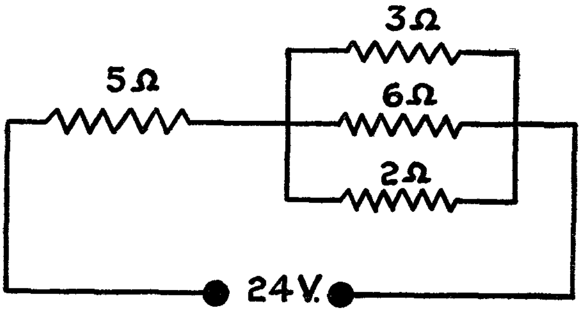


FIG. 16.2. Mixed circuit

In finding the overall resistance of such a circuit, the parallel elements should be tackled first and reduced to their single equivalent resistances. After that, the circuit becomes a simple series arrangement. A common problem is one which requires the determination of the sub-currents in a parallel element. Take the circuit of fig. 16.2 as an example. It is easy to see that the main current will be 4 amps ; how will this current sub-divide in the 3, 6 and 2 ohm system ?

Solution :—

The volts drop across the parallel element
= total resistance × total current
= 1 × 4
= 4 volts.

The currents are therefore

$$\frac{4}{3}, \frac{4}{6} \text{ and } \frac{4}{2} \text{ amperes, reading downwards.}$$

The sum of these sub-currents is of course equal to 4 amps, the main current.

The reader should undertake the solution of a number of circuit problems by way of fixing the ideas involved, and he should also test their properties practically in the laboratory.

Electromotive force and potential difference

We have already remarked that a cell or battery will have internal resistance ; in the foregoing circuit theory we have escaped this point by assuming that internal resistance was negligible, but we must now examine its effects carefully.

A cell is really a combination of two things

(i) a source of E.M.F.

(ii) a conductor, with resistance

and we can show these two components separated, as in fig. 17.2.

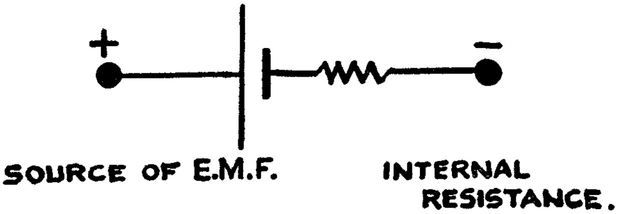


FIG. 17.2. Cell showing internal resistance

Let us now connect this cell in a simple circuit. Remembering that A and B (fig. 18.2) are the terminals of the cell, whose constant natural voltage is E volts, the voltmeter V will read the P.D. at the terminals of the cell, the voltage available for working into the external resistance R. Let the current be I amps. Consider the potential variations in passing from B (low potential terminal) to A; these will be

(i) a volts drop Ir

(ii) a volts rise E

and therefore an actual voltage rise of

$$E - Ir \text{ volts}$$

which is the voltage V which will be registered by the voltmeter. This expression tells us that the greater the current delivered, the less will be the P.D. across AB. On the other hand, if the current is zero, the internal volts drop is zero, and the terminal voltage is E.

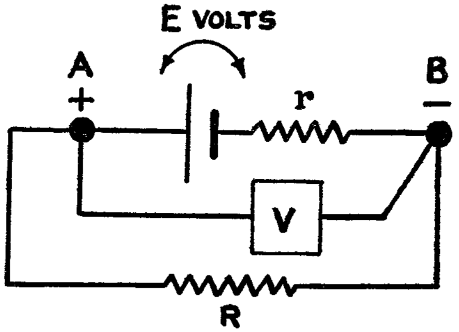


FIG. 18.2. Circuit to show effect of internal resistance on terminal voltage

It is this open circuit voltage which is numerically equal to the ELECTROMOTIVE FORCE of the cell. It is the constant, natural voltage produced by the appropriate combination of two conductors and an exciting fluid. If, however, the cell is driving current, the terminal voltage (V) is less than the E.M.F. by the internal volts drop. So far as this consideration affects circuit theory we may observe that the current in a circuit may be calculated as

$$\text{either} \quad \frac{\text{E.M.F.}}{\text{Internal} + \text{external resistances}}$$

$$\text{or} \quad \frac{\text{Terminal voltage}}{\text{External resistance}}$$

$$\text{that is,} \quad I = \frac{E}{R + r} \text{ or } \frac{V}{R}$$

It will now be realised that it is important to test the voltage of a cell while it is maintaining the load current. Measurements of the voltage of the cell on and off load will give an indication of its condition.

CHAPTER III

THE EFFECTS OF A CURRENT

From the engineer's point of view the important question is " What will an electric current do ? ". Reduced to its simplest terms the answer is that it will do three things ; the three effects of a current, we say, are :—

(i) *The heating effect :* heat is generated in a conductor in which current is flowing.

(ii) *The chemical effect :* electricity is driven through fluid conductors, the molecules being usually broken down into simpler systems ; this is a process we shall call ELECTROLYSIS.

(iii) *The magnetic effect :* the space around a conductor in which there is a current resembles the space around a magnet ; pieces of iron or steel experience forces which tend to move them about in this space ; we say that a magnetic field has been formed around the conductor.

The heating effect—energy considerations

Everybody is familiar with the effects of friction ; sliding down a rope may burn the hands ; sparks fly from the brake of a locomotive when it is pressed against the wheel ; friction is always accompanied by heat. Something analogous to friction takes place inside a conductor when electricity is being driven through the tiny inter-molecular spaces, and in the process heat is formed. The applications of this effect are manifold ; electric lamps, electrically heated clothing for pilots, soldering irons, furnaces, are commonplace enough, but where does this heat come from ?

One of the greatest contributions to scientific thought has been the concept of energy ; we believe that all physical processes are accompanied by, if not caused by, changes in the energy of the systems taking part in the process. Like human beings, some systems are energetic ; they are able to perform a lot of work, others have little energy. The energy of a body or system is its ability to do work.

Take a simple job like raising a weight from the ground ; work is done, or energy is expended by the person who lifts it. His muscles contained energy in the form of *chemical energy*, and as a result of volition on his part, some of this energy is released to do the lifting job. But we do not believe that the energy has been destroyed. Indeed, a very important principle in physical science called the conservation of energy states that " energy can neither be created nor destroyed " ; it can only be changed from one form to another.

In the case of the lifted weight, the energy is regarded as stored in the weight in the form of *potential energy*. If the weight were allowed to fall again it could be made to do a useful job (like a clock weight) out of this store of potential energy.

A bullet shot from a rifle is another case ; energy in the cordite of the cartridge (chemical energy) is released and converted into *kinetic* (*motion*) *energy* in the bullet. As the bullet is driven through the air, much of this energy is converted into heat owing to friction. If the bullet hits a target its remaining kinetic energy is turned partly into heat and otherwise used up in deforming the bullet or tearing through the target.

The bullet provides a good parallel to our electric particles driving their way between the fixed atoms of a conductor ; just as the energy of the bullet becomes converted into heat in passing through the air, so the energy of the electric particles is converted into heat in overcoming the resistance of the conductor ; just as the bullet is unchanged in its flight through the air, so are the electric particles unchanged in passing through the circuit. In both cases all that has happened is that energy has been given up in the form of heat. The bullet gets its original store of energy from the cartridge ; the electric particles derived their energy from the battery or generator feeding the circuit.

The analogy of a water pump is a valuable one.

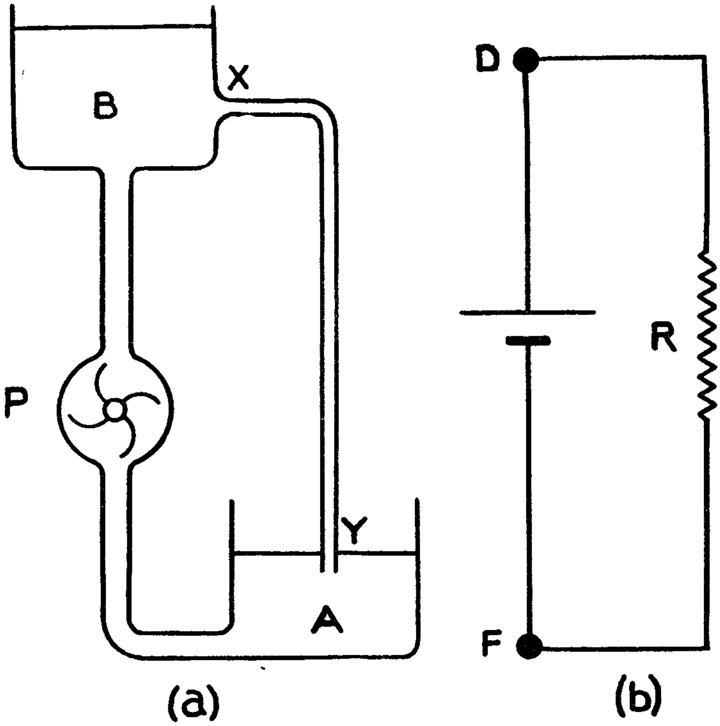


FIG. 1.3. Water analogy

P is a pump [fig. 1.3 (a)] which raises water from the lower tank A to the higher one B. The water in B has more energy per pound than that at A, the energy having been provided by the agent working the pump. If a conduit XY be opened so that B is joined externally to A the water will flow through it to A, giving up energy on the way in overcoming friction, but when it gets to A it will be caught up again by the pump, given more energy again and raised to B. The net result of this circulation is that energy is introduced into the water circuit by P, only to be expended as heat, while the water goes round and round unchanged.

Similarly in fig. 1.3 (b), a cell raises electricity (positive coulombs in our convention) from F the lower potential pole, to D the high potential pole. The electricity at D has more energy per coulomb than that at F, the energy having been provided by the chemical actions within the cell. If a conducting path R is joined externally from D to F, electricity will flow through it giving up energy on the way in overcoming resistance and being converted into heat, but when it reaches F it is caught up by the cell, through which it is driven (by an agency called its electro-motive force), given a fresh store of energy, and passed again to D.

The net result of this action is that energy is drawn from the cell and converted into heat in the resistance, while the electricity, which is merely the vehicle of the energy, remains unchanged.

Still considering the electric circuit, it is clear that we might have put some element in the external circuit other than mere resistance, which would have abstracted the energy of the electricity in some form other than heat. For example, an electric motor would have converted the electrical energy to kinetic energy ; a lamp would have converted some of it to light ; an electric bell would have converted some of it to sound. The form in which the energy is abstracted from the electricity is purely a matter of what kind of device it is made to traverse in its passage from the high potential (or high energy) condition to the low potential condition. One thing always happens, however, no matter how careful we are in the design of our " load ". Some energy will be wasted as heat, just as in all mechanical devices some energy is always wasted in overcoming friction.

We can give these ideas quantitative expression. The amount of energy required to lift one pound of matter vertically up through one foot is called a *foot pound,* thus if B is 12 feet above A [fig. 1.3 (a)], every pound of water at B has 12 foot pounds of energy more than that at A, and each pound passing down XY will give up 12 foot pounds of energy. We can say that between A and B there exists a difference of energy, per pound, of 12 units.

Now in the case of the electric circuit, the cell raises coulombs of electricity from places of low to places of high potential and in so doing gives them each a store of energy. The amount of energy required to raise one coulomb up through a potential difference of 1 volt is called a JOULE ; in the more familiar terms of lifting weights, a joule is the amount of energy required to lift 100 grams vertically through 102 centimetres.

In the electric circuit, therefore, if the cell has a voltage of 2, every coulomb at D [fig. 1.3 (b)] has 2 joules of energy more than a coulomb at F, and in passing down through R every coulomb will give up its 2 joules of energy. We can say that a " difference of energy per coulomb " of 2 units exists between F and D.

This is very important as throwing light on the meaning of potential difference ; if there is a P.D. of V volts between two points in a circuit, it means that coulombs situated at the point of higher potential possess V joules of energy more than they possess at the point of lower potential.

When we speak of the potential at any point we are really making a statement about the energy of the electricity at that point. We agree to take as our zero of potential, points in conducting connection with the earth. That does not mean to say that electricity so situated has no energy ; it means that we agree to take that condition as our datum level and to regard any point at which there is more energy per coulomb as having positive potential, and any point where there is less energy per coulomb as having negative potential.

Another important conception is that of POWER. The rate at which a machine (or an electric circuit) converts energy from one form to another is called its power. In mechanical engineering when we speak of a machine of one horse-power, we mean it can convert 33,000 foot pounds from one form of energy to another in the course of one minute, the idea being that this is how fast a horse can work.

In electrical engineering we estimate power in WATTS. A circuit which converts one joule every second from one form to another is said to have a power of one watt. A 60 watt lamp for example takes 60 joules of energy every second from the electricity flowing through it and gives it out as heat and light.

It will be quite simple to calculate the power (or wattage as it is often called) of a circuit on the basis of the ideas already discussed. Suppose a P.D. of V volts exists between two points and the current between the two points is I amps. Then every coulomb which passes between the points gives up V joules ; but I coulombs pass every second, therefore, every second the energy output will be :—

$$I \times V \text{ joules per second}$$

$$\text{or the power will be } I \times V \text{ watts.}$$

This is perfectly general and applies equally to a whole circuit as to a part of a circuit.

Since by Ohm's Law :—

$$I = \frac{V}{R} \text{ or } V = IR$$

we can restate the power formula in two other ways :—

$$\text{(i)}\quad \frac{V}{R} \times V = \frac{V^2}{R} \text{ watts.}$$

$$\text{or (ii)}\quad I \times IR = I^2R \text{ watts.}$$

Heat being a form of energy there is a strict equivalence between the joule and the heat unit (called a CALORIE).

Very accurate experiments have shown that 4·2 joules of energy go to make up 1 calorie. A calorie is the amount of heat which has to be given to 1 gram of water to raise its temperature by 1° Centigrade. If, then, a conductor of resistance R ohms carries a current I amps., then I^2R joules are converted to heat within it per second, or the rate of heat production is :—

$$\frac{I^2R}{4 \cdot 2} \text{ calories per second.}$$

It is interesting to observe that the rate of heat production is proportional to the square of the current ; if the current is doubled, the heating effect is quadrupled.

All this means, of course, that an Electric Supply Company really supplies energy to its customers ; the energy is conveyed from the Company's generators to the customers by means of electricity, which itself remains unchanged in the process.

Such companies charge for energy in terms of a commercial unit called a Board of Trade Unit or a Kilowatt-hour. It is equal to 3,600,000 joules and is defined thus :—

> " A Kilowatt-hour of energy is the amount of energy given out by a machine of output power one Kilowatt working steadily for one hour ".

To find the number of B.O.T. units absorbed by any electrical system, we have to evaluate the relation :—

Power of system in kilowatts × number of hours used.

or,

$$\frac{\text{watts} \times \text{hours}}{1{,}000}$$

A very common power problem concerns the output of small generators, e.g. :—

A 500 watt generator is designed to give a terminal voltage of 14. What is the maximum current it can give ?

$$\text{Watts} = \text{volts} \times \text{amps.}$$

$$500 = 14 \times I$$

$$I = \frac{500}{14} = 35 \cdot 7 \text{ amps.}$$

Notice that the generator can give any current up to 35 amps ; if it were put in a circuit so that Ohm's Law required more than 35 amps to flow, its voltage would fall and in fact the generator would cease to behave with any regularity.

The above relations are quite true, whatever form the energy has before and after transformation. They are as true for a chemical cell producing heat as for a steam driven turbo-generator supplying energy to work trains, lifts and other machinery.

The chemical effect—chemistry of the accumulator

It was pointed out before (page 4) that when certain substances are dissolved in water their molecules are ionised, that is to say a molecule is broken into two more simple structures, one having a positive electric charge and the other a negative ; in consequence they can be sorted out by the application of potential difference, the positive ions moving towards the region of low potential and the negative ions moving in the opposite direction.

Referring again to fig. 3.1 the common salt molecule (NaCl) is ionised into Na + and Cl —, merely by the act of solution. If two plates are dipped into the solution and joined externally to a source of E.M.F. the Na + ions move towards the low potential plate and the Cl — ions move towards the high potential plate. On arrival at the plates the ions give up their charges and revert to normal chemical atoms and as such they react with either the solution or the material of the plate, or both, until their chemical affinity is satisfied. All this complicated process is called the chemical effect of a current and its importance to us lies in its application to the accumulator.

In ordinary commerce, perhaps its most important bearing is in the process of electroplating.

The lead-acid accumulator

If two lead plates are placed in dilute Sulphuric Acid, they will not constitute a cell, because to do so there must be two different metals, but let us try this experiment :—

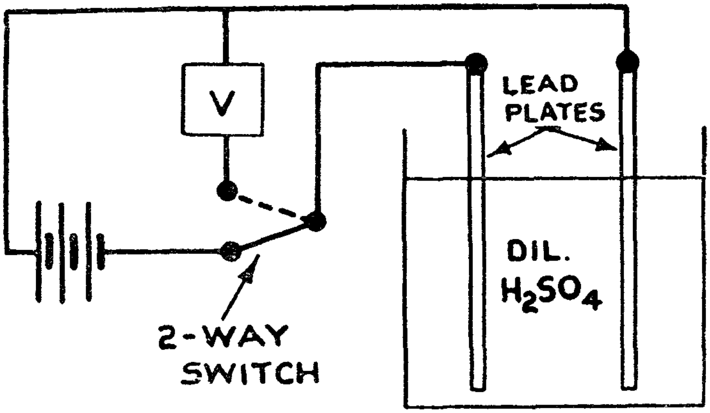


FIG. 2.3. Illustrating the principle of the accumulator

Connect the lead plates immersed in dilute H_2SO_4 (fig. 2.3) to an external circuit and pass a current through the acid using them as electrodes. After a moment, break the external (charging) circuit and connect a voltmeter across the plates ; the instrument will register a voltage. This voltage will soon disappear for a variety of reasons, but the important fact is that the passage of the current has altered the plates in some way so that they are no longer the same substance, but different substances, and are able to show simple cell action for a short time. Obviously the difference, whatever it is, is not very permanent. Here is the principle of the secondary cell however ; the passage of a current between suitable plates immersed in a suitable electrolyte so alters them (by the chemical effect) that the arrangement becomes a kind of cell. It is called a secondary cell because it requires to be charged before its action commences. When the charged secondary cell is made to drive a current on its own account, the difference in the nature of the two plates (again due to the chemical effect of the current through the cell) disappears and the cell is discharged.

There are two types of accumulator in Service use, Lead-Acid cells and Nickel Alkali cells. The first of these types has the disadvantage of being very heavy and requiring expert attention ; the second type is lighter and will stand hard usage, but its voltage is low and its first cost rather high.

The plates of a lead-acid cell consist of a kind of lead grid, into the openings of which are pressed certain chemical substances in paste form. The positive plates are pasted with red lead (Pb_3O_4) ,while the negative plates are filled with a paste of litharge (PbO). The pasted plates are not yet ready for use, but they have to go through a process which converts the litharge to spongy lead and the red lead to lead peroxide. This is done by immersing the pasted plates in dilute sulphuric acid, and passing a current through the arrangement as indicated in fig. 3.3. Current from an external battery or generator is passed through the bath so that it enters by the red-lead plates and leaves by the litharge plates. It may be useful for future reference to note that the plate by which current enters an electrolytic bath is called the ANODE, and the plate by which it leaves the bath is called the CATHODE ; whenever hydrogen or metallic elements are produced in electrolytic actions, they always appear at the cathode—a rule which is useful if it is necessary to determine the polarity of a supply at any time.

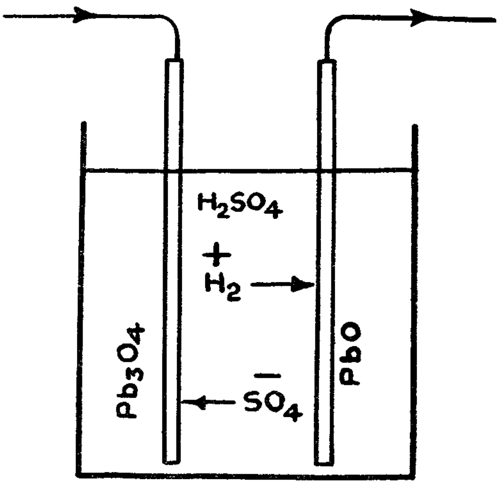


FIG. 3.3. Formation of pasted plates

In the case we are discussing the ions of sulphuric acid will move to the plates as shown in fig. 3.3 ; the hydrogen ion arriving at the cathode gives up its charge to the plate, reverts to normal hydrogen, reacts with the litharge and forms spongy lead and water :—

$$H_2 + PbO = Pb + H_2O$$

The sulphion arriving at the anode gives up its charge and reacts with the water, re-forming sulphuric acid, and evolving oxygen, which combines with the red lead to give lead peroxide :—

$$2SO_4 + 2H_2O = 2H_2SO_4 + O_2$$
$$O_2 + Pb_3O_4 = 3\,PbO_2$$

In time the whole of the oxides are changed in this way, and when the action is complete we have a modified form of simple cell with plates of spongy lead and lead peroxide in place of the original copper and zinc. The first charge is complete and the cell is now ready for use.

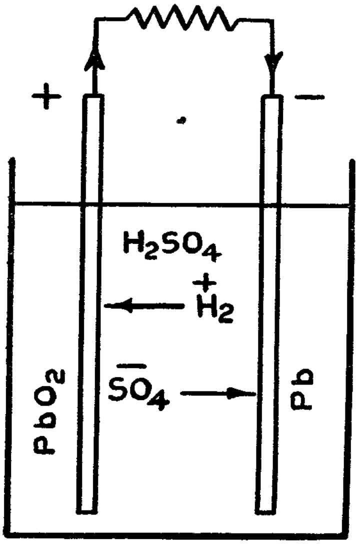


FIG. 4.3. Discharge action

Suppose we now let the cell drive current in an external circuit ; the current within it will now be in the opposite direction. Hydrogen appears at the peroxide plate and sulphion at the lead plate, as indicated in fig. 4.3.

The discharge reactions are :—

$$PbO_2 + H_2 \longrightarrow PbO + H_2O$$
$$PbO + H_2SO_4 \longrightarrow PbSO_4 + H_2O$$

The net result at the positive plate then, is the formation of lead sulphate (which is insoluble and remains in the grids), and the replacement of heavy sulphuric acid by water. This latter causes the density of the electrolyte to fall during discharge and it affords a handy method of testing the extent of the discharge.

At the other plate we have :—

$$SO_4 + Pb = PbSO_4$$

i.e. lead sulphate is also formed at this plate, and when discharge is complete, both plates will be pasted with insoluble lead sulphate. This action again removes a heavy sulphion and reduces the density of the electrolyte.

The discharge cell may now be connected to an external source and a charging current passed in the opposite sense to the discharge current, as shown in fig. 5.3.

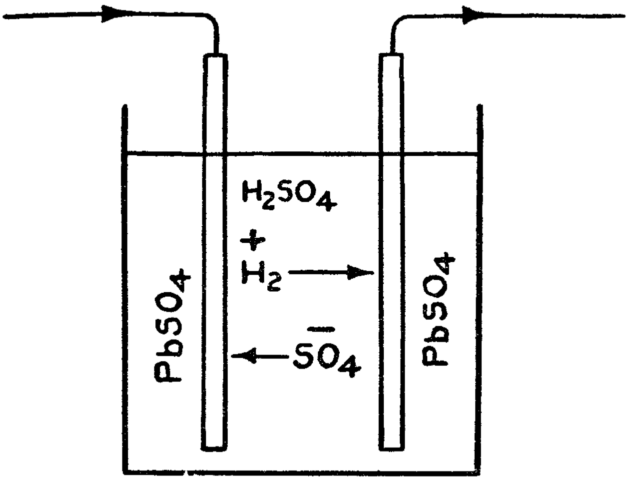


FIG. 5.3. Charging action

The actions now are :—

Positive plate $SO_4 + PbSO_4 + 2H_2O=2H_2SO_4 + PbO_2$.

i.e. the lead peroxide reappears and water is replaced by sulphuric acid, with a consequent rise in the density of the electrolyte.

At the negative plate $PbSO_4 + H_2 = H_2SO_4 + Pb$.

Spongy lead is re-formed and the cell returns to its charged condition again. During discharge, the peroxide and the lead are converted to lead sulphate which occupies a larger volume than either, and consequently a heavy discharge will set up stresses which may buckle the plates and damage the cell. The makers state the maximum discharge for which the cells are designed ; this depends on the strength of the retaining grid to accommodate the swelling and shrinkage of the active material.

The voltage of a charged lead-acid cell in good condition is 2·2 ; it gradually falls during discharge but should never be allowed to fall below 1·8.

The nickel-alkali cell

In this type of accumulator the positive plate consists of nickelic hydroxide held in a specially constructed grid of nickel. The electrolyte is concentrated potassium hydroxide solution, while the negative plate is of specially prepared iron.

The actual chemical changes which go on during charge and discharge are very complex, but may be generally expressed by saying that during discharge nickelic hydroxide is changed to nickelous hydroxide, while the iron is changed to ferrous hydroxide. On charging the chemical actions are reversed.

The lid of the cell has a valve which allows the escape of gas, but which prevents the entry of atmospheric carbon dioxide which would spoil the electrolyte by changing it to potassium carbonate. The E.M.F. of such a cell is about 1·3 volts and its efficiency is lower than that of the lead-acid type. Its great advantages lie in its lightness and its ability to stand not only hard mechanical treatment, but overcharging and over-discharging. Unlike a lead-acid cell it does not deteriorate when left standing uncharged for long periods.

Capacity of a battery

The amount of electricity a fully charged cell can drive round a circuit till its voltage falls to 1·8 volts, is called its *capacity*. Capacity would thus be measured in coulombs (or ampere-seconds) since a current of 1 ampere for 1 second means the passage of 1 coulomb, but for convenience we raise our unit to 3,600 coulombs or 1 ampere-hour.

Thus in theory a 10 amp.-hour cell ought to be expected to deliver 10 amps. for 1 hour or 600 amps. for 1 minute or 1 amp. for 10 hours ; in each case it would have passed 10 × 3,600 coulombs, its rated capacity. In practice this is not so and most statements of capacity are based on the 10 hour rate, meaning that the discharge should be at such a value that the battery will fall to 1·8 volts/cell after 10 hours' use.

Charging arrangements

Just as an accumulator must not be discharged at more than a stated rate, so as charging current must not exceed a given value which is usually printed on the site of the cell.

Suppose we had a cell of internal resistance ·01 ohm, discharge down to 1·8 volts, and that we propose to charge it from a 35 volt D.C. generator. Further, let the charging rate of the cell be 10 amps. If we join the cell directly to the generator terminals (positive to positive) the current flowing would be that required by Ohm's law :—

$$I = \frac{\text{Generator voltage} - \text{accumulator voltage}}{\text{accumulator resistance}}$$

$$= \frac{35 - 1{\cdot}8}{{\cdot}01} = 3{,}320 \text{ amps.}$$

A current of this value is quite out of the question, so we must limit it by means of a resistance ; the value of this resistance (R) to bring down the charging current to the correct value (10 amps) can be found thus :—

$$10 = \frac{35-1{\cdot}8}{R+{\cdot}01}$$

In nearly all cases it will be possible to neglect the cell resistance when :—

$$10 = \frac{33{\cdot}2}{R} \text{ or } R = 3{\cdot}32 \text{ ohms.}$$

In practice we should take an adjustable resistance of about 5 ohms, big enough to carry 10 amps, and having connected it in series with the generator, accumulator and ammeter (fig. 6.3), adjust its value until the correct charging current flows.

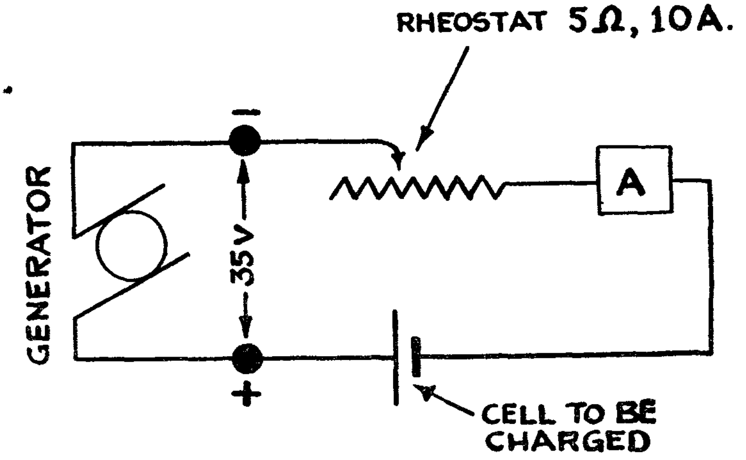


FIG. 6.3. Charging circuit

As the charging process goes on, the voltage of the accumulator rises and therefore the total voltage acting in the circuit (generator voltage—accumulator voltage) falls and the current gets smaller in consequence. It is then desirable to reduce the value of the series resistance to restore the current to its correct value.

Magnetism—The magnetic effect

Our normal experience tells us that if we wish to move an object we must exert force upon it, either directly, or through some intermediate mechanism which has direct physical contact with it.

A permanent magnet seems to contradict this experience, and is able to move pieces of iron or steel situated some distance away and apparently without any connecting mechanism. The space around a magnet through which its influence can be detected is called a magnetic field.

Place a bar magnet on the table, cover it with a sheet of cardboard and sprinkle it with iron filings. Tap the cardboard gently and observe that the filings fall into a definite line pattern (see fig. 7.3). The magnetic properties appear to be concentrated in two regions of the magnet termed the NORTH and SOUTH POLES. A freely suspended bar magnet comes to rest with the North pole pointing North.

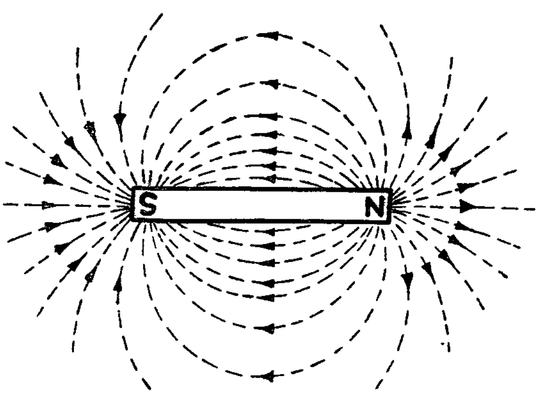


FIG. 7.3. Field of a bar magnet

Repeat the process with two magnets, first with like poles adjacent, then with unlike poles adjacent (fig. 8.3).

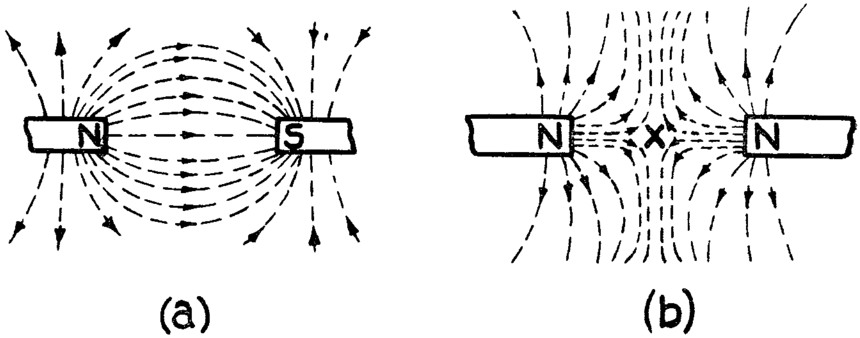


FIG. 8.3. Fields between magnetic poles

The significance of these iron filing maps is that they indicate that a magnetic field has got a structure of some kind, and as a first guess we might suppose it to consist of a mass of invisible filaments (lines of force) whose spatial arrangements are revealed when iron filings are scattered among them in this way. More than this, if we imagine the filaments to be like stretched elastic threads tending to shorten longways and widen sideways, the repulsion of two similar poles [fig. 8.3 (*b*)] and the attraction of two dissimilar poles [fig. 8.3 (*a*)] would appear to be a necessary consequence of the line disposition in the two cases.

If we carry this investigation further and examine a great number of magnetic fields by the iron filing method, we find that this result seems to be quite general, and we therefore accept the elastic-line theory of the magnetic field as a working hypothesis. We agree to regard the lines as emerging from the north pole of a magnet and entering by the south pole after having traversed the surrounding space and converting it into a magnetic field.

This convention, of course, is equivalent to defining the direction of a line of force and in this connection it is important to notice that if a small light freely suspended magnet is placed in a magnetic field it will move until its magnetic axis (the line joining its poles) coincides with the direction of the lines of force of the field. This property of a suspended needle has its most important application in the magnetic compass. The reader will be able to deduce for himself that it is the S–N direction of the needle which coincides with the direction of the lines (or, as we sometimes say in the direction of the field) in the immediate vicinity.

The iron filing experiment then enables us to determine the general layout of the lines of force in a field, and the suspended magnetic needle tells us the direction of the lines at any point.

It may seem far fetched to talk of counting these lines, but by an extension of our fundamental convention, we can introduce quantitative ideas, and we shall often speak of *magnetic flux* ; by the magnetic flux through any region, or through any closed curve described in the field, we mean the total number of lines of force passing through that region or through that curve.

If it is said for example that a certain bar magnet has a magnetic flux of 500 lines, that is one way of expressing its strength as a magnet. In general, the effects of a magnetic field are most marked where the lines of force are most closely packed together ; the term *flux density* is used to denote the degree of packing, and it means the number of lines of force which are passing through an area of one square centimetre, drawn at right angles to the lines.

Iron and steel and some of their alloys react most readily to magnetic fields, and on that account are called magnetic substances. When a fragment of such a substance is placed in a magnetic field, it becomes a magnet, complete with flux and poles. When removed from the field, the extent to which it retains this " induced " magnetism depends on what is called its *retentivity*. Steel has a high retentivity and is therefore used in the manufacture of permanent magnets ; soft iron has a low retentivity.

In 1820 a Danish scientist, Oersted, made the important discovery that the space around an active conductor was similar to the space around a magnet ; pieces of iron and steel experienced forces, and suspended magnetic needles were deflected while the current was flowing and the effects disappeared when the current was switched off. It is this production of a magnetic field by an electric current which we call its *magnetic effect.*

Magnetic field of a straight conductor

The conductor AB fig. 9.3 is a piece of thick copper wire passing at right angles through a sheet of cardboard CD. If a current of about 20 amperes is passed through the wire, and at the same time iron filings are sprinkled on the cardboard, it will be seen that the filings fall into a pattern of concentric circles whose common centre is on the axis of the conductor AB.

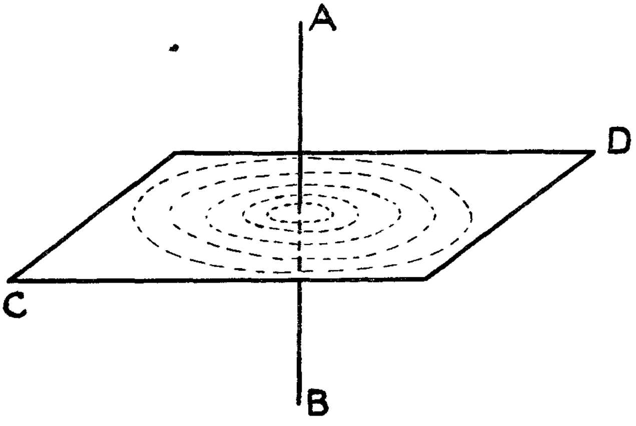


FIG. 9.3. Field of a straight wire

By this experiment we learn that the magnetic effect of a straight wire is relatively feeble, because we have to have a fairly heavy current to produce any effect, and further, that the lines of force take the form of concentric circles about the conductor.

This type of magnetic field differs from that of a bar magnet in one very important respect ; the lines of force neither converge towards nor diverge from any particular point and for this reason the field has no regions which behave like poles. In spite of this however, the lines will have direction, and this can be determined by placing a small charm compass on the cardboard when the current is flowing. The compass needle will take up a position as a tangent to one of the circular lines of force and its S–N direction will give the direction of the lines at that point.

Such an experiment will show that the direction of the lines is related to that of the current in the same way as the direction of turn of a right-handed screw is related to the direction in which its point moves. These results are indicated in fig. 10.3.

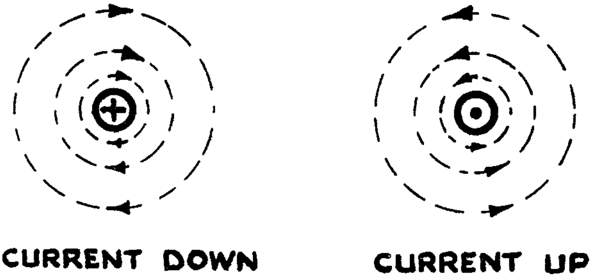


FIG. 10.3. Direction of straight wire field

Near to the surface of the wire the flux density is greatest : as we move away from the wire at right angles to its length, the flux density falls off fairly rapidly. It is interesting to note that not only do circular magnetic lines exist outside the wire but they are also present within the substance of the conductor itself. This internal flux density falls from a maximum value at the surface to zero at the axis. This flux distribution is illustrated by the graph of fig. 11.3.

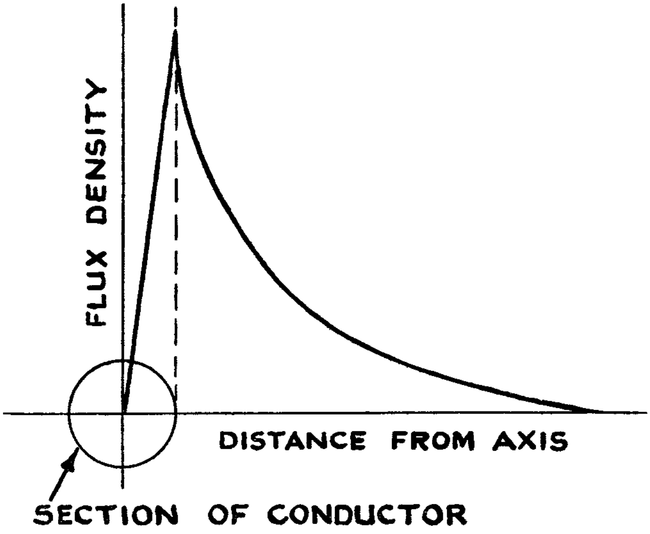


FIG. 11.3. Flux distribution about a straight wire

Magnetic field of a solenoid

A solenoid is a uniform helix with its neighbouring turns spaced just enough to avoid short circuiting. The magnetic field due to current in a solenoid has the same general characteristics as the field arising from an open helix, but the reader is warned that this similarity does not persist when the quantitative properties of the two circuits come up for consideration. In this section, however, we shall deal only with general characteristics, and we shall not have to consider whether the coil is close wound or not.

Investigation by the iron filings method reveals a field similar to that of a permanent bar magnet ; in this case, however, we are able to follow the flux lines through the interior of the coil, where we notice that they are parallel and evenly spaced. Such a field we call a uniform field ; fig. 12.3 shows the type of field arising from a solenoid, the latter being shown in section.

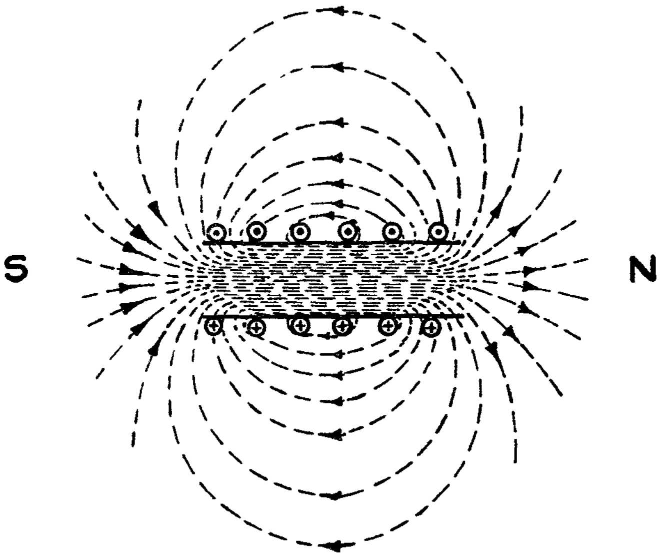


FIG. 12.3 Field of a solenoid

Tests with a compass needle will show that with the current direction as indicated, the direction of the lines of force will be as marked with the arrows. Since the lines appear to diverge from the right-hand end, that end will behave like a north magnetic pole, and for a similar reason, the left-hand end will behave like a south pole. If the direction of the current is reversed, the polarity of the solenoid reverses also.

There are several rules relating the direction of the current with the polarity; we shall give two of them :—

(1) If, looking endways at a solenoid the direction of the current is clockwise, that end is the south pole.

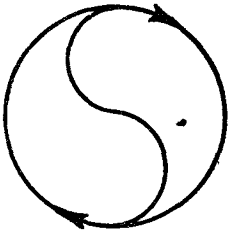


FIG. 13.3. The S rule

This rule is most conveniently remembered by the sketch of fig. 13.3.

It is drawn by describing a letter S with arrows at each end, and then completing the circle. The arrows indicate the direction of the current around the solenoid and the S denotes the polarity.

(2) Place the right hand over the top of the coil with the fingers pointing in the direction of the current; then the thumb points towards the north pole of the solenoid.

Experiment shows that the flux created by a solenoid is proportional to the current flowing in the windings. This is of the greatest practical importance because it affords us a magnetic " source " whose strength is completely under control.

In terms of a formula the flux through a close wound solenoid can be shown to be :—

$$\Phi = \frac{4\pi NiA}{10l}$$

where N = total number of turns.
i = current (amps).
A = area of cross section (sq. cm.).
l = length (cms).

This formula shows that, other things being equal, it is the product $N \times i$, or the AMPERE-TURNS which decides the strength of the flux; it means that the magnetic effect of a solenoid can be doubled either by doubling the current or by doubling the number of turns, provided A and l are unchanged.

Iron cored solenoid

In the above discussion, the core (the material occupying the inside of the coil) was assumed to be air. If the core is a bar of soft iron the magnetic effect of a solenoid becomes very much more powerful. We have already noted that when a magnetic substance is placed in a magnetic field, it becomes, at least temporarily, a magnet, and adds its own acquired flux to that of the neighbouring field. The iron core of a solenoid finds itself in a magnetic field as soon as the current is switched on and therefore becomes a magnet, adding its own flux to the flux due to the solenoid alone.

If Φ_s is the flux due to the solenoid alone and Φ_i is the flux due to the iron, the total flux is now $\Phi_s + \Phi_i$.

The ratio

$$\frac{\Phi_s + \Phi_i}{\Phi_s} = \mu$$

is regarded as a measure of the ability of the iron to increase the flux in its immediate neighbourhood, and is called the PERMEABILITY (μ) of the iron. This definition of μ is exact provided that the solenoid is several times as long as it is wide, or if it is what is called a " ring solenoid " or its equivalent.

As all practical cases with which the reader will have to deal will come within the terms of this proviso, we shall adopt the definition for μ as given, and, in so doing, avoid a theoretical discussion of considerable difficulty. So far as we are concerned then, permeability is the " flux multiplying property " of a substance.

Owing to its lack of retentivity, the iron core loses its magnetic flux as soon as the current in the solenoid ceases. An iron cored solenoid is often called an " electromagnet " : it provides a powerful source of magnetism which can be created or destroyed at will by merely turning a switch.

Hysteresis

The iron flux Φ_i set up in the iron core of a solenoid is caused by Φ_s the flux due to the current flowing in the windings : Φ_s is the cause, Φ_i is the effect. We can carry out experiments in which the value of Φ_s and the corresponding Φ_i can be measured, and their corresponding values can be plotted on graph paper to show how they are related. The graph of fig. 14.3 (a) is the result of an experiment of this sort.

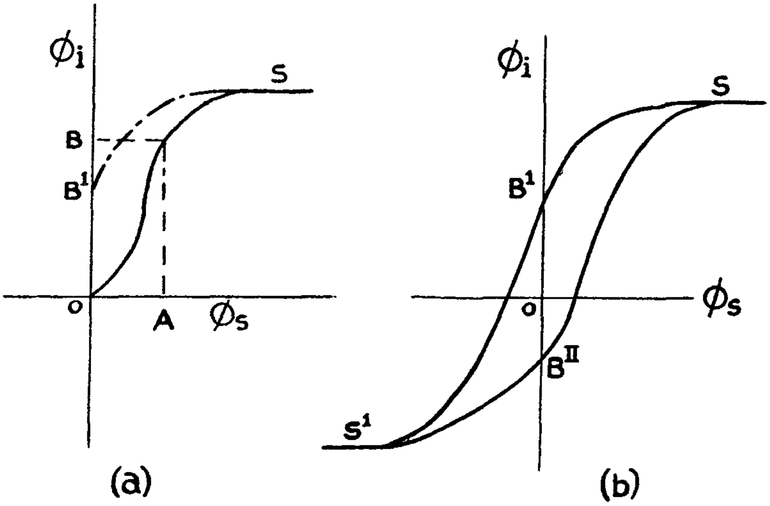


FIG. 14.3. Hysteresis curves

While Φ_s increases from zero to OA, Φ_i increases from zero to OB, and the points lie on a curve showing that the two quantities are not directly proportional to one another. Further Φ_i reaches a maximum value, called the *saturation value*, and this value now remains constant no matter how much Φ_i increases. If we now decrease Φ_i step by step and observe the corresponding Φ_s, we shall find that the latter decreases much slower than the former : the process of removing Φ_i is represented by the chain line SB^1 which shows that when Φ_s has fallen to zero, Φ_i still retains the value OB^1 which is therefore a measure of the RETENTIVITY of the iron of which the core is made. This lag of Φ_i behind Φ_s is called HYSTERESIS.

If we continue the experiment by now creating Φ_s in the opposite sense and increasing it step by step we shall eventually arrive at another saturation value of Φ_i, in the opposite direction. This is indicated by S^1 in fig. 14.3 (b) : next let us reduce Φ_s step by step to zero and we shall find Φ_i retains the value B^{II}. Lastly, let us create Φ_s in the original direction, and gradually increase it until we reach saturation again. We describe this process as having gone through a magnetic cycle, from positive saturation to negative saturation and back again, and the closed loop of fig. 14.3 (b) depicting the relation of Φ_s and Φ_i throughout the cycle, is called a hysteresis loop. The area of the loop can be shown to be a measure of the energy which has had to be expended in changing the magnetism of the specimen. Clearly, if there were no hysteresis, B^I and B^{II} would both coincide with O, the area of the loop would be zero and no energy would have been used up.

The iron parts of electromagnetic machinery when in action are continually undergoing cycles of magnetisation such as we have here described : it is obviously desirable that the iron used to make such parts should show as little hysteresis as possible, otherwise the energy wasted would be considerable and since this energy appears as heat, the parts would get very hot and inefficiency would result.

CHAPTER IV

THE ELECTRIC MOTOR—MEASURING INSTRUMENTS

If the pole of a magnet is brought near a conductor in which there is a current, the conductor will tend to move : it will not be attracted to or repelled from the magnetic pole in the familiar way, but it will tend to move in a direction at right angles to the magnetic axis. Further investigation would show that the force producing this movement is greatest when the conductor is at right angles to the lines of force proceeding from the magnet, and that it is zero when the conductor is parallel to the lines. In what follows we shall confine our considerations to the first of these cases, because in the practical application of this effect, we shall always arrange our conductors to be disposed at right angles to the lines of force of the surrounding field.

Let us show first that this effect is to be expected from what we already know about magnetic fields and lines of force.

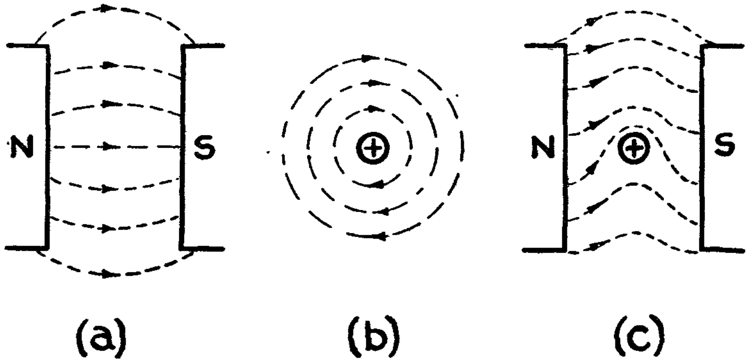


FIG. 1.4. Distorted field

Fig. 1.4 (a) represents a magnetic field developed between two large pole faces N and S: fig. 1.4 (b) represents the field due to a straight wire carrying current downwards at right angles to the plane of the paper. Fig. 1.4 (c) represents the field which would result if the two previous fields were superimposed. Here the conductor is placed in a magnetic field at right angles to the lines of force, and since on one side of it the directions of the fields coincide, while on the other they are in opposition, we get a concentration of lines on one side and a thinning out on the other.

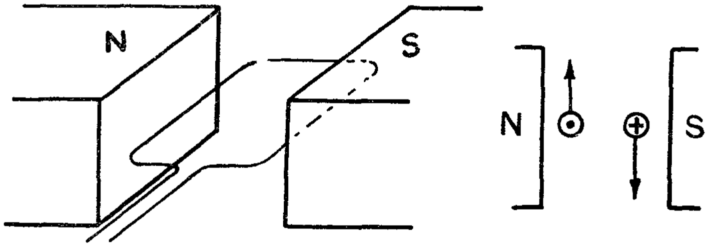
Fig. 1.4 (c) is known as a distorted field, and it can easily be reproduced by the iron filing method. Remembering that lines of force are in a state of longitudinal tension, the diagram of the distorted field at once indicates that the conductor will be pulled downwards, just as a telegraph wire will be pulled downwards by a piece of string thrown over it and then drawn tight.

The development of a force on an active conductor in a magnetic field is termed the *Motor Principle* because it is upon force developed in this way that the action of an electric motor depends. The relative directions of force, current and magnetic field are related by the left hand rule :—

Extend the thumb, fore and second fingers of the LEFT hand all at right angles, then the

thuMb represents the direction of Motion
Fore-finger points along the lines of Force
seCond finger points in the direction of the Current.

We now proceed to develop this idea a little further: instead of having a single active conductor in an external magnetic field, let us take a conductor in the form of a rectangular loop, as shown in fig. 2.4: (a) shows a general view of the arrangement and (b) shows a sectional view. The current direction is indicated by the usual dot and cross convention, and an application of the left hand rule will show that one side of the rectangle will be forced upwards and the other downwards as shown by the arrows in fig. 2.4 (b).



(a) Fig. 2.4. Rotating loop (b)

These oppositely directed forces constitute a torque (twisting effort) which tends to turn the rectangular coil about its long axis. If the coil was suitably mounted it would rotate under this torque, but it will readily be seen that it will only go on so doing for a quarter of a revolution. Fig. 3.4 shows what would then be the state of affairs; the two forces would still be up and down but they are now acting in the same straight line and their joint turning effect is nil.

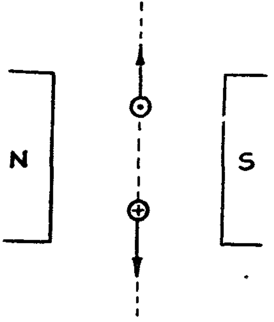


Fig. 3.4. Coil after one quarter revolution

Thus the coil comes to rest with its plane vertical, in which position it will be seen that the coil embraces as much magnetic flux as possible. It is a useful rule to remember that an active coil, free to move in a magnetic field, will always tend to take up a position in which as many lines of force pass through it as possible.

The simple motor

We now have to consider how the coil can be maintained in revolution instead of merely executing a quarter turn and coming to rest. First of all, let us give the coil some weight, so that, having executed the quarter turn it will " run over " the position shown in fig. 3.4 to that shown in fig. 4.4 and at the same instant let us reverse the current in the coil by some means or other.

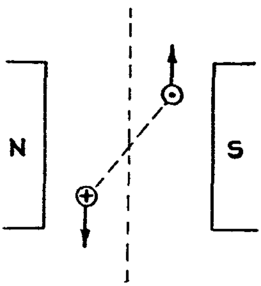


Fig. 4.4. Coil in " run over " position

This reversal will (left-hand rule) have the effect of reversing the direction of the forces (see fig. 5.4) and the coil, instead of returning to the vertical position as it would under the action of the unreversed forces, will continue to revolve for another half revolution.

The reader can now see for himself that another current reversal will take the coil round for yet another half revolution.

The torque acting on the coil is proportional to the flux density in which it moves. For this reason the coil (which may consist of many turns) is always wound in longitudinal slots on a soft iron cylinder. Fig. 6.4 shows a sectional view of the arrangement and we shall in future refer to the combination of coil and cylinder as an armature.

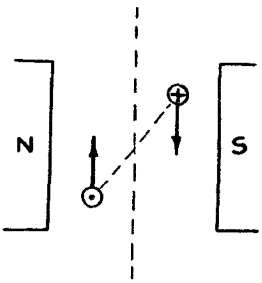


Fig. 5.4. Coil after current reversal

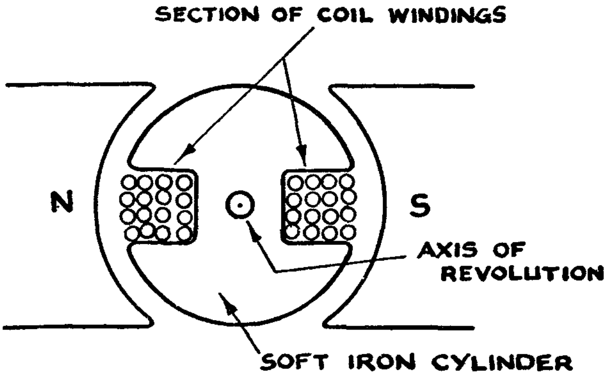


Fig. 6.4. Section of armature

A further modification consists in making the magnets which produce the external fields of such a shape that the armature is in a kind of a tunnel, with as little clearance as possible between it and the pole faces of the field magnets.

The soft iron of the armature and pole-pieces serves to intensify the magnetic flux by reason of its high permeability. This is helped further by special shaping which reduces the air gap to a minimum. It is a point to be noted that air gaps in magnetic systems always tend to reduce the magnetic flux, just as high resistance elements in an electric circuit tend to reduce the electric current.

We shall not discuss how the revolving coil can be kept in contact with an external source of supply. This problem is implicit in the wider question as to how the revolving coil can be supplied from an external source with current which changes direction at the critical instants, which occur just after the coil passes the positions where its plane is vertical.

The commutator

The necessary reversals of current in the armature are brought about by means of a *commutator* which is connected to the external source of supply by means of *brushes.*

Fig. 7.4 shows a simple commutator with brushes; S is the shaft carrying both armature and commutator; E is an ebonite cylinder keyed to the shaft and carrying two half-cylinders of copper or brass (C) rigidly fixed to its surface and insulated from each other by longitudinal gaps (G). Carbon brushes, B_1 and B_2, press lightly against the commutator, and as the latter revolves a brush makes contact first with one segment of the commutator, then with the other.

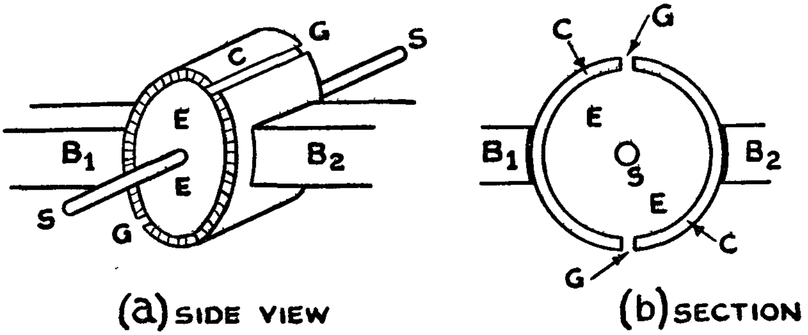


FIG. 7.4. The commutator

To see how this system produces the necessary current reversal, let us follow what happens in more detail; the sketches at fig. 8.4 represent a magnified end view of the two segments of the commutator, and inside it, is diminished sectional view of the coil. One half of the commutator is shaded, so that it may be identified in various positions, and it is assumed that the magnetic field in which the coil revolves is directed from left to right throughout. The condition represented in fig. 8.4 (a) shows the current entering the shaded half of the commutator and flowing down the side of the coil to which it is connected. After passing round the coil the current leaves by the unshaded half of the commutator. The torque will be such as to turn the coil (and the commutator which is fixed on the same shaft) anticlockwise, and after a quarter revolution it will reach the dead-centre position, through which it will run owing to its inertia, to reach the position indicated in fig. 8.4 (b). Here it will be seen that the current is now entering the coil by way of the unshaded segment, so that the current direction in the coil has been reversed, and the torque is again in the direction of the initial rotation. Of course there was an awkward instant when the brushes actually covered the gaps between the commutator segments; this occurred just before the condition represented by fig. 8.4 (b), and the supply was short circuited. This difficulty is one which is inseparable from such a simple motor as we are now considering; it is avoided in larger machines by special design. Apart from the difficulty of this one position, the reader will see that as a coil side passes to the left of the vertical line AB, it is carrying a current into the plane of the paper, and therefore it is being forced downwards; similarly, as a coil side passes to the right of AB, it is being forced upward.

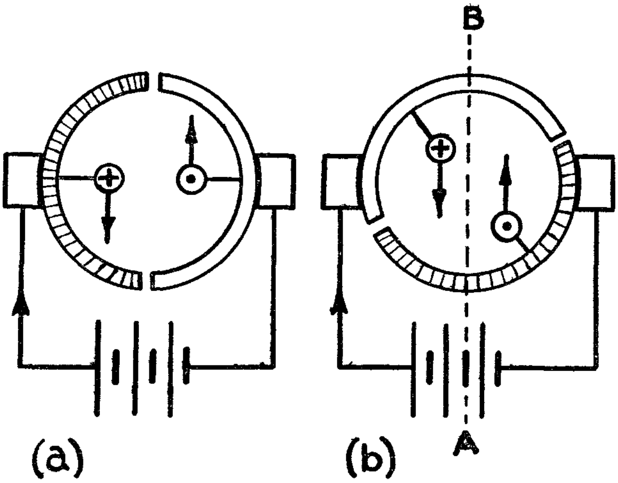


FIG. 8.4. To show the action of a commutator

A motor such as we have described, with a single coil on the armature and a two part commutator, would not be met in practice except as a toy or a model of some sort. Modern motors have a large number of coils disposed in slots all round the armature, and a corresponding number of segments on the commutator, and although such an armature looks a very different thing from the simple one we have dealt with, the essential principle of action is the same. The armature of a large motor is not made of solid soft iron for reasons which will be discussed in a later chapter ; it is built of suitably shaped discs or laminations of soft iron separated from each other by a thin layer of insulation, and pressed together.

Field excitation

Permanent magnets are rarely used to provide the external field in a motor ; the field magnets are excited electrically from the source providing the armature current. According as to whether the field windings are in parallel or series with the armature, the motor is called a " shunt " or " series " motor ; a mixture of the two types is sometimes used and is called "compound" winding. Fig. 9.4 shows these types of winding.

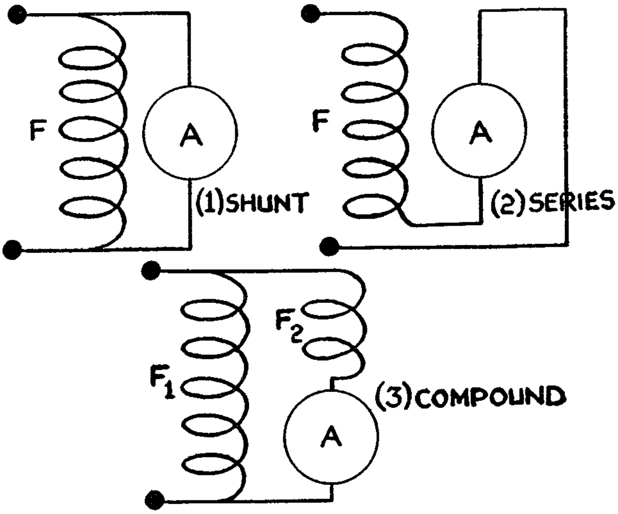


FIG. 9.4. Types of field winding

The field winding in a shunt machine is usually of fine wire with a large number of turns ; in a series machine the field winding is of low resistance. The shunt motor is a constant flux machine and should run at practically the same speed under all normal loads. In practice the speed drops slightly as the load increases ; a series wound machine operates on a varying flux and will therefore run at high speeds on small load ; as the load increases, the speed falls. Series motors are not usually employed where there is a possibility of their running light because they might attain a speed which is beyond the safety limit of the materials of which they are built. The great advantage of a series motor is the enormous starting torque ; for this reason it is used in tramcars and electric trains.

The usual compound-wound machine may be regarded as a modified shunt motor with a minor series field coil which enables it to run at constant speed in spite of a variable load. Less common is the type that approximates to a series motor with a minor shunt winding to prevent racing when the load is removed.

Shunt motor starter

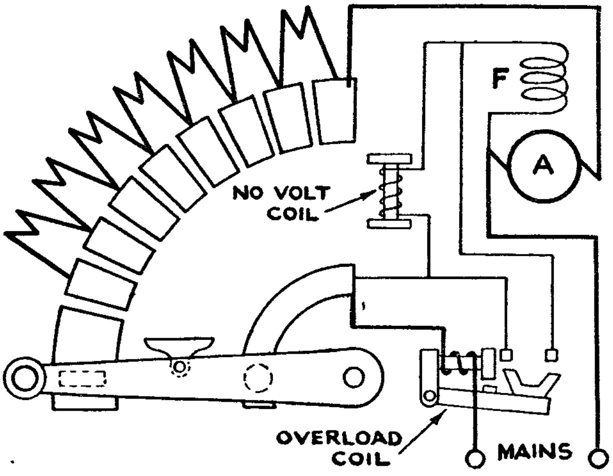


FIG. 10.4. Shunt motor starter

When a supply is switched on to a shunt motor, the only resistance in the circuit is the very low resistance of the armature. A dangerously high current would flow unless we arranged to have some extra resistance in the circuit ; this resistance can be cut out as the motor speeds up because as it does so other things occur which keep the armature current down to safety level (see Chapter V, p. 51) ; the faster the armature revolves, the less current it takes. A shunt motor starter is really a complicated switch ; instead of joining the supply direct to the armature it joins through a number of resistance coils in series. As the starting handle is moved over it cuts these coils out one by one until, when the machine is at full speed, the armature is directly connected to the supply. A spring is fitted which tends to restore the handle to the position shown in fig. 10.4. When the motor is running steadily, the starting handle is held in position by the attraction of the " no-volt " coil ; if the supply fails the starter flies right back so that when the supply is restored the motor has to be started again. Another addition is the " overload coil " ; if the supply current exceeds a set value, this coil brings into operation a switch which short circuits the no volt coil and stops the motor.

Some starters used in the Service are automatic.

Moving coil instruments

The motor principle has another important application ; it is the basis of the action of a class of instruments called moving coil instruments. Suppose a current is passed through a rectangular coil situated in a magnetic field, in the position indicated in fig. 2.4 (a) : the coil will turn about its long axis and if nothing stops it, it will go on so doing until it reaches the vertical position shown in fig. 3.4. But suppose we attach some kind of a spring, like a very weak clock spring to the axis of the coil, so that as it turns it winds up the spring ; then the coil will turn until the torque due to the motor effect is just balanced by the opposing torque due to the wound-up spring. The stronger the current in the coil, the stronger will be the motor torque and the greater will be the angle through which it will turn before it is balanced by the spring torque. In such a system the angle through which the coil turns is a measure of the strength of the current flowing in the coil ; such a system affords a convenient method of measuring a current. Further, since the current in a circuit is directly proportional to the voltage acting in it, a moving coil system (with a large constant series resistance to avoid drawing a large current from the source under test) also serves as a basis for voltage measurement.

It is desirable at this stage that the reader should examine a dismantled moving coil instrument ; the following description of some of the constructional details will then be much easier to follow.

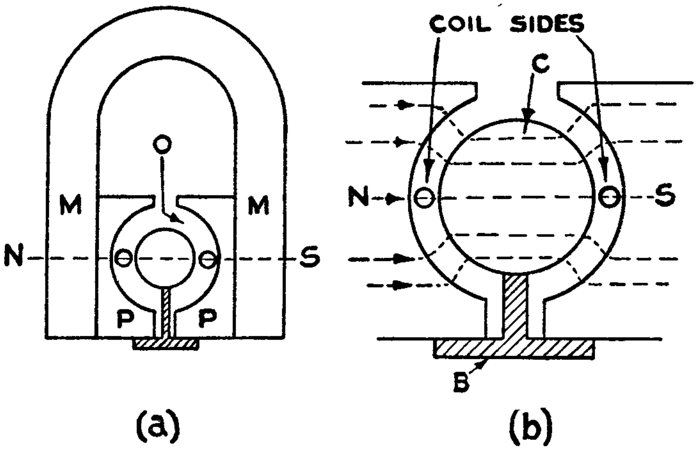


FIG. 11.4. Moving coil instrument (1)

The external field is provided by a strong permanent magnet M (fig. 11.4 (a)) provided with soft iron pole faces P, machined to form a hollow cylinder O on their inner surfaces.

A soft iron core is suspended in the hollow cylinder by means of a brass plate B (fig. 11.4 (b)), so as to leave an annular space in which there will be a strong and uniform magnetic flux.

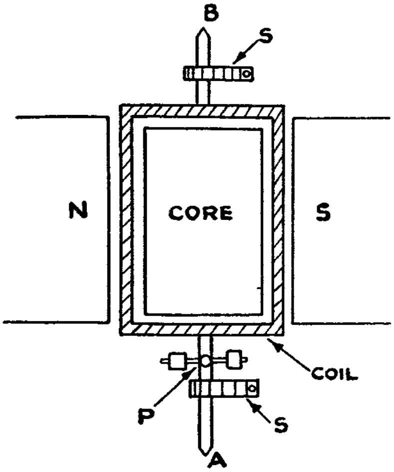


FIG. 12.4. Moving coil instrument (2)

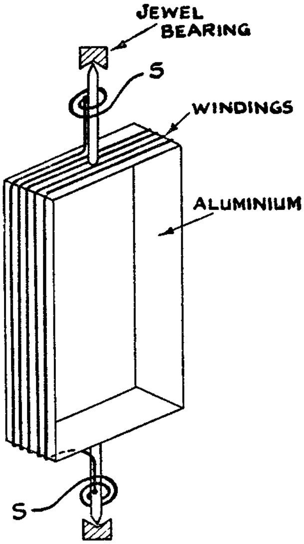


FIG. 13.4. Moving coil instrument (3)

A plan of a horizontal section through the line N.S. (fig. 11.4 (a)) is shown in fig. 12.4; it also shows how the coil is mounted in the annular space, together with the springs S which apply a torque as it turns.

The ends of the spindle A and B are mounted in jewelled bearings. P is the pointer fixed to the spindle, with two counterbalance weights. The coil consists of a large number of turns of fine insulated wire, mounted on an aluminium former and the current enters and leaves by the springs S, fig. 13.4.

The aluminium former on which the moving coil is wound serves, for reasons we shall discuss later (see Chapter V, p. 51), to damp the motion of the coil, and so to make it read quickly, without undue oscillation about its rest-point. An instrument whose needle behaves in this way is said to be aperiodic or dead-beat; it is just a matter of good fortune that moving coil instruments can be rendered dead-beat by so simple a method as we have mentioned.

Now although the angle through which the coil turns against the restoring torque of the springs is a measure of the current in the coil, it is clear that we shall not be able to use this device for the measurement of very large currents without some modification. In the nature of things, the coil windings must be of fairly fine wire, which would not carry currents of more than a few milli-amperes at the outside.

To meet this difficulty, a moving coil ammeter is provided with a shunt, that is to say, a conductor is joined in parallel with the moving coil, and its resistance is chosen so that a fraction only of the total current passes in the coil, the rest going through the shunt which is designed to carry its load without overheating.

Let us take an example to show how the value of a shunt is found. Suppose a certain moving coil instrument, complete except for its shunt, is required to cover a range of 0–10 amperes; the moving coil, let us say, has a resistance of 8 ohms, and a test shows that the instrument is fully deflected when the current in the coil is 5 milliamperes. We base our calculations on the full-scale deflection (FSD); if the relative resistances of coil and shunt are correct for one setting they will be correct for the whole range and FSD conditions provide a very simple solution

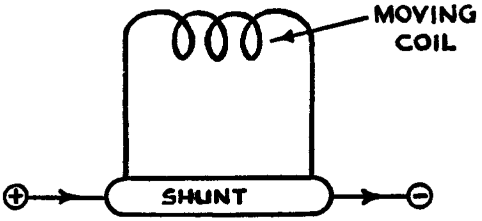


FIG. 14.4. The shunt

(1) Since coil and shunt are in parallel (fig. 14.4), coil current (·005A) + shunt current = 10A, therefore shunt current = 9·995A.

(2) Again, because shunt and coil are in parallel, volts drop across shunt = volts drop across coil, i.e.

$$9{\cdot}995 \times \mathrm{R} = {\cdot}005 \times 8,$$

$$\text{therefore } \mathrm{R} = \frac{{\cdot}04}{9{\cdot}995} \text{ ohms}$$

$$= {\cdot}004 \text{ ohms approximately.}$$

In practice a piece of manganin sheet would be taken having a resistance as near as possible to ·004 ohms, and it would be soldered in place in the instrument. The latter would now be put in series with a master ammeter, and a current would be passed through the two instruments ; if they both read the same, our shunt is correct. If our instrument reads less than the master, it means that not enough current is passing through its coil ; this implies that the shunt resistance is too low. The resistance is raised by gently scraping the shunt with a pen-knife or small file until the reading of the instrument coincides with that of the master. If, on the other hand, our instrument reads more than the master, we must lower the resistance of the shunt. This is usually done by the use of additional high resistance shunts.

It has been mentioned above that the moving coil system can be used as a device for voltage measurement on account of the direct proportionality of current and voltage. We must take care however to ensure that the instrument is adjusted to such a high resistance that it draws no appreciable current ; it is usual to join a fixed high resistance in series with the moving coil to achieve this end.

The series coil in a voltmeter is called a " multiplier ", and by suitably choosing its resistance the instrument can be made to cover any desired range ; the following calculation shows how we might have made up the same movement which we used to form a 0–10 ammeter, to form a 0–100 voltmeter ; instead of a shunt we shall have a series resistance.

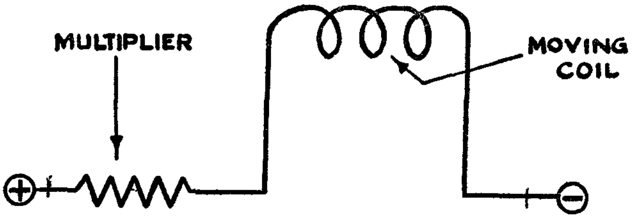


FIG. 15.4. The multiplier

Fig. 15.4 shows the circuit diagram ; again using FSD conditions :—

(1) because the two coils are in series, current in multiplier = current in moving coil = ·005A,

(2) again because the coils are in series, volts drop across multiplier + volts drop across coil = volts drop across instrument = 100,

$$\text{i.e. } {\cdot}005\,\mathrm{R} + {\cdot}005 \times 8 = 100,$$

$$\text{whence } \mathrm{R} = 19992 \text{ ohms.}$$

Instruments, both ammeter and voltmeter, can be made with several ranges ; the reader will easily see that this is only a matter of providing several shunts or multipliers, and employing enough terminals to allow external connections to be made to the appropriate points.

If the scale of a moving coil instrument is examined it will be found that the scale spacing is even, that is, the needle turns through the same angle when it moves from zero to 1 ampere, as it does when it moves from 9 to 10 amperes ; this is a great advantage, in that it gives quite reliable readings at the commencement of its range.

Another important thing to notice about moving coil instruments is that their deflections reverse with reversal of external connection ; this means that unless the instrument is a centre zero type, it will only register when the external connection is correct. If the latter is incorrect, the needle will be forced up against the stop and may be damaged.

Moving iron instruments

These more robust instruments are made in both attraction and repulsion types. A short solenoid CC (fig. 16.4) attracts an eccentric iron disc mounted on a spindle S.

This also carries a pointer and a piston swinging within an air dash-pot, giving pneumatic damping. Spring or gravity control may be used.

The repulsion type has a short solenoid in which are placed two parallel soft-iron strips. One is fixed and the other is attached to a pointer. When current flows, both irons become similarly magnetised and move apart against the action of a spring. Both scale-shapes are non-uniform.

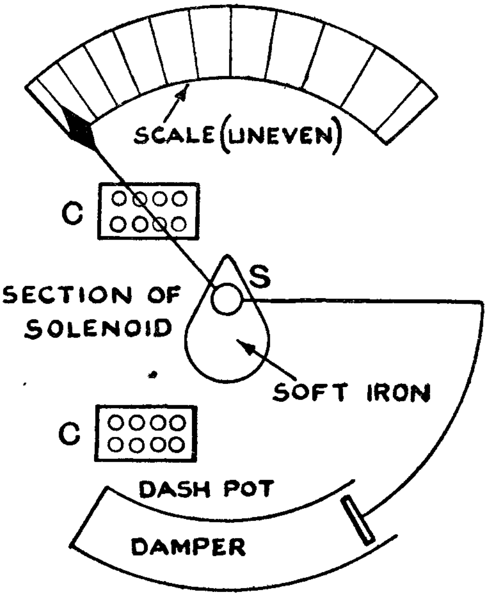


FIG. 16.4. Moving iron instrument (attraction type)

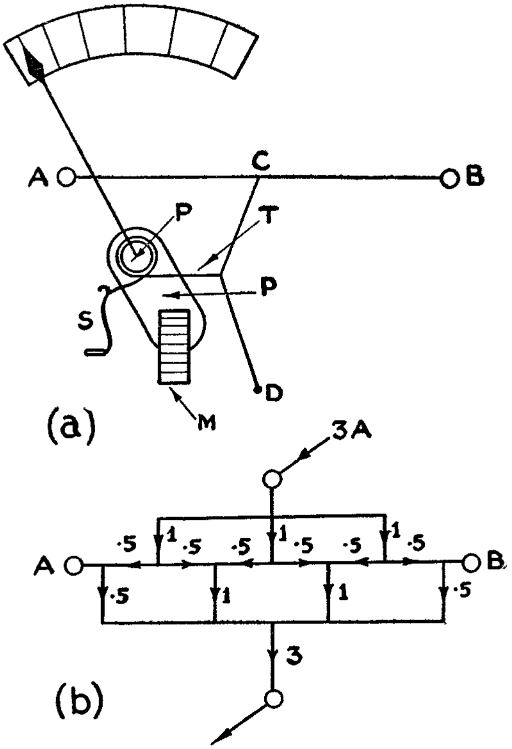


FIG. 17.4. Hot wire instrument

Hot wire instruments

Current passing in a wire generates heat which in its turn causes the wire to expand ; this principle is employed to measure current in this type of instrument.

AB [fig. 17.4 (a)] is a fine wire through which the current to be measured passes. A phosphor-bronze wire CD is fixed to the centre of AB and anchored to the frame of the instrument at D ; at the middle of CD is attached a silk thread T, which passes round a light grooved pulley P, and is fixed to a spring S. The pulley, which rests in jewelled bearings, also carries the needle of the instrument. As AB is heated by the current, it expands and sags ; the spring S then takes up the slack in CD by pulling on the silk thread thus turning the pulley and moving the needle. Also attached to the pulley is a thin sheet of aluminium P, which, moving between the poles of a permanent magnet M, provides the necessary damping.

Owing to the fact that AB must be fairly thin, special arrangements are necessary for the measurement of heavy currents ; one such arrangement is shown in fig. 17.4 (b), in which a 3-amp. current is introduced so that at no part of AB is the current greater than ·5 amp.

The H.W. instrument has a cramped scale, is sluggish in action, uses a lot of energy, and is easily damaged by overloads.

Thermo-junction instruments

Another type of instrument we shall meet depends on the fact that if the junction of two dissimilar metals is heated, an E.M.F. is produced.

The current to be tested passes by the terminals T-T′ (fig. 18.4) through a small heating coil ; this coil warms the thermo-junction and develops a P.D. across its ends AB which is recorded on a suitably sensitive instrument. By passing known currents the scale can be calibrated. This type of instrument is much used in wireless work. The E.M.F. produced is roughly proportional to the square of the current, so that the indicating instrument calibrated in amperes has an uneven scale cramped at the beginning.

The table below is a summary of the outstanding points of the various types of instrument. The electrostatic type will be described in Chapter VI, p. 63.

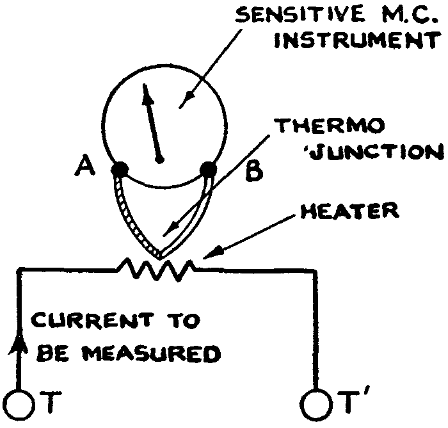


FIG. 18.4. Thermo-junction instrument

Comparative table (instruments)

Type	Power consumption.	Scale shape.	Mechanical features.	A.C. or D.C.	Accuracy.
Moving coil.	Very low.	Uniform.	Moderately robust.	D.C. or rectified A.C.	Very high.
Electro-static voltmeter.	Almost zero	Square law ; end crowded.	Delicate in low voltage models.	D.C. or A.C. (wide frequency range).	Moderate
Thermal type (hot wire).	High.	Square law.	Easily burnt out.	D.C. or A.C. (wide range).	Low.
Thermo-junction type.	High.	Square law.	Easily damaged by overload.	D.C. or A.C. (wide range).	Can be very good.
Moving iron.	Moderate.	Square law (can be different).	Robust.	D.C. or L.F. A.C. (25–100 c/s).	Moderate to good.

Resistance Measurement

If we have suitable instruments we can always find the resistance of a conductor by applying a known voltage and measuring the current set up ; the ratio voltage/current gives the resistance directly. This method, however, can only be regarded as approximate, unless a careful selection of ammeter and voltmeter is possible.

We usually find the resistance of a conductor by a comparative method, which can easily be improvised with very simple apparatus, and it gives quite good results.

Fig 19.4 shows the arrangement ; R_1 is a known resistance R_2 the unknown ; they are joined in series, and in parallel with the pair is a uniform wire (usually a metre long) AB. A galvanometer is connected to a point between R_1 and R_2, and its other terminal is connected to a sliding contact C. A battery is joined to the ends of the slide wire (or to the two resistances in series, whichever way it is regarded). The slider is now moved up and down the wire AB until the galvanometer indicates zero deflection ; when this balance point is found the distances AC, CB are measured off on a scale. The reader should have no difficulty in showing that when the galvanometer is undeflected the following relation holds

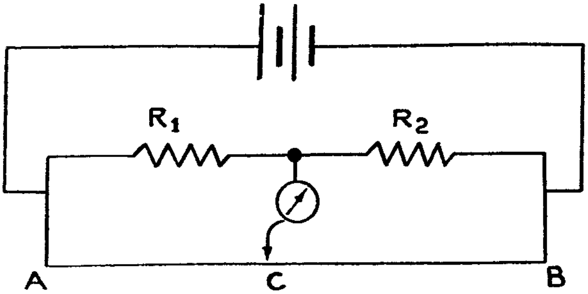


FIG. 19.4. Metre bridge

$$\frac{R_1}{R_2} = \frac{AC}{CB}$$

$$\text{or} \quad R_2 = \frac{R_1 \times CB}{AC}$$

The strength of this method lies in its simplicity and its independence of accurate measuring instruments ; the galvanometer here is merely an indicator, not a measurer.

When we are required to measure resistances of a very high order, in the regions of thousands to millions of ohms, the bridge method above cannot be used ; in such cases we employ an instrument called a " Megger ". Fig. 20.4 illustrates the main features of a megger. Two coils A and B are rigidly fixed together at right angles and pivoted so that they can turn in the field of a powerful magnet. A generator (operated by hand) affords a high potential which is applied directly to one coil through a fixed resistance R. The same potential operates on the other coil which has in series with it the high resistance to be tested. The leads to the coil are fine ligaments exerting practically no torque ; the coils then swing until the two torques (which are in opposition) balance ; and the needle indicates the resistance directly.

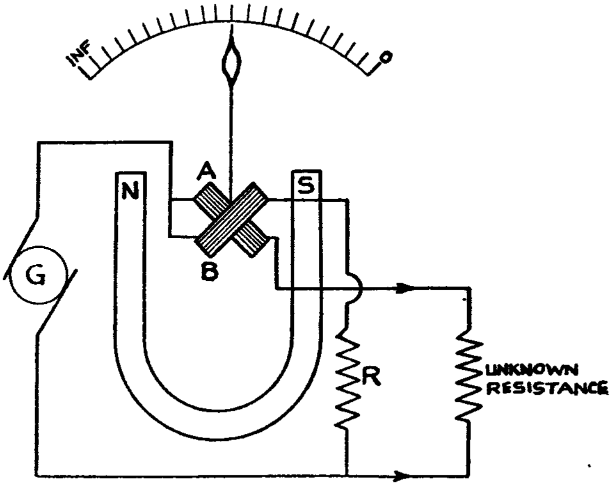


FIG. 20.4. The principle of the Megger

CHAPTER V

ELECTROMAGNETIC INDUCTION

An important advance in electrical science occurred in 1831 when Faraday discovered that electromotive force could be generated by the interaction of a circuit and a magnetic field.

In the original experiment an iron ring R (fig. 1.5) was provided with two windings of insulated wire: one of these P, the primary, could be joined to a battery by means of a switch, while the other, the secondary winding S, was connected to a sensitive galvanometer. When the primary circuit was "made" the galvanometer needle gave a momentary deflection, but when the primary current was established this deflection disappeared. A similar result occurred when the primary circuit was broken but the deflection of the galvanometer was in the opposite direction.

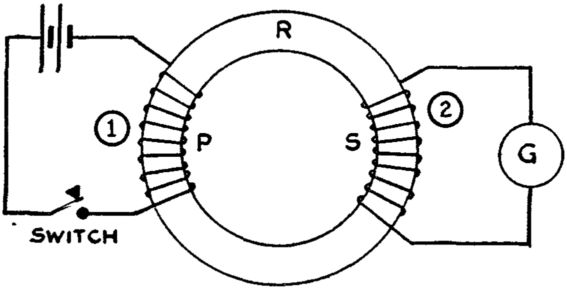


Fig. 1.5. Faradays' experiment

The implication of this important experiment is that while the primary current is growing, an E.M.F. is set up or induced in the secondary, and while the primary current is decaying E.M.F. is again induced in the secondary but in the opposite direction.

Since the only connection between the two windings is the magnetic flux generated by the primary current and threading through the secondary circuit, it would appear that the observed effects might be due to this flux.

Confirmation for this idea is readily forthcoming as the result of the simple experiments about to be described, which the reader should perform for himself.

A solenoid of many turns is connected to a sensitive galvanometer and a permanent magnet is plunged into the coil (fig. 2.5). While the magnet is in motion the galvanometer needle is deflected, but as soon as the magnet comes to rest the needle returns to zero. As the magnet is withdrawn from the coil similar results are obtained but the galvanometer deflection is in the opposite direction.

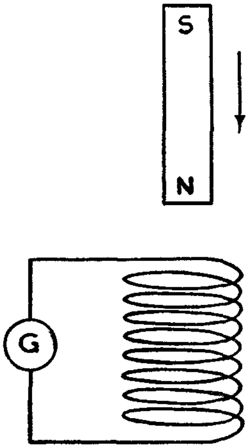


Fig. 2.5. Experiment in Electromagnetic induction

In this experiment it is quite clear that by inserting and withdrawing the magnet we have *altered the number of magnetic lines* threading through the solenoid, and as a result of this action, have set up or induced an E.M.F. in its windings. The same thing happened in the iron ring experiment: switching on the primary current established a magnetic flux through the secondary winding (as well as through the primary) and this change of flux was accompanied by an induced E.M.F.

Subsequent experiments all confirm this result and so we are entitled to make this statement:—

> "Whenever the magnetic flux threading a circuit changes, an E.M.F. is induced in the circuit."

Like all experimental results, this statement is provisional and we must not overlook the possibility that some day it may have to be revised. Up to the present, however, no observation has been made which throws any doubt on its accuracy.

The magnitude of an induced E.M.F. can easily be seen to depend on the speed at which the inductive action is carried out : the faster the magnet is plunged into the solenoid the greater is the deflection of the galvanometer. If it is inserted very slowly, the deflection is very small, but whether inserted rapidly or slowly, the total flux change is exactly the same : the magnitude of the effect therefore appears to depend on the *rate of change of flux.*

Further experiments show that the value of the induced E.M.F. is directly proportional to the rate of change of flux, so we can make another statement :—

" Induced E.M.Fs. are proportional to the rate of change of flux "

which again is supported by all the appropriate experimental evidence.

It is usual to combine these two statements of experimental fact, and to raise them to the status of a physical law, known as Faraday's Law.

> " Whenever the magnetic flux through a circuit changes, an E.M.F. is induced in the circuit : this E.M.F. is proportional to the rate of change of flux."

The phrase " rate of change of flux " needs a word of explanation. If, at a given instant of time the flux through a circuit is 100 lines and exactly one second later the flux is 50 lines, then the *average* rate of change of flux is 50 lines per second. The average rate of change of flux during a given time interval " *t* " seconds (which, of course, may be, and generally is, a fraction of a second) is calculated thus :—

$$\frac{\text{Final flux — original flux}}{t} \text{ lines per second.}$$

If this is negative, the flux is decreasing, if it is positive the flux is increasing. It is this average rate of change which is implied in all cases in this book, but the reader might observe that there is a big difference between an average rate of change and an instantaneous rate of change. A motor car might do a 100-mile journey in 3 hours, when its average speed (or rate of change of position) would be 33⅓ miles per hour, but as a matter of fact, its speed would always be changing : at one moment it would be 10 miles per hour, at another 40 miles per hour and so on. It is this speed, the reading of the speedometer at any selected instant, which is called the instantaneous speed at that instant. Similarly with a flux change, the average rate of change might be 50 lines per second over a given interval of time, but the instantaneous rates of change at selected instants within that time interval might be something totally different. The importance of this lies in the fact that the induced E.M.F. at any instant is proportional to the instantaneous rate of change of flux at the same instant, but, as noted above, we shall avoid this difficulty by supposing that all the flux changes we have to deal with are uniform and that their rates of change are equal at all instants to the average value throughout the interval considered.

Influence of number of turns

If the circuit through which a flux change takes place consists of one single turn, a certain E.M.F. will be induced in the turn, but if the circuit has two turns so close together that the flux change acts through both alike, then the second turn will also have E.M.F. induced in it, and the *total* E.M.F. set up in the circuit will be double what it was with a single turn. Just like cells in series, each convolution in a circuit of many turns through which flux is changing will contribute its quota, and the terminal E.M.F. will be the sum of the contributions from each turn.

All these considerations are expressed in a formula due to Neumann, which gives the actual value in volts of the E.M.F. induced in a circuit of N turns :—

$$e = \frac{\text{N} \times \text{rate of change of flux}}{10^8} \text{ volts.}$$

The factor 10^8 is really a unit-conversion factor and is put in to convert the E.M.F. from what are called electromagnetic units to volts, the practical unit of the engineer. The formula shows that the flux would have to be changing at a rate of 100 million lines per second through a single turn circuit in order to generate an E.M.F. of 1 volt.

Reverting for a moment to the original Faraday ring experiment, the rate of change of flux through the secondary coil cannot be controlled by the operator in the same sense as he could control the rate of change brought about by thrusting a magnet rapidly or slowly into a solenoid. Having switched on the primary circuit, the rate at which the flux grows is dependent upon the properties of the circuits themselves. We shall return to this point later.

So far as the direction of an induced E.M.F. is concerned, we have merely observed that the reversal of any inductive operation results in the reversal of the induced E.M.F.

LENZ'S LAW is a general statement from which we can deduce the direction of an induced E.M.F. in any given case :—

> " Induced E.M.Fs. act in such a direction as to oppose the action to which they are due."

Now, whatever may be the mechanics of the process, the immediate cause of an induced E.M.F is a change of flux. Lenz's Law therefore means that such E.M.Fs. will try to drive a current in such a direction in the circuit that the flux produced thereby will try to neutralise the impressed flux change.

Let us apply this general result to the case of a conducting ring held with its plane vertical and acted upon by magnetic poles approaching towards or receding from its faces as indicated in fig. 3.5.

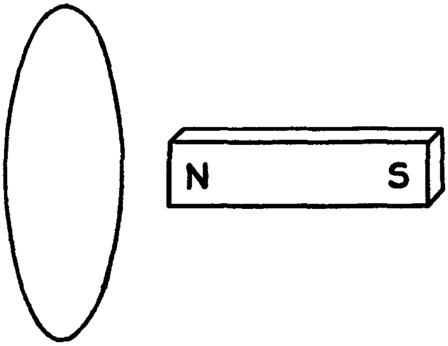


FIG. 3.5. Direction of E.M.F. induced in a ring by a moving magnet

The possible operations with their results are tabulated below. The flux change is called the " action " while the flux which would be generated in the ring by the induced current is called the "reaction". From the direction of the reaction flux, that of the induced current and of the induced E.M.F. can be deduced by the ordinary solenoid rules (p. 27). In order to avoid tiresome phraseology we shall call the " anti-clockwise viewed from the right " direction round the ring positive, and the opposite direction will be regarded as negative.

Operation.	Action.	Reaction.	Direction of E.M.F.
N approaching R S ,, L or both simultaneously.	Increase of flux R to L	Formation of flux L to R	+
N receding from R S ,, ,, L or both simultaneously.	Decrease of flux R to L	Formation of flux R to L	—
S approaching R N ,, L or both simultaneously.	Increase of flux L to R	Formation of flux R to L	—
S receding from R N ,, ,, L or both simultaneously.	Decrease of flux L to R	Formation of flux L to R	+

Alternative view of induction

Consistent results are obtained if, instead of attributing inductive effects to the change of flux through a circuit, we regard them as due to the *cutting of lines of force* by the conductors forming the circuit.

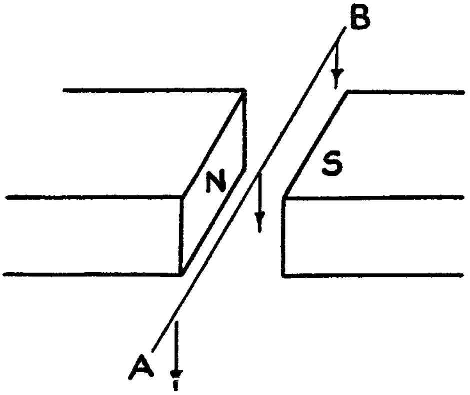


FIG. 4.5. Conductor cutting lines of force

Thus, if the wire AB (fig. 4.5) is allowed to fall through the field between the N and S poles of a magnet, it will cut the lines of force (which must be supposed to join up again after the wire has passed), and in so doing will generate an E.M.F. in itself.

From this point of view we may state :—

> " Whenever a conductor cuts lines of force an E.M.F. is set up in it, proportional to the rate at which the lines are cut."

The value of the E.M.F. is equal to :—

$$\frac{\text{rate of cutting lines}}{10^8} \text{ volts.}$$

Its direction within the moving conductor is given by Fleming's right hand rule :—

> " Extend the thumb, fore and second fingers of the right hand so as to be mutually perpendicular, then the
> Fore-finger gives the direction of the lines of Force ;
> seCond finger gives the direction of the Current, and the
> thuMb gives the direction of Motion of the wire."

Both the idea of flux change and that of cutting lines of force lead to correct results. In using the latter, however, care must be taken always to consider the directions of the induced E.M.Fs. because two parts of a circuit might both be cutting lines simultaneously but the induced E.M.Fs. may be oppositely directed, giving no resultant value at all, as for example in the case of a ring falling under gravity with its plane at right angles to a uniform horizontal magnetic field.

Self-Induction and Inductance

Consider the case of a solenoid, with an iron core for preference. It is in circuit with a switch and a battery : when the switch is made there is a sudden growth of flux in the core. This means that an E.M.F. must have been induced in the solenoid and that during the interval of flux growth two E.M.Fs. have been acting :—

(i) The battery E.M.F. creating current and hence flux ;
(ii) Induced E.M.F. opposing the flux growth (Lenz).

This kind of induction which is due fundamentally to change of current in the circuit itself is called *Self-induction.* The sequence of events is actually :—

> Changing current—changing flux—induced E.M.F., but we agree for convenience to forget the middle step and attribute the induced E.M.F. directly to changing current.

Since the self-induced E.M.F. will act in such a direction as to oppose the action to which it is due, a decrease in current will give rise to an E.M.F. acting in the direction of the current and tending to prevent the decrease. This is why, when a circuit, such as that of an electromagnet, is broken, a large spark is formed between the points of the switch. The spark is due to the E.M.F. set up as a result of the destruction of the current, and it is so big that it is able to arc across the switch contacts.

Conversely, when a circuit containing a source of P.D. is closed, the rising current sets up an induced E.M.F. in the opposite sense to the current, whose establishment is thus delayed. The graph of fig. 5.5 indicates how the current rises and falls in a circuit, taking into account the effects of self induction.

Of course, these inductive effects only last a very short time : unless the circuit is specially designed, their duration is of the order of thousandths of a second.

A convenient way of looking at inductive effects of this sort is to regard the circuit itself as having the ability to offer opposition to current changes, just as a massive body offers opposition to any attempt to change its speed, by virtue of a property called *inertia.*

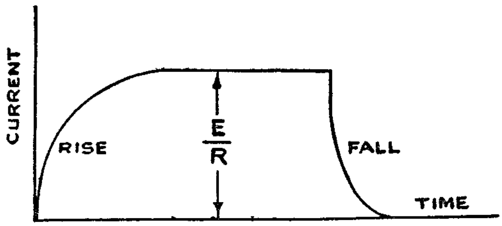


FIG. 5.5. Rise and fall of current in inductive circuit

This property of a circuit enabling it to oppose current changes is called its SELF-INDUCTANCE, or more usually its INDUCTANCE. Just as resistance is the property of a circuit by which it opposes the flow of current, so inductance is the property by which it opposes changes of current.

We are able to attach a numerical value to this property by a process similar to that which we used in the case of resistance. Then, it will be remembered that because I was proportional to V, we could write for any given circuit :—

$$\frac{V}{I} = \text{a constant (see p. 11).}$$

We agreed that this numerical constant should be taken as the numerical value of the resistance in ohms.

In the case of induction, experimental results show that the self-induced voltage is directly proportional to the rate of change of current, so that for any given circuit we can write :—

$$\frac{\text{Induced E.M.F. (volts)}}{\text{Rate of change of current (amps/sec.)}} = \text{a constant}$$

and we agree that this constant shall be taken as the numerical value of the inductance ; a circuit will have one unit of inductance (one HENRY) if, when the current changes at the rate of 1 ampere per second, the self induced E.M.F. is 1 volt. Transposing the above relation, the voltage (E) induced in a circuit of L henries, when the current is changing at S amperes per second is

$$E = L \times S.$$

A highly inductive circuit is one which, for a given steady current, produces a very large magnetic flux ; to be more exact, it is a circuit for which the product

$$\text{(flux per amp.)} \times \text{turns}$$

is as large as possible. Thus a large number of turns of insulated wire on a soft iron core will constitute a circuit af high inductance. Such a circuit is called a *choking coil* or simply a *choke.*

A non-inductive circuit can be made by ensuring that at any point, equal currents are flowing in opposite directions in conductors situated very near to one another. Under these circumstances the magnetic effects of the two currents cancel one another and no matter what the value of the current is, the resultant flux is zero. From this it follows

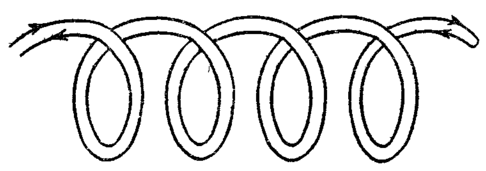


FIG. 6.5. Non-inductive winding

a change of current will not be accompanied by a change of flux, and inductive effects will be absent. Standard resistance coils are wound non-inductively: before the wire is wound on to the former it is doubled as indicated in fig. 6.5, and, therefore produces no resultant magnetic field.

Circuits of variable inductance are widely used in wireless work. One of the easiest ways of achieving such a variation is to have some sort of switching arrangement whereby the number of turns included in the circuit can be varied. Fig. 7.5 (a) shows a device of this kind. The moving arm can make contact with a series of studs tapped into the coil at various points along it. Another way is to have a core of some magnetic material whose position within the coil can be varied (see fig. 7.5 (b)). Special kinds of magnetic materials have been developed almost free from hysteresis and eddy current (see pp.29, 51) effects which make this simple type of variable inductance reliable and effective.

Another form of variable inductance is called a variometer: two coils are arranged so that one can be turned inside the other as indicated in fig. 8.5.

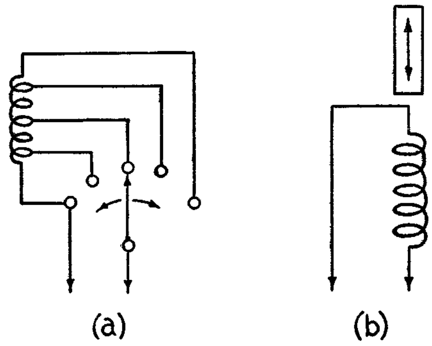


FIG. 7.5. Variable inductances

The coils may be connected either in series or in parallel. When the two coils are lying in parallel planes their magnetic fields may be directed in either the same or in opposite directions. Suppose they are acting in the same direction, then the magnetic fluxes reinforce each other and the inductance of the circuit is high. If the inner coil is turned through 180° the fluxes cancel each other and the circuit is practically non-inductive. Thus a continuous range of inductance, zero to maximum, is obtained by turning the inner coil through half a revolution. If the coils are each of inductance L henries then, with series connection the range of the variometer is 0 to 4L henries: in parallel, the range is 0 to L henries.

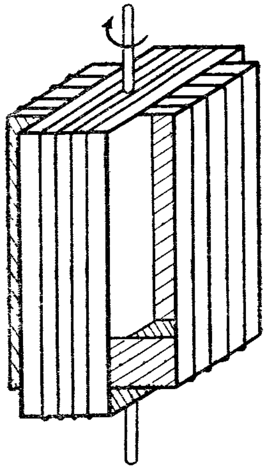


FIG. 8.5. The variometer

Time constant

It is clear from what has already been said that the time necessary for the establishment of current in a circuit will depend upon the inductance: the greater the inductance, the longer time must elapse between the closing of the switch and the establishment of full current. Closer investigation than we are in a position to undertake shows that the inductance is not alone in regulating the rate of rise of current: the resistance comes into the story as well, with the net result that the quotient

$$\frac{L}{R}, \text{ i.e. } \frac{\text{inductance}}{\text{resistance}}$$

appears to be the determining factor.

L/R is called the " time-constant " of the circuit and is equal to the time in seconds which must elapse between the closing of the switch and the establishment, approximately, of two-thirds of full current. Notice that the *higher* the resistance the quicker the current rises, provided that the inductance remain constant. Similar reasoning applies to the rate of fall of current.

Mutual Induction and Inductance

Another special case of induction is that in which E.M.F. is set up in a circuit as a result of current changes in a neighbouring circuit. Here the series of operations is :—

(i) Current change in circuit 1, the primarary circuit. (See fig. 1.5.)
(ii) Flux change through circuit 2, the secondary circuit.
(iii) Induced E.M.F. in secondary,

but, as in the case of self-induction we forget the middle step and attribute the secondary E.M.F. directly to the current changes in the primary circuit.

This process is called *mutual induction,* and the pair of circuits is said to possess MUTUAL INDUCTANCE. By a train of reasoning similar to that previously followed, the reader will be able to arrive at the following results :—

(i) $$\frac{\text{E.M.F. (volts) induced in secondary}}{\text{rate of change of primary current (amps/sec.)}} = \text{mutual inductance of the pair of circuits (henries)}$$

(ii) Two circuits have a mutual inductance of one henry if an E.M.F. of one volt is set up in the secondary when the primary current changes at the rate of one ampere per second.

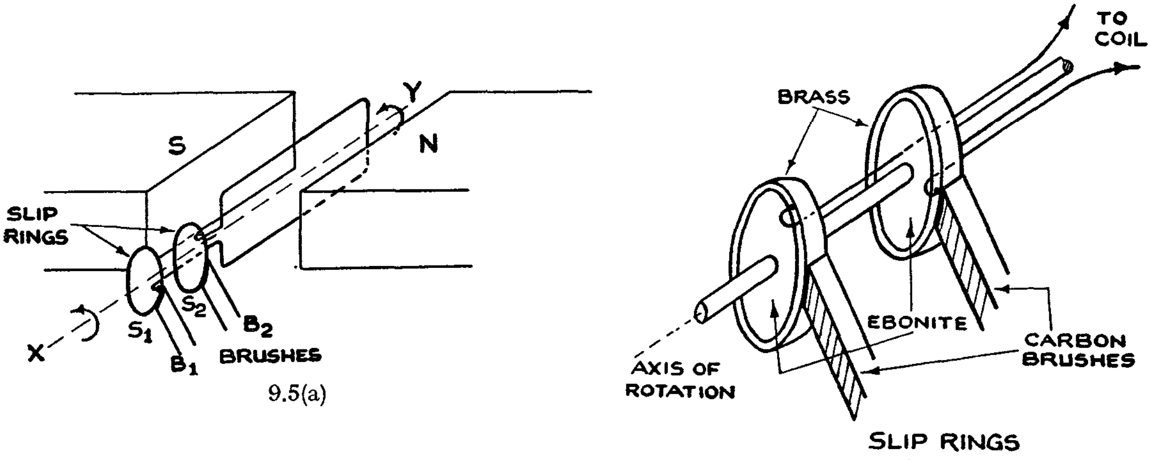
Transposing the above relations, we get that the voltage E induced in the secondary of two circuits having a mutual inductance of M henries when the primary current is changing at S amps. per second is

$$E = M \times S$$

The mutual inductance of a pair of circuits can most easily be varied by altering their relative positions. When they are close together, so that practically all the flux produced by current in the primary links with the secondary circuit, their mutual inductance is a maximum ; a wireless engineer would say that they were " tightly coupled ".

The dynamo principle

The phenomenon of induction has a most important application in the dynamo or generator, which reduced to its simplest terms is a machine which keeps the magnetic flux threading a circuit in a continual state of change, and by so doing is constantly generating an E.M.F.

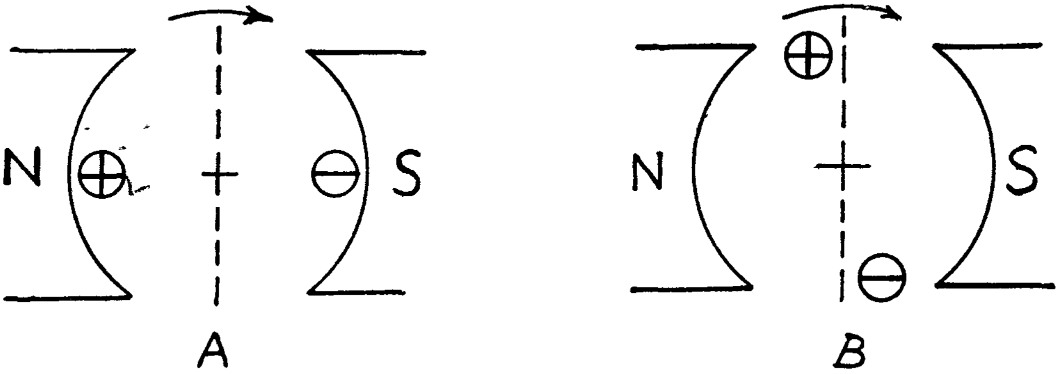


9.5(a)

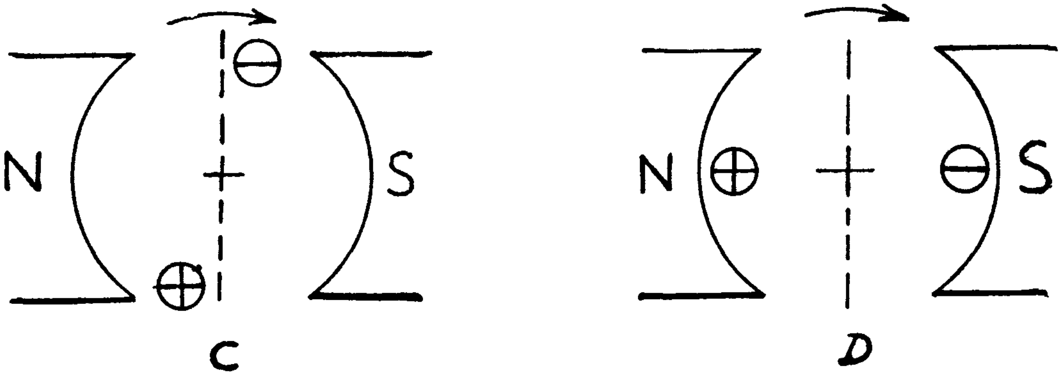
9.5(b)

FIG. 9.5. Simple dynamo and slip rings

The structural features of a dynamo are the same as those of a motor. The field magnets produce an intense magnetic field in which a soft iron laminated armature, carrying coils of insulated wire in slots on its surface, is kept in rotation by some external agency. Let us consider the action of a very simple form of dynamo consisting of a rectangular loop of wire rotating between the pole faces of a magnet, as indicated in fig. 9.5 (a). In order that we may examine the electrical actions going on in this coil, we will suppose that its ends are brought out to slip rings fixed to the rotating shaft over an insulating bush, and upon these slip rings press lightly a couple of carbon brushes (see fig. 9.5 (b)). By means of this device we can keep in electrical contact with the coil as it revolves and we will suppose that connected to the brushes there is a central zero galvanometer or voltmeter which will register any E.M.F. set up as the loop (fig. 9.5 (a)) rotates in a clockwise sense about the axis XY.



FIRST QUARTER REVOLUTION



SECOND QUARTER REVOLUTION

Fig. 9.5 (c).

The direction of the induced voltage can be predicted using Fleming's Right Hand Rule (p. 43), and we arrive at the results shown on the diagrams of fig. 9.5 (c). Starting from the horizontal position A, there will be no flux threading through the coil, but as soon as the latter commences to move, a rapid increase of flux occurs and a large induced E.M.F. is generated. The E.M.F. gradually falls to zero as the position B is approached. Between B and C the voltage is reversed, and an increasing E.M.F. is now generated, which reaches a maximum at the position D, after which the cycle is repeated. It is not usual to draw the voltage graph starting at a maximum value, and to agree with convention, it is customary to start at the position C, and work through the same cycle of operations. Fig. 10.5 shows the resulting graph, in which position (1) coincides with C on fig. 9.5 (c) and the others follow in cyclic order.

Thus we see that as the coil revolves, the induced E.M.F. will be in one direction for half the revolution and in the other direction for the other half, and the needle of our galvanometer will swing to and fro as the coil goes round. An E.M.F. which reverses its direction in this way is called an alternating E.M.F. We learn therefore that our rotating coil does not produce a direct E.M.F. like that of a cell but gives an alternating E.M.F. If we had suitable instruments we could register the value of the induced E.M.F. at every instant as the coil went round, and we could then plot a graph showing the relation between the E.M.F. at any instant and the corresponding angular displacement of the coil. We should get a graph like that of fig. 10.5, which represents one CYCLE of the alternating voltage.

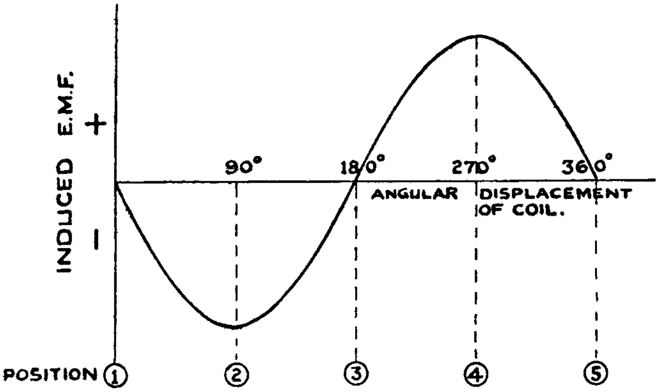


Fig. 10.5. Alternating voltage

This is the principle underlying the action of all dynamos: it is true that the drum wound laminated armature of a modern generator with its many coils bears little superficial resemblance to the simple system described above, but the essential principle of action is the same.

By means of a commutator we can convert the output of our simple generator, not to direct voltage, but to a uni-directional pulsating voltage. The diagram of fig. 11.5 indicates how the commutator functions. Just as when we dealt with the corresponding problem in the case of the motor the diagrams of fig. 11.5 show an enlarged 2-part commutator and a diminished section of the coil. One half of the commutator is shaded in for the purposes of identification. Employing the usual convention, ⊕ signifies that the E.M.F. is acting away from the reader, and ⊙ signifies the opposite direction.

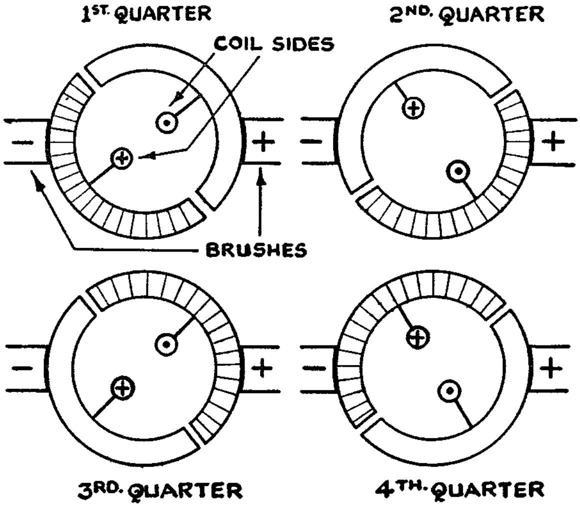


Fig. 11.5. Commutator action

It will be seen that the direction of the E.M.F. induced in the coil reverses every time the coil plane passes through a vertical position, but at the same instant, by the action of the commutator, the connections between the coil and the external circuit are also reversed, with the result that the E.M.F. acting in the external circuit is always in the same direction. The rotation is assumed clockwise, and the direction of flux from left to right. In other words, as a coil side moves from one side to the other of the vertical plane the E.M.F. acting in it reverses and it is automatically connected to the opposite brush.

Although the output is thus rendered uni-directional, it is still far from steady. It is a pulsating voltage, rising and falling twice every revolution, as shown in the graph of fig. 12.5.

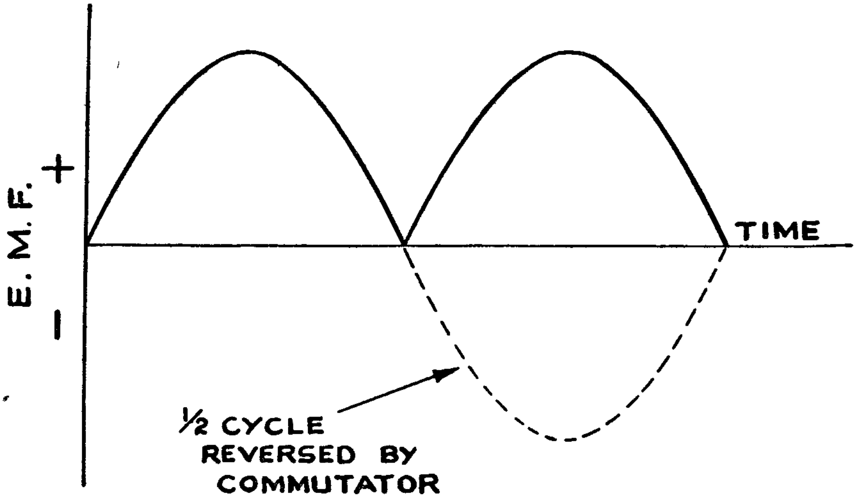


Fig. 12.5. Uni-directional output due to commutator action

A single turn coil such as we have been considering would constitute a most inefficient dynamo. The first obvious modification would be to have a coil of many turns of insulated wire carried in slots on the surface of a soft iron armature, as indicated in fig. 13.5 which is a sectional drawing of a simple H-type armature.

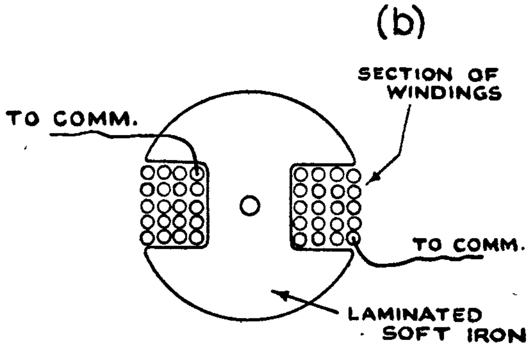


Fig. 13.5. Simple armature

Modern generators have many coils evenly spaced on the surface of the armature (drum winding) and the commutator has a correspondingly large number of segments. Not only does this secure a big output, but it smooths out the pulsations and makes the voltage almost steady. There is always a slight residual pulsation which is called " commutator ripple ".

Since induced E.M.F. depends on rate of change of flux, the faster the armature of a dynamo rotates, the higher will be the voltage generated. When the dynamo is driven by a machine whose speed is maintained constant, this effect is not important, but it is serious in cases where the dynamo is operated by a machine whose speed must vary, as, for example, in the case of the engine driven generator on an aeroplane or that of a motor car dynamo. In such cases the field is provided with special compensating windings arranged so that if the terminal voltage rises above normal, the field excitation is reduced and the rise in voltage is checked. Another constant voltage device is the Tirrill Regulator* which automatically puts a resistance in the field circuit if the output voltage gets too high.

Motor generator

A generator may be driven in a variety of ways : steam turbines are extensively used in commercial practice. Water turbines are employed where water power is plentiful. Small generating sets suitable for charging accumulators are usually driven by a petrol engine. A generator can also be driven by an electric motor. The motor and dynamo may be separate machines with their axles suitably coupled together, or, as in the case of the

---

* See Naval Electrical Manual (H.M.S.O.) 1928, p. 356.

motor generator, a single armature may carry two separate windings, each with its commutator and brush gear, one acting as a motor driving the other round in their common magnetic field and causing it to act as generator. Usually the motor winding is designed to work on a 12 volt supply, while the generator winding is designed to give about 1,200 volts, sufficient to operate a wireless transmitter. A motor generator is used to change the voltage of a D.C. supply to another value. A typical motor generator in service use is the 80 watt type : the motor input is about 12 volts 10 amps from the 12 volt supply of an aeroplane, while the output is about 1,200 volts 70 milliamperes.

Field excitation

The flux in which the armature of a dynamo revolves can be provided in a variety of ways. Small generators (e.g. magnetos) use strong permanent magnets ; in addition we have :—

(*a*) Separately excited machines.

(*b*) Shunt wound machines.

(*c*) Series wound machines.

(*d*) Compound wound machines.

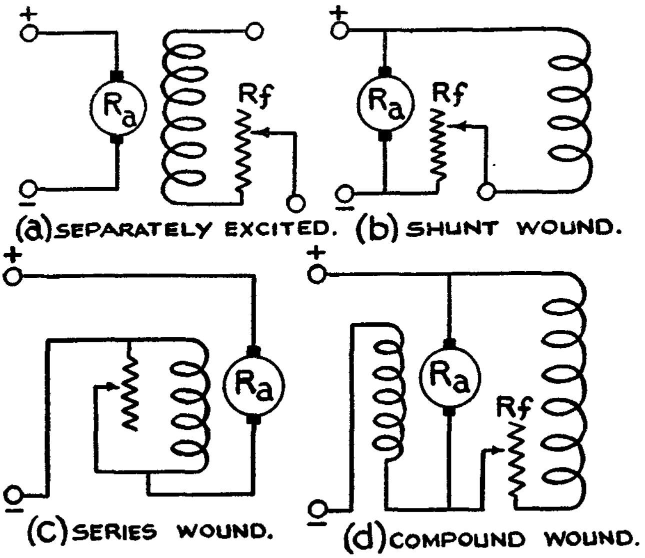


Fig. 14.5. Various types of winding

The sketches of fig. 14.5 indicate the arrangements of the windings in each case.

In a separately excited machine, the flux will be almost independent of the current in the armature, so that at constant speed the E.M.F. generated will be independent of the current delivered : the terminal voltage, however, will fall as the current rises owing to the internal resistance of the armature. The fall in terminal voltage from no load to full rated load, we call the " regulation " of the machine.

In a shunt wound machine, the initial voltage on starting up is due to the small amount of residual flux in the field magnets : as the armature voltage builds up, so does the shunt field current until steady conditions are reached.

In starting up such a machine care must be taken to ensure open circuit : if the machine is asked to deliver current at the outset, it may not build up. As more and more current is drawn from a shunt wound machine, the terminal volts fall and if the load becomes excessive the machine shuts down altogether.

A series wound machine has what is called a " rising characteristic " : the more current we draw from it, the higher is the E.M.F. induced in the armature, and the higher the terminal volts. This is because the armature current also goes through the field windings and until magnetic saturation is reached, the higher the current, the stronger the flux in which the armature is rotating.

As mentioned above, a combination of shunt and series windings can be used to provide a machine with a characteristic which is almost level. For further details on electrical machines see Admiralty Handbook of Wireless Telegraphy (H.M.S.O.), Chapter IV, p. 152 *et seq.*

Note on the electric motor

We are now in a position to see why an electric motor takes less current as its speed increases. The windings of the armature are revolving in a magnetic field and therefore they are not only doing their proper job of developing a torque to maintain the revolution of the armature, but they are also having E.M.F. induced in them according to the dynamo principle, and moreover this E.M.F. must oppose the applied E.M.F. to which it is originally due. As the motor speeds up, therefore, it will develop an internal E.M.F. in opposition to the applied voltage, so that the total voltage acting on the armature coils is now their *difference,* and the current falls. Put in the form of an equation, for a motor running light, that is, doing no external work,

$$\text{Applied voltage (V)} - \text{induced voltage (E)} = \text{volts drop across armature } (r\,i)$$

$$\text{thus } i = \frac{V - E}{r}$$

The motor will speed up until this condition is reached, and the current, at first large, will drop to the value i given by this equation. E, of course, will also depend on the strength of the field in which the armature is revolving. If we want a motor to go faster we must weaken the field, so that the speed must get higher to reach the condition mentioned. Resistance must be *put in* the field circuit to increase the speed of the motor.

Eddy currents—laminations

So far we have only considered inductive effects in circuits consisting of linear conductors. Suppose we brought about a flux change through a mass of metal, as for example, when we bring a powerful magnet up to a sheet of copper. The copper sheet could be considered to be built up of an enormous number of rings, rectangles or closed circuits of any shape. Through each one of these a flux change would occur and an E.M.F. would be set up round its boundary. The induced currents which flow in a solid conductor owing to this kind of action are called " Eddy currents " and by their nature, oppose the cause to which they are due. If the changing flux is due to the motion of the metal surface through the field the eddy current reaction will be to introduce a drag opposing that motion. On page 35 we mentioned that the coils of M.C. instruments were usually wound on thin aluminium or silver formers ; the reason for this is now clear. As the coil moves through the field, eddy currents are induced in the metal former, and their effect is to steady its motion just as if it were moving in oil, and to damp out any vibration.

Eddy currents present a serious problem in the case of the armature of a motor or of a dynamo. Consider first a soft iron cylinder revolving about its axis in a transverse magnetic field, as shown in fig. 15.5.

Take a " slice " of this cylinder contained between the two planes AB and CD parallel to each other and to the axis of the revolution, disregarding the rest of the cylinder for a moment. As this rectangular slice revolves, its edges will behave just like the rectangular coil of the simple dynamo : alternating E.M.F. will be set up in it and alternating current will flow around it. Similar arguments apply to any other parallel slice of this sort, so that as the cylinder goes round, eddy currents will flow on its surface in paths parallel to the axis of rotation as in fig. 15.5(b). Not only will these currents introduce a mechanical drag on the revolving cylinder, but they will heat its surface causing a waste

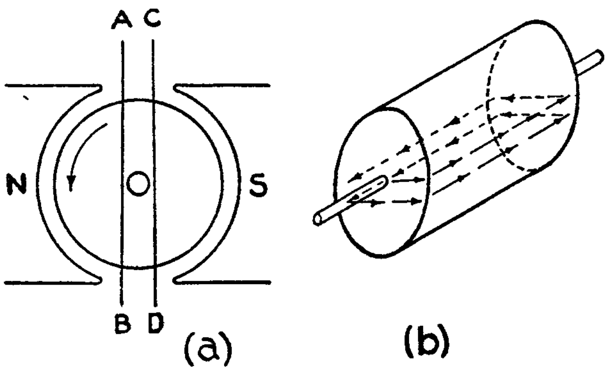


FIG. 15.5. Illustrating eddy currents

of energy, and possibly burning the insulated windings the cylinder might have to carry.

To meet this difficulty, an armature is not made of solid iron: thin sheets of iron (laminations) of the correct shape, each separated from its neighbour by a thin film of shellac, are threaded on to the steel shaft, and when a sufficient number are in position the laminations are pressed together to form a solid mass. Although when a laminated armature revolves in a magnetic field small E.M.Fs. will be generated just as in the case of a solid cylinder, *no appreciable current will flow* because the path along which the eddy currents want to travel is intersected with strips of insulation—the exposed films of shellac situated between the laminations. All the iron parts of electromagnetic apparatus subjected to changing flux are laminated in this way and for the same reason.

CHAPTER VI

ELECTROSTATICS

We learned in Chapter III that a permanent magnet or an active conductor created in the surrounding space a condition which we called a magnetic field. The outstanding characteristic of this space was its ability to generate forces on pieces of iron or steel situated in it. We sought to explain these forces by assuming that the magnetic field was a collection of elastic filaments called lines of force which, owing to the distribution shown by iron filing maps and their own elastic properties, would in fact generate the observed forces.

The magnetic field is not the only " distance-force " field; the gravitational field of the earth (or of any massive body) is another; any piece of matter is attracted to the earth's centre with a force proportional to its mass, and called its weight, but we are so familiar with the gravitational field, that we seldom pause to think that it is really quite a mysterious sort of thing.

Yet another field of this type is the electric field; the space around an electric charge is one in which fragments of almost any substance experience forces tending to move them in some way or other; indeed the attraction of light bodies by a piece of amber which has been suitably rubbed is the earliest electrical experiment on record; the amber has become the seat of an electric charge, and the space surrounding it an electric field in which light objects are moved under the action of the forces set up.

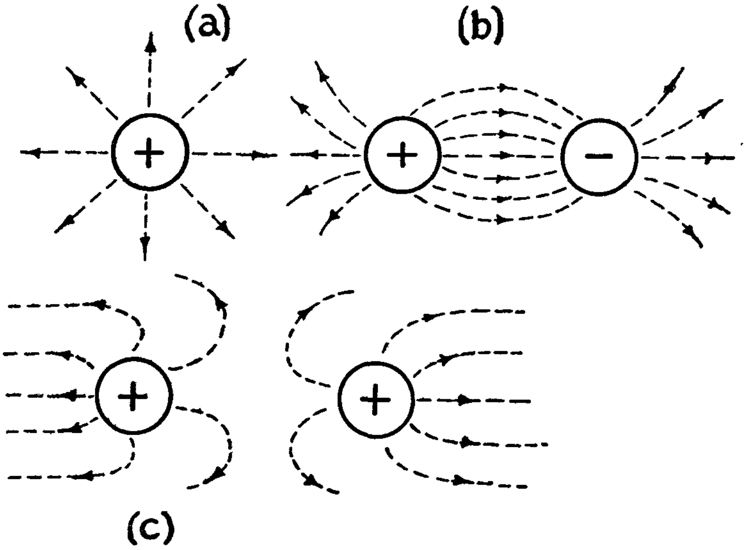


FIG. 1.6. Examples of electric fields

The explanation of the observed facts of the electric field follows the same general lines as in the case of the magnetic field. Positive electricity is regarded as the origin of "*electro-lines*", which have the usual elastic properties; these lines end on a negative charge, and if there is no free negative electricity about for the line to terminate on it will "induce" a negative charge for itself on the most convenient conducting object.

The distribution of these lines can be plotted by scattering certain crystalline substances in the field, just as we scattered iron filings in the magnetic field. Fig. 1.6 (a) shows the field surrounding a positively charged isolated sphere. The lines are radial and their distribution indicates that the charge behaves as if concentrated at the centre of the sphere; as this particular point will be useful later on, let us notice in passing that it is capable of rigid mathematical confirmation.

Fig. 1.6. (b) shows the line system due to two equal and opposite charges and clearly results in attraction; fig. 1.6 (c) shows the lines for two equal and like charges causing repulsion. The ends of the lines which appear to be unattached will travel on to the boundary of the system and end in an induced charge of the appropriate sign.

The introduction of uncharged conductors into electric fields has some interesting features. Fig. 2.6 shows an uncharged sphere (B) placed in the field of a charged sphere (A). The lines from A seem attracted to B and some of them actually terminate on B with their induced charge terminations; induced charges always happen in equal and opposite pairs, so that these lines reappear on the other side of B, growing apparently from the induced positive charge. If A were removed from the neighbourhood of B these induced charges would disappear and B would again be free of charge; but if, while still under the influence of A, B were momentarily earthed, the positive induced charges would have been removed by the earth connection and the removal of A would have left B with a negative charge; this is known as "charging by induction".

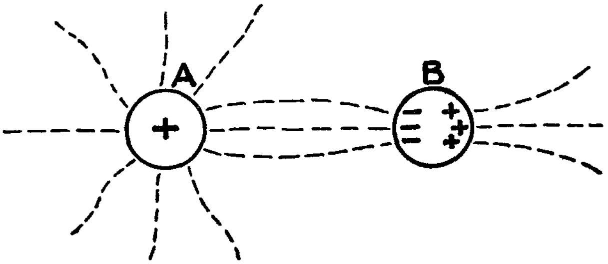


FIG. 2.6. Effect of an uncharged conductor

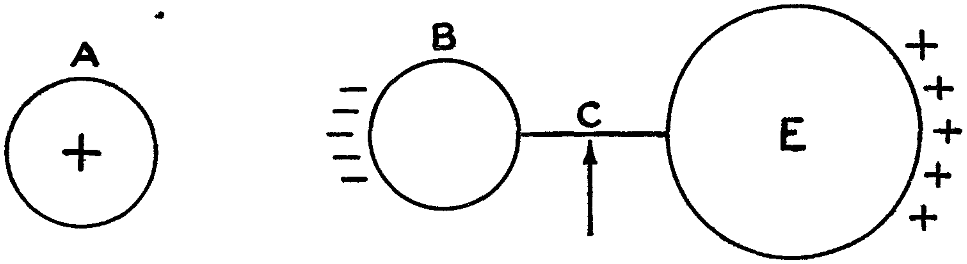


FIG. 3.6. To illustrate charging by induction

A (fig. 3.6) is a positively charged conductor, B another conductor connected to earth. Regarding earth for a moment as just another enormous conductor, we shall have the state of affairs represented in the figure; negative charges will appear on B (attracted by the positive charge in (A) and positive charges will be repelled to earth. If now the connecting wire is broken at C, B will be left with a negative charge, but the small positive charge acquired by such a huge body as the earth will have no effect on its electrical condition. B is said to have been charged (negatively) by induction.

Action of points

When an isolated sphere is given a charge, the latter spreads out evenly all over its surface; the electric *surface density*, we say, is uniform. The same is not true for conductors having sharp points; the charge appears to accumulate on the points, and so dense does it become on this account that the air surrounding the point becomes charged as well. The repulsion between the charged point and the charged air causes the latter to stream away and to take much of the charge with it. For this reason electrical apparatus used in static experiments has no sharp points; rounded knobs as far as possible form the free ends of any conducting member of the apparatus.

The action of a lightning conductor is worth mentioning in this connection. A charged cloud (A) fig. 4.6 passes over a building, and the electrolines terminate on the nearest available object, the point of the lightning conductor, with terminal induced charges, of the opposite kind to that on the cloud.

The surface density of the induced charge is so high on the point that the surrounding air is charged and driven off. Being of the opposite kind to the charge in the cloud it is attracted to the latter, and the two charges neutralise each other harmlessly.

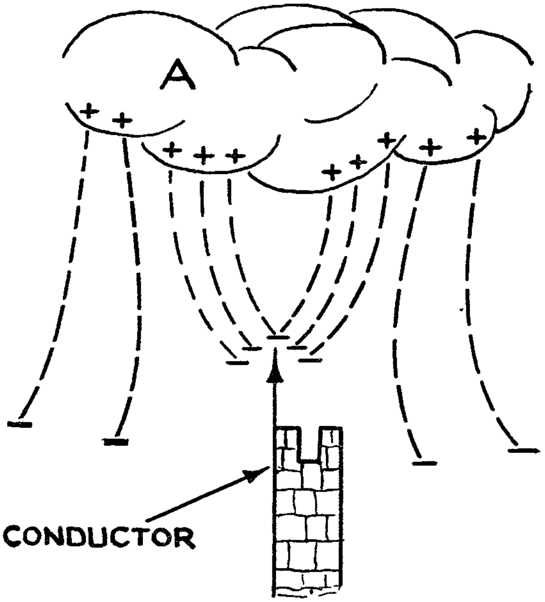


FIG. 4.6. Lightning conductor

Potential

In studying the electrostatic field, we make use of a scheme of ideas already described in Chapter III, involving the conception of potential. In current electricity the potential difference in volts between two points in a circuit is merely the difference in the energy (joules) associated with 1 coulomb as it passes from one point to the other; but it is important to remember that in this case the potentials at the various points are maintained constant by the batteries or generators, and are unaffected by the movements of electricity.

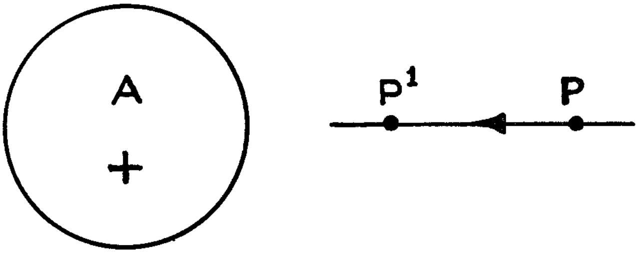


FIG. 5.6. To illustrate electrostatic potential at a point

Consider the field round the positively charged sphere fig. 5.6.

A small positive charge at P is repelled by the charge in A, and in order to move it up to P^1 work will have to be done. It would appear, therefore, that P^1 is at a higher potential than P, but this makes the very big assumption that the field of A and the potential of the various points in it *have been unaffected by the new charge* we have pushed from P to P^1; there is only one way to ensure this; the exploring charges which we push about must be so exceedingly small that they do not disturb the original field. If P and P^1 were points on a circuit fed by a battery, then without any reservation, the potential difference between P and P^1 is the energy lost or gained by one coulomb passing from one point to another, but if as we supposed earlier, P and P^1 are points in the electrostatic field due to A, then the placing of a coulomb at P, not to mention its transference to P¹, would so upset the field that we should no longer be dealing with the original field due to A alone, but to the field due to A and our exploring coulomb combined.

In estimating the potential difference between two such points as P and P¹, therefore, we will suppose that we move an exceedingly small charge (say a micro-micro coulomb) from one to the other, find the work done, and multiply the answer by 10^{12} and whenever we speak of the " transfer of a coulomb " in an electrostatic field this is the process implied. We therefore define the P.D. between two points in an electrostatic field thus :—

> " The potential difference in volts between two points in an electric field, is the work done in joules in transferring a positive coulomb from one point to the other *assuming no modification of the original field.*"

Earth connected points are by agreement at zero potential, so that the potential at a point, in volts, in an electrostatic field will be :—

> " the work done in joules in transferring a positive coulomb from earth to that point ". These ideas are difficult ; let us try to represent them diagrammatically.

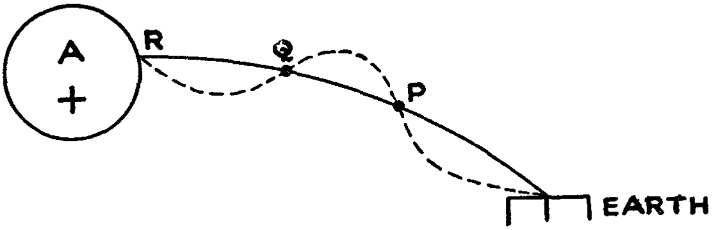


FIG. 6.6. To illustrate electrostatic potential of a conductor

A, fig. 6.6, is a positively charged conductor ; the work done by the transfer of a positive coulomb from earth to P is " the potential at P " ; the work done in transferring it from P to Q is " the potential difference between P and Q " and because work is done *on* the coulomb, Q has the higher potential.

The work done in transferring the coulomb from earth to R (a point on the surface of A) is " the potential of A", because since the surface of A is a conductor, no difference of potential can persist on it, and the potential of the whole surface is the same as that of any point on it.

The path taken by the exploring charge is quite immaterial ; the work done at any stage depends merely on the terminal points and not on the path ; the transfer could have taken place along either the full line or the dotted line ; the work done would be the same.

All this discussion about the transfer of coulombs is necessary for the understanding of the subject, but the mathematicians have shown us that the potential at points in the field of a point charge can be worked out quite simply by arithmetic ; in fact, the potential in volts at any selected point is simply :—

$$\frac{\text{Value of the point charge in coulombs} \times 9 \times 10^{11}}{\text{Distance from the point charge in cms.}} = \frac{Q \times 9 \times 10^{11}}{d}$$

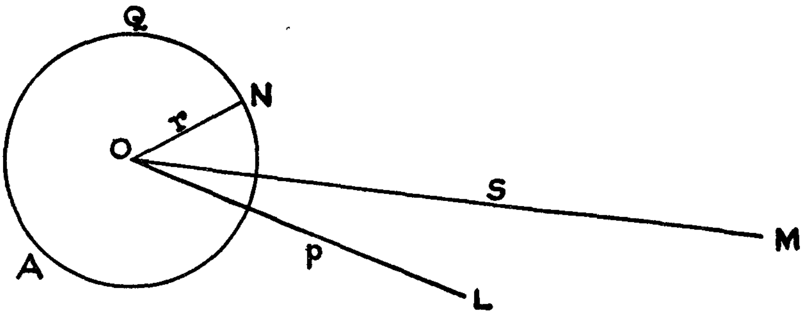


FIG. 7.6. Potential due to charged sphere

If the sphere A of radius r cm. (fig. 7.6) has a charge of Q coulombs on its surface, it behaves (see p. 53) as if it were concentrated at its centre O, and the potential at N on its surface (" the potential of A ") is :—

$$\frac{Q \times 9 \times 10^{11}}{r} \text{ volts.}$$

The potentials in volts at L and M are respectively

$$\frac{Q \times 9 \times 10^{11}}{p} \text{ and } \frac{Q \times 9 \times 10^{11}}{s} \text{ and so on.}$$

Capacity

In the previous paragraph we saw that the potential of a sphere (V) was related to the charge Q thus :—

$$V = \frac{Q \times 9 \times 10^{11}}{r}$$

or by transposition

$$Q = \frac{r}{9 \times 10^{11}} . V$$

The factor $r/9 \times 10^{11}$ is a constant so far as the sphere is concerned, and, by taking $V = 1$, this constant is clearly equal to

> " The quantity of charge (coulombs) which must be given to the sphere to raise its potential by 1 volt".

It is called the CAPACITY of the sphere and is expressed in farads ; thus a sphere of radius 9×10^{11} cm. (about 5½ million miles) would have a capacity of 1 farad, which is rather amazing ; the Earth considered as a sphere of radius $6{\cdot}4^6 \times 10$ metres would have a capacity of

$$\frac{6{\cdot}4 \times 10^8}{9 \times 10^{11}} = 7{\cdot}1 \times 10^{-4} \text{ farads.}$$

In practice we employ the microfarad, one millionth part of a farad, as the unit of capacity and the capacity of the earth would thus be about 710 microfarads.

It does not matter what shape a conductor has, a similar relation holds between its charge and its potential, namely :—

$$Q = CV$$

C standing for the capacity of the conductor in farads, Q its charge in coulombs and V its potential in volts. As we depart from the simple case of the sphere to more irregular shapes, the expression for C in terms of dimensions gets more and more complex.

We can regard C as some physical property of the conductor ; suppose for a moment we had a conductor of zero capacity, then because

$$Q = CV \quad \text{or} \quad V = \frac{Q}{C}$$

the provision of even the smallest charge element would raise its potential to an infinite value,

$$\text{for } V = \frac{Q}{0} = \text{infinity, no matter how small } Q \text{ is.}$$

This implies that infinite energy would be required to add the next element of charge and in consequence the charging process could not go on. Such a conductor, therefore, *could not acquire a charge,* and by converse reasoning we may say that the capacity of a conductor is its *ability to acquire a charge of electricity.*

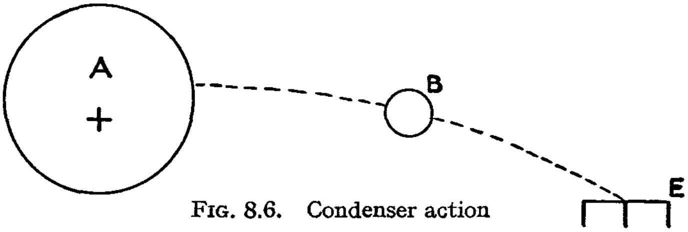


FIG. 8.6. Condenser action

The condenser

A, fig. 8.6, is a positively charged conductor, B is uncharged and insulated. In transferring a positive charge from the earthed point E to B, we shall have to do work in overcoming the repulsive forces arising from A; therefore B, although uncharged, has a definite potential. Potential acquired by reason of a neighbouring charge is called "induced potential" to distinguish it from the "free potential" due to a charge on the body itself. Continuing our transfer of positive charge from B to A, clearly more work has to be done—which means there is a potential difference between A and B. The nearer B gets to A the less the potential difference becomes until when they touch, the potentials of A and B become equal and the P.D. disappears. A system of this sort, 2 conductors separated by an insulating medium is called a *condenser*. In wireless work we normally use a condenser consisting fundamentally of 2 parallel plates A and B fig. 9.6.

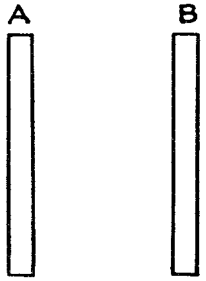


FIG. 9.6. Simple parallel plate condenser

If A is charged, a P.D. is developed between A and B; again the mathematicians have helped us by working out a formula for the potential difference; if the area of overlap of the plates is A sq. cm. and the distance between them is d cm., then with air as the separating medium,

$$V = \frac{4\pi d \times 9 \times 10^{11}}{A} . Q$$

or transposing,

$$Q = \left[\frac{A}{4\pi d \times 9 \times 10^{11}} \right] . V$$

The expression in brackets is a constant as far as the two plates are concerned, and we can denote it by "C" and we get

$$Q = CV$$

which is very much like the equation on page 56 relating the charge and potential of a single conductor, but Q in this latter equation means the charge on *one* plate and V the *potential difference* between them.

Taking V = 1, we obtain Q = C; that is, C denotes the charge in coulombs necessary to raise the P.D. between the plates by 1 volt; C is called the *capacity of the condenser*, is expressed in farads and represents the ability of the condenser to acquire a charge.

This inter-conductor capacity is very much greater, size for size, than the single conductor capacity discussed earlier. An isolated sphere would have to have a radius of 9×10^5 cms. (about 5½ miles) to provide 1 microfarad of capacity, but wireless condensers of several microfarads constructed on the parallel plate system, are small enough to go into the waistcoat pocket.

Faraday discovered that the medium between the plates of a condenser had a great effect on its capacity. Taking the capacity of an air condenser as standard, it was found that with sulphur as the separating medium the capacity was multiplied by about 4; oil multiplies the capacity by about 2·5, mica by 6. The numbers are called the *Specific Inductive Capacity* or the *Dielectric constant*, of the substance (generally denoted by K). The formula for our parallel plate condenser can now be amended to apply to cases where the separating medium has a dielectric constant K.

$$C = \frac{K\,A}{4\pi d \times 9 \times 10^{11}} \text{ farads.}$$

Multiplate condensers are almost universally used in wireless; instead of using a pair of plates of very big area to obtain a large capacity, several pairs of plates are joined together, each pair contributing capacity in accordance with the above formula.

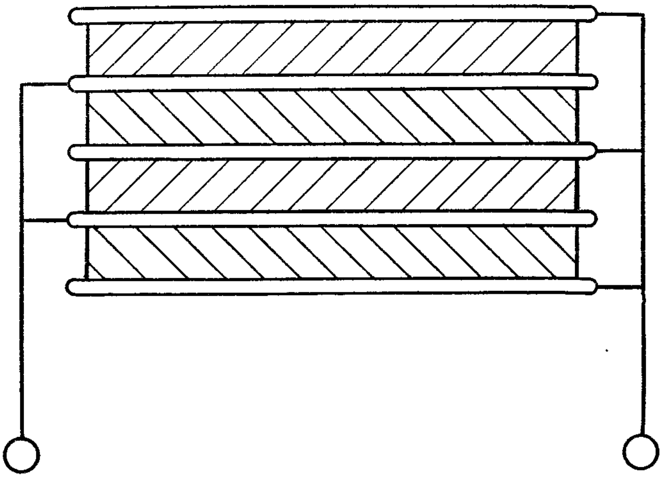


FIG. 10.6. Multiplate condenser

In fig. 10.6 a sketch of such a condenser is shown ; its capacity will be 4 times that of one of its pairs of plates ; in general if there are N dielectrics, the capacity is N times that of one pair of plates, and the final formula for the capacity of a multiplate condenser is

$$C = \frac{K\,A\,N}{4\pi d \times 9 \times 10^5} \text{ microfarads.}$$

where N = number of dielectrics.

Variable condensers, i.e. condensers of which the capacity can be altered at will, usually work on the principle of altering the area of overlap of the plates ; fig. 11.6 shows a variable air condenser ; as the spindle carrying the moving plates is turned, the area of overlap is changed and the capacity varies, being proportional to the angle through which the spindle is rotated. When the movable vanes are entirely " out of mesh " with the fixed ones, and the area of overlap is zero, we might expect the capacity to be zero ; this is not quite true because even in this position, we still have two conductors separated by a dielectric forming a condenser of very small capacity.

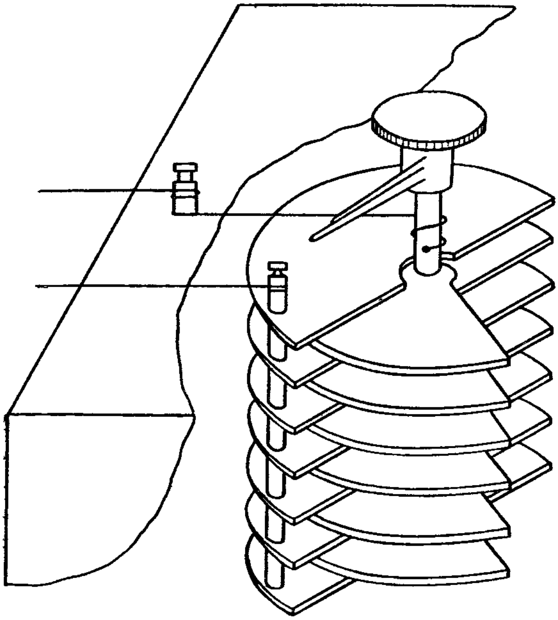


FIG. 11.6. Variable air condenser

Fixed condensers are usually made of sheets of tinfoil separated with paraffined paper, or in more expensive types, with mica. In this form of construction a large capacity can be developed in a small space.

Condenser dielectrics must have a high *dielectric strength* in addition to good insulating properties. This means that they should be able to withstand high potentials without a spark passing and consequent failure of insulation. The reader should consult the pamphlets of the various condenser manufacturers for practical constructional details.

Grouping condensers

Like cells and resistances, condensers can be grouped in series or parallel.

For a *parallel* grouping the capacities add up [fig. 12.6 (a)]. For a series grouping we have reciprocal formula, like the one for resistances in parallel [fig. 12.6 (b)] so that if 2 condensers are put in series the resultant capacity is less than that of either component.

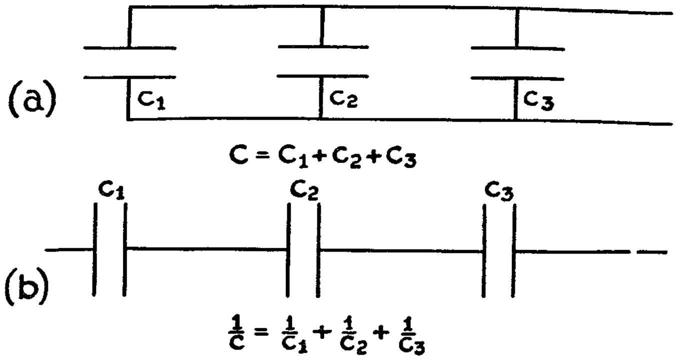


FIG. 12.6. Grouping of condensers

Field energy

Having reviewed some of the features of both the magnetic, (Chapter III), and the electrostatic field, we can now discuss a feature which they have in common : both are seats of energy.

This is really implied when we give the explanation of the phenomena in terms of elastic lines of force ; a piece of stretched elastic contains energy, and it will be only reasonable to imagine that our lines of force contain energy also. Take the electric field first, because the problem is somewhat easier to handle than in the magnetic case.

A conductor of capacity C is to be charged with Q units of electricity ; the charge shall be brought up in instalments $\frac{Q}{n}$, each so small that its transfer (from an earthed point) does not seriously upset the field ; the first instalment will not require work to be done on it, because A is without charge and exerts no repulsive action. The last instalment will have to be forced up against the full charge Q, and the work done will be, charge element × final potential of A

$$= \frac{Q}{n} \times \frac{Q}{C}$$

so that the average work per excursion (because charge and potential are directly proportional) is :—

$$\frac{1}{2} \cdot \frac{Q}{n} \cdot \frac{Q}{C}$$

and since the number of excursions is n

$$\therefore \text{ the total work} = \frac{1}{2} \cdot \frac{Q}{n} \cdot \frac{Q}{C} \times n = \frac{1}{2} \cdot \frac{Q^2}{C} = \frac{1}{2} \cdot \frac{(CV)^2}{C} = \frac{1}{2} CV^2 \text{ joules.}$$

Where is the energy ? The obvious answer is that it is the energy in the stretched lines of force or in short, this is the energy in the electric field of the charged conductor.

Similar reasoning applies to a charge condenser ; the strained dielectric is a seat of energy equal to

$$\frac{1}{2} CV^2 \text{ joules.}$$

If the plates of the condenser were shorted, this energy would be transformed into heat, generally in the form of a spark.

The magnetic case cannot be deduced so simply as this, but it is clear that when a circuit containing a battery is closed, magnetic flux begins to grow and an induced E.M.F. opposing the battery is set up. The battery has to supply energy to overcome this counter E.M.F. and this energy goes out to form the magnetic field and as before we can regard it as being responsible for the tension and pressure of the lines of force. When the circuit is broken, this field energy returns to the circuit, through the mechanism of self induction, and is dissipated, usually as heat, often again in the form of a spark.

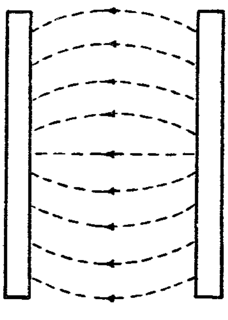


FIG. 13.6. Field between charged parallel plates

If L is the self inductance of the circuit in henries, and *i* the current in amperes the field energy is $\frac{1}{2}$ Li^2 joules. These considerations afford us another sidelight on the physical nature of capacity and of inductance. The capacity of a condenser is its ability to store energy in the form of an electrostatic field, and the inductance of a circuit represents its ability to store energy in the form of a magnetic field around it.

Charge and discharge of a condenser

So far we have spoken of charging a condenser without any very exact statement as to how it should be done; it can be done by holding one plate in the hand (thereby earthing it) and conveying charge to the other from a static electrical machine or even from a rod or ebonite rubbed with dry cat-skin. A better method, however, is to connect the plates to the terminals of a battery; a momentary charging current flows, dragging free electricity off one plate, through the battery, and piling it up on the other plate; this process will go on until the P.D. developed across the condenser by virtue of the charge it has acquired, is just equal to the E.M.F. of the battery.

A very elegant experiment can be performed to show this process, by slowing it down so much that it can be followed and observations taken which allow us to plot graphs about it.

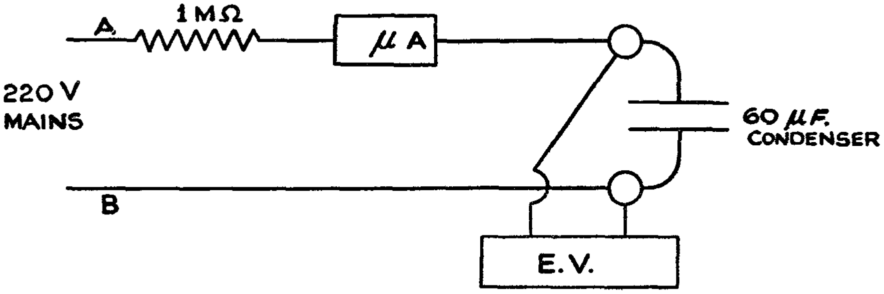


FIG. 14.6. Experiment on slow charge of condenser

A 60μF condenser is connected as shown in fig. 14.6, with an electrostatic voltmeter across its terminals, and to the mains via a 0–300 microammeter and a megohm resistance.

Switching on, the charging process begins; the initial value of the charging current will be

$$\frac{220}{10^6} \text{ amps.} = 220 \text{ microamperes}$$

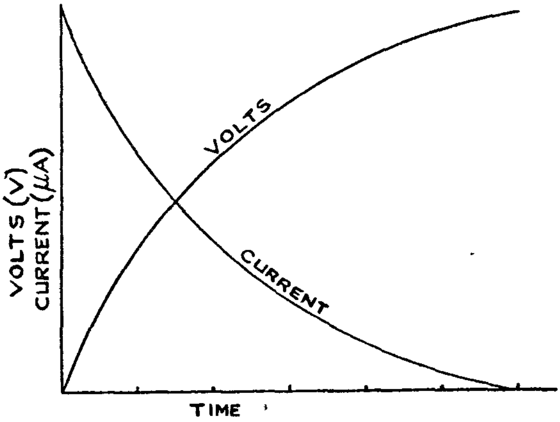


FIG. 15.6. Charging graphs

which is *the same current as would flow if the condenser were shorted,* but it will fall off as the charging goes on. At the same time the voltage across the condenser will mount up, until finally it reads 220 volts.

Current and voltage readings can be taken every half minute and plotted on a graph which will be something like that shown in fig. 15.6.

Although the operation has been slowed down so that we can observe it, a similar process takes place when the condenser is joined directly across the mains, but instead of taking 5 minutes to complete the process is all over in an instant;

naturally the bigger the condenser, the longer it takes to change so that the time taken is governed jointly by the resistance in the circuit and the capacity of the condenser; in fact the product

$$C \times R$$

is called the " time constant " of the circuit and taken as a measure of the time necessary to charge (or discharge) the condenser through a resistance; actually CR (seconds) is the time to complete 2/3 of the full charge or discharge. If the microammeter in fig. 14.6 is a centre-zero instrument, the mains can be disconnected and a shorting wire connected across AB; the condenser now discharges, and again we can read the current and voltage every half minute; the time graph of both current and voltage now follows the " falling " shape of the current curve in fig. 15.6.

The discharge of a condenser assumes a character of very great importance to wireless engineers, when it is allowed to discharge, not through a high resistance as described above, but through an inductive coil with as little resistance as possible.

In this case the discharge is of an oscillatory nature, the charge swings backwards and forwards from one plate of the condenser to the other, and would continue to do so indefinitely if it were not for certain unavoidable energy losses.

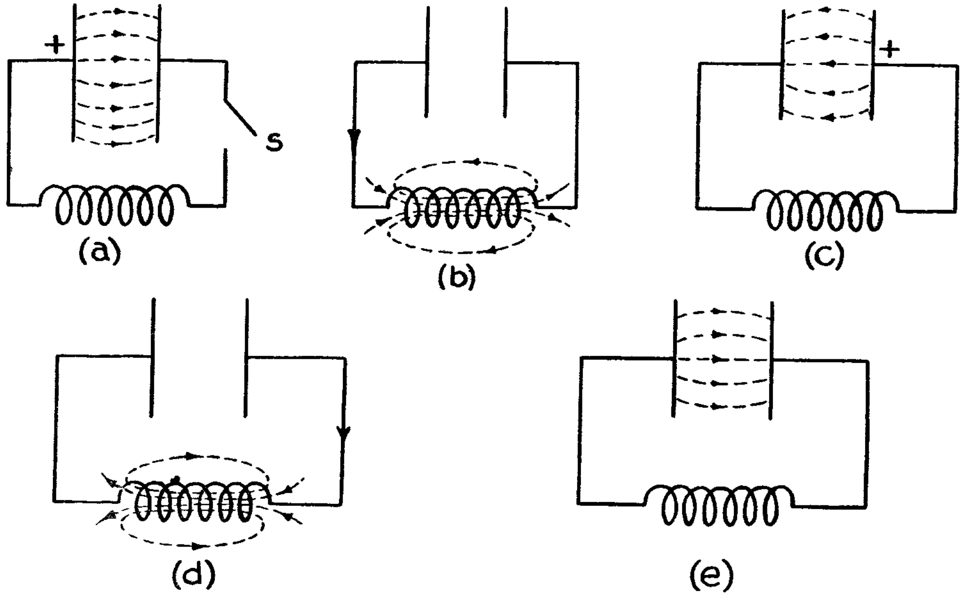


FIG. 16.6. To illustrate the action of an oscillatory discharge

The sketches in fig. 16.6 represent what happens when the discharge is oscillatory in character; (*a*) shows the condenser charged, with its electrostatic field; when the switch S is closed, discharge commences and a magnetic flux is built up in and around the coil, and when this flux is at its maximum, the condenser is completely discharged (*b*). It may seem strange that when the driving voltage is nil the current is a maximum; that is one of the curious things about this type of discharge, and it must be accepted as a fact, unless one is prepared to enter upon some fairly difficult mathematics.

At stage (*b*) then, we have maximum current and no voltage to keep it going, so the current begins to fall, and with it the flux, but owing to the inductance of the coil, the fall of current is opposed, and the electricity is kept on the move by the induced E.M.F. till it piles up on the other plate of the condenser and stage (*c*) is reached, in which the condenser is recharged in the opposite sense. The same process now repeats itself, passing through the magnetic field (*d*) and back to the electric field (*e*), which is just the same as we started with.

In theory, this cycle of operations should continue indefinitely, but of course energy losses are unavoidable. The energy in the original condenser charge was $\frac{1}{2}$ CV^2 and in the first magnetic field (*b*) $\frac{1}{2}$ Li^2, and again in theory, these two should be the same, i.e.

$$\frac{1}{2}\,CV^2 = \frac{1}{2}\,Li^2$$

$$i = V\sqrt{\frac{C}{L}}$$

where i is the maximum current and V the original voltage, but owing to resistance in the wire some energy is lost as heat ; other energy losses we shall discuss later, but the net result of them all is that these " peak energies " get progressively less and less ; the oscillations, we say, become *damped* and die away just like the oscillations of a pendulum, unless we inject a little energy at every swing to compensate for these losses.

It is instructive to represent these oscillations on a time graph. The horizontal scale fig. 17.6 represents time, the vertical scale represents current (full line) or voltage (chain line) upward measurements indicate that the left-hand side of the condenser is positive, or that the current is anti-clockwise. Downward measurements have the opposite meaning.

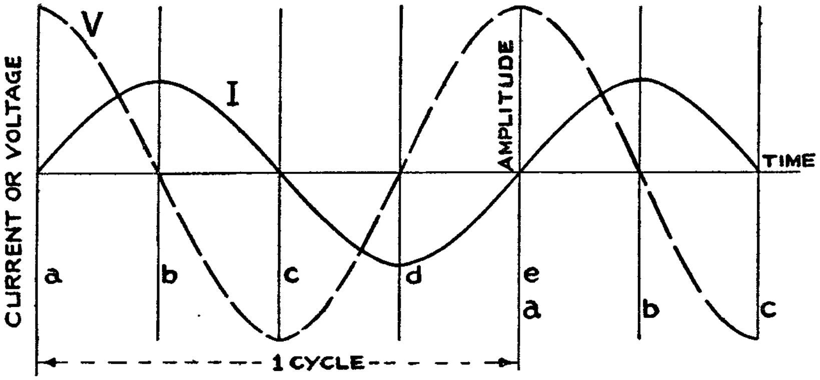


FIG. 17.6. One cycle of undamped oscillations

The curves of fig. 17.6 represent one and a half cycles of undamped oscillation ; the curves of fig. 18.6 indicate how these oscillations fall in amplitude, owing to the various energy losses ; they are graphs of a damped oscillation.

The letters *a*, *b*, etc., on the graphs serve to relate them to the sketches of fig. 16.6, thus sketch (*c*) fig. 16.6 corresponds with the state (or *phase*) represented at *c* on the graphs ; the current is zero, and the voltage is at a reverse maximum.

Notice that the " states " or " phases " of the current and the voltage do not coincide ; when one is a maximum the other is zero ; we say they are " out of phase ". Further, if we divide the cycle up into 360 equal time intervals (which we shall call degrees), we see that corresponding phases (say the zeros) of current and voltage are 90° apart ; this is expressed by saying that the voltage and the current are 90° out of phase *with the voltage leading* because starting from the origin 0, the voltage curve is the first to reach its maximum.

The oscillations in a circuit of this type take place very rapidly; the FREQUENCY (*f*) of the oscillation (cycles per second) is given by the formula

$$f = \frac{1}{2\pi\sqrt{LC}} \text{ cycles/sec.}$$

where L represents the inductance of the coil in henries and C the capacity of the condenser in farads.

To take a simple case; if L = 1,600 micro-henries and C = ·04 micro-farads

$$f = \frac{10^6}{2\pi\sqrt{64}}$$

$$= 19{,}150 \text{ cycles/sec.}$$

$$\text{or} = 19{\cdot}15 \text{ kc/s. (kilocycles per second).}$$

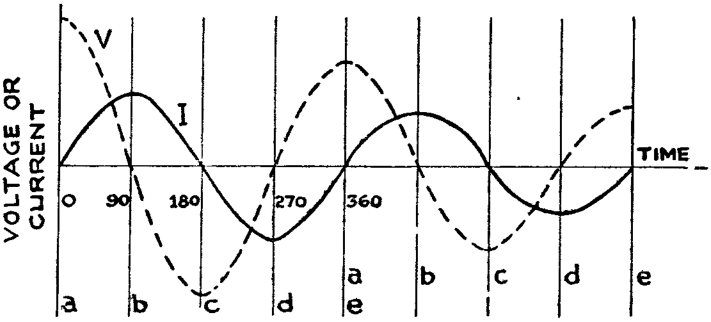


FIG. 18.6. Damped oscillations

Open oscillatory circuit—radiation

The type of oscillatory circuit we have described is known as " closed ". Its fields are concentrated and at every oscillation all the energy associated with a field returns to the circuit. Some of it is lost in the circuit of course, but none is wasted in the fields.

If, however, we open out the plates of the condenser so that the lines of force of the electric field are diffused over a wide space, the circuit becomes an " open " oscillatory circuit and a new effect occurs which is of very great importance to us; *the circuit emits electromagnetic radiation.*

Just as a vibrating tuning fork emits waves of sound in the air, so an open oscillatory circuit emits wireless waves in the ether of space. The generation of such waves means an outpouring of energy, and constitutes another damping factor, but whereas we try to avoid other damping factors, we encourage this one, because, using the open oscillatory circuit as the basis of a wireless transmitter, we want to throw out as much energy as possible in this way. At the same time, of course, we have to devise methods of injecting energy into the circuit as fast as it loses it, so that the oscillations do not die away.

Electrostatic voltmeter

The strained electro-lines between the plates of a charged condenser, set up a mechanical force tending to draw the plates together; this effect is employed in the electrostatic voltmeter, where the voltage applied to two plates of a condenser is measured in terms of the forces so generated.

Figs. 19.6 (a) and (b) illustrate two ways in which this may be done. In (a) we have a single movable plate A, forming one pole of the instrument, and a fixed plate B forming the other pole. The electrostatic forces set up tend to rotate A in a clockwise direction about the spindle C, thus moving the pointer, which is attached to A, across the scale. The restoring force is often provided by gravity, but sometimes a weak spring control is used. The scale is cramped at the beginning of its range. In fig. 19.6 (b)

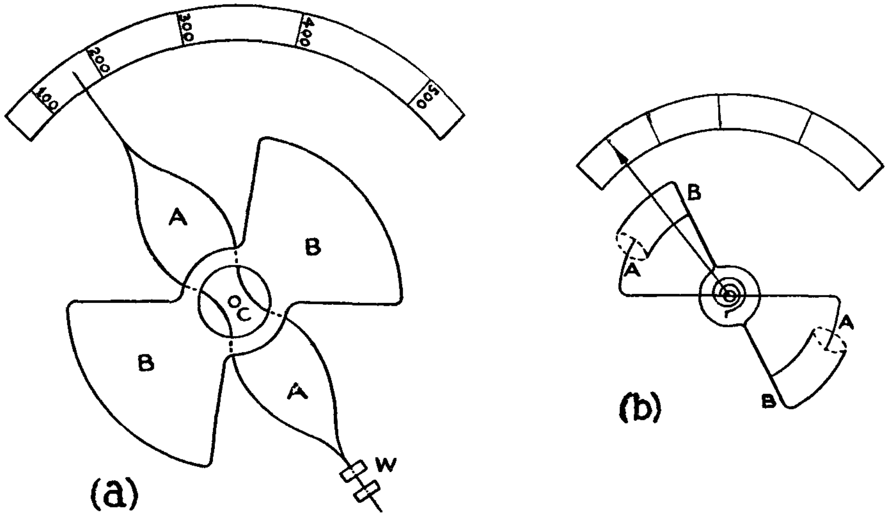


FIG. 19.6. Principle of electrostatic voltmeter

another arrangement is indicated in which the ends of the movable member (A) are drawn by the electrostatic field into the cylindrical conductors (B) at the end of the fixed member against a weak spring control.

Electrostatic instruments such as these are very insensitive for low voltages ; their great advantage is that they absorb no power.

CHAPTER VII

ALTERNATING CURRENTS (1)

In Chapter V we described a simple A.C. generator consisting of a rectangular loop or wire, turning about an axis in a magnetic field, and furnished with slip rings and brushes; we saw that the E.M.F. generated by the rise and fall of flux through the coil was directed one way round the loop for half a revolution and the other way for the succeeding half revolution.

If we plot a graph of this E.M.F. on a time basis we get a curve like that shown in fig. 1.7.

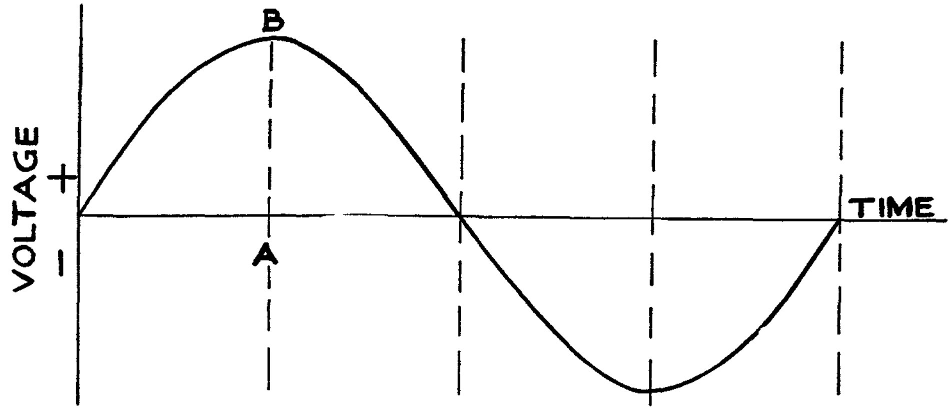


FIG. 1.7. Sinusoidal E.M.F.

The shape of this graph, like those drawn in the previous chapter to illustrate the behaviour of the oscillatory circuit, is not an accident, it is a curve well known to mathematicians, who call it a " sine curve ". An E.M.F. or current whose time graph is of this shape is called a " sinusoidal " E.M.F. or current and for the sake of simplicity we shall assume that the alternating electrical quantities with which we have to deal are truly sinusoidal in character.

Let us define our terms. The *amplitude* or *peak value* of the A.Q. (alternating quantity) is the maximum value it attains in the *cycle*, which is one complete series of values. The graph fig. 1.7 represents a cycle and the length of the line AB is the amplitude.

The number of cycles which occur in a second is called the " frequency ", while the time in seconds in which one cycle takes place is known as the " time period ".

Suppose we joined a centre zero moving coil voltmeter to the terminals of an A.C. generator, and ran it, just at a very slow speed—say one revolution per second—and then gradually increased the speed till the frequency rose to 50 cycles/second. At first the needle of the voltmeter would " follow " the voltage, swinging above and below the zero in time with the rise and fall of the voltage ; as the speed increased, the voltmeter would cease to follow and soon the needle would be completely unaffected. A moving coil instrument is therefore useless for A.C. work. Let us try a moving iron instrument. As the speed of the generator increases this instrument settles down to a steady deflection but its indication must be an average of some kind ; it is in fact what we call the Root Mean Square or R.M.S. value of the alternating voltage, which in the case of sinusoidal quantities is :—

$$\frac{\text{Peak value}}{\sqrt{2}} = \text{peak value} \times \cdot 707.$$

It is this value which we refer to when we speak of A.Qs. unless otherwise stated ; for example, if we describe a commercial supply as 220 V 50 c/s A.C. we mean that the R.M.S. value of the supply voltage is 220, and in consequence the peak value is

$$220 \times 1 \cdot 414 = 311 \cdot 08 \text{ volts.}$$

The *phase difference* between two A.Qs. is the peak separation in degrees, as we saw on p. 62.

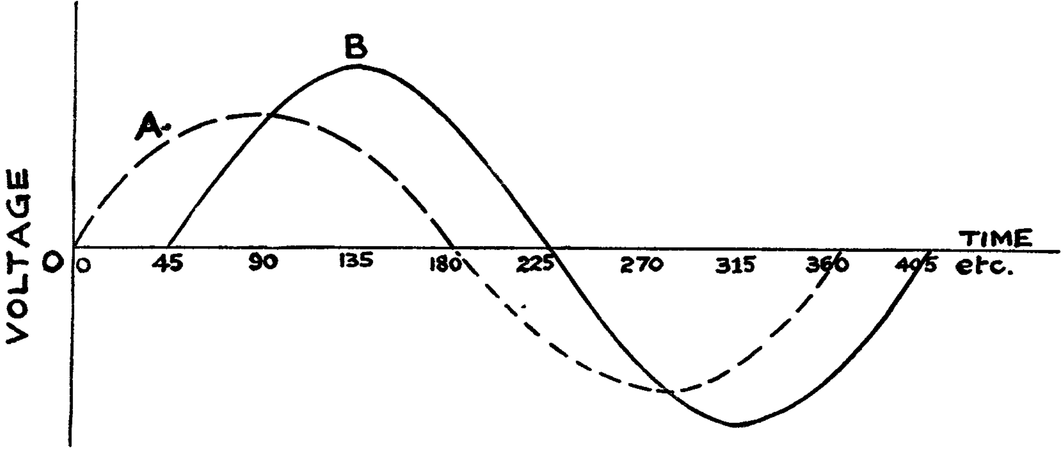


FIG. 2.7. Illustrating phase difference

A and B (fig. 2.7) represent the time graphs of the voltage of two generators ; they are out of phase by one eighth of a cycle, that is by 45°, and A is leading (because, proceeding along the time axis from O, A reaches its maximum first). Suppose these two generators (A 200 V. R.M.S., B 300 V. R.M.S.) are connected in series ; what would the terminal voltage of the combination be ?

With two cells in series it would be easy ; we should simply add the voltages of the separate cells, but with alternating voltages it is not quite so simple because *we must take the phase difference into account.*

We do this by a simple geometrical method ; draw a line OB (the lagging component) representing 300 volts (fig. 3.7) ; three inches will do, then the scale is 1 in. = 100 V. R.M.S. Now from O draw a line inclined to OB by an angle equal to the phase difference, and mark off on it a distance OA representing the other (the leading) voltage. Complete the parallelogram OACB, then the line OC (on the same scale of voltage) represents the voltage of the two generators in series ; it can be scaled off and converted to volts directly. Observe the difference the phase separation makes to this answer. If they were in phase the answer would have been 200 + 300 = 500 V. ; if they were in antiphase (180° out) the answer would have been 300 − 200 = 100 V. ; any other phase difference would have given an answer between these two extremes. Further we can learn from our diagram that the resultant voltage *leads* on the one represented by OB, by the angle BOC. It also lags behind the voltage represented by OA, by the angle COA.

The lines OB and OA which we have drawn are called the " vectors " of the two voltages and the figure of fig. 3.7 is called the " vector diagram " of the arrangement.

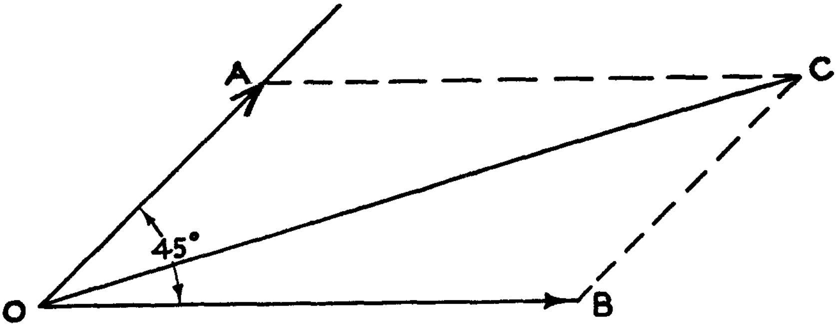


FIG. 3.7. Vector addition

Circumstances may arise in which we have to represent perhaps a dozen A.Qs. in a diagram at the same time ; each line must have a length representing its A.Q. on the agreed scales and the angles between the lines must represent their phase differences, remembering the convention that *leading* quantities have their vectors further round in an *anticlockwise* direction. It does not matter in what direction on the paper the first vector is drawn ; it is essential however that all the others should be drawn at the proper angles with respect to the first one.

One of the curious things about a circuit in which alternating voltage is operating, is that the voltage and the current may be out of phase ; in fact, unless we make special provision they usually are out of phase. An example occurred in the discussion of the oscillatory circuit ; the current and the voltage are 90° out of phase, with the result that when the voltage is zero the current is a maximum.

Fig. 4.7 shows a vector diagram corresponding to the time graph of fig. 17.6.

The series A.C. circuit

Following the same general lines as in Chapter II we must now develop the voltage-current relations which in A.C. theory correspond to Ohm's Law in D.C. work. Alternating quantities are difficult to think about, because they are always on the change, rising, falling, reversing and so on, but we can forget all this if we deal always in R.M.S. values, represent the quantities by their appropriate vectors, and remember that all operations which in D.C. would be pure arithmetical addition and subtraction, must be carried out by the geometrical construction described above, called the " vector parallelogram".

First of all, suppose we had a circuit consisting of resistance and nothing else; the current at every instant would be such that the volts drop (RI) was just equal to the applied voltage, and further this voltage and the current would be exactly in phase. Expressing this as a vector diagram we have just two coincident lines which in fig. 5.7 have been separated slightly for the sake of clearness.

Secondly, let us take a circuit having inductance only; it might be objected that this is impossible, but by winding a solenoid of sufficiently thick wire, we are able to make the resistance negligibly small, and that is good enough. As the current alternates in this circuit self-induction will come into play and voltage will be induced, proportional at every instant to the rate of change of current; the current amplitude will be such that the instantaneous counter voltage is equal to the applied voltage.

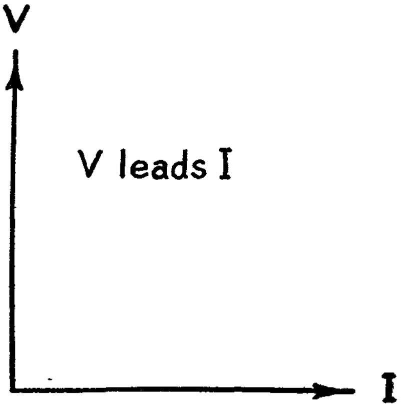


FIG. 4.7. Vector diagram; 90° phase difference

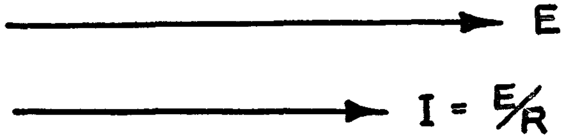


FIG. 5.7. Vector diagram of circuit of pure resistance

The actual value of the induced voltage is

$$2\pi f L I \text{ volts,}$$

where f is the frequency of the supply,
L is the inductance of the circuit in henries, and
I is the current in R.M.S. amps.

The applied voltage (E) is equal and opposite to this, and leads the current by 90°.

The vector diagram for such a circuit is shown in fig. 6.7.

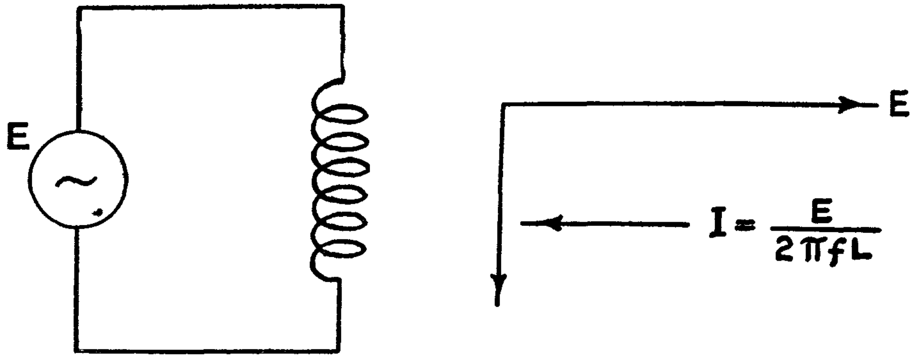


FIG. 6.7. Circuit of pure inductance

A more common way of expressing the phase relation in a purely inductive circuit is to say that the "current lags behind the applied voltage" by 90°, which is of course the same thing. We usually denote the expression $2\pi f$ by the symbol ω, and we therefore have

$$I = \frac{E}{\omega L}$$

which is something like Ohm's Law, where the product ωL replaces the resistance; ωL in fact expresses the opposition arising in an A.C. circuit due to inductance and is called the *Inductive Reactance* of the circuit and is expressed in ohms.

Notice that this opposition is proportional to the frequency; the higher the frequency, the more difficult it is to drive alternating current through an inductance. An inductive circuit is called a " choke ", and it follows on this account that a coil of low inductance will have as much choking effect in a high frequency circuit, as one of high inductance in a low frequency circuit.

Thirdly, let us consider a circuit consisting only of a condenser. Here the equilibrium state is represented by saying that the charge on the condenser will vary in such a way that the voltage it produces across the condenser is always equal and opposite to the supply voltage. The value of the voltage set up across the condenser is

$$\frac{I}{2\pi f C} = \frac{I}{\omega C} \text{ volts (R.M.S.)}$$

where I is the current (R.M.S. amps.) determined by the alternating charge, and C is the capacity in farads.

The current leads on the applied voltage E by 90°.

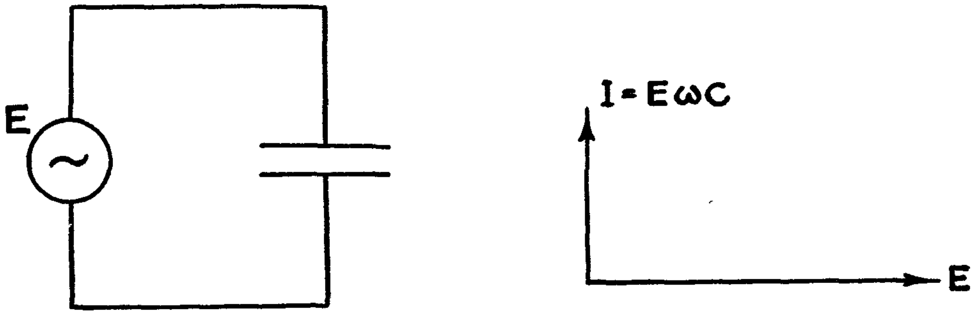


FIG. 7.7. Circuit of pure capacity

The usual way of expressing the phasing here is to say that in condensers the current leads the applied voltage E by 90°. (Fig. 7.7).

Again, we can write

$$I = \frac{E}{\frac{1}{\omega C}} = E\omega C$$

and observe that the opposition due to [capacity capacity reactance (ohms)] is expressed as $\frac{1}{\omega C}$; it varies inversely with both the capacity and the frequency ; the higher the frequency the easier does A.C. flow in a condenser. Further, a small condenser at high frequency will provide as easy a passage to A.C. as a large condenser at low frequency.

Lastly, let us suppose we have a circuit with all three of these things, R, L and C in series. Observe that as in a D.C. series circuit, the current will be the same (i.e. have the same magnitude *and be in the same phase*) all through the circuit. Fig. 8.7 shows the circuit with the various volts drops represented by vectors above, and the current by vectors below.

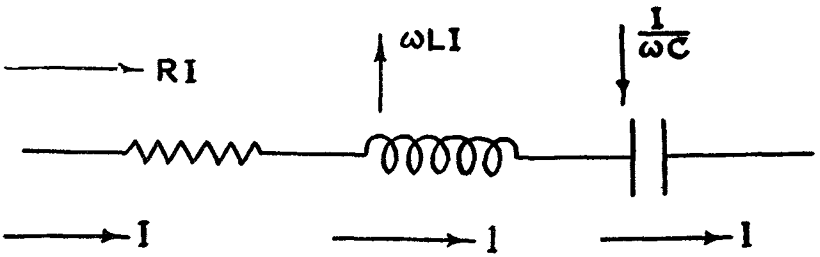


FIG. 8.7. The general circuit

The applied voltage will be the sum (in the geometrical sense) of these three volts drops. Assuming $\omega L\,I$ to be greater than $I/\omega\,C$, the three voltages to be compounded are shown in the vector diagram of fig. 9.7(a).

Fig. 9.7(b) shows the first stage in compounding the vectors ; ωLI and $I/\omega C$ being in the same straight line their resultant is equal to their difference acting along the greater of the two. Fig. 9.7(c) shows the final stage ; OB is the resultant of the three drops and therefore represents the applied voltage.

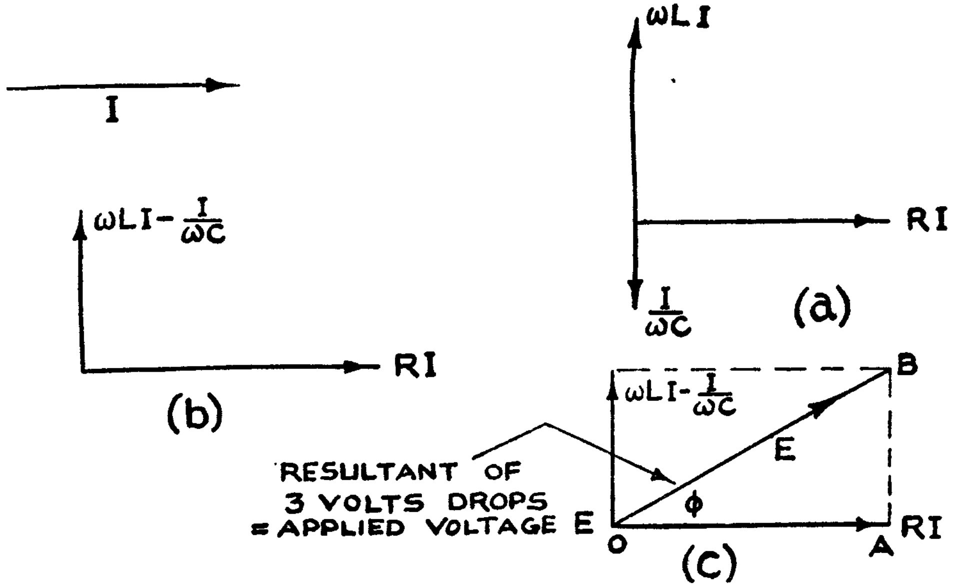


FIG. 9.7. Vector addition of volts drops

The angle AOB (ϕ) is the phase difference between the current and the voltage in the circuit ; here the current lags because there is more inductive than capacity reactance ; had the opposite been the case, there would have been a leading current.

The triangle AOB is a right-angled triangle,

$$\therefore\ OB^2 = BA^2 + OA^2$$

i.e.

$$E^2 = \left(\omega LI - \frac{I}{\omega\,C}\right)^2 + (RI)^2$$

$$E^2 = I^2\left[\left(\omega L - \frac{1}{\omega C}\right)^2 + R^2\right]$$

or

$$I = \frac{E}{\sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}} \quad \cdots\cdots \quad (1)$$

Further, the phase angle ϕ is such that

$$\tan\phi = \frac{\omega L - \frac{1}{\omega C}}{R} \quad \cdots\cdots \quad (2)$$

These complicated expressions 1 and 2 replace our simple Ohm's Law of D.C. ; notice that the total reactance of a circuit is the difference between the inductive and capacity reactances ; the symbol X is used for the total reactance, while X_L and X_C stand for inductive and capacity reactance respectively. The denominator in expression (1) represents the total opposition in the circuit, both resistive and reactive, and it is called the *impedance* of the circuit; it is measured in ohms and is usually denoted by Z. We can write therefore for A.C. circuits :—

$$I = \frac{E}{Z}$$

$$\text{where Z, the impedance,} = \sqrt{R^2 + X^2}$$

$$\text{and also } \tan \phi = \frac{X}{R}.$$

To find the R.M.S. current in an A.C. circuit then, all we have to do is to divide the R.M.S. volts by the impedance; we can work out the value of the impedance from the constants of the circuit if it is not given in ohms directly.

It must not be forgotten that an A.C. problem is never complete unless the phasing is worked out; to say the current in a circuit is 5 amps. is incomplete; the statement must also tell us whether this 5 amps. lags or leads with respect to the voltage, and by how much.

Notice that our method of compounding A.Q.s may result in the *sum* of two of them being less than either, which seems unreasonable and an affront to our arithmetical sense, but the point to remember is that A.Q.s do not obey the simple arithmetical rules; 2 + 2 in A.C. do not necessarily make 4; they may make anything between 4 and 0 according to the phase difference. Conversely, and this is a trick we use in wireless, we can split an A.Q. up into parts which individually may be greater than the thing we started with; we can, for example, design a circuit so that the volts drops across some parts of it are much greater than the applied voltage. We shall refer later to the " magnification " of a circuit and express these ideas more exactly.

Resonance in the series circuit

Consider a series A.C. circuit of the type shown in fig. 8.7. The volts drop across the inductance is ωLI, assumed to be greater than that across the condenser $I/\omega C$.

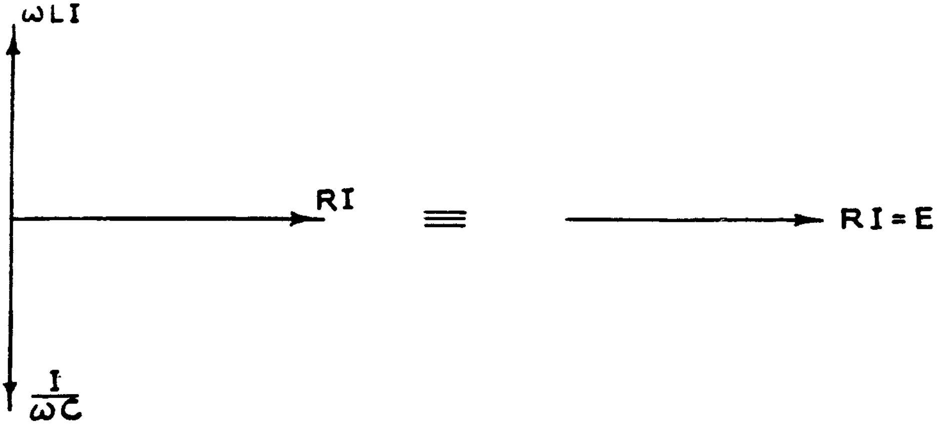


FIG. 10.7. Vector diagram for resonance

Now let the frequency of the applied voltage be lowered; ωLI will get less and $I/\omega C$ greater and by careful adjustment of the frequency we can make them equal. This will have a very big effect on the circuit; take the volts drop vector diagram first (fig. 10.7).

The reactive volts drops just cancel each other and we get the simple result that the applied voltage E is just equal to RI; in other words the only opposition is resistance, and the current is given by

$$I = \frac{E}{R}$$

This result also follows from the impedance formula

$$Z = \sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}$$

$$\text{if } \omega LI = I/\omega C\text{, then } \omega L = \frac{1}{\omega C} \text{ and } \left(\omega L - \frac{1}{\omega C}\right)^2 = 0 \text{ and } Z = R.$$

There is a particular frequency then, for which the reactances just balance one another, and at which the impedance reduces to the resistance ; this frequency is called the ***resonant frequency*** and the circuit is said to be ***in resonance*** with the supply.

The vale of the resonant frequency can easily be found for :—

$$\omega L = \frac{1}{\omega C}$$

$$LC\omega^2 = 1$$

whence, $$\omega = \frac{1}{\sqrt{LC}} = 2\,\pi f \text{ or } f = \frac{1}{2\pi\sqrt{LC}}$$

If we compare this result with the expression for the natural frequency of an oscillatory circuit it will be seen that the resonant frequency is the same as that at which the circuit would oscillate alone, under suitable conditions.

This principle of resonance, of course, is common enough. Troops marching across a slender bridge must break step, lest their step coinciding with the natural vibration frequency of the bridge should set it into oscillations of a dangerous amplitude ; a loose panel-bolt in a motor car will be set into strong vibration when revolution frequency of the engine coincides with the natural frequency of the panel ; at speeds above and below, the panel is unaffected ; it resonates at its natural frequency.

The reader should try for himself the effect of applying the same voltage at various frequencies to a series A.C. circuit.

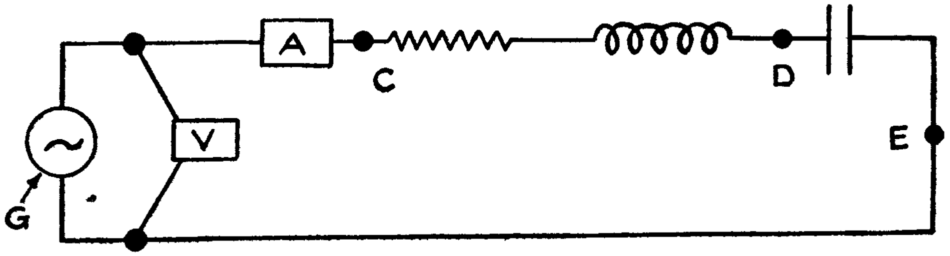


FIG. 11.7. Circuit for experiment on resonance

Fig. 11.7 shows the circuit ; G is a variable frequency generator ; V a voltmeter and A an ammeter ; the circuit consists of a choke CD (its resistance is shown separately in fig. 11.7) and condenser DE. Suitable values are :—

Frequency range of generator	20–100 c/s.
Inductance	·3 henry.
Condenser	30μF

Starting with a frequency of about 20 c/s, observe V and A for various frequencies throughout the range. Record thus :—

Frequency (1)	Voltage (2)	Current (3)	$\frac{\text{Current}}{\text{Voltage}}$ (4)

The 4th column is necessary because it will not be possible to keep the voltage constant as the machine speeds up ; the ratio I/V gives the current which would flow, at any particular frequency, under 1 volt.

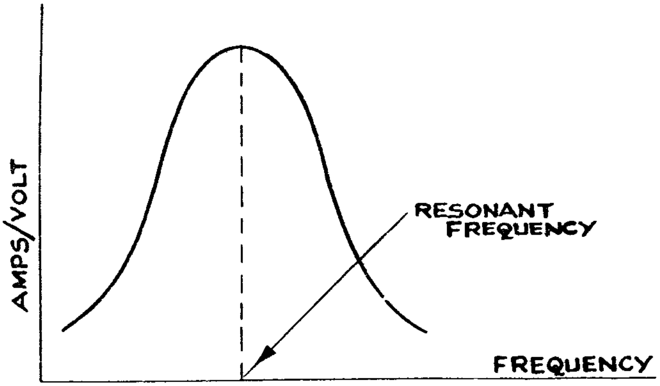


FIG. 12.7. Resonance curve

If a graph be plotted connecting columns 1 and 4 we shall get what is called the " frequency response curve " or " resonance curve " of the circuit, (fig. 12.7). As the frequency increases, the current goes up till, at the resonant frequency, the current is a maximum, equal to E/R, after which it falls again.

The experiment should now be repeated, with double the dead resistance introduced into the circuit. The resonance curve will not only be halved at peak value, but it will be much flatter; this is an important fact which we have to watch carefully in wireless circuits.

Magnification

Consider a series circuit at resonance ; the inductive and capacity reactances balance and the current is equal to E/R, R being the effective resistance of the circuit. Now the volts drop across the inductance is—

$$2\pi f L I$$
$$= \frac{2\pi f L E}{R} = \frac{\omega L}{R} E.$$

i.e. it is $\omega L/R$ times bigger than the applied voltage ; it is this factor $\frac{\omega L}{R}$ which we call the *magnification* of the circuit, and we generally denote it by the letter "Q".

This type of series circuit we shall call an " acceptor circuit " because it " accepts " current most easily at resonant frequency ; it is extensively used in wireless work.

Skin resistance

Unlike D.C. which makes use of the whole of the cross section provided, alternating currents tend to flow on the surface of a conductor. The higher the frequency the more marked does this tendency become, and in consequence the resistance of a wire is greater for A.C. than for D.C. A coil of 4 ohms D.C. resistance may run up to 50 ohms for high frequency A.C. just simply because the A.C. is trying to crowd into the surface layer, and so effectively reducing the cross sectional area of the wire. We can lessen this effect by making our conductors have as much surface as possible ; a large number of thin insulated wires braided together will have much less high frequency resistance than if they were fused into one thick wire ; the type of wire called " Litz ", much used in H.F. A.C. work, is made in this way.

The A.C. parallel circuit

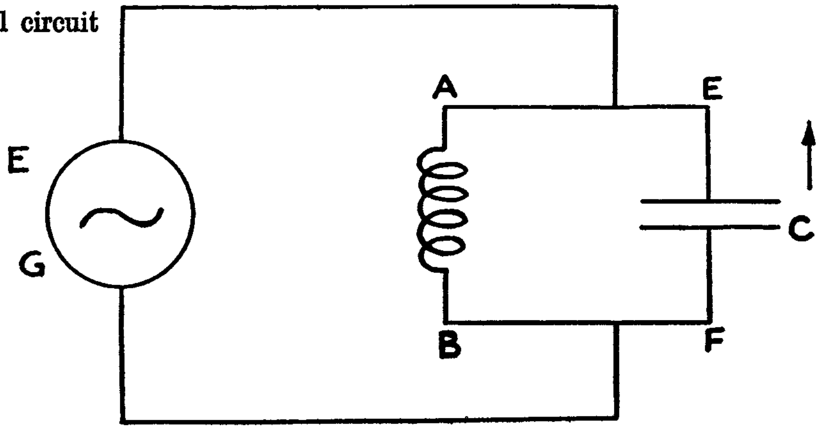


FIG. 13.7. Circuit for parallel resonance

An inductance of negligible resistance AB, fig. 13.7, is joined in parallel with a condenser EF, and the combination is supplied from an A.C. generator G. AB and EF are both under the same voltage because they are in parallel, but the main generator current will be the vector sum of the currents in AB and EF.

We can easily find these sub-currents.

(1) in AB, $I_1 = \frac{E}{\omega L}$ and lags 90° on E.

(2) in EF, $I_2 = E\omega C$ and leads 90° on E.

The vector diagram will be as shown in fig. 14.7, assuming that I_1 is the greater :—

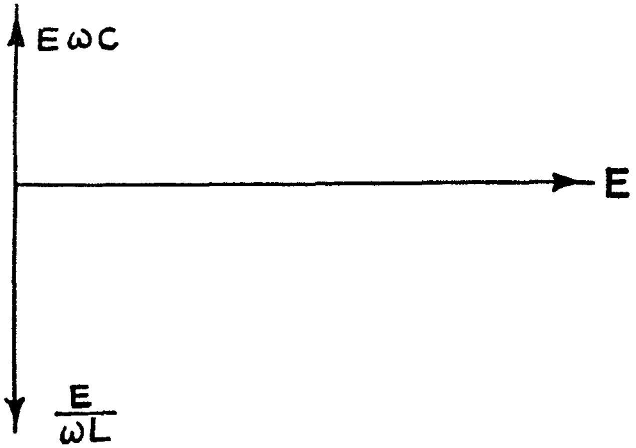


FIG. 14.7. Vector diagram for parallel circuit

Now as the two currents are just antiphased; the final current I is their arithmetical difference.

$$I = \frac{E}{\omega L} - E\omega C = E\left(\frac{1}{\omega L} - \omega C\right)$$

and the impedance $\left(\frac{E}{I}\right)$ is

$$\frac{1}{\frac{1}{\omega L} - \omega C} \text{ ohms.}$$

Resistance has been neglected here, because its introduction makes the problem quite difficult ; we will discuss its effects later. Notice, however, that if the frequency is such that

$$\frac{1}{\omega L} = \omega C \text{ or } LC\omega^2 = 1 \text{ or } f = \frac{1}{2\pi\sqrt{LC}}$$

the impedance becomes *infinite.*

This is resonance in a parallel circuit ; such a circuit offers infinite impedance at its resonant frequency and is called a " *rejector circuit* ".

Now for resistance ; if, as is clearly unavoidable, the inductive coil has some resistance (it may be high frequency resistance) the impedance at resonance is not infinite but the circuit behaves in all respects as a resistance of value :—

$\frac{L}{Cr}$ ohms, where r is the effective ohmic resistance of the coil

or, since $\omega = \frac{1}{\sqrt{LC}}$, this may be written, $\frac{\omega^2 L^2}{r}$ or $\frac{1}{\omega^2 C^2 r}$

this value is called the *dynamic resistance* of the circuit, and it is made very large by

(1) increasing the ratio L/C,
(2) making the resistance r very small.

A frequency response curve for such a parallel circuit can easily be plotted by suitable rearrangement of the components used in the series resonance experiment. It will be something like the graph of fig. 15.7.

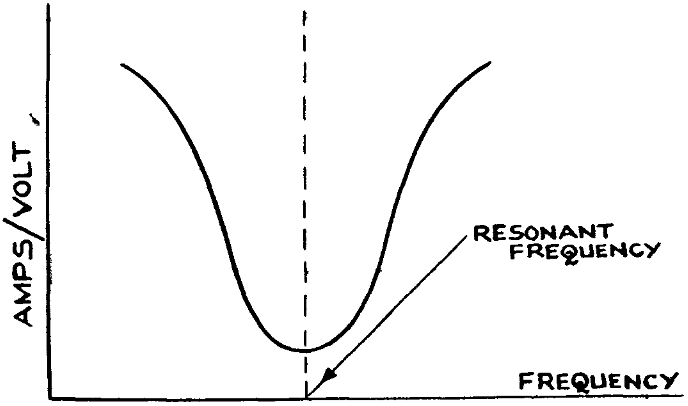


FIG. 15.7. Rejector resonance curve

From the graph, the dynamic resistance may be found, since it is equal to

$$\frac{\text{volts applied}}{\text{current at resonance}} \text{ (ohms).}$$

Returning for a moment to the ideal case (no resistance) ; although at resonance the impedance is infinite, there is still current flowing in the two branches of the circuit ; in the inductance we have $\frac{E}{\omega L}$ amps. and in the condenser $E\omega C$ amps., but as they are equal and antiphased they do not appear in the external circuit ; they constitute together a " circulating current " in the parallel system, which is now behaving like the oscillatory circuit, described in Chapter VI.

If the circuit does contain some resistance there will still be this circulating current, out to compensate for the I^2R losses, there will be a " make-up " current in the external leads equal to $E\Big/\frac{L}{Cr}$ amperes, in phase with the applied voltage.

Here then, is one way of keeping an oscillatory circuit in operation ; we shall refer to this type of oscillation as "forced" oscillation ; that which occurs when we charge a condenser and allow it to discharge through an inductance unaided, we shall call " free " oscillation.

Coupled circuits

A and B, fig. 16.7, represent two circuits, so placed that the magnetic field arising from the inductance in A links with that in B ; any current variations in A therefore are reflected in B owing to mutual induction ; the circuits are said to be, *coupled* magnetically. Coupling can be of various kinds, electrostatic coupling is indicated in fig. 17.7 (a), while resistance coupling is shown in fig. 17.7 (b). If the resistance of fig. 17.7 (b) is replaced by a condenser, we have common capacity coupling, fig. 17.7 (c).

If we plot a frequency response curve for a pair of coupled circuits, separately tuned to the same frequency, we shall get a double hump curve as shown in fig. 18.7 (a).

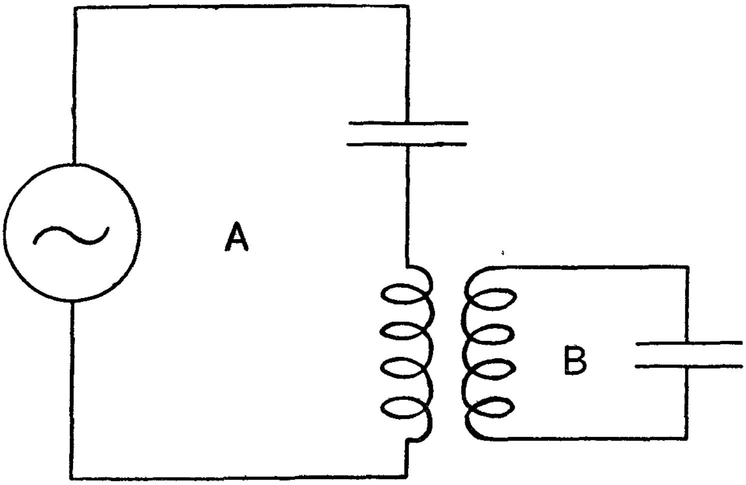


Fig. 16.7. Magnetically coupled circuits

As the coupling of the two circuits becomes weaker the humps tend to become less pronounced and a response curve like the one shown in fig. 18.7 (b) results.

This fact is largely used in wireless receivers ; a single tuned circuit has a response curve like that shown by the dotted line fig. 18.7 (b) and such a circuit responds to a comparatively narrow range of frequency ; it is, we say, *selective*. The coupled circuits respond to a much wider frequency range ; they give *band pass* tuning, because they respond more evenly to a considerable band of frequencies : this kind of response is of special value in the reception of radio telephony signals.

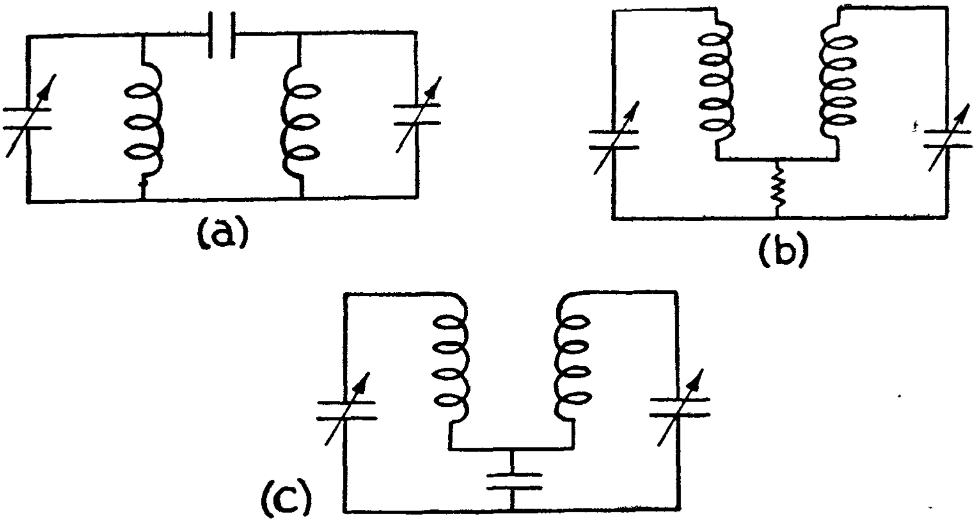


Fig. 17.7. Other forms of coupling

Transformers

An important case of coupled circuits is the transformer ; alternating current is set up in P (the primary) by means of a generator, and by magnetic coupling E.M.F.s are set up in S (the secondary), fig. 19.7. Transformers fall into three classes, according to the frequency range with which they are required to deal.

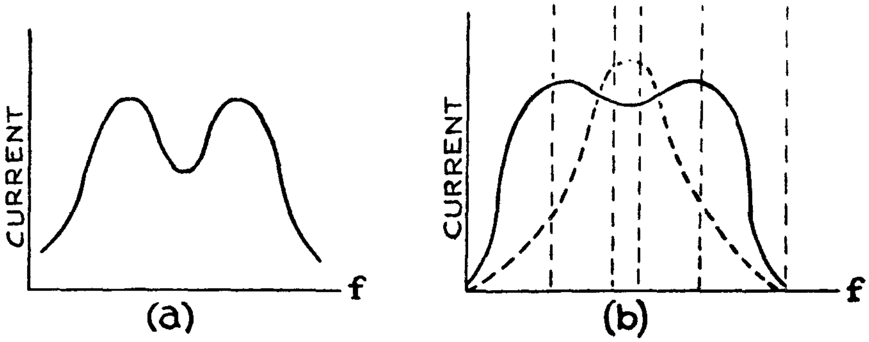


FIG. 18.7. Coupled circuit response curves

(1) Power transformers ; these are iron cored fixed frequency (usually 50 cycles/sec.) transformers, designed to have a high efficiency and to handle large powers.

(2) Low or " Audio " frequency transformers ; these are required to deal with what we call speech frequencies from about 100 to 8,000 cycles/sec. Their efficiency is not very important but they are required to amplify evenly voltages of all frequencies in the above range and they should not respond to one better than to another. These again are iron cored.

(3) High or " Radio " frequency transformers. These transformers are air cored and have to deal with frequencies of the order of 50 kc/s and above.

The most interesting, and useful property of a transformer is that it can step up, or step down, voltages in accordance with the relative number of secondary and primary turns.

The ratio :—
$$\frac{\text{Turns on secondary}}{\text{Turns on primary}} = t$$
is called the " turns ratio."

In a close-coupled transformer (types 1 and 2), the voltages are proportional to the turns ; if a primary of 40 turns is supplied by a generator at 200 volts then a secondary of 400 turns will provide a terminal voltage of 2,000, i.e.,
$$\frac{\text{Secondary voltage}}{\text{Primary voltage}} = t$$
and that would be a *step up* transformer. Of course we could only expect to draw 1/10th of the current flowing in the primary from the secondary circuit, because the transformer cannot create energy, and the power on both sides must be equal, apart from certain small and unavoidable losses.

The transformer provides an interesting case of what we call an " equivalent circuit," a simple circuit replacing a more complicated one, such for example as the single resistance equivalent to a number of resistances in parallel in D.C. theory.

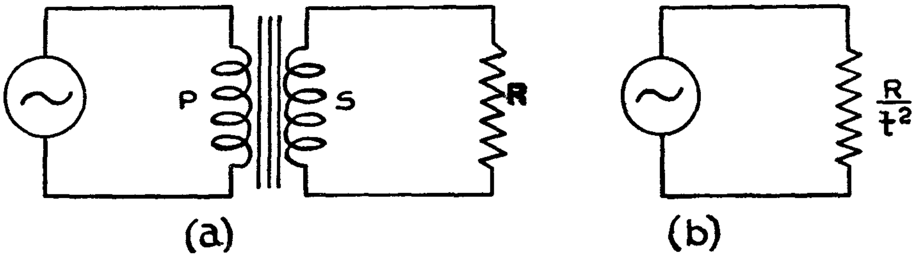


FIG. 19.7. Transformer

If the secondary is loaded with a resistance R ohms [as in fig. 19.7 (a)] the whole transformer from the point of view of the generator is equivalent to a circuit like fig. 19.7 (b). The generator behaves as if it is working into a simple resistance equal to R/t^2.

This property of a transformer is used in what we call *matching*. Let us take a simple example. A certain A.C. generator works most efficiently if the load is 5 ohms ; it is required to operate a device of resistance 20 ohms ; can it do this and still work at its highest efficiency ?

The trick, of course, is to put a transformer between the generator and the 20 ohms load and adjust its turns ratio until $\left(\frac{20}{t^2} = 5\right)$ i.e., $t = 2$, i.e., till the equivalent resistance is equal to the resistance most favourable to the generator ; in this case there must be twice as many secondary as primary turns. Iron cored transformers have their cores built of thin sheets of iron ; between each is a thin insulating layer. This lamination reduces the losses due to eddy currents. Special materials and alloys are also employed to reduce hysteresis losses to a minimum. Such alloys are nickel iron, mumetal and radiometal.

CHAPTER VIII

ALTERNATING CURRENTS (2)

High frequency circuits

We have discussed some of the more important points of A.C. theory, without any reference to the range of frequency through which our results could be expected to hold true.

As a matter of fact these relations are universally true, but unfortunately the components, coils resistances condensers and so on, begin to behave irregularly at frequencies of the order of 1,000 kc/s.

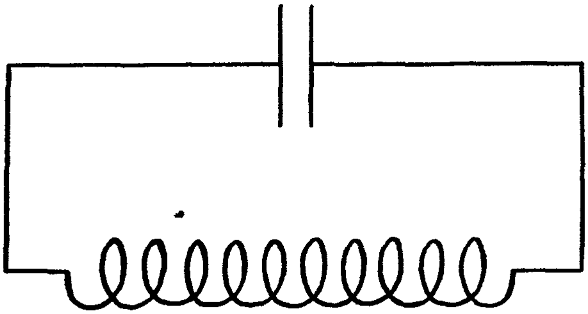


FIG. 1.8. Self-capacity of a coil

Consider an inductive coil first ; when steady (direct) current is passing there will be steady volts drops between any one turn of the coil and any other, and the pairs of turns behave like the plates of a condenser, having an electrostatic field, of which the lines are disposed in a very complex way between the various turns. The net result is that the coil actually has capacity as well as inductance ; it is equivalent to a coil with a small condenser in parallel (see fig. 1.8). The capacity of the equivalent parallel condenser is called the " self-capacity " of the coil.

Self-capacity can be reduced by winding the coil in certain ways, but it is always present to some extent and when the coil is subjected to high frequency voltages the high frequency alternating current passes over via the self-capacity rather than by overcoming the inductive reactance of the coil. A coil with self-capacity can behave like a parallel resonant circuit, without an additional condenser, and for frequencies above the resonant frequency it will take a leading current, and actually behave as if it were a condenser.

We have already mentioned skin resistance ; as our coil is subjected to voltages of higher and higher frequency, so its effective resistance goes up. Self-capacity then, and high frequency resistance are two of the imperfections of a coil, when we pass into high frequency work.

Again, a wire wound resistance will have self-capacity for reasons given in the previous paragraph; it will have a H.F. resistance also. Modern high resistances, however, are fairly free from serious defects, although a certain amount of capacity is unavoidable.

Note on resistances for use at high frequencies

Two forms of resistance have been evolved which are fairly free from the effects of (1) stray capacity (reduces total impedance); (2) skin effect (increases impedance). The first is avoided by using physically small components, the second by using materials of high specific resistance or by the use of good conductors in extremely thin layers.

Modern composition resistances consist of powdered carbon mixed with some insulating cement, the proportion of each ingredient depending on the resistance value required. The paste formed is moulded and baked into short rods with wire ends.

Another pattern uses a small rod of insulating material sprayed or " sputtered " with a very thin layer of metal (the thickness deciding the resistance). This is enclosed in a porcelain tube with metal ends and wire leads.

Condensers suffer from *dielectric losses*; just as a piece of iron in an alternating magnetic field gives rise to hysteresis losses (see p. 29) so a dielectric in an alternating electric field causes dielectric losses; energy is lost in the form of heat owing to the imperfection of the dielectric itself and in the supports of the plates. The effect of this is that the equivalent circuit of a condenser at H.F. is a condenser with a high resistance in parallel.

A tuned rejector circuit will therefore actually behave as if it were equivalent to the circuit of fig. 2.8.

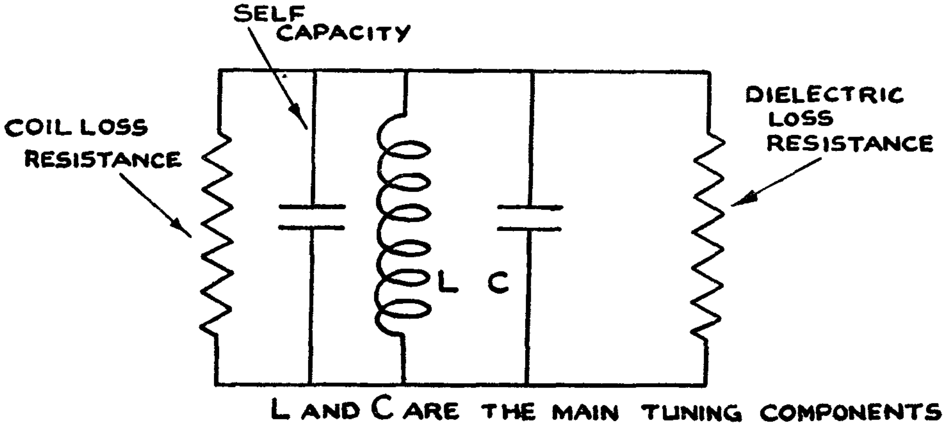


FIG. 2.8. Rejector circuit showing losses

The self-capacity does not matter very much—it is in parallel with the main condenser and behaves as part of it, but the resistance has the effect of flattening the resonance curve of the circuit, and damping the oscillatory current. Suppose, for instance, that the coil itself has negligible series resistance, then the equivalent dielectric loss resistance (R) represents the effective resistance of the whole parallel tuned circuit at resonance, because the other two branches constitute an infinite resistance and the circuit will behave as if its dynamic resistance is R, that is :—

$$\frac{L}{Cr} = R$$

where r is the effective series resistance injected into the coil, owing to the parallel dielectric loss resistance R, i.e.,

$$r = \frac{L}{CR}$$

For a given parallel resistance, therefore, the equivalent series resistance will depend on the ratio L/C.

Similar considerations apply to an acceptor circuit; it will be equivalent to the circuit shown in fig. 3.8.

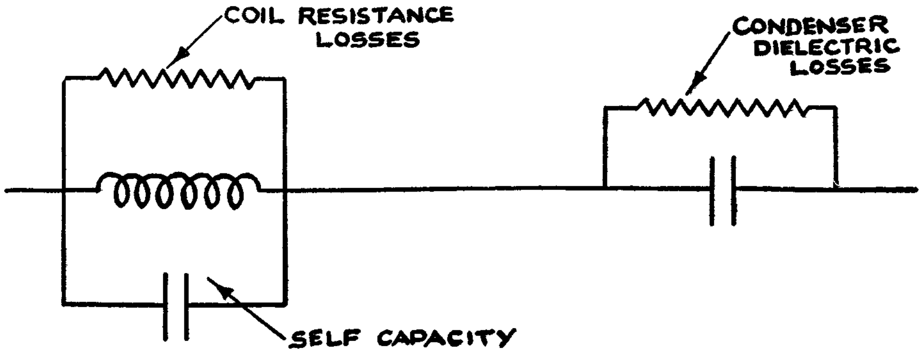


FIG. 3.8. Acceptor circuit showing losses

Here the self-capacity of the coil forms an H.F. by-pass, and it must be reduced as much as possible by careful coil construction; the dielectric loss resistance (R) can be regarded as an additional series resistance (r), whose value can be found thus :—

Provided that R is very large compared with the capacity reactance the voltage across the condenser is, with the usual symbols I/ωC, so that the watts lost in the resistance,

$$= \frac{\left(\frac{I}{\omega C}\right)^2}{R} = \frac{I^2}{\omega^2 C^2 R}$$

The same loss would be caused by a series resistance r, if

$$I^2 r = \frac{I^2}{\omega^2 C^2 R}$$

and since at resonance,

$$LC\omega^2 = 1$$

$$r = \frac{L}{CR},$$ just as in the case of the rejector circuit.

The reader can, by using the same method, show that a parallel resistance across the coil reduces to a series resistance in the same way. This fact is important, because when we actually use these circuits in wireless apparatus we connect things across the coil, such as valve holders and so on, which constitute parallel resistances at high frequency, the general effect being to damp the circuit, to flatten the response curve and depress the gain. Note that a high resistance in parallel is equivalent to a low resistance in series with the coil or condenser.

Selectivity

We have already referred to selectivity in a general way; the steeper the resonance curve of a circuit, the more quickly does the current fall off for " off resonance " frequencies. Consider the graphs of fig. 4.8; response curve A is for a carefully designed circuit with all its losses reduced to a minimum; its equivalent resistance is as small as possible.

At resonance (f cy./sec.) it passes fP amps.; at "off resonance" frequencies f_1 and f_2 it passes f_1Q and f_2R amps., *about half the current at resonance.*

For circuit B, however, the current only falls to about seven-eighths of that at resonance for the same "off-resonance" frequencies. Circuit A is said to be more selective than circuit B. The peakiness of a curve of this sort depends entirely upon the *magnification* which we have already discussed; the bigger the ratio L/r, the bigger the gain and the sharper the response curve.

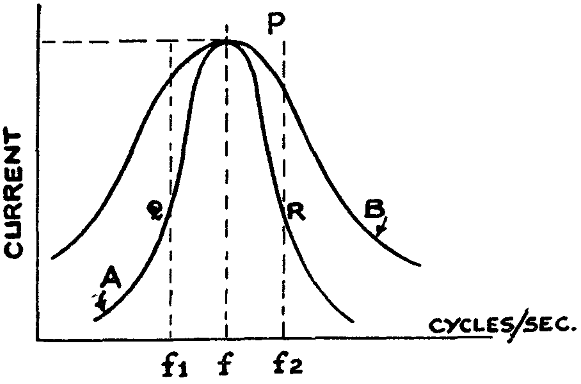


FIG. 4.8. Resonance curves

Radiation from H.F. circuits

It has already been mentioned that an oscillating circuit will radiate electromagnetic waves; in the previous chapter we saw that such a circuit could be kept in continuous undamped oscillation by excitation from a generator at resonant frequency; although it is not a very practical method, let us suppose that we have such a system, and that we want the circuit to radiate as much energy as possible.

The closed oscillatory circuit, having a close wound coil and with condenser plates close together will not radiate very well; if, however, we open out the condenser plates the circuit becomes "open" and radiation is increased; the circuit (together with its power supplies) would become a simple "transmitter" of electromagnetic (wireless) waves.

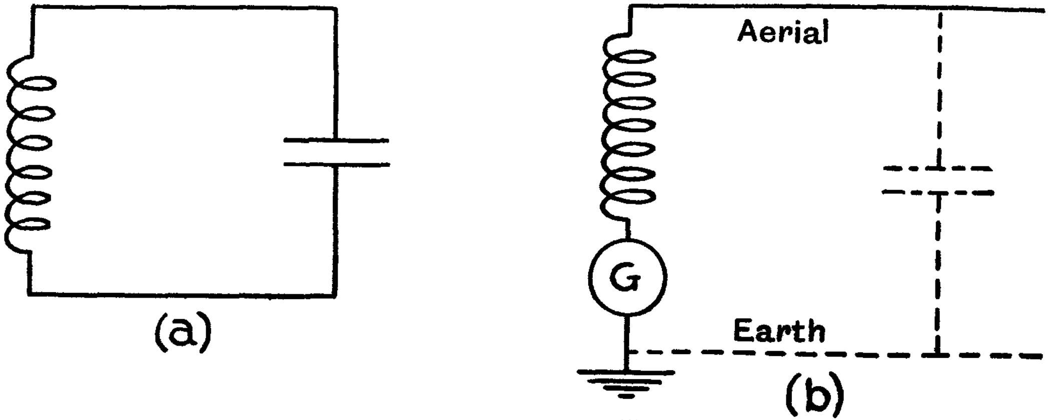


FIG. 5.8. The open oscillator

Fig. 5.8 shows how the closed circuit (a) can be so modified; one plate of the condenser consists of an insulated wire slung up in the air, and the other the surface of the ground below it; here are our "aerial" and "earth", equivalent to the condenser shown in dotted lines in diagram (b).

With the small capacity represented by our aerial-and-earth condenser the resonant frequency will be very high, many kilocycles per second, and our generator would have to be of very special design—but let us take it for granted for a moment. When the "condenser" is fully charged there will be a system of vertical electrolines between aerial and earth as shown in fig. 6.8.

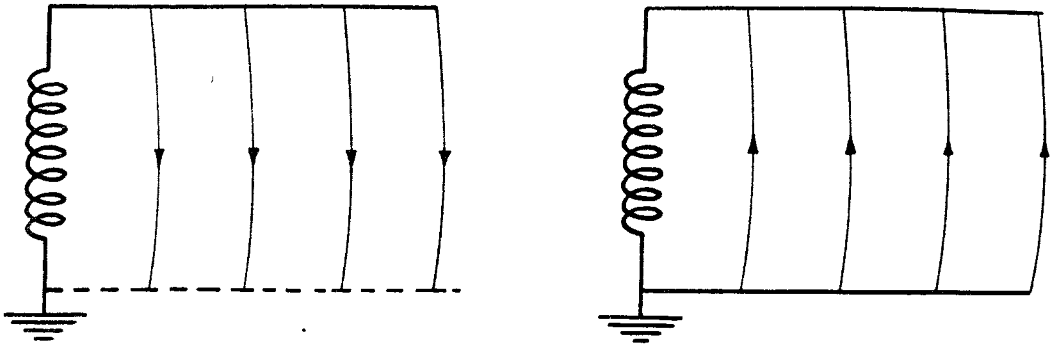


FIG. 6.8 Showing the vertical electric field

The condenser discharges, but owing to the rapidity of the change not all the electrolines are able to collapse into the circuit before the reverse field is created, and some of them are " cut off " as it were and travel out into space with a velocity of 3 × 10^8 metres per second. Travelling lines of this sort constitute electromagnetic radiation or wireless waves. The generation of these waves will depend on the frequency of the oscillation ; for slow oscillations all the electrolines will collapse into the circuit ; the more rapidly the circuit oscillates the more lines are cut off ; more energy is thrown out as radiation. That is why in wireless work we employ high frequency transmitter circuits ; low frequency oscillations up to a few thousand per second would not give rise to any radiation worth mentioning.

Electrolines in motion have another curious property—they always give rise to magnetic lines which travel along with them. The electric and magnetic lines are in time phase and at right angles to each other.

An observer, (who for the moment we will imagine can " see " wireless waves), stationed at some distance from the radiator, would be conscious of the arrival of vertical electrolines and horizontal magnetic lines, becoming stronger or weaker and reversing in phase together.

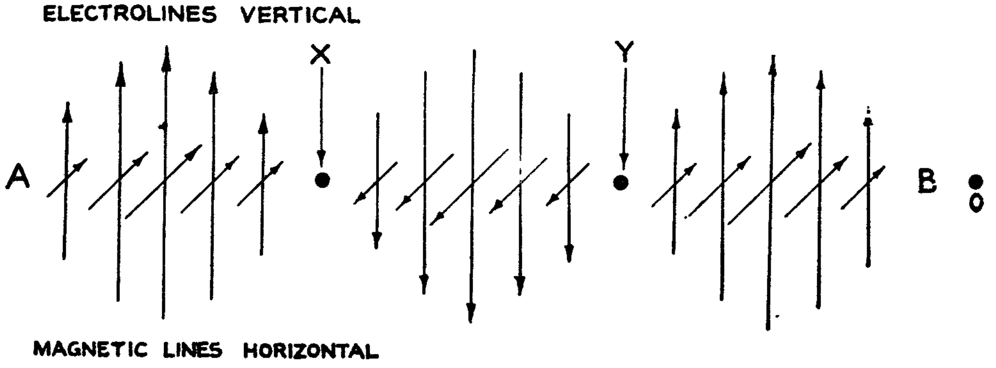


FIG. 7.8. Representation of an electromagnetic wave

Let the observer at O, fig. 7.8, remain stationary and let the system of electric and magnetic lines portrayed between A and B sweep past him with their natural velocity of 3 × 10^8 metres per second ; our supernatural observer would then be conscious of electromagnetic radiation, or in more practical language he would " pick up the radiation " of the distant oscillator. The radiation is not confined to a line or a beam ; it flows out all around the oscillator unless we use special types of aerial system, and observer O would " see " just the same type of radiation if he were anywhere on the circumference of a circle with the transmitter at the centre.

The distance between two points on the wave at which the components are in the same phase (e.g. points A and Y, fig. 7.8) is called the *wave length* of the radiation, usually denoted by the symbol λ. One oscillating cycle of the radiator sends out one complete wave, so that in one second it sends out n waves, n being its frequency in cycles per second, or a " length " $n \times \lambda$ of radiation, since 1 wave length occupies λ metres.

The distance which the radiation goes in 1 second is 3×10^8 metres

$$\therefore n \times \lambda = 3 \times 10^8,$$

$$\text{or} \qquad n = \frac{3 \times 10^8}{\lambda}$$

Although we generally speak of the frequency of the radiation rather than its wave length, this latter is expressible by a very simple formula.

$$\text{Since} \quad n = \frac{1}{2\pi\sqrt{\text{LC}}} = \frac{3 \times 10^8}{\lambda}$$

the reader will be able to deduce that :—

$$\lambda = 1885\sqrt{\text{LC}} \text{ metres}$$

if L is in microhenries and C in microfarads.

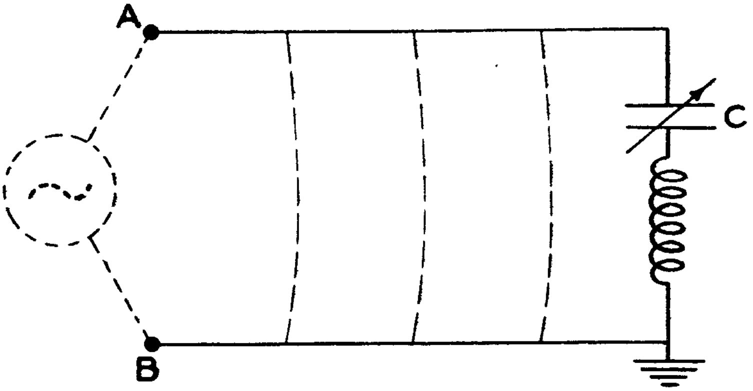


FIG. 8.8. Receiving circuit

Instead of our observer O let us see if we can devise a simple electrical method of receiving the radiation ; let us in fact put up an oscillating circuit consisting of aerial and earth, with a loading inductance, but whose frequency can be adjusted ; the adjustment may be done by altering the capacity of the circuit with a variable condenser C (fig. 8.8) ; otherwise the circuit is identical with the transmitting circuit.

The electrostatic field of the incoming radiation at the receiving aerial will cause alternating current to surge back and forth through the circuit, and to all intents and purposes the effect is the same as if a tiny A.C. generator were connected across A B (fig. 8.8). By adjusting the condenser C we can bring the circuit into resonance with the incoming wave (we call this *tuning* the circuit) when the current set up in it will be a maximum. Further, by suitably designing this receiving circuit the voltage drop across the inductance can be made several times greater than the small voltage produced by the incoming electrolines.

Here is a simple receiver then ; this type of open aerial receiver, makes use of the electroline part of the radiation.

A closed loop of wire held with its plane vertical in the line joining the observer to the transmitter will undergo flux changes as the radiation flows past, and alternating E.M.F. will be set up in it ; a loop aerial makes use of the magnetic component of the radiation. The reader can work out for himself what would happen if the loop were held so that its plane were vertical and at right angles to the line joining loop and transmitter.

All this is straightforward ; we see how a maintained oscillating circuit can be made to radiate some of its field energy in the form of wireless waves ; further, we see how we might design a circuit to receive this radiation, but what we want to do is to convey information or to " signal " by means of wireless waves. This involves two questions :—

(1) How are we going to impress signals on the outgoing waves at the transmitter ?
(2) How are we going to extract these signals from the waves at the receiving station ?

Both these questions form the subject of a future chapter of this book, and this is a convenient stage at which to leave the considerations of high frequency radiation and to take up the discussion of the telephone and the microphone, whose operation depends on audio frequency electric currents.

Sound

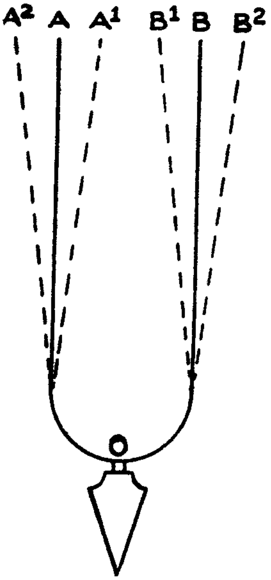


FIG. 9.8. Tuning fork

Fig. 9.8 represents a tuning fork ; AOB is its undisturbed configuration, and when it is vibrating its prongs swing between the positions A^1B^1 and A^2B^2.

As the right-hand prong travels outward from B^1 to B^2 it compresses the air to the right and rarefies the air to the left ; the next movement of the prong from B^2 to B^1 leaves a rarefaction on the right and a compression on the left, and this process goes on so long as the fork vibrates. The compressions and rarefactions travel outwards as indicated in fig. 10.8, and constitute what we call a " sound wave ". The velocity with which the pulses travel depends on the temperature and the nature of the medium, but in air at normal temperature (15° C.) the velocity is 1,100 ft. per second.

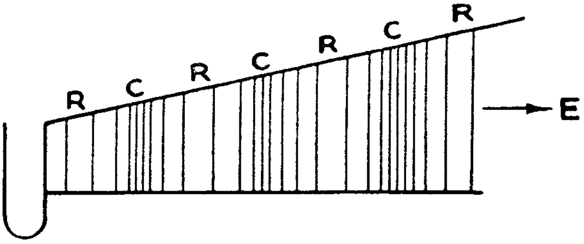


FIG. 10.8. Sound wave

If an observer places his ear at E, fig. 10.8, his ear drum will respond to the pulses, being driven in slightly by the compression (C) and drawn out slightly by the rarefactions (R), and will be set in vibration of the same frequency as that of the fork. If the frequency of the fork is above 50 cycles/sec., the brain will cease to receive each separate pulse sent to it along the nerves of the ear, and will begin to record a musical note ; the observer " hears " the note of the tuning fork.

The human ear is a " sound receiver " and it will receive sounds from 50 to 10,000 cycles per second ; the frequency of the vibrations correspond to the pitch of the note ; 50 cycles/sec. is a very low pitch ; 10,000 cycles/sec. a very high one. Middle " C " on the piano has a frequency of 256 vibrations/sec. The graph of fig. 11.8 is a distance pressure graph ; it corresponds with the diagram below it, in which the pressure and rarefaction pulses of a sound wave are represented by the closeness of the vertical lines. When this pressure curve is a simple sine graph we say that the sound is " simple " or " pure ". The difference between notes of the

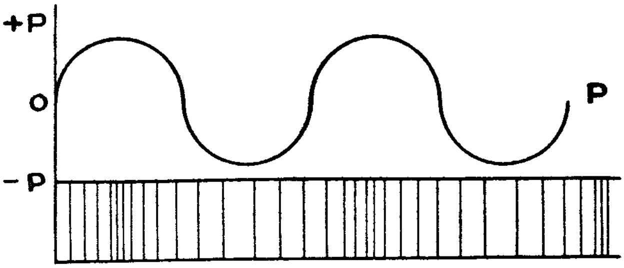


FIG. 11.8. Pressure graph

same pitch on different instruments (e.g. on a violin and on a trumpet) depends on the shape of their respective pressure graphs; the *timbre* (quality) of the note depends on the *wave form.* Fig. 12.8 shows the pressure graphs corresponding to two notes of the same pitch, but of different quality; this difference in quality corresponds to the difference in the wave forms.

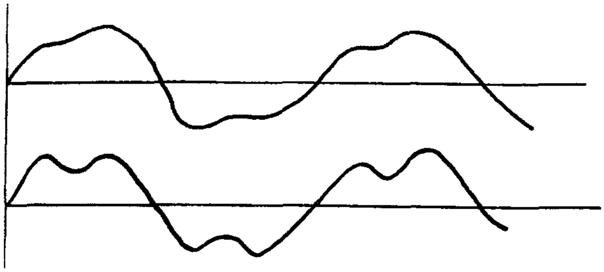


FIG. 12.8. Pressure graphs of dissimilar sounds

When we come to consider sounds which are no longer simple tones, but are a combination of all sorts of wave forms and many different frequencies (such as, for example, the sound produced by an orchestra), we find that the wave form becomes extremely complex; but the fact remains that the ear responds to all the complicated pressure changes involved and the brain interprets these responses as a musical sound.

The *loudness* of a note is decided by the amplitude of the vibration.

The human voice is a sound source of great flexibility; it can produce simple sounds, as for example when a single note is sounded for some seconds; it can make the note louder or softer; it can change its pitch; it can change its timbre.

Further than this, the tongue, the lips and the teeth can be used to " chop up " the sound emitted by the vocal chords, into an orderly and recognisable sound-pattern, which we call speech. This point is worth a little consideration, because it has an almost exact parallel in the field of wireless communication; just as a person's voice emitting one continuous note can only convey to a listener the fact that he is emitting that note, so a maintained oscillatory circuit, steadily emitting radiation, can only inform a distant receiving station that it is so radiating. The sound, or the radiation, as the case may be, has to be chopped up into some kind of pattern, which can be recognised by the hearer or by the receiver, as conveying certain information.

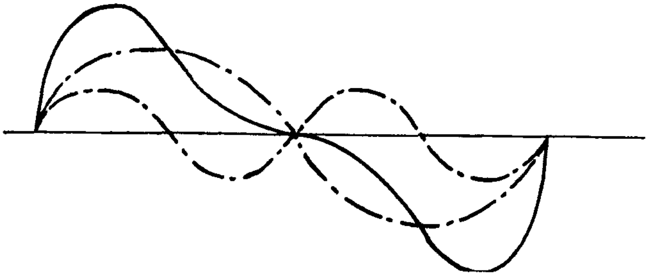


FIG. 13.8. Analysis of non-sinusoidal wave form

When the pressure graph of a note departs from the simple sinusoidal form, it is said to contain overtones or *harmonics*; in other words, it is no longer one single note, but a mixture of notes of frequency f, $2f$, $3f$, etc. The main component or fundamental (f) is accompanied by others, the second harmonic ($2f$), third harmonic ($3f$), etc.

Fig. 13.8 shows how a non-sinusoidal note (thick line) can be regarded as built up of two sinusoidal notes, fundamental and second harmonic (chain lines). As a matter of fact this is quite general; a mathematical theorem due to Fourier, tells us that no matter how complicated a wave form may be, it can always be resolved into a fundamental and a series of harmonics of appropriate amplitudes and the theorem is applicable to all kinds of waves, sound waves, light waves, water waves or electric waves.

We shall have occasion in later chapters to speak of *distortion*; when a sound suffers a change in quality, and therefore in waveform, it is said to have become distorted. In the light of the present discussion it is to be noted that distortion consists of the introduction of harmonic components not originally present in the sound or alteration of the relative amplitude of harmonies already present. Certain forms of radio-receiving circuits for example, introduce what is called second harmonic distortion ; in its simplest terms this means that if the broadcasting station sent out a pure sinusoidal middle C, the receiver will produce not only middle C, but its upper octave as well.

The microphone

A thin diaphragm A (fig. 14.8) rests lightly upon a small chamber (B) containing carbon granules ; the granules, the diaphragm A and the metal base C of the containing chamber form part of an electric circuit, connected externally as shown, to a battery (D). When a sound wave impinges on the diaphragm, it responds to the pressure and rarefaction pulses by moving in and out, and in so doing it alters the resistance of the circuit by changing the packing of the granules.

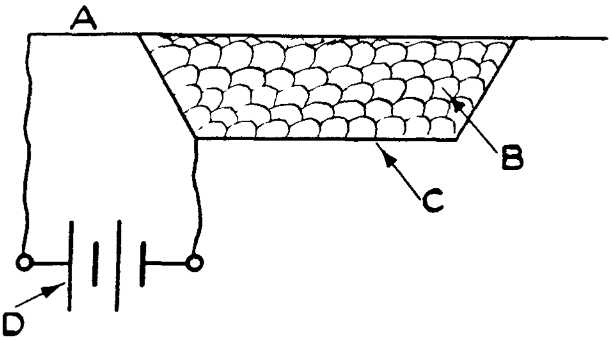


FIG. 14.8. Principle of the microphone

Thus the current in the circuit will vary and the electrical fluctuations will be a reproduction of the pressure pulses acting on the diaphragm. This arrangement, which is called a *microphone*, will transform the sound impulses into electrical impulses which can be conveyed over very long distances along suitable conducting wires.

Fig. 15.8 shows the general construction of a common form of microphone.

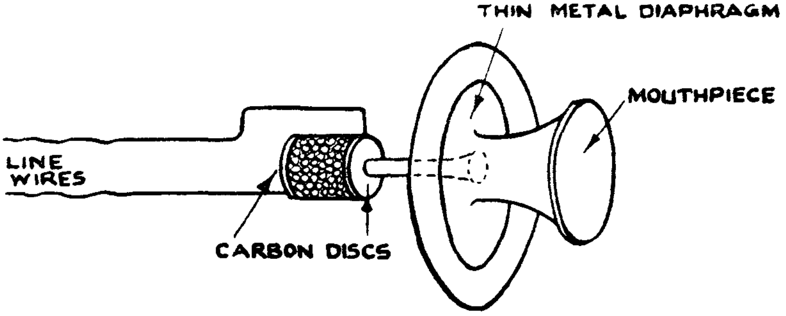


FIG. 15.8. Common form of microphone

The current in a microphone will be fluctuating D.C. ; it will consist effectively of a D.C. component with an A.C. of complicated wave form (" speech current ") superimposed.

Thus the fluctuating current shown in fig. 16.8 is equivalent to a direct current at value OD, together with an alternating current consisting of the variations above and below the D.C. level marked by the line DC on the diagram.

If it is desired to eliminate this D.C. a transformer is used as shown in fig. 17.8.

The resistance of the microphone is of the order of 200 ohms, and may vary under the influence of speech by about 10 per cent. ; if the lines were of very high resistance, the overall variation of lines and microphone would be very small indeed, so a step up transformer of about 1 : 25 would be used as shown in fig. 17.8.

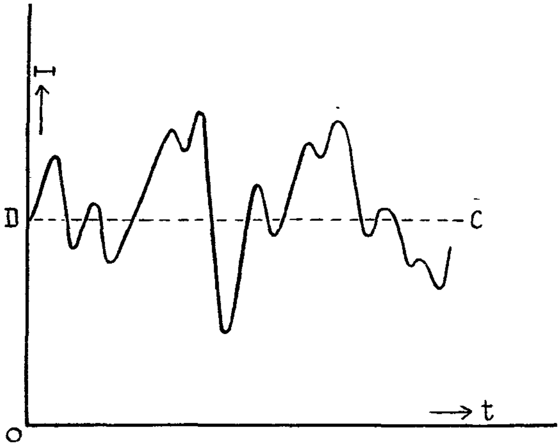


FIG. 16.8. Speech current

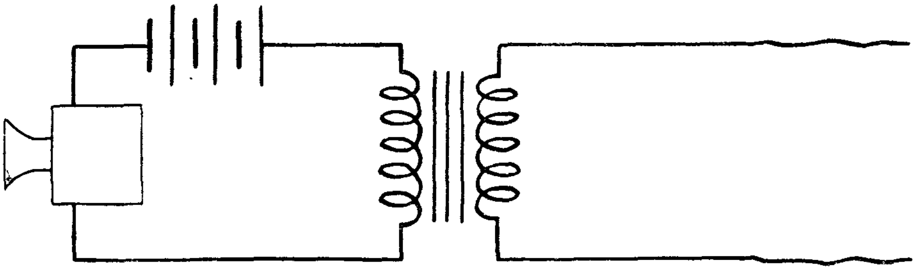


FIG. 17.8. Use of line transformer

The reader should handle a microphone in the laboratory and sketch its component parts.

The type of microphone used in the Post Office hand telephones is very sensitive, but it does not respond evenly to sounds of different frequency. Its response extends only from 200 to 3,000 cycles/sec.

There are other types of microphone, moving coil, ribbon, condenser and crystal types; for a description of these the reader must refer to a standard text-book. The carbon-granule microphone is the one most frequently met with in the Service at present.

The telephone

When the speech currents from a microphone reach the end of a transmission line they must be reproduced as audible sounds; the telephone or the loudspeaker enables this to be done.

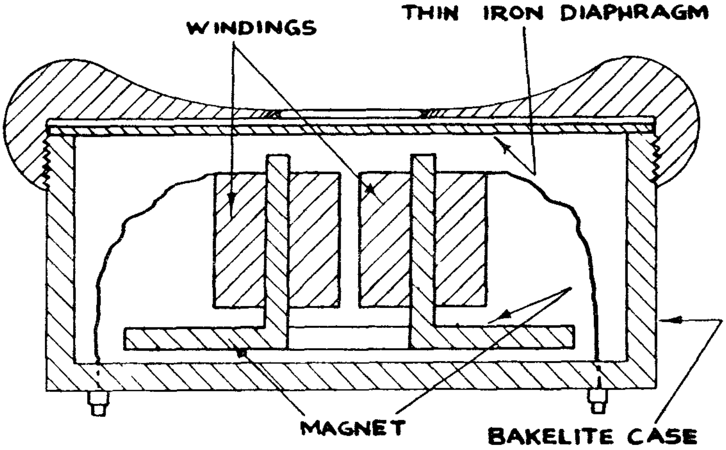


FIG. 18.8. Telephone receiver

The speech currents pass through the windings (fig. 18.8), arranged on the soft iron pole pieces of a permanent magnet. These currents cause the strength of the magnet to fluctuate and cause greater or less attraction on the iron diaphragm, which therefore vibrates to and fro and reproduces in the neighbouring air the sound waves which originally fell on the microphone.

The coils may consist of a few turns of wire of low resistance (say 60 ohms) or of many turns of fine wire giving up to 2,000 ohms resistance. Such a high resistance telephone would have to be transformer operated.

The permanent magnet gives extra sensitivity and also avoids a serious form of distortion; without it alternating current of a given frequency would produce audible sounds of double frequency.

The loudspeaker

A telephone receiver is not designed to give very strong signals; if loud signals are required the feeble speech currents delivered from the microphone must be amplified and reproduced by a *loudspeaker*. The moving coil loudspeaker, which is one of the best of modern types, employs the motor principle.

A soft iron cylinder (*a*), fig. 19.8, has a central pole piece ; over this pole piece, and inside the cylinder, is a magnetising coil (*b*).

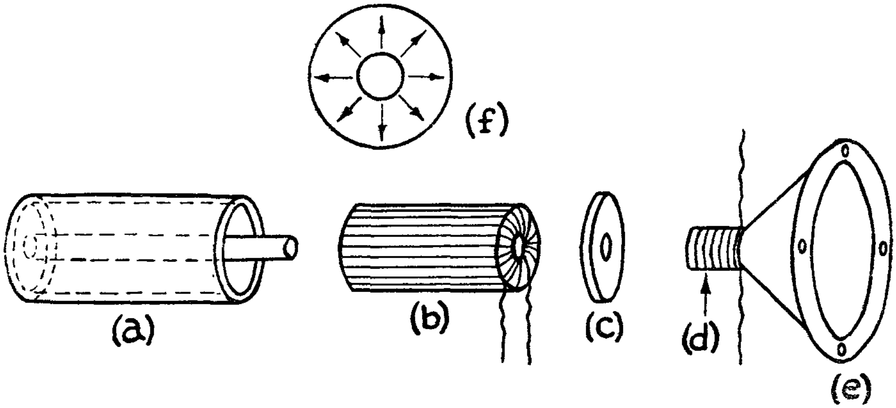


FIG. 19.8. Moving coil loudspeaker

Direct current must be supplied to this coil, which causes a radial field (*f*) to be set up between cylinder and central pole. A light coil of wire [the speech coil (*d*)] is suspended within the field ; speech currents in this coil cause it to move up and down parallel with the axis of the central pole. By doing so it causes the cone, to which it is attached, to vibrate and reproduce the sound which actuated the distant microphone. An iron disc (*c*) has a central hole rather larger than the coil, and serves to complete the magnetic circuit, leaving only a narrow radial gap in which the coil can move.

The speech coil is usually of low resistance (2 to 15 ohms), and it will require to be matched by means of a transformer to the line, or to the amplifier which operates it.

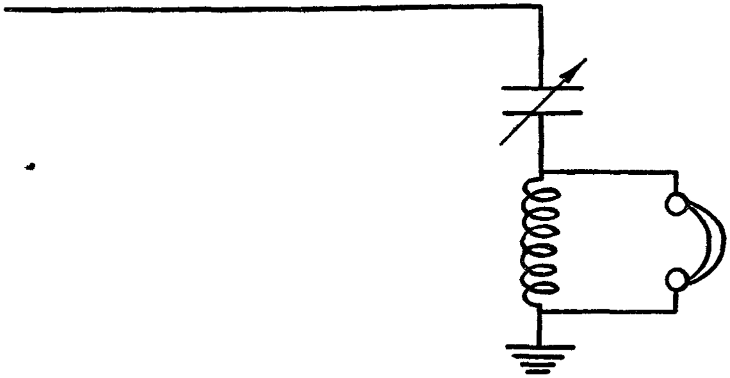


FIG. 20.8. Illustrating telephone cut-off frequency

Frequency limits

Like the microphone, the telephone has a definite frequency " cut off " ; below 200 cycles and above 5,000 cycles per second it fails to respond. This is a very important point ; the telephone or loudspeaker is almost invariably the instrument used to translate electrical signals into audible signals, and it follows that if these signals are at radio frequency the telephones are useless. If, for instance, in our simple wireless receiver mentioned on p. 82 a pair of telephones were placed in the receiving circuit as shown in fig. 20.8, they would be unaffected ; because the frequency (radio frequency) is far beyond the upper cut off of the telephones, and the ear does not hear at such a frequency.

This is another problem to which we shall return later.

Separation of A.C. and D.C.

In wireless circuits we often get cases where the same conductors are carrying D.C. and A.C. as well—or, what is the same thing, a direct current with pulsations. Further, we may have both low frequency and high frequency pulsations at the same time ; the problem is how to separate these component currents. The solution is based in general on the reactance of chokes and condensers to A.C. of different frequencies.

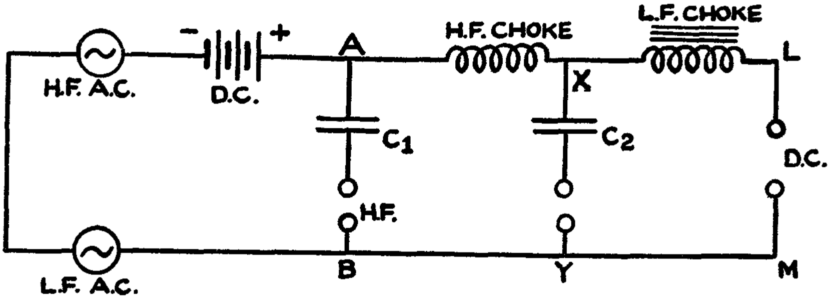


FIG. 21.8. Filter circuit

Suppose the frequencies to be separated are

H.F. A.C.	1,000 kc/s.
L.F.	50 c/s.

together with some direct current.

Condenser C_1 (about 0·001 μ F.), fig. 21.8, is of low reactance and the H.F. choke is of high reactance to the high-frequency component which therefore takes the path A–B. D.C. cannot pass along A–B, nor can L.F. A.C. owing to the smallness of condenser C_1. The L.F. A.C. gets through the H.F. choke and the large condenser (say 10 μ F.) C_2, but is stopped by the L.F. iron-cored choke. The L.F. component therefore passes along X–Y ; the D.C. cannot get along X–Y because of the condenser C_2, but it can pass through the L.F. choke and takes the path L–M.

This is the principle underlying the operation of " smoothing circuits ". The D.C. mains voltage or the output of a rectifying unit is pulsating, and to smooth it out the circuit shown in fig. 22.8 can be used.

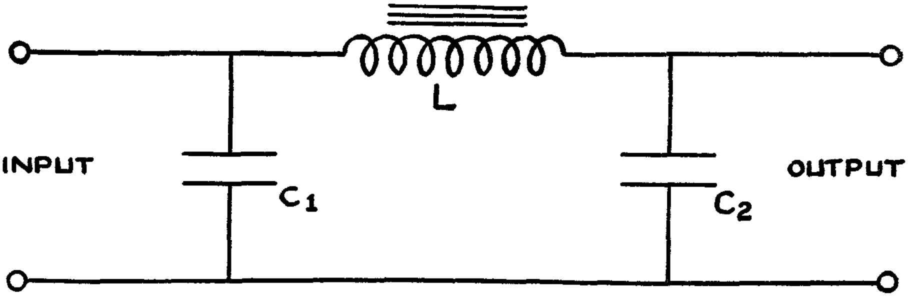


FIG. 22.8. Smoothing circuit

The alternating component of the input is kept from the output by the L.F. choke L and it short circuits through the large condenser C_1 ; C_2 is another condenser designed to pass any alternating current which might have got through the choke.

CHAPTER IX

THE THERMIONIC VALVE

We cannot get much further in our study of wireless without the help of the thermionic valve, which we shall discuss in this chapter. This device grew almost by accident out of the researches of O. W. Richardson, Fleming and de Forest, and it has lifted high-frequency work from the uncertainties of the experimenter's laboratory to the exactitudes of practical engineering. Its development is still incomplete, and the extent to which it may still be applied to the various branches of engineering can only be a matter of conjecture, but so far as modern wireless is concerned, it is the essential instrument upon the action of which all operations depend.

Diode ; diode rectification

When a piece of platinum or tungsten wire is raised to white heat, its conduction electrons are endowed with so much energy that some of them are able to escape from the inter-atomic spaces and break right through the surface of the hot metal, just as a " hot " molecule of water will escape from the surface and become a molecule of water vapour in the process of evaporation. An escaping electron (a thermion) will leave the hot wire positively charged, and presently an equilibrium will be reached when a cloud of escaped electrons around the wire (called a *space charge*), together with the corresponding positive charge in the wire cause any emitted electrons to return to the filament owing to electrostatic repulsions and attractions. Consider a system such as that represented in fig. 1.9 ; a tungsten filament F is made white hot by means of an electric current maintained by the battery B. Electrons are emitted from F, but instead of allowing them to collect around it they come under the action of a potential difference supplied by battery A, rendering a metal plate P positive with respect to the filament.

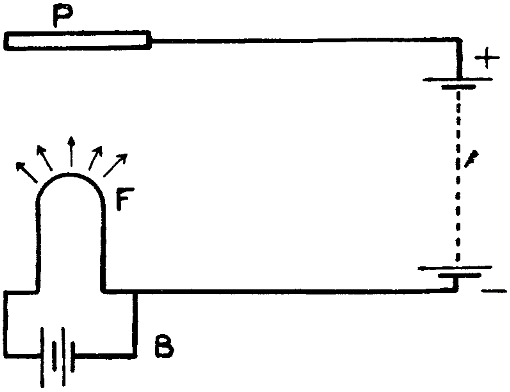


FIG. 1.9. Thermionic emission

This P.D. causes the emitted electrons to move towards P, where they are captured and returned to F via the battery A.

This is all very well in theory, but the closely packed air molecules between F and P will tend both to prevent the escape of electrons and to hinder their passage across to P ; the next step then is to put the system in an evacuated glass envelope and we get the simplest kind of valve called a DIODE, because it comprises two electrodes, the filament and the plate or anode.

The construction of a valve is of course a highly specialised process, but in simple terms we can regard the structure, i.e. anode and filament, with their leads as being assembled first ; the structures are usually made of nickel, while the filament may be of tungsten, a substance with a high melting point, so that it may be raised to a very high temperature and give a copious supply of electrons. If the tungsten filament is treated with barium or calcium oxide, electrons are emitted in large numbers at quite low temperatures ; this is the principle of the dull-emitter valve. A glass envelope is then sealed over the structures, an opening being left for attachment to a vacuum pump, and evacuation proceeds.

while the structures are maintained at a red heat in an induction furnace, or eddy current heater, to drive off occluded gases. When this is complete the pump attachment is finally sealed off and a small piece of magnesium inside the bulb is heated (again by eddy currents) till it finally evaporates, clearing out the last traces of gas and covering the inside of the glass envelope with a thin film of metal ; this is called " gettering ".

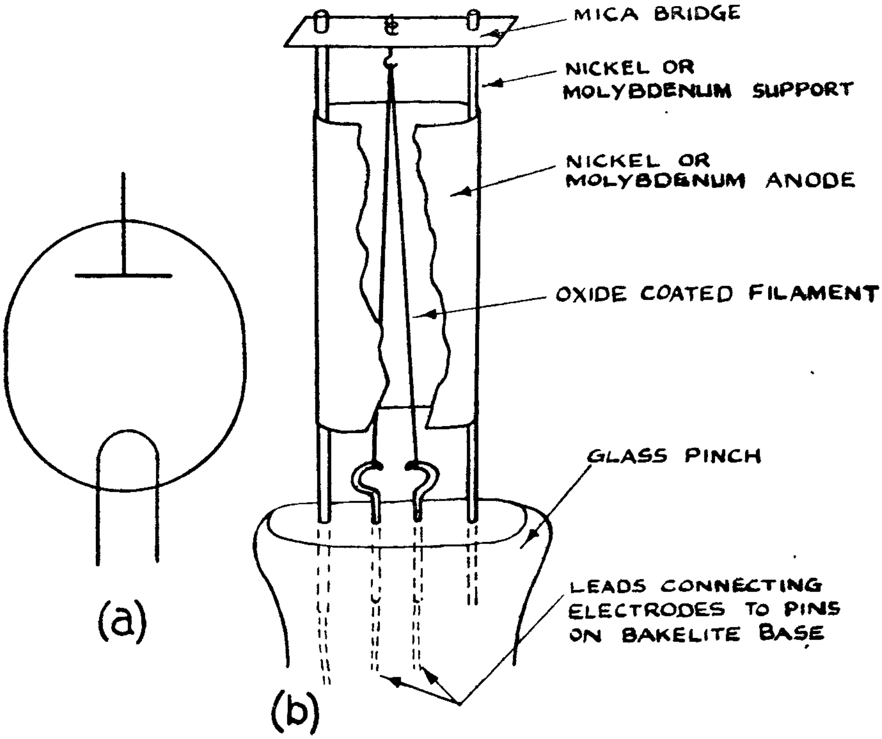


FIG. 2.9. The Diode

Fig. 2.9 (a) shows the conventional representation of a diode, while fig. 2.9 (b) is a diagram of a practical diode, the glass bulb being omitted.

The first thing to notice about this system is that, with the filament maintained hot and supplying electrons, it will conduct by an electron stream from filament to plate, but *not from plate to filament,* because the plate is not emitting thermions.

It is this " one-way conductor " property which gave it the name " valve ", because its action is like that of a bicycle tyre valve, through which air can pass in one direction only.

The electrical properties of a diode are best studied by a practical test in the laboratory ; the diagram of fig. 3.9 indicates the kind of circuit which should be used, and the reader should proceed to plot the volts-amps. characteristic, while the filament is heated by a battery whose voltage is equal to that for which the filament is designed.

A (fig. 3.9) is the diode whose filament is heated by the battery B ; C is the high-tension battery ; the negative terminal is joined to the filament and the positive terminal is joined through a milliammeter to the anode. The positive tapping is variable so that any desired voltage, up to 200 V., measured by the voltmeter V, can be applied to the anode. A series of readings should be taken and plotted in the usual way, as shown in fig. 4.9. It will be seen that this volts-amps. characteristic differs radially from that of an ordinary conductor such as we dealt with in

Chapter II, p. 10. It is no longer a straight line through, and on both sides of, the origin. It is a curve which ceases at a very small distance on the negative side of the origin and reaches a steady value for a certain anode voltage. The steady value (a), fig. 4.9, is called the " saturation value " and corresponds to the condition in which all the emitted electrons are dragged across to the anode, and none return to the filament or remain to augment the space charge. An increase of filament temperature provides more electrons and results in a higher saturation current as shown at (b) in fig. 4.9.

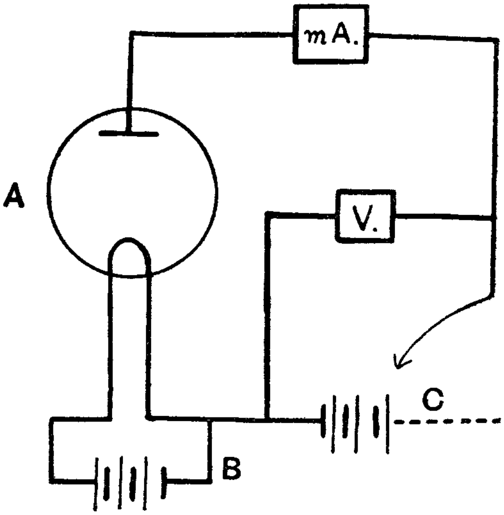


FIG. 3.9. Diode testing circuit

We saw in Chapter II that conductors with a straight characteristic obeyed Ohm's Law ; the diode, having a curved characteristic, is a " non-ohmic " conductor, and we cannot hope to find a simple relation like $I = \frac{V}{R}$ to enable us to work out the current through the valve in any given case. The behaviour of the valve has to be exhibited by graphs or *characteristics* on this account ; assuming that the filament temperature always has the rated value, we must refer to the characteristic curve to find out what the current corresponding to any given anode voltage will be.

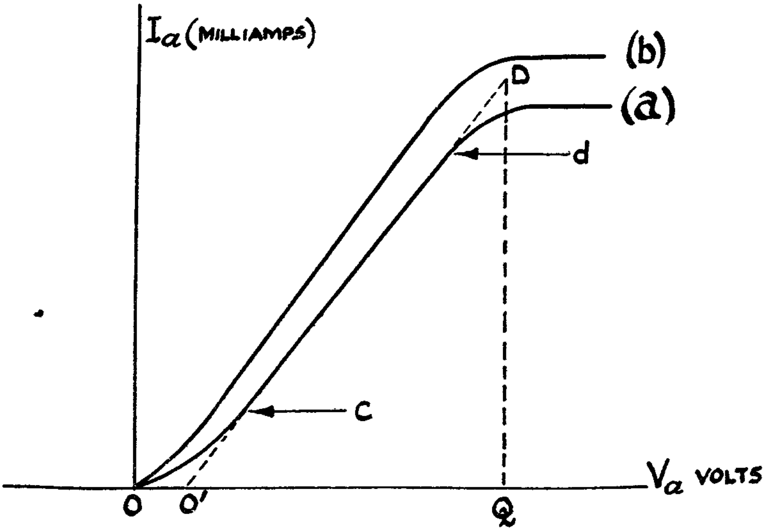


FIG. 4.9. Diode I*a* V*a* characteristic

In Chapter II, p. 11, we defined the ratio $\frac{\text{applied volts}}{\text{current}}$ as the resistance of a conductor ; the constancy of this ratio constitutes, in fact, Ohm's Law. If we work out a series of values $\frac{\text{volts}}{\text{amps.}}$ in the case of our diode we shall find that the ratio is by no means a constant ; but this does not prevent our coming to some agreement as to what we mean by the internal resistance of the valve ; a large part of the characteristic is straight, between *c* and *d* for example, in fig. 4.9 (a), and for this region of the graph, therefore, *increases* of current are proportional to *increases* of voltage or, stated naother way

$$\frac{\text{increases of voltage}}{\text{increases of current}} = \text{constant},$$

and it is this constant which is called the *slope resistance* R or the *A.C. resistance* of the valve for this straight portion of the characteristic.

Its value is found quite simply ; the straight portion is produced to meet the voltage axis at O^1, and from some convenient point such as D ; a perpendicular DQ is dropped ;

$$\text{then the ratio } \frac{O^1Q \text{ (volts)}}{Q\,D \text{ (amps.)}}$$

is the A.C. resistance in ohms.

It follows that the *steeper* this graph, the *less* is the A.C. resistance, or looked at from the opposite point of view the *steeper* the graph the *more* is the A.C. conductance of the valve.

Let us now examine the effect of applying an alternating voltage to a diode, under various conditions. Consider the circuit of fig. 5.9 (a).

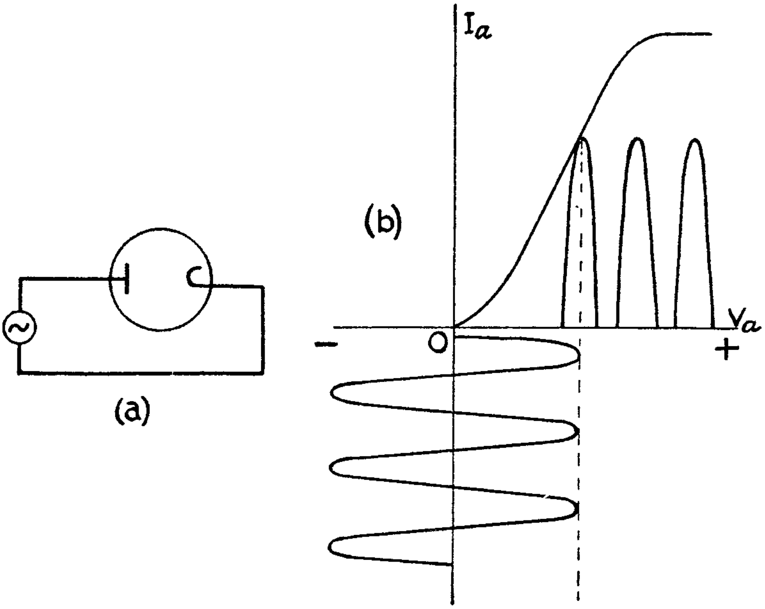


FIG. 5.9. Diode rectification

In this circuit the filament heating battery has been omitted ; this omission will be standard practice in future as it makes for simplicity in the diagrams. When the voltage of the generator makes the plate positive, current will flow ; during the negative epochs of voltage no current will flow, and there will be a unidirectional pulsating current in the circuit as shown in fig. 5.9 (b).

We shall make considerable use of the type of graph shown in fig. 5.9 (b) and it will be advisable to consider it a little more fully. It is fundamentally the I_a V_a graph of the diode ; the sine curve below the V_a line represents the input voltage variations, while the loops standing to the right represent the pulses of current flowing in response to the positive halves of the input voltage.

The input voltage graph is symmetrical about the line representing zero volts ; we describe this state of affairs by saying that the *operating point is zero.* By putting a battery in the circuit we can make the operating point more or less what we like. Fig. 6.9 (a) shows how we can make the operating point negative, and the graph 6.9 (b) shows the kind of current pulses which result : if the operating point were so far down in the negative region that the peaks of the alternating input never rose beyond zero, there would be no current at all.

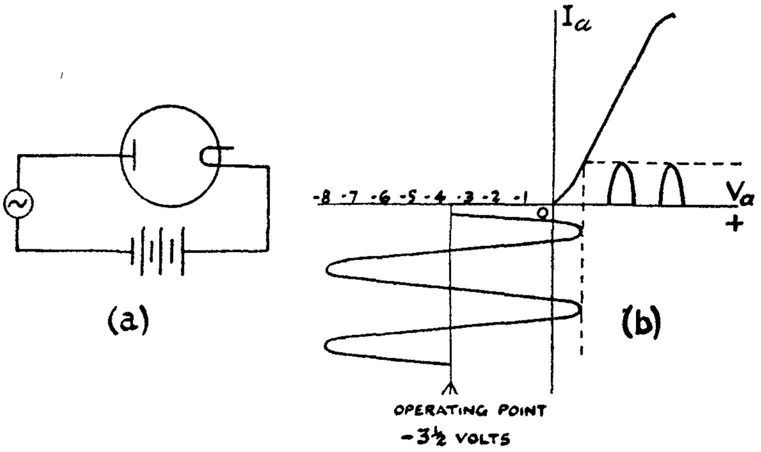


FIG. 6.9. Diode with negative operating point

Functioning in this way, the diode acts as a *rectifier* ; in common and not very correct language, it turns A.C. into D.C. ; expressed more correctly, in response to an alternating voltage, the valve allows unidirectional current impulses to flow. If we simply wanted to charge an accumulator, we could do so by connecting it in circuit with such a valve and an A.C. generator and this is in fact one of the methods we use when accumulators have to be charged from A.C. mains.

In most wireless applications of a rectifier valve we do not use the unidirectional current pulses themselves ; we generally put a high resistance (a load) in the anode circuit and the rectified current pulses set up across the load, voltage pulses which we can employ more conveniently. Fig. 7.9 (a) shows the circuit with a resistance load (R), generally of the order of 1 megohm ; fig. 7.9 (b) shows how the applied voltage, acting at zero volts, causes pulses to be set up.

If a suitable D.C. voltmeter were joined across R, it would indicate a mean value of some sort, because it would be too sluggish to follow the pulses individually ; this does not mean, however, that there is a steady voltage developed making the anode negative with

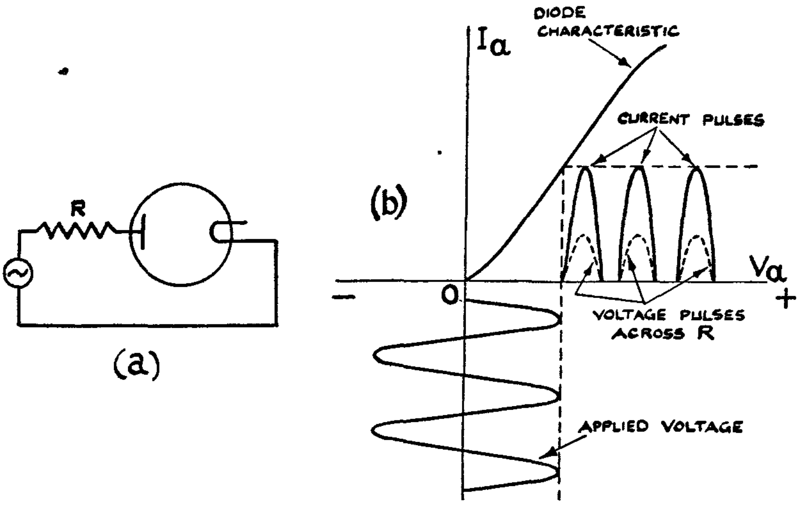


FIG. 7.9. Diode with resistive anode load

respect to the filament ; the operating point is still zero, as indicated in fig. 7.9 (b). If a condenser were substituted for the resistance the first few current pulses would charge it up and develop a potential across it ; it would then act like the battery in fig. 6.9 (a) and the operating point would gradually " go negative " until the steady negative potential was just equal to the peak value of the applied voltage.

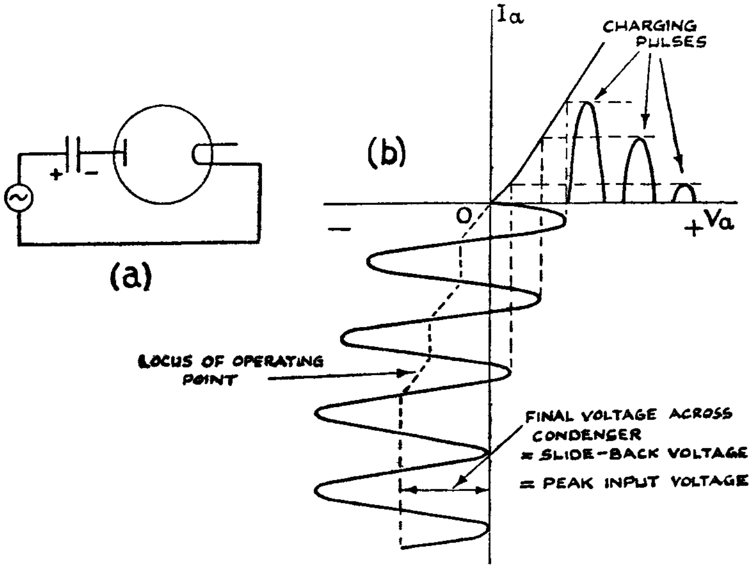


FIG. 8.9. Diode with condenser in anode circuit

Under these conditions no current flows at all even on the positive half cycles ; this action is illustrated in the graph of fig. 8.9 (b).

A combination of a condenser with a high resistance in parallel for load, is a very common one in wireless technique. Just as before, the first few positive pulses will charge up the condenser, but during the negative half cycles the charge on the condenser will leak away through the resistance and its P.D. will fall slightly, only to recover on the next positive half cycle. The slide-back voltage (which means the same thing as the voltage

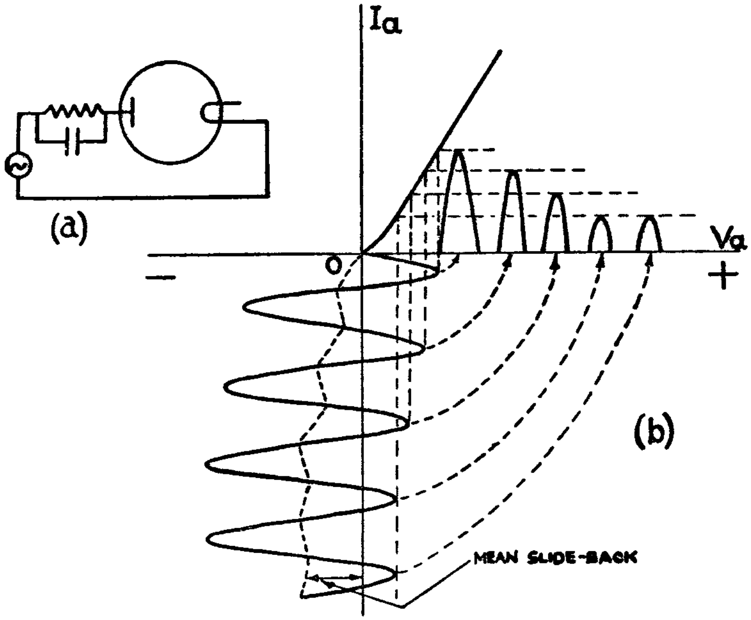


FIG. 9.9. Diode with condenser and leak resistance

developed across the load) will never become quite constant as in the case of the condenser alone ; it will rise and fall slightly at the input frequency, and its mean value will be less than the peak of the input ; how much less, depends on the value of the resistance ; the higher the resistance, the more closely will the mean voltage approach the input peak value. Fig. 9.9 (a) and fig. 9.9 (b) illustrate this important action.

The frequency of the input voltage does not enter into the above considerations fundamentally ; the physical actions are the same whether we are dealing with very high frequencies or 50 cycle commercial A.C., but it will be useful for our further work to note that the voltage developed across a " resistance and condenser " load consists of a steady voltage somewhat less than input peak, together with a " ripple " of input frequency.

We might finally enquire how this steady voltage depends on the amplitude of the input ; the problem could be approached by practical experiment, but it is simpler to note that for a wide range of values the I*a* V*a* characteristic of a diode is a straight line except for a small region at the origin and again at saturation. It is only a matter of drawing graphs to show that the load voltage is directly proportional to the amplitude of the input, provided the latter is big enough to swamp any effect due to the initial curvature of the characteristic. It is on account of this proportionality that the diode is called a *linear rectifier ;* we return to this point in a later chapter.

The triode

It has been remarked above that a swarm of electrons situated around an incandescent filament, will, by its electrostatic repulsion reduce or even prevent further emission. This is the space-charge effect, and it obviously must have a big effect on the value of the current in the case of the diode. Of course, as the plate voltage is raised, space charge electrons are removed and emission from the filament is facilitated ; in fact when the plate voltage reaches a sufficiently high value, no electron can " hang about " as it were, the wire emits freely and saturation conditions will supervene if the filament is not already destroyed. But there is a better way of controlling the space charge—something like putting a policeman on duty in a crowded street to keep people on the move and prevent congestion. If a third structure is put in the tube quite near the filament, it serves as a powerful control on the anode current, accelerating or retarding the electrons according to the sign of its potential with respect to the filament.

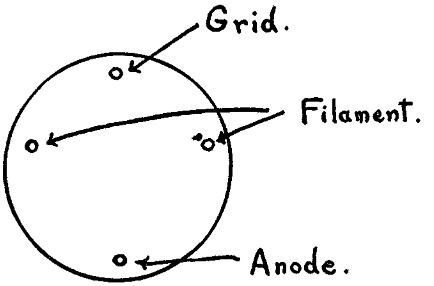


FIG. 10.9. Base of triode

This third structure (the control grid) is so called because it is made usually of a spiral of wire ; it is connected to the outside of the valve in the usual way through the pinch, and can therefore be given any desired potential with respect to the filament. Quite small positive potential on this grid breaks up the space charge and causes relatively large increases in the plate current ; volt for volt it has a much greater effect on the current than the anode, which is so much further away from the filament.

Such is the TRIODE ; we now have the anode, grid and filament to deal with ; there will be four contacts on the base of the valve, two of them being filament leads ; they are usually arranged as shown in fig. 10.9.

Fig. 11.9 (a) shows the conventional representation of a triode whilst fig. 11.9 (b) is a diagram of a typical triode.

Again we proceed to investigate the voltage-current relations, assuming that the filament is maintained at the constant temperature for which it is designed ; but the electron stream (or anode current, I*a*) is dependent on two things (1) the anode—filament voltage V*a* and (2) the grid-filament voltage V*g*.

First we plot a series of graphs relating I*a* and V*a*, for various constant values of V*g* ; these characteristics we shall call the I*a*–V*a* characteristics ; they really tell us all we want

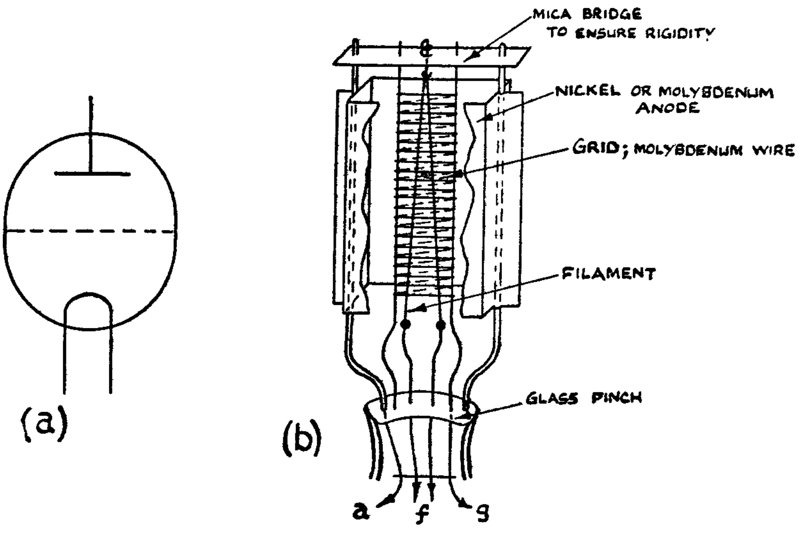


FIG. 11.9. Triode

to know about the valve, but it is customary to plot also what are called the " mutual characteristics " or Ia–Vg characteristics relating anode current to grid voltage for fixed values of anode voltages. These curves should be plotted in the laboratory; fig. 12.9 shows the type of circuit which should be employed.

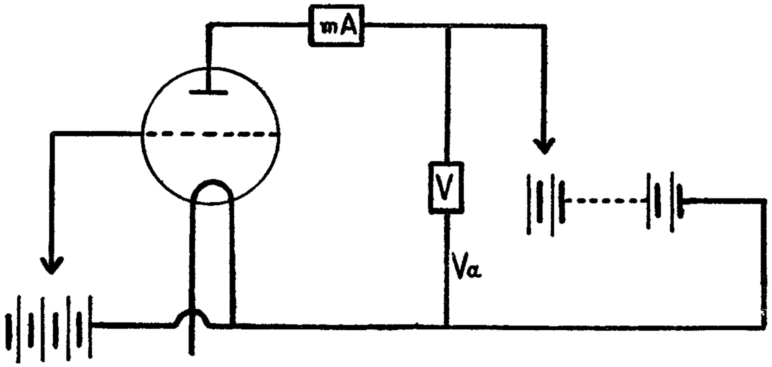


FIG. 12.9. Triode testing circuit

First, with $Vg = 0$, etc., plot a range of values of Ia against Va; repeat the process with $Vg = -2, -4, -6$, etc., till a whole family of curves is obtained. [Fig. 13.9 (a).]

Secondly, with $Va = 80$ volts, plot a range of values of Ia against Vg from $Vg = -12$ $Vg = +4$; repeat with $Va = 70$, 60 and 50 volts and obtain a family of mutual characteristics. The curves should be something like those shown in fig. 13.9 (b).

The first family of curves (the Ia–Va characteristics) show how the anode current is controlled by the anode voltage for a fixed value of grid voltage and they are of the same type as those we met in the case of the diode.

They afford us immediately a value of the internal slope resistance of the valve. By the same convention as used in the diode case this resistance is taken to be

$$\frac{\text{small changes in voltage } (Va)}{\text{corresponding changes in current } (Ia)}$$

taken over the straight part of the graph. A convenient shorthand for " small changes in " is the Greek letter Δ (delta); thus a small change in the anode voltage would be written $\Delta\, Va$ and a small change in the anode current, $\Delta\, Ia$, and with this notation the slope resistance of the valve R*a* would be written

$$\mathrm{R}a = \frac{\Delta\ \mathrm{V}a}{\Delta\ \mathrm{I}a}.$$

The same convention can be applied when a *curved* part of a characteristic is examined, on condition that the changes taken are so small that the graph is approximately straight over the range considered.

The second family of graphs are called mutual characteristics and they tell us how, for a fixed anode voltage, the anode current is controlled by the grid voltage ; we can read off immediately how much increase of current follows a given small increase of grid voltage. The ratio of these two changes is called the *mutual conductance* (g_m) and is expressed in milliamps per volt

$$g_m = \frac{\Delta\ \mathrm{I}a}{\Delta\ Vg}.$$

The reader should determine from the graphs obtained in his experiments the value of both R*a* and g_m for the valve used.

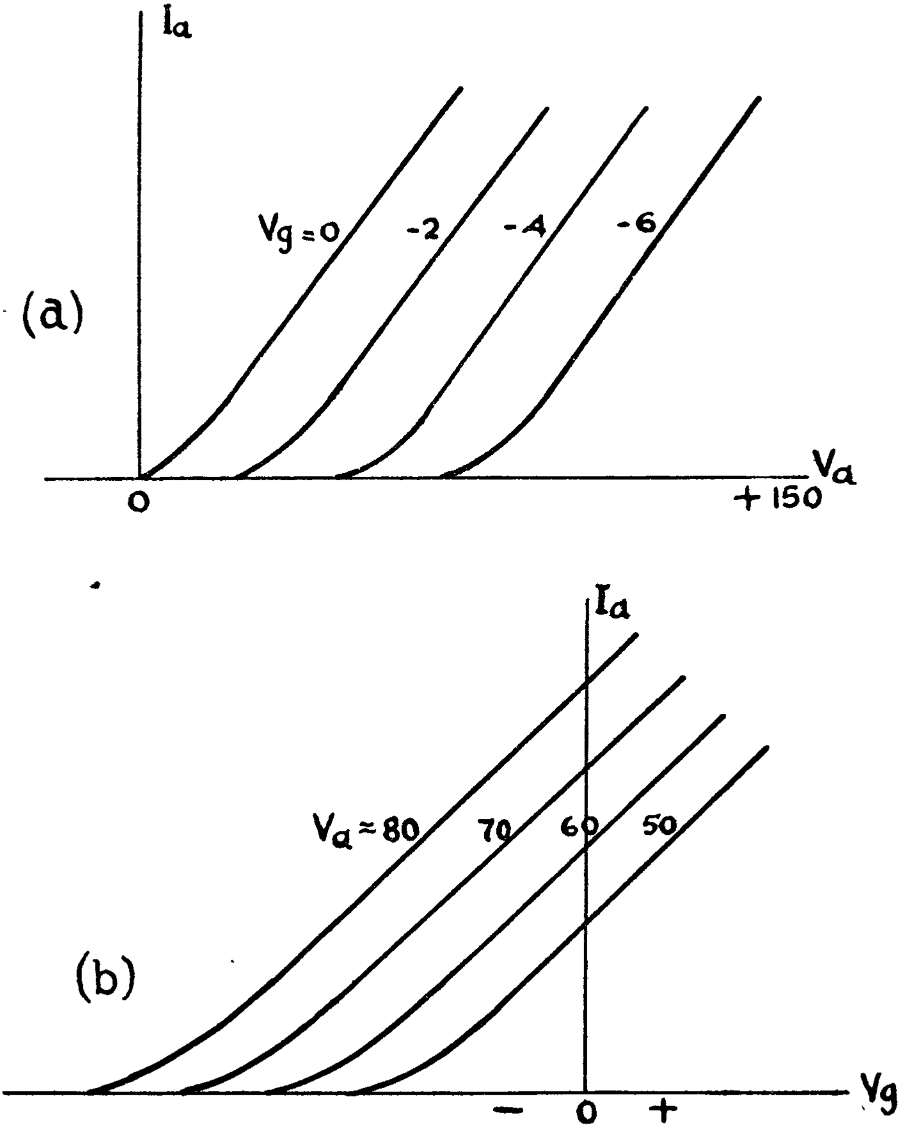


FIG. 13.9. Triode I*a*-V*a* and I*a*-V*g* characteristics

It will be observed that the anode current can be increased by either (1) increasing V*a* or (2) increasing V*g* (i.e. making V*g* more positive), and of the two methods the second is more effective.

The relative effect of changes of V*a* and V*g* is expressed by the amplification ratio (factor) or μ of the valve. This is defined as the ratio :—

$$\mu = \frac{\text{Small change of } Va}{\text{Small change of } Vg},$$

the changes being such as to keep constant the anode current, i.e.

$$\mu = \frac{\Delta\, Va}{\Delta\, Vg}; \ (Ia \text{ constant}).$$

It will be left to the reader to show that $\mu = g_m \times Ra$.

Triode and alternating grid voltages

Let us concentrate on the mutual characteristics and consider the effect on the anode current of alternating voltage on the grid, i.e. alternating voltages applied between grid and filament. We will not apply this voltage directly, but in series with some direct voltage, which we shall call the *grid bias*.

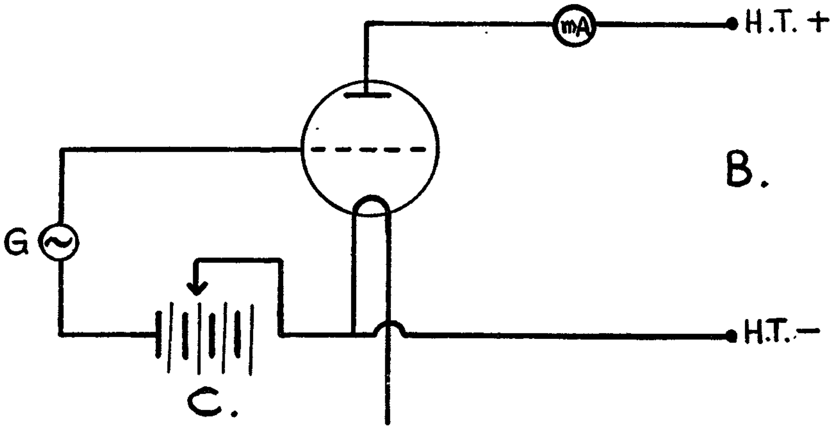


FIG. 14.9. Triode with alternating grid input

Fig. 14.9 represents this state of things ; the H.T. battery B maintains a steady P.D. between anode and filament ; the generator G supplies alternating voltage to the grid-filament circuit, and the bias battery C maintains the average grid potential at any set value by suitable adjustment of the tapping. Using the nomenclature already employed in the case of the diode, we can say that the bias battery allows us to fix the " operating point " at whatever value we want. It will simplify matters if we take the amplitude of our generator at 1 volt. Inspection of the graphs, fig. 15.9, will show immediately that the results of the alternating grid voltage on the anode current will be very different according to the value of the bias. Two general cases emerge (1) if the operating point is so chosen that the whole of the grid swing falls under the straight part of the I*a*–V*g* curve and (2) if the operating point is such that all or part of the grid swing is under the curved foot of the characteristic.

We proceed to discuss case (1).

Triode as Class A amplifier

The important thing to notice here is that the fluctuations in anode current, resulting from the alternating grid input, are symmetrical with respect to its value before the grid swings were applied. It is usual to call the steady, undisturbed anode current the " pre-signal" current because in practical cases the alternating grid voltage is provided by a tuned circuit whose oscillations constitute a signal and the expression " pre-signal current level" is commonly used to denote the value of the undisturbed anode current. Fig. 15.9 (a) shows a simple form of circuit in which the tuned circuit TC, deriving its alternating current from an aerial Ae, takes the place of the A.C. generator of fig. 14.9.

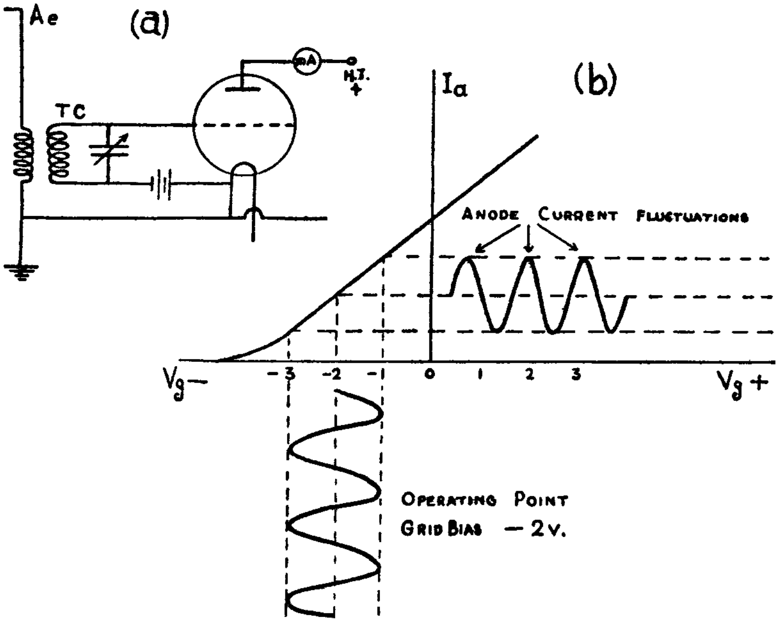


Fig. 15.9. Triode as Class A amplifier

The milliammeter in the anode circuit will read just the same with or without the alternations on the grid and further the fluctuations of anode current are an exact picture of the grid voltage variations ; there is no distortion of the input under these circumstances. Note also that the grid never becomes positive and therefore no grid current flows ; as pointed out before, this is a distinct advantage because grid current would load the tuned circuit and make it behave as if shunted with a resistance, resulting in loss of gain and selectivity.

If we put a resistance in the anode circuit, fig. 16.9 (a), the fluctuating anode current will set up an alternating volts drop across it. Let us examine the magnitude of this volts drop, and compare it with the input voltage.

We saw on p. 98 that x volts in the grid circuit are equivalent to μx volts in the anode circuit, μ being the static amplification ratio of the triode.

If, therefore, our grid swing has an amplitude of V*g* volts, we are effectively applying μ V*g* alternating volts in the anode circuit of the valve. Fig. 16.9 (b) shows the equivalent circuit ; here we see μ V*g* volts acting on the anode circuit consisting of the load resistance R and the valve resistance R*a* in series.

We are only considering alternating quantities in this discussion ; the steady components of anode current and voltage do not concern us. It is easy to see from the equivalent circuit that the alternating voltage developed across the load resistance is $\frac{R}{R + Ra} \cdot \mu Vg$, which is $\frac{\mu R}{R + Ra}$ times bigger than the input voltage V*g*. For this reason the expression $\frac{\mu R}{R + Ra}$ is called the *stage gain* of the system.

A little consideration will show that the greater we make R the more nearly will the stage gain approach μ, the static amplification ratio ; but increasing R will also increase the D.C. volts drop across it and in consequence we shall have to employ a much higher H.T. voltage to operate the valve at all ; the choice of the value of R must therefore be a compromise between these two considerations.

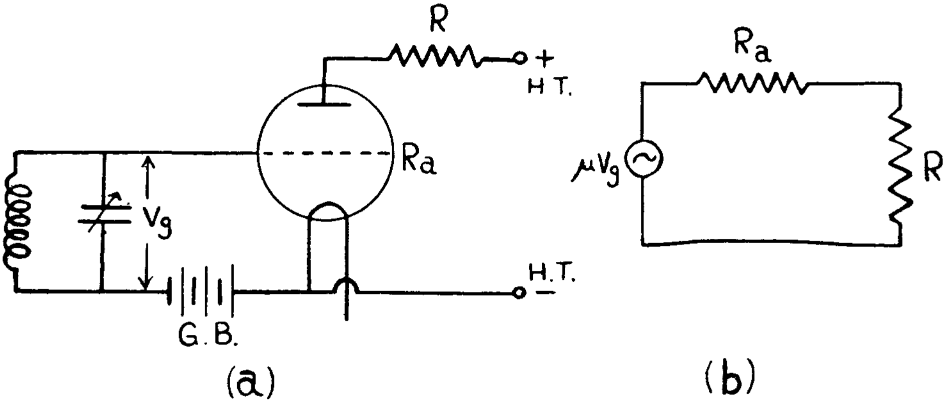


FIG. 16.9. Triode with anode load

This type of amplification in which the grid swing is limited to the straight part of the I*a*–V*g* characteristic, we shall refer to as Class A amplification. We shall discuss in later chapters types of amplification in which this limitation does not apply.

Triode as Rectifier—Anode rectification

Case (2) in which the grid swings are wholly or partly under the curved part of the I*a*–V*g* curve is illustrated in fig. 17.9. The input is shown with a mean operating point of

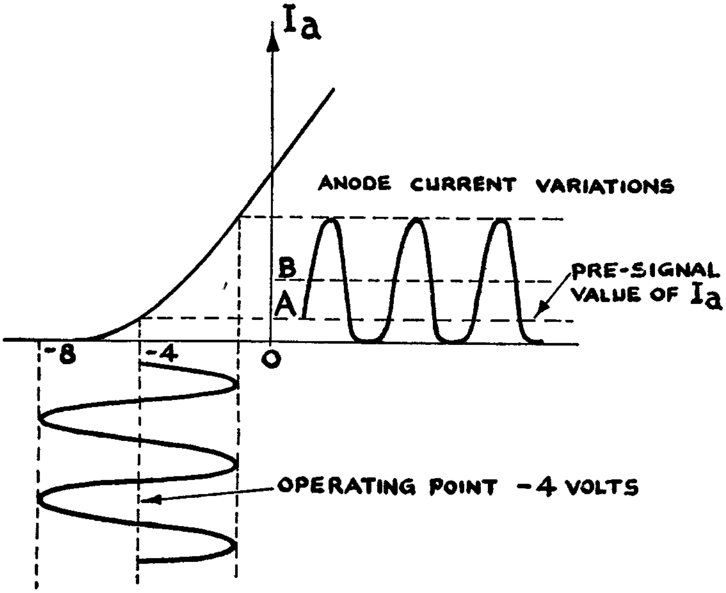


FIG. 17.9. Anode rectification

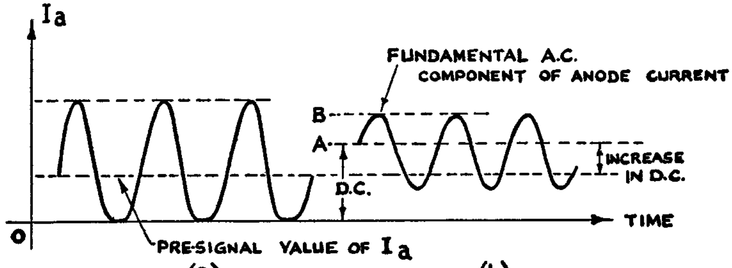
—4 volts ; before the input operates the value of I*a* is OA ; the signal causes the anode current to rise and fall but the increase of I*a* due to positive pulses is greater than the decrease of I*a* due to negative half cycles of grid voltage, so that the mean value of the anode current rises from OA to OB.

This is a type of rectification called *anode rectification*; the rectification will become more pronounced as the operating point moves left, and generally the bias is adjusted so that the anode current pre-signal value is nearly zero; in wireless language this condition would be described by saying " the valve is biassed to anode current cut-off." We could put a load, say a resistance, in the anode circuit of a valve acting in this way, and the rectifying action would produce across it two things :—

(1) An increase in direct volts drop due to the mean increase in anode current.

(2) An alternating voltage at signal frequency.

This is illustrated by the curves of fig. 18.9; the anode current varies as shown in fig. 18.9 (a); this variation which is a distorted A.C., is equivalent to D.C. of value OA and an A.C. of amplitude AB, fig. 18.9 (b). Strictly speaking, of course, there will also be alternating components of twice, thrice, etc., the frequency of the input signal, but for the sake of simplicity we have omitted them in this case. It is really another application of Fourier's theorem, which we have mentioned previously; the distorted A.C. can be analysed into a D.C. plus alternating components of fundamental frequency and harmonics.



(a) (b)

FIG. 18.9. Analysis of " distorted " A.C.

We can easily separate out these two effects a radio frequency choke will pass the D.C. but will send the signal frequency A.C. through the condenser to earth (fig. 19.9.) In such a system the alternating input produces an increase of the *direct* voltage across the load.

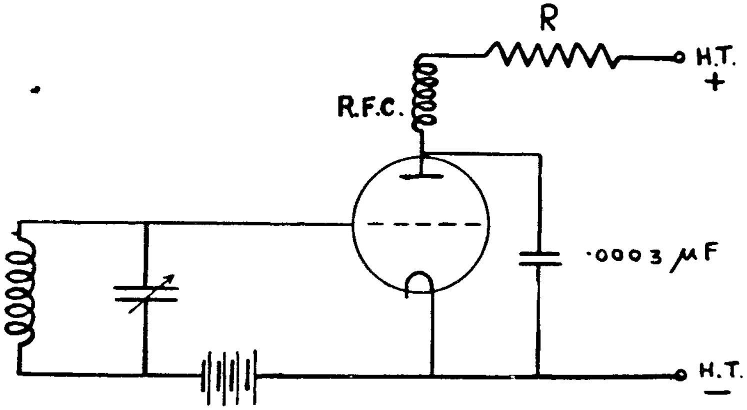


FIG. 19.9. Circuit for separating products of rectification

The relation between this direct voltage produced across the load and the amplitude of the input is not so simple as in the case of diode rectification. It is proportional to the square of the input amplitude, and for that reason anode rectification is what is called square law rectification. A little consideration will show that for this reason the rectified voltage will fall off very rapidly as the signal amplitude goes down; an anode rectifier we say is insensitive to weak signals. For very large signals, rectification is linear.

The valve as an oscillator

We noted in Chapter VI that an oscillatory circuit could not maintain its oscillations for more than a few cycles, owing to losses of various kinds. In Chapter VII we saw that one way of keeping such a circuit in oscillation was to excite it, as a parallel rejector circuit, with alternating voltage at its resonant frequency. This method is extremely limited, because it is so difficult to design a generator with a sufficiently high frequency output to excite the kind of oscillatory circuit used in wireless as the basis of transmitters. The triode solves the difficulty for us at once. As we saw on p. 99, a triode can function as an amplifier developing voltage variations across an anode load which are a magnified copy of the voltages applied between grid and filament; the energy necessary for the magnification, of course, has been supplied by the high tension battery in the anode circuit.

Now if a portion of the magnified output could be returned to the grid, the valve would be supplying its own input, independently of any external supply and under these conditions it would be acting as a sustainer of oscillations, or less correctly, as an oscillator. Consider the circuit of fig. 20.9.

L_1C_1 is a closed oscillatory circuit, in the anode circuit of the triode T; L_2 is a coil, coupled magnetically to L_1, and joined between grid and filament of the valve. If L_1C_1 were set in oscillation by some means, then by mutual induction between L_1 and L_2, oscillations of the same type will be impressed on the grid circuit; these oscillations by amplifier action will tend to reproduce themselves in the anode circuit, and can be made to sustain the oscillations in L_1C_1. Let us examine this action a little more closely.

We saw on p. 61 that if a charged condenser were connected to an inductance of low resistance, the discharge would be oscillatory, but on account of things like radiation, resistance and so on, energy is lost and the oscillations are quickly damped out. To maintain the oscillation at a steady amplitude, we might arrange to join a battery across the condenser for an instant in every cycle, just to bring it up to its original potential again. To do this mechanically one hundred times per second, not to mention several million times per second, is quite out of the question, but a triode provides an easy way of doing it. If the grid is positive, anode current flows (see graph on p. 97) and the condenser C_1 is effectively connected to the H.T. battery through a high resistance, the conducting valve. If the grid is made sufficiently negative then no anode current can flow, the valve has assumed a very high resistance, and the LC circuit has in effect been disconnected from the H.T. battery, and the condenser will begin to discharge and set up oscillations, which will decay unless the condenser is re-charged. This is where the coil L_2 comes into play; every cycle of oscillation in L_1C_1, by mutual induction drives the grid through a positive and negative cycle of voltage, and we so arrange L_2 that just when the condenser is returning to its original state, at the close of one cycle of oscillation, L_2 makes the grid sufficiently positive to allow anode current to flow and raise the condenser

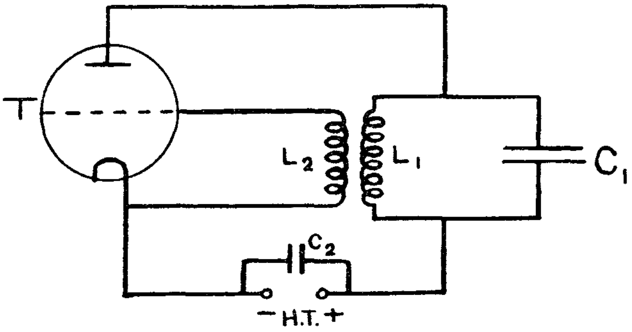


FIG. 20.9. Simple oscillator

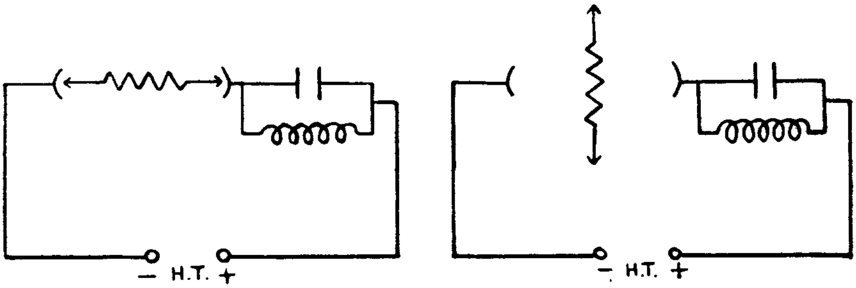
charge to its original value. In this way oscillations are maintained and the system is called a valve oscillator. This interaction between the anode circuit of a valve and its own grid circuit is called *reaction* or *feedback.*

The frequency of the oscillations is fixed by the L_1C_1 circuit; there is no need to design a generator to drive the circuit; it drives itself by returning a portion of its oscillation to the grid of the valve, which acts as an automatic switch, opening the circuit during negative half-cycles, and closing it during positive half-cycles. The condenser C_2 in fig. 20.9 is across the H.T. battery; we shall meet this arrangement over and over again in later chapters, and in all cases its function is to provide a low reactance path for the alternating component of anode current, which otherwise would waste energy in the internal resistance of the H.T. battery.

Notice that, in an oscillating valve the energy used up in the oscillating circuit as heat or radiation, or whatever it is, is supplied by the H.T. battery, which is called upon to make up the condenser charge every cycle. If the losses are very low we can arrange for the charging to be done once every two or three cycles; this is the principle used in a valve-operated frequency multiplier.

Again, if a large negative grid bias is used, the anode current will be zero for a large portion of the cycle, and during such time the valve losses will be zero and the efficiency of the system will be improved. It is often arranged in practice that the anode current flows for less than one-quarter of each cycle.

The switching action of a valve can be followed from the diagrams of fig. 21.9.



(a) V_g Positive (Switch Closed)
• Condenser Charges

(b) V_g Negative (Switch Open)
Condenser Completes Cycle
Of Discharge.

FIG. 21.9. Illustrating switching action

The grid coil L_2 must, of course, be arranged so that the grid becomes most positive just when the condenser returns to its original state of charge; with L_2 one way round this condition is fulfilled; with L_2 the other way round oscillation will not be maintained; the switch will be out of synchronism as it were. It is a matter of simple trial and error, in practice, to find which is the correct orientation of the grid coupling coil.

This type of oscillator in which the feed back from the anode circuit to the grid circuit is accomplished through mutual induction is known as the Meissner oscillator; other types will be dealt with in a later chapter; the present description is intended as a typical example of the way in which a triode can sustain electrical oscillations.

Automatic grid bias for oscillators

The efficiency of an oscillator is increased by the use of negative grid bias, because, as we have already noted, negative bias will reduce the average anode current and lower the valve losses. If we wish to increase the power output from a valve maintained oscillator we must raise the anode potential and at the same time increase the negative bias and the feed-back ; we therefore seek some automatic device whereby the valve is correctly biassed under any value of anode potential.

The grid leak and condenser method previously mentioned on p. 94 provides us with all we want ; during those parts of the cycle when the grid is positive [fig. 22.9 (a), between O and O^1], electrons flow to the grid and charge the condenser as shown in fig. 22.9 (b) ; this action goes on until charge is accumulating in the condenser just as fast as it is leaking away through the resistance R, and a steady negative bias is applied to the grid.

If the anode voltage is raised the feed-back is raised, and the grid swing goes further into the positive region ; this causes more charge to flow per cycle into the condenser and increases the bias ; in this way the bias is self-adjusting.

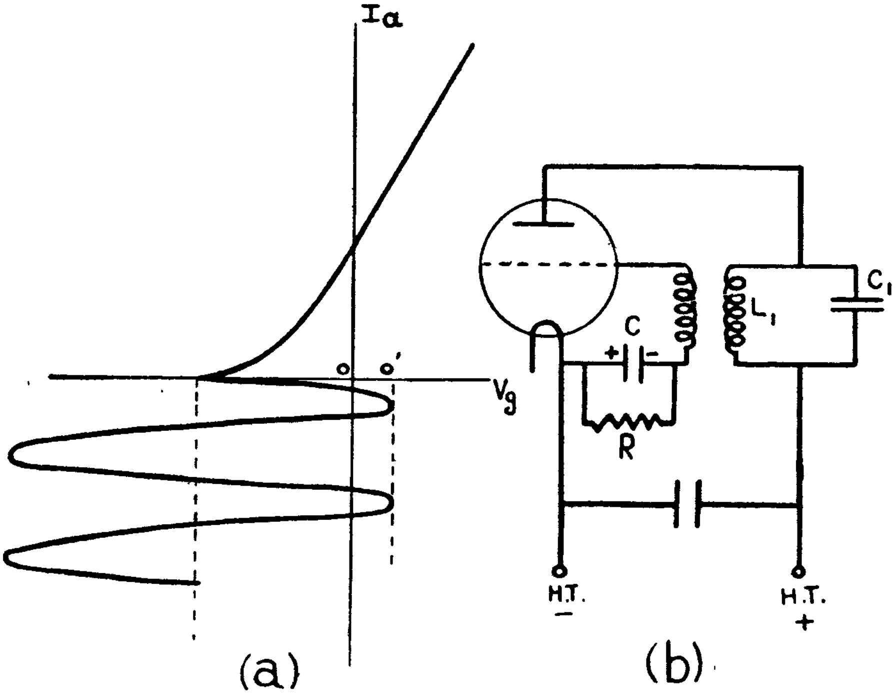


FIG. 22.9. Automatic grid bias

Of course the resistance must not be so big that the bias increases to such an extent that oscillation stops ; the reader should reason out for himself what would happen if the leak resistance were removed altogether.

Finally, should something occur in the anode circuit to stop the oscillations, the automatic grid bias disappears (because it depends on feed-back from the anode circuit) and the valve finds itself with full H.T. and zero grid bias ; under such circumstances it is likely to be destroyed by overheating of the anode on account of excessive current, unless, of course, the fault is due to the anode circuit having become opened.

Inter electrode capacity—tetrode

The anode and grid of a valve, together with the leads connecting them to the external connecting pins, constitute a condenser ; allowing for this internal capacity a triode can be represented as shown in fig. 23.9.

The anode-grid capacity, small though it is, has a very big effect on the operation of the valve at high frequencies ; attempts to minimise it have given rise to the 4-electrode valve or *tetrode*. A conductor can be screened from the electrostatic influence arising from other conductors by the interposition of an earthed screen ; in the case of the valve we could interpose an earthed screen (fig. 24.9) between the control grid and the anode, and thus eliminate the electrostatic action of one on the other, but a screen at zero potential would also screen the electrons emitted by the filament from the accelerating effect of the anode, and the valve would cease to function ; so we actually maintain our screen at about 60 or 70 volts positive to the filament, and at the same time take the anode connection out through the top of the valve to avoid capacity in the leads through the pinch. Fig. 24.9 is a conventional diagram of a TETRODE.

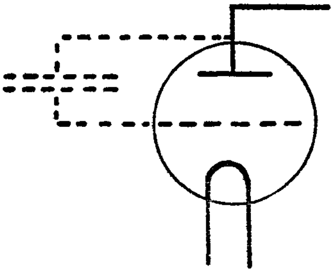


FIG. 23.9. Internal anode-grid capacity

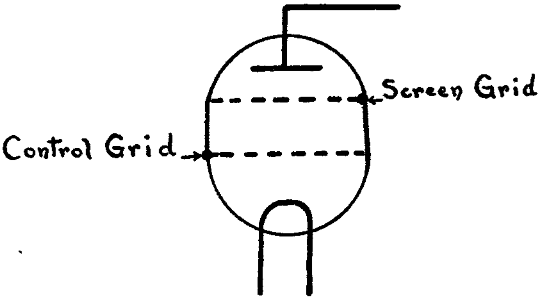


FIG. 24.9. Screen-grid valve (tetrode)

The characteristics of a tetrode have some interesting features ; the I*a*–V*a* characteristics are shown in fig. 25.9.

The fall in these curves is due to what is called " secondary emission " ; when an electron, which has gained sufficient velocity by its passage through a potential difference, hits a metallic surface, it actually dislodges some of the conduction electrons in the surface ; it is something like a stone thrown into a pond ; the stone represents the original electron and the drops of water in the " splash " represent these " secondary emission " electrons. The tetrode I*a*–V*a* curves indicate that up to about 18 volts (line AB, fig. 25.9) the velocity of the electrons is not sufficient to cause secondary emission at the anode, but at about 20 volts this effect sets in, and because the *screen is at a higher*

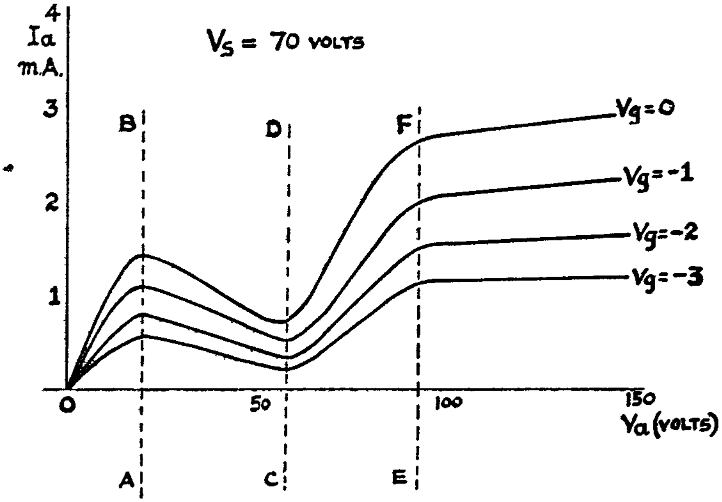


FIG. 25.9. Tetrode I*a*-V*a* characteristics

potential than the anode, the secondary electrons pass to the screen and reduce the anode current ; this effect lasts till the anode potential is higher than that of the screen, when the curves rise again in the region of the line CD.

The low slope of these curves beyond EF indicates that the screen grid valve has a very high A.C. resistance. The mutual characteristics are very much like those of the triode and the mutual conductance is of the same order. It follows from this that the amplification factor of a tetrode is much higher than that of a triode.

The pentode

There is a difficulty with the tetrode ; if the anode potential should fall below that of the screen, the valve is working on a " falling " part of its characteristic (i.e. between the lines AB and CD) and its behaviour becomes very erratic.

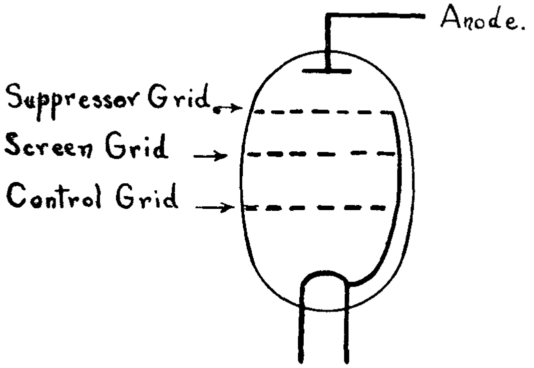


FIG. 26.9. Pentode

Since this falling characteristic is due to secondary emission from the anode, it is to the elimination of this effect that we must look for improvement, and it is done by the addition of still another electrode, making the valve into a PENTODE (see fig. 26.9).

A suppressor grid of very open construction, connected internally to the filament is inserted between anode and screen grid. Any secondary electrons emitted from the anode are thus protected from the attractive influence of the screen, and they pass back in due course to the anode from which they were displaced. Thus the effect of secondary emission is overcome, and the I_a-V_a characteristics of a pentode have no negative resistance kink.

The additional grid has its effect on the A.C. resistance of the valve ; it is of a higher order than that of the tetrode.

The variable mu valve

On p. 99 the stage gain m, afforded by a triode valve of amplification ratio μ and A.C. resistance Ra working into a resistive load of R ohms is given as

$$m = \frac{\mu R}{R + Ra}$$

The same expression holds for any type of valve with a purely resistive anode load, as, for example, a screen grid amplifier with a tuned anode circuit, for a tuned circuit behaves as a resistance at resonance.

Now μ for an average tetrode is about 500 and Ra is about 1 megohm under working conditions. The dynamic resistance of a normal tuned circuit is about 100,000 ohms. R is only about one-tenth of Ra and so we neglect it in the denominator of the last equation in comparison with Ra, and the expression may be written

$$m = \frac{\mu}{Ra} \cdot R$$

But
$$\frac{\mu}{Ra} = g_m$$

$$\therefore\ m = g_m \cdot Ra$$

This brings out the important fact that the stage gain afforded by a screen grid amplifier depends on the mutual conductance of the valve—the larger the slope the bigger the stage gain.

The effective slope of a valve characteristic depends on the grid bias applied, being low for large negative values and high in the region $Vg = 0$. Therefore the stage gain of a valve amplifier can be varied by varying the grid bias applied to it, and in doing this we have a convenient form of gain control.

In a valve of normal design the I*a*–V*g* characteristic ends sharply (see curve B, fig. 27.9) and there is an abrupt change of slope over a small change of V*g*. This type of valve is unsuitable for gain control purposes. It is necessary to have a gradual change of slope over a fairly wide range of grid bias values.

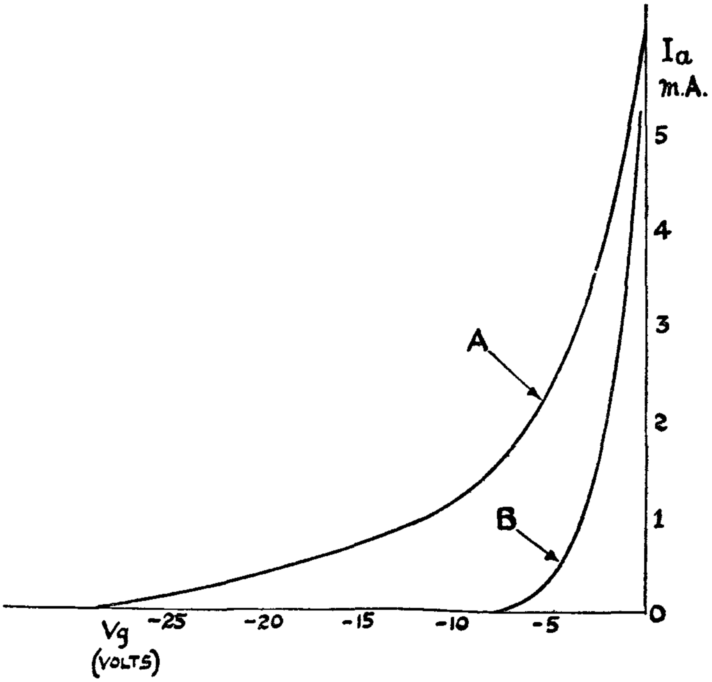


FIG. 27.9. Variable-μ valve I*a*-V*g* characteristics

An I*a*–V*g* characteristic of the desired shape may be obtained by winding the control grid in an irregular manner, i.e. by leaving gaps in it at intervals along its length. The result of this is that some electrons from the space charge are able to reach the anode, even with a large negative grid bias, and the I*a*–V*g* curve takes a shape similar to that of curve A in fig. 27.9. By suitable choice of grid bias, any one of a wide range of slope values can be obtained, and the amplification given by the valve varies accordingly. The corresponding characteristic of a valve with a normal grid is given in curve B for comparison purposes.

A valve which has an I*a*–V*g* characteristic similar to that of curve A, fig. 27.9, is called a VARIABLE–MU valve. This is really a misnomer, for the μ of a valve is fixed by the geometrical design of its electrodes, and it is g_m which is varied.

Certain pentode valves have similar characteristics.

Compound valves

It still remains to describe several valves, such as the " double diode triode " the " hexode " and the " heptode ", which are really combinations of two or three fundamental valves in one glass envelope. These valves will be described later, in connection with those wireless circuits in which the special properties of compound valves find their application.

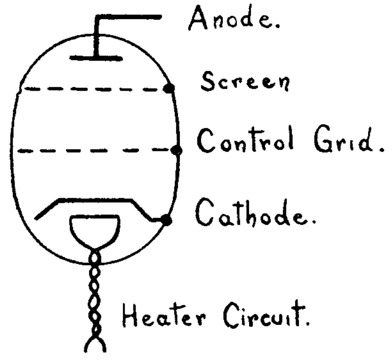


FIG. 28 9. Indirectly heated valve

Indirectly heated valves

Some modern valves do not use a simple hot filament as the emitting electrode. A small tube of nickel, coated with emitting substance, carries an insulated " hairpin " of wire inside it, and this loop is heated directly by means of low tension current. The heat thus generated soon raises the tube to its emitting temperature, when it supplies electrons to operate the valve. The advantage is that the heater can be operated with A.C. if necessary, because it does not take part in the actual operation of the valve. It is usual to call the indirectly heated emitter the *cathode.*

The conventional representation of an indirectly heated valve is shown in fig. 28.9.

CHAPTER X

THE SIMPLE TRANSMITTER

In Chapter VIII we saw how an oscillatory circuit with a suitable " open " condenser consisting of aerial and earth could radiate electromagnetic waves into the surrounding space.

The frequency of this radiation was the same as that of the oscillating circuit and, other things being equal, the higher the frequency the more was the energy radiated every cycle.

In the next chapter (Chapter IX) we saw how an oscillating circuit could be maintained in continuous (undamped) oscillation by means of a triode valve, operating under suitable conditions of feed-back from the anode to the grid circuit.

The combination of these two considerations brings us to the conception of a simple valve transmitter. If we " open " the condenser in the tuned circuit of the valve oscillator, making it into the usual aerial and earth system, we shall have established one of the conditions for radiation, namely a diffuse electrostatic field. In so doing, of course, we shall enormously reduce our capacity and thereby bring about a big increase in frequency. It may well be, however, that the substitution of aerial and earth for condenser may stop the oscillation altogether in so simple a circuit as we are considering : we therefore proceed to introduce some refinements, which will ensure that the transmitter functions properly and with some degree of flexibility.

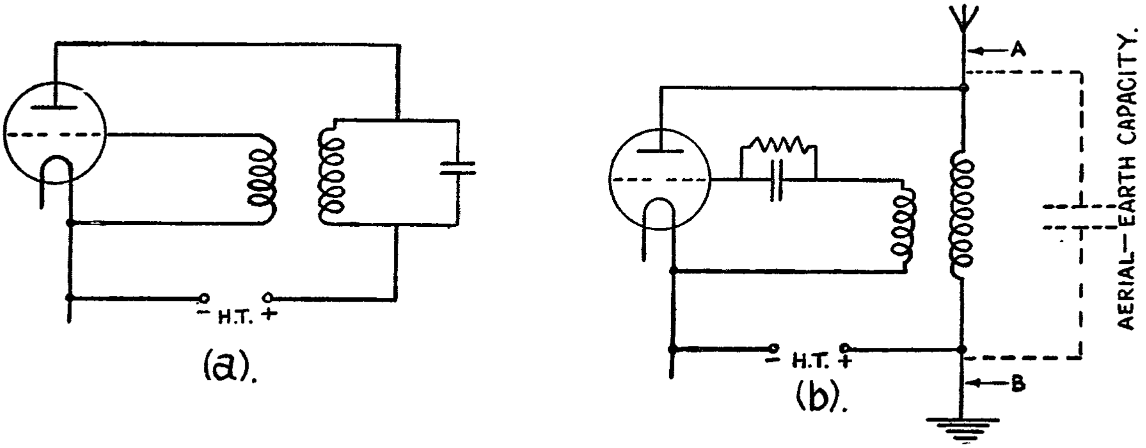


FIG. 1.10. Simple valve oscillators

Fig. 1.10 (a) is the simple valve oscillator previously described.

Fig. 1.10 (b) represents the first attempt to turn such an oscillator into a transmitter ; even assuming that the circuit continues to oscillate, it is obvious that it now contains certain undesirable features :—

(1) The whole aerial is at high potential.

(2) There is a risk that the H.T. supply may be shorted, because its positive end is definitely earthed and its negative end is connected to the filament battery, which is likely to be earthed.

(3) The high-frequency pulsations of anode current will have to pass through the H.T. source, which will certainly present a fairly high impedance.

(4) There is no device for matching the tuned circuit to the valve. A valve, like any other generator, works best (i.e. gives maximum output) with a certain value of load impedance. In the above circuit there is no way of altering the impedance of the aerial system to obtain the best output conditions.

(5) There is no way of varying the frequency of the oscillations. As it stands the frequency is that of an oscillating circuit with the aerial coil as inductance and the aerial-earth capacity as condenser, and as such it has a fixed value.

These difficulties can be surmounted without much trouble :—

(1) A large condenser ($\cdot 01\ \mu$ F.) inserted at A will cut the aerial off from H.T.+ and will not interpose any serious reactance to high-frequency currents.

(2) A similar condenser inserted at B will cut off the H.T.+ from earth, and again, will not offer any serious opposition to R.F. currents.

(3) A $\cdot 01\ \mu$ F. condenser in parallel with the H.T. supply will provide a suitable R.F. by-pass.

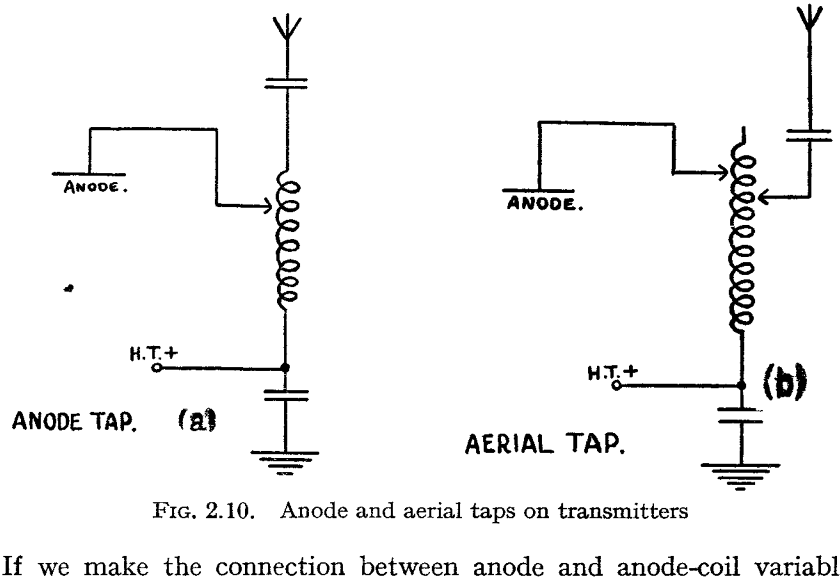


Fig. 2.10. Anode and aerial taps on transmitters

(4) If we make the connection between anode and anode-coil variable [see fig. 2.10 (a)] the point of contact can be varied so that just enough inductance is included to allow for matching to give the best output. This device is called an *anode-tap*. It is really an " auto-transformer ", that is, a transformer using only one winding of which the primary uses some few turns and the secondary the whole lot, in the step up case. In the anode-tap of fig. 2.10 (b) the part of the coil between anode-tap and earth is the primary and the part between aerial-tap and earth, the secondary. Just as in an ordinary transformer, we adjust the turns ratio to obtain optimum output conditions.

(5) In order to vary the frequency of an oscillatory circuit we must alter either inductance or capacity, or both. In this, the capacity (Ae to E) is fixed, and we have no choice but to alter the inductance. This is done by means of a variable connection to the aerial. It allows us to alter the number of active turns on the coil and so varies the inductance, and through it the frequency of the oscillations , this is the *aerial-tap* [fig. 2.10 (b)].

Finally, it might be said that the circuit affords no indication of what is going on some indicating instruments are required, for example :—

(*a*) A D.C. milliammeter in the valve anode circuit.
(*b*) A thermo-ammeter in the aerial circuit, to read the aerial current.
(*c*) A voltmeter across the H.T. supply.
(*d*) An ammeter to show the filament current.

These instruments are shown in the completed circuit of fig. 3.10.

We will suppose that such a circuit (fig. 3.10) has been properly set, i.e. the aerial tap has been adjusted to give the desired frequency (determined by a wave meter, see p. 116), and that the anode tap has been adjusted for maximum efficiency and output ; the circuit will then be sending out what we call continuous wave (C.W.) radiation. By means of a key in the H.T. to filament circuit (G, fig. 3.10) we can interrupt this radiation and send it out in long or short bursts corresponding with the dashes and dots of the morse code. This is the method of C.W. telegraphy. The arrangement of the key and the grid leak in the diagram of the simple transmitter (fig. 3.10) is a matter of some importance.

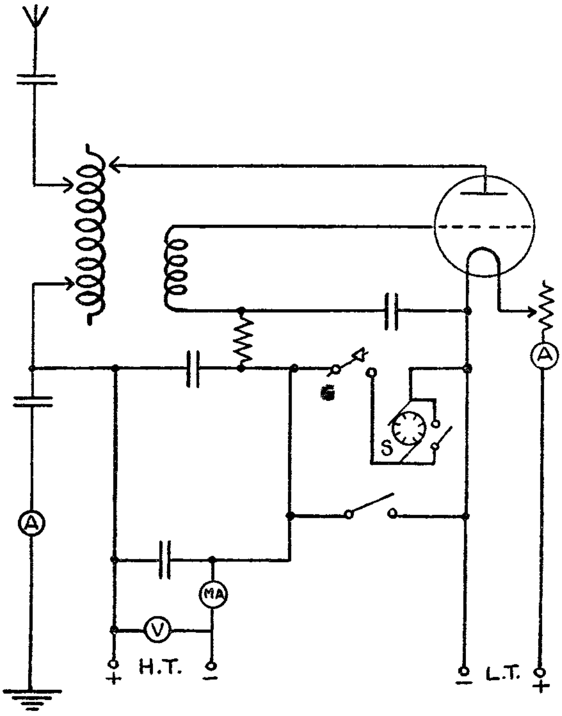


Fig. 3.10. Practical transmitter circuit

When the key is " down ", the grid leak is in parallel with the grid condenser, and this combination, by rectification of the grid input, provides automatic bias (see p. 104, Chapter IX). When the key is " up ", however, it will be seen that the H.T. positive is connected to the anode, and the H.T. negative to the grid via the grid leak, the grid being negative with respect to the anode.

At the same time the filament (still heated and emitting electrons) is isolated from the H.T. negative pole ; it therefore assumes a potential equal to that of the anode. It is a little difficult to see how this may be ; take a simplified case as indicated in fig. 4.10. A, B and C are 3 metallic conductors corresponding to anode, grid and filament, respectively ; A and B are joined to a generator as shown, making A positive with respect to B.

XY is another conductor joining A (anode) and C (filament), but missing B. This corresponds to the electron stream in the valve which can pass from filament to anode but not from grid to anode (because the grid is not emitting) or from filament to grid because of the negative bias. Thus C must have the same potential as A, simply because it is in conducting contact with it. The net result is that the grid is negative with respect

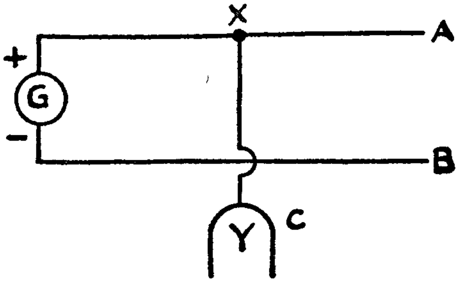


Fig. 4.10. Equivalent circuit of valve when keyed

to the filament, to the extent of the H.T. supply, directly the key is opened.

This action results in a very rapid cessation of anode current, and therefore of radiation, directly the key lifts. Further than this, the placing of the key on the " earthy " side of the H.T. supply protects the operator from electric shock.

Other methods of keying such an oscillator exist ; " plate keying " in general consists in interrupting the power supply to the anode ; grid keying operates on the principle of altering the bias on the grid by means of the key, so that when the latter is open a big negative bias is introduced and plate current is cut off.

The keying described, in the case of the simple transmitter, is a combination of these tw methods.

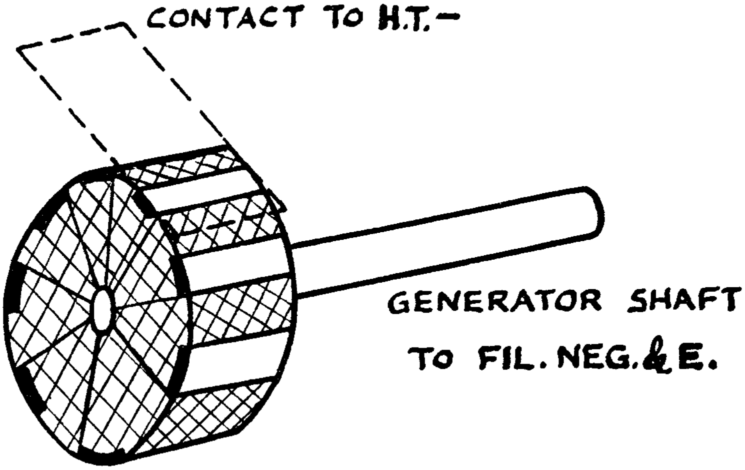
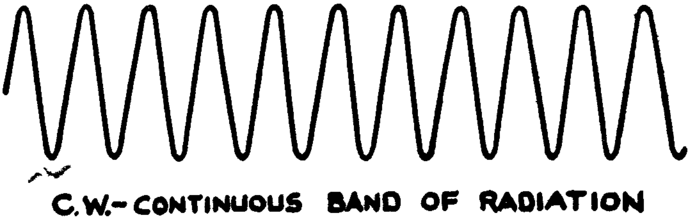


Fig. 5.10. Interrupter or tone wheel

Another type of transmission is what is called Interrupted Continuous Wave radiation or I.C.W. To produce such radiation we have a rotary switch (S, fig. 3.10) in addition to the morse key, which makes and breaks the circuit many hundreds of times per second ; at each " make " the circuit sends out a train of radiation, and during the " break " it is quiescent. The rotary switch is called a " tone wheel " and is generally driven off the shaft of the motor generator supplying the H.T. voltage for the valve.

The difference between these two types of radiation is indicated in fig. 6.10 ; complete pictorial representation is impossible, because the radiation frequency is so high compared with the pulse frequency of the I.C.W. The difficulty might be compared with that of drawing a lump of sugar to show all its component molecules. The sine-waves in fig. 6.10 merely *indicate* radiation.



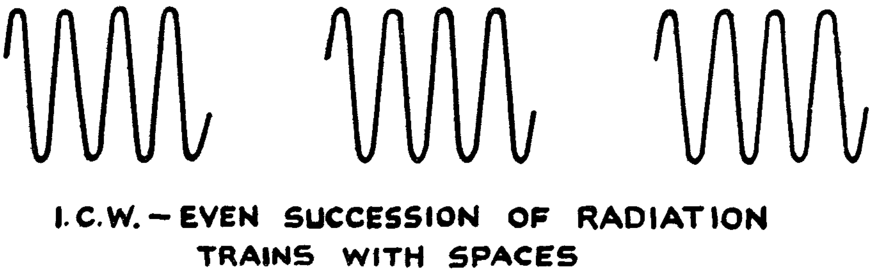


FIG. 6.10. Illustrating C.W. and I.C.W.

This difference will be appreciated better when we come to look at C.W. and I.C.W. from the receiving end ; here we are more concerned with the methods employed to produce them with the consideration of their application in the practice of signalling.

Adjustment of transmitter

The adjustment of a simple transmitter such as described above involves two operations, assuming that the filament current is correct and the H.T. supply functioning :—

(1) Adjust the aerial tap until the transmitted frequency is correct.

(2) Adjust the anode tap until a satisfactory power output is obtained. This adjustment will also have a very slight effect on the frequency and it may necessitate a further slight re-adjustment of the aerial tap.

In practice, both adjustments are made until :—

(1) A satisfactorily large aerial current is obtained showing good radiation.

(2) The steady D.C anode current is low indicating a reasonably high efficiency.

Aerial Coupling

The aerial shown in fig. 3.10 is a " directly excited " aerial ; if so desired, an aerial can be coupled to the oscillating circuit by means of a double wound transformer ; this device isolates the aerial from the H.T. supply and avoids the use of the blocking condenser.

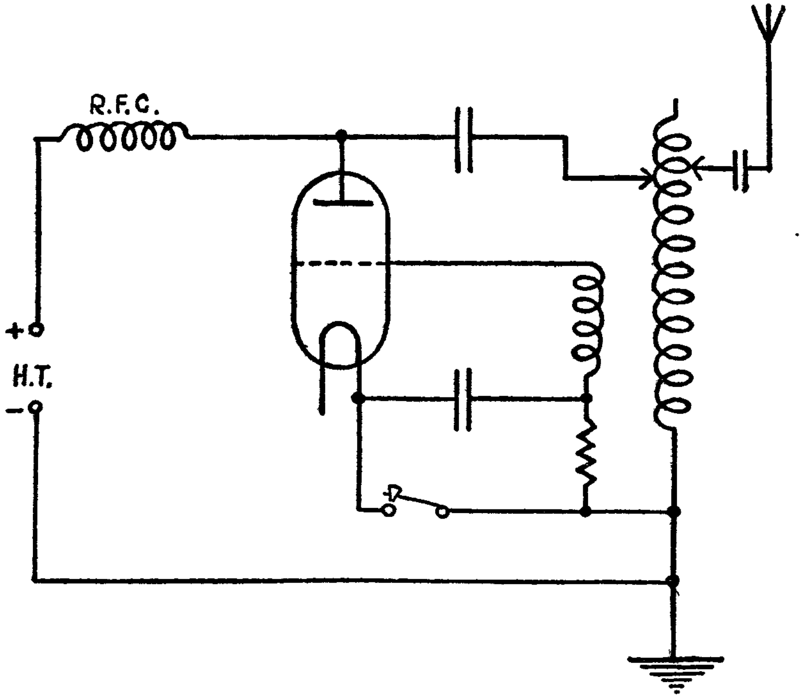


FIG. 7.10. Simple parallel-fed transmitter

Feed

The " feed " of an oscillator refers to the method in which the H.T. supply, oscillating circuit and anode are connected ; in fig. 3.10 we pass from anode through the oscillating circuit to H.T. in series ; for this reason it is called a *series feed* oscillator or transmitter.

The same transmitter arranged for *parallel feed* is shown below in fig. 7.10. An important modification is the R.F. choke which is designed to prevent the radio frequency oscillation from getting through to the H.T. supply. Here the oscillating circuit is in parallel with the steady D.C. anode and H.T. circuit. A blocking condenser is necessary in the anode lead to avoid an H.T. short circuit and also to prevent the aerial from being " live ".

Other oscillators

Up to the present we have only considered the Meissner oscillator, in which the feedback from anode to grid circuits is achieved by mutual induction. There are many other types of oscillator which could be used in transmitting circuits, and a number of standard oscillator circuits can be designed on the following lines :—

(1) Draw the oscillator circuit.

(2) Connect the valve anode to one end of the oscillating circuit and the grid to the other.

(3) Join the centre of either inductance or condenser to earth.

(4) Arrange power supply. Under this heading the following points must be observed.

- (*a*) If direct voltage is applied to any point which is not " earthy " in respect of high frequency voltages, an R.F. choke must be inserted.
- (*b*) The grid must be isolated from H.T. by a condenser of low H.F. reactance.
- (*c*) Some biassing system must be included.
- (*d*) The aerial must be suitably isolated from the direct H.T. supply.

The following diagrams indicate the development of the Hartley oscillator and the Colpitts oscillator, from the principles laid down above.

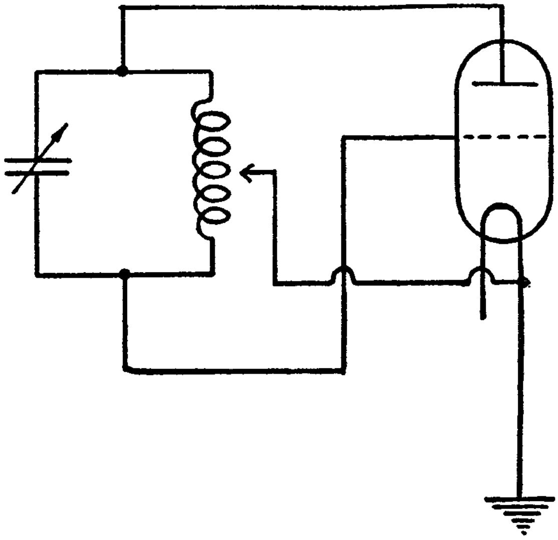


FIG. 8.10(a). Typical oscillator circuit : A.C. conditions only.

No power supply shown

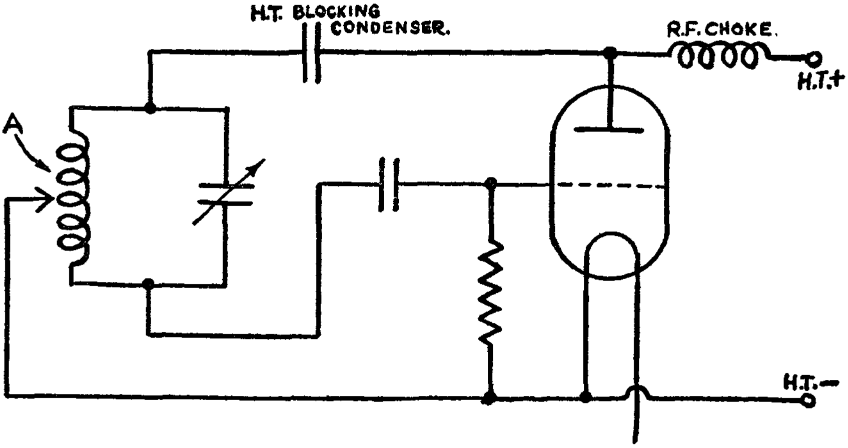


FIG. 8.10(b). Shunt-fed Hartley oscillator

In this type of oscillator the grid " feed-back " is obtained by tapping off a portion of the R.F. voltage in the inductive branch of the oscillating circuit. The amount of grid driving voltage (or " excitation ") is controlled by altering the position of the filament tap A. Normally this tap should be rather nearer the grid end than the anode end of the coil. When series feed is used, the circuit becomes as shown below.

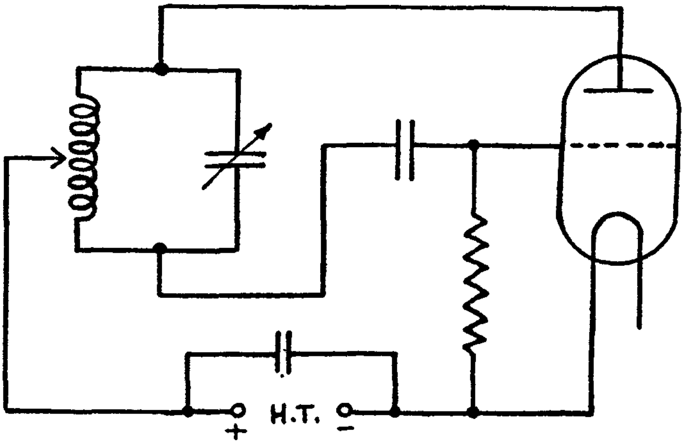


FIG. 8.10(c). Series-fed Hartley oscillator

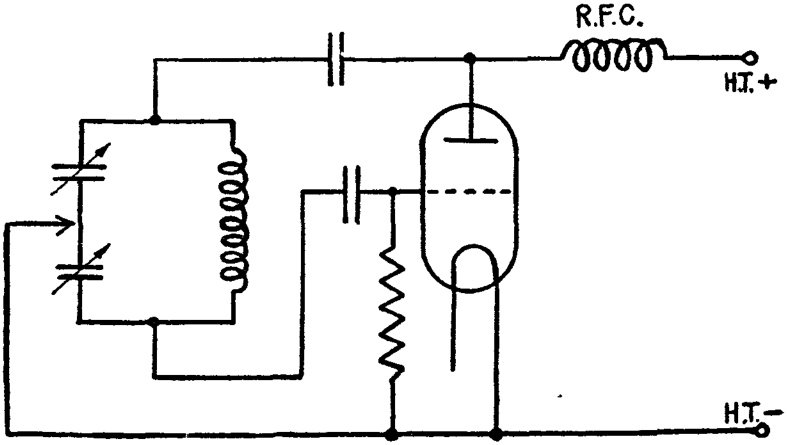


FIG. 8.10(d). Shunt-fed Colpitts oscillator

The feed-back [fig. 8.10 (d)] is obtained by tapping off a part of the R.F. voltage across the capacity branch of the oscillating circuit.

The oscillation in a Colpitts circuit is controlled by the relative value of the two condensers ; normally the grid condenser will have about twice the capacity of the anode condenser.

Another common type of oscillator is the tuned plate tuned-grid oscillator (TPTG). It is shown in fig. 9.10. In this type of oscillator the feed-back is through the internal anode-grid capacity of the valve itself. The frequency is controlled chiefly by the constants of the plate circuit ; the grid circuit is tuned only approximately to the oscillation frequency ; just sufficient tuning is employed to maintain the desired amplitude of the oscillation in the plate circuit ; the tuning of the grid circuit in fact, controls the excitation.

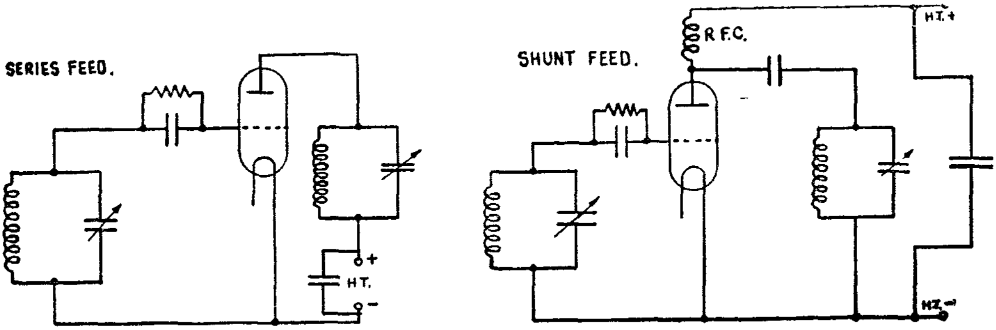


FIG. 9.10. Tuned plate tuned grid (T.P.T.G.) oscillator

The reader should take each of these oscillating circuits and open them out into practical transmitting circuits, just as, at the beginning of this chapter, the Meissner oscillator was developed into a simple CW or ICW transmitter.

Frequency—stability of oscillation

In high quality transmissions it is essential that the frequency of the radiation shall remain constant. A little examination will show at once that the simple transmitter discussed above will fail, in some degree, to achieve a high degree of frequency stability ; the aerial is part of the oscillating circuit, and if it blows about in the wind for instance, the capacity will alter, and with it, the frequency. Generally, the causes of instability fall into two classes :—

(1) Mechanical. (2) Electrical.

Under the first heading fall the results of the vibration of components, including the aerial, and of the effects of temperature. This latter is peculiarly noticeable if a simple transmitter is used under the wide temperature ranges to which aircraft sets are necessarily exposed. Coil and condenser spacings vary owing to thermal expansion and their electrical constants vary as a result. The valve elements heat up and the gradual alteration of their spacing introduces a noticeable frequency creep. Electrical instability includes the results of the alteration of the valve characteristics during operation ; anything which alters the valve impedance as, for example, an H.T. supply with pronounced commutator ripple or a poorly smoothed rectifier supply, gives rise to what is called " frequency flutter ", which is very undesirable.

At ground stations, frequency stability is achieved by what is called the " Master Oscillator " system which has also been applied in aircraft transmitting sets ; M.O. control is described in a later chapter of this book. A much simpler method of control for small power transmitters is by means of a quartz crystal.

A thin slice of quartz exhibits a curious electric phenomenon which we call the Piezo (pressure) electric effect ; if the crystal is compressed it develops a potential difference across its opposite faces, and if a potential difference is applied to the faces of the crystal, the crystal becomes strained.

If a slice of quartz is put between the plates of a condenser under an alternating voltage, the crystal is set in vibration. As the crystal vibrates it will be strained and develop an alternating potential difference. The resultant vibration is at the natural frequency of the quartz crystal, i.e. the latter takes control and maintains oscillations of·very stable frequency.

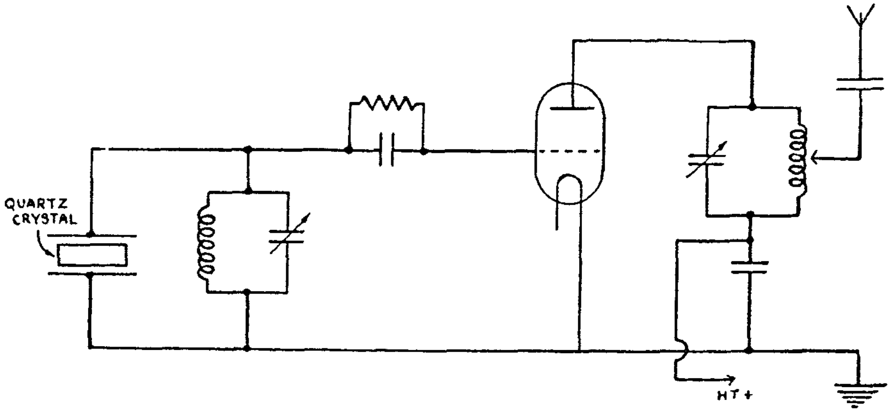


FIG. 10.10. Quartz crystal control

The simplest way of using a quartz crystal in a transmitter is to connect a crystal of the appropriate frequency across the anode or grid coil of the valve oscillator (see fig. 10.10) ; the crystal exercises a strong control over the generated frequency, and a high degree of stability is attained.

In Chapter XIII we shall deal with quartz crystal control when discussing master oscillator systems.

Wavemeters

To determine the frequency of the radiation being sent out by a transmitter, we use a " wavemeter ". There are several types of this instrument and at this stage we shall content ourselves with a study of the simplest type. The others will be described when the principles underlying their actions have been dealt with (see pp. 140, 141).

In essence a wavemeter is a closed oscillatory circuit, in which either the inductance or the capacity can be varied to give a range of accurately determinable frequencies, together with some device to indicate when the current in the circuit is a maximum. When loosely coupled to an oscillator whose frequency is to be determined, alternating current is set up in the wavemeter circuit whose variable element (say the condenser) can be adjusted until the current is a maximum ; the resonant condition indicates that the frequency of the radiation under examination is the same as the natural frequency of the wavemeter circuit.

The frequency corresponding to any setting of the condenser is indicated by a scale fixed to the instrument ; the scale may be calibrated directly in frequencies or it may be marked off in degrees, when it will be necessary to refer to the calibration chart of the instrument to find the frequency corresponding to a certain angular setting of the adjustment spindle.

Maximum current may be indicated in a variety of ways, e.g. by a low voltage pea-lamp or by a thermo-galvanometer.

Fig. 11.10 represents the wavemeter circuit and shows how a pea-lamp can be coupled to it by an R.F. transformer ; the lamp may be wired directly in series with the tuned circuit, instead of using transformer coupling.

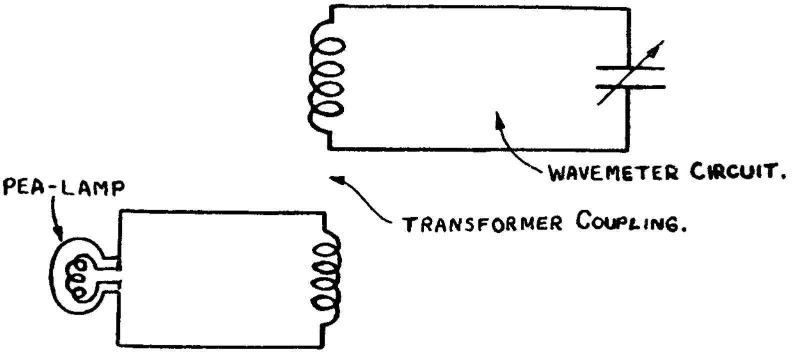


FIG. 11.10. Simple absorption wavemeter

If a coupling coil is used, a small battery is sometimes put in series with the lamp to warm the filament up to the point of a dull glow. A small R.F. current then raises the filament to brightness and a higher sensitivity is obtained. The lamp, however, is not a very good " maximum indicator " and does not lend itself to accurate measurement.

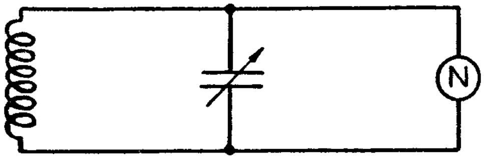


FIG. 12.10. Neon lamp wavemeter

A neon tube is a small glass bulb containing two electrodes and is filled with very low pressure neon gas. An E.M.F. applied to the low pressure neon sets up ionisation and produces the familiar scarlet glow. Such a tube would be connected across the wavemeter as in fig. 12.10, because it is operated by the high voltage which is developed across the tuned circuit at resonance.

With sufficiently loose coupling the tube will flash just for an instant as the adjusting device passes through the resonant point. The neon tube like the pea-lamp, however, is not a good maximum indicator ; the glow has roughly the same intensity at some distance on either side of the maximum.

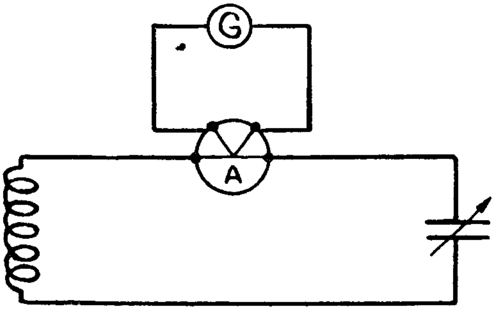


FIG. 13.10. Absorption wavemeter (thermal type)

An indicating device of a higher order of accuracy is the thermo-ammeter or thermo-galvanometer. This instrument was described on p. 38. Its usual disposition in the wavemeter circuit is shown in fig. 13.10.

The R.F. current heats the junction A, which is the meeting point of the elements of a thermo-couple. The galvanometer indicates the current produced by thermo-E.M.F. It has a square law scale and the maximum deflection can be easily determined.

When the thermo-junction is contained in an evacuated glass bulb the instrument becomes independent of external draughts and is capable of a high degree of accuracy. A vacuum thermo-couple, however, is rather fragile and easily burnt out if the wavemeter is roughly used.

Wavemeters can be made to operate over several bands of frequencies by the insertion of fixed parallel condensers.

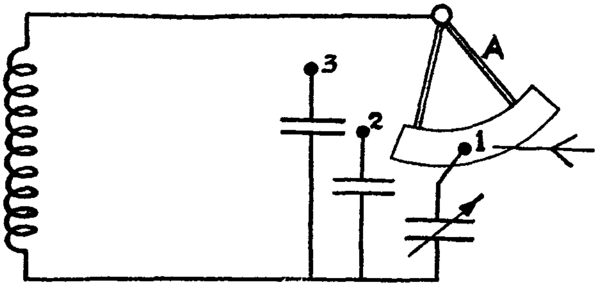


Fig. 14.10. Absorption wavemeter. Range switching

Fig. 14.10 indicates how this is done ; with the range switch A set as shown, the frequency band is determined by the limits of the variable condenser. If the switch is turned clockwise the metal sector makes contact on studs 1 and 2, bringing in the first fixed condenser, and the variable condenser now covers another (lower frequency) range. A further turn brings in fixed condenser No. 2 and gives a third frequency range.

The sharpness of tuning of a wavemeter will depend on the magnification factor " Q " of its tuned circuit which is equal to

$$\frac{1}{R}\sqrt{\frac{L}{C}}$$

where R = equivalent resistance.
L = inductance.
C = capacity.

In order to make the gain as large as possible we can increase the ratio $\frac{L}{C}$ or decrease R. The first alternative is not easy ; a small capacity is unsuitable in a wavemeter on account of stray circuit capacity. The wavemeter capacity must be large enough to swamp the effect of stray capacity.

The chief source of resistance is the indicating device ; its effect is offset by loose transformer coupling, so arranged that the resistance " reflected into " the circuit is less than the actual resistance of the indicator.

Modulation—radio-telephony

On page 110 the process of keying a transmitter was described. In effect the process imprints, as it were, information on the radiation, which carries it out into space with its characteristic velocity. In this section we shall describe how we can imprint speech signals on the radiation, instead of the simple discontinuities of the morse code. The process of imprinting information or " signals " on the radiation is called *modulation* ; the constant undisturbed output of the transmitter is called the *carrier wave* and the process of modulation may consist in either the modification of the carrier amplitude (amplitude modulation) or of its frequency (frequency modulation).

Only the first of these methods will be dealt with here ; the second is of very limited application and is rarely used.

Firstly, let us dispense with a common difficulty in regard to radio-telephony. It might be argued that the audio-frequency output from a microphone could be magnified and applied directly to an aerial-earth system ; it would then appear that the aerial would emit radiation of the appropriate acoustic frequency which could be picked up by distant observers with the greatest of ease. That would be very convenient, but it is quite impossible because the *aerial will not radiate unless the frequency of oscillation is of the order millions of cycles per second* ; the audio-frequencies forming the output of the microphone lie within the range 50–10,000 cycles/sec., and are therefore well below the lower limit of radiation possibility which extends from 100 kc/s to 100 Mc/s.

It is on account of this that the technique of radio-telephony consists in impressing audio-frequency variations on the R.F. carrier wave of a transmitter, and at the receiving end of extracting the audio-frequency variations from the modulated R.F. radiation. This latter process is variously called de-modulation, detection, or rectification, and forms the subject of the next chapter.

The unmodulated output from the transmitter is merely C.W. radiation of constant amplitude, and it is represented by the diagram fig. 15.10(a).

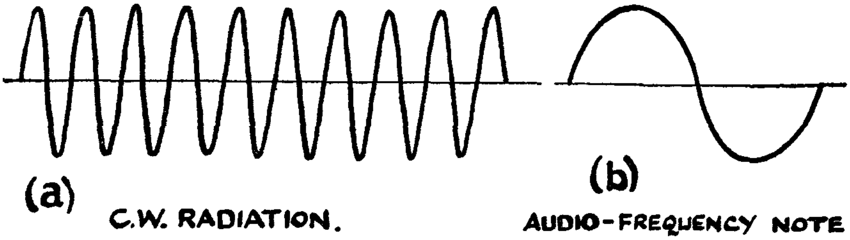


FIG. 15.10. Carrier and audio frequency note

Let us suppose that we want to impress on this carrier a simple audio-frequency note represented by fig. 15.10(b). Observe that when we speak of a " note " in this connection we refer to the electrical counterpart of an audible note, i.e. the alternating electrical output of a microphone operated by the musical note referred to. For simplicity we have supposed the amplitude of the R.F. and A.F. waves to be equal.

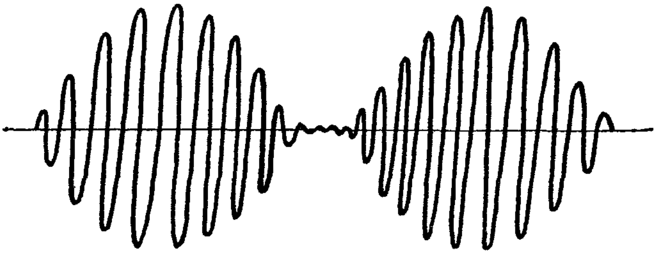


FIG. 16.10. Amplitude modulated wave (100 per cent. modulation)

By methods to be described shortly, we make our audio-frequency oscillation control the amplitude of the carrier, so that the envelope of the carrier becomes a picture of the audio-frequency wave, as in fig. 16.10, which represents the carrier (*a*) of fig. 15.10, modulated by the A.F. note corresponding to fig. 15.10(b). The amplitude of the carrier is varying at modulation frequency.

Now in this particular case, the two amplitudes were equal ; the resultant R.F. amplitude therefore varies at audio-frequency between zero and twice the amplitude of either component ; such a case we refer to as 100 per cent. modulation. Let us make it a little more general ; if the carrier amplitude is A, and the A.F. amplitude is B, the amplitude of the modulated wave varies between $A + B$ and $A - B$; the *depth of modulation* is B/A, and the percentage modulation is $B/A \times 100$. Thus, if $B = \frac{1}{2}A$, the percentage modulation is

$$\frac{\frac{A}{2}}{A} \times 100 = 50 \text{ per cent.}$$

Such a case is depicted in fig. 17.10.

The above description is limited to the case when the modulating sound is a simple sine-wave sound ; but speech, or the sound of an orchestra, are the things we want to transmit and fortunately the process is exactly the same, except that the modulating signal is no longer a simple sine-wave, but a wave of tremendous complexity. The highest sound, however, has a frequency so very much lower than the carrier, that the latter responds with complete faithfulness to the amplitude variations, and its envelope (dotted in fig. 17.10) becomes an exact copy of the complex sound waves it is desired to transmit.

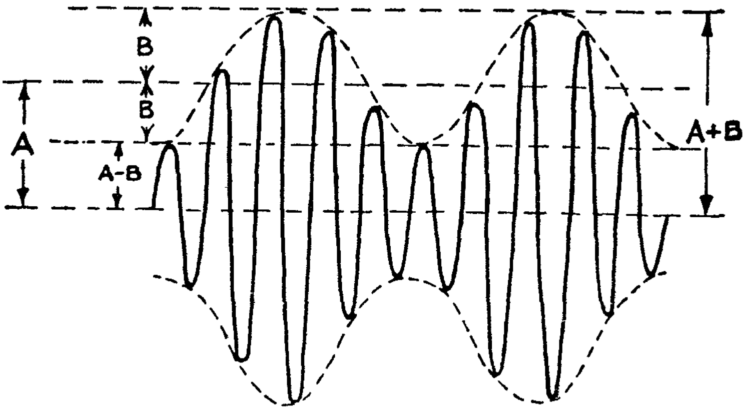


FIG. 17.10. Amplitude modulated wave. (50 per cent. modulation).

Methods of effecting modulation

In the present section we limit our discussion to the case in which the C.W. output is due to a simple valve-oscillator transmitter. We shall take up the subject again in Chapter XIII, when the methods of modulating the output of master-oscillator controlled transmitter will be described.

The amplitude of the output from an oscillator depends chiefly on two things (*a*) anode voltage, (*b*) grid voltage.

Generally speaking, the amplitude increases with rising anode voltage and decreases with increasing negative grid bias.

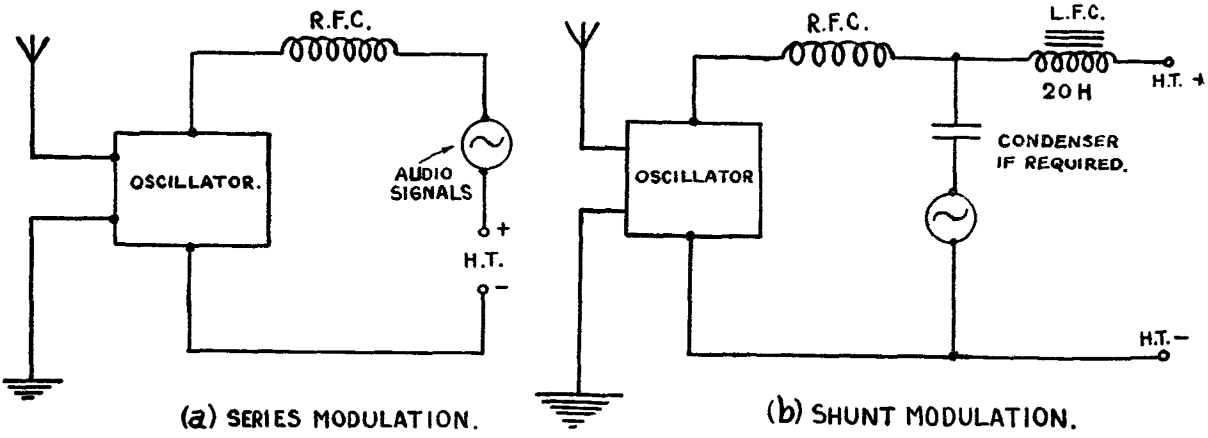
The problem of amplitude modulation therefore becomes that of controlling either the anode or the grid voltages by means of audio-frequency power.

(*a*) Anode-modulated valve oscillator

The amplitude of the R.F. output from a valve oscillator using leak and condenser bias is directly proportional to the H.T. voltage. To obtain a modulated output, therefore, all we have to do is to join the source of audio-frequency voltage in series with the H.T. supply and the amplitude of the R.F. output will then rise and fall with that of the A.F. input. If the peak of the A.F. input is equal to that of the unmodulated R.F. output we shall have the state of affairs represented in fig. 16.10 (100 per cent. modulation).

This method is a very good one, but it suffers from the disadvantage that, if high-powered transmitters are used, enormous audio-voltages are required to produce full modulation. It is interesting to note that the audio-output required to modulate the final amplifier stages of the American station WLW. is 250 kilowatts.

Theoretical circuits



(a) SERIES MODULATION.

(b) SHUNT MODULATION.

FIG. 18.10. Anode-modulation of valve oscillators

In series modulation the audio source is in series with the H.T. supply, and the development of R.F. voltages across this supply is prevented by the use of the high frequency choke, R.F.C.; in shunt modulation the development of both R.F. and A.F. voltages across the H.T. supply is prevented by the use of two chokes; a high frequency choke, R.F.C., and an audio frequency choke L.F.C.

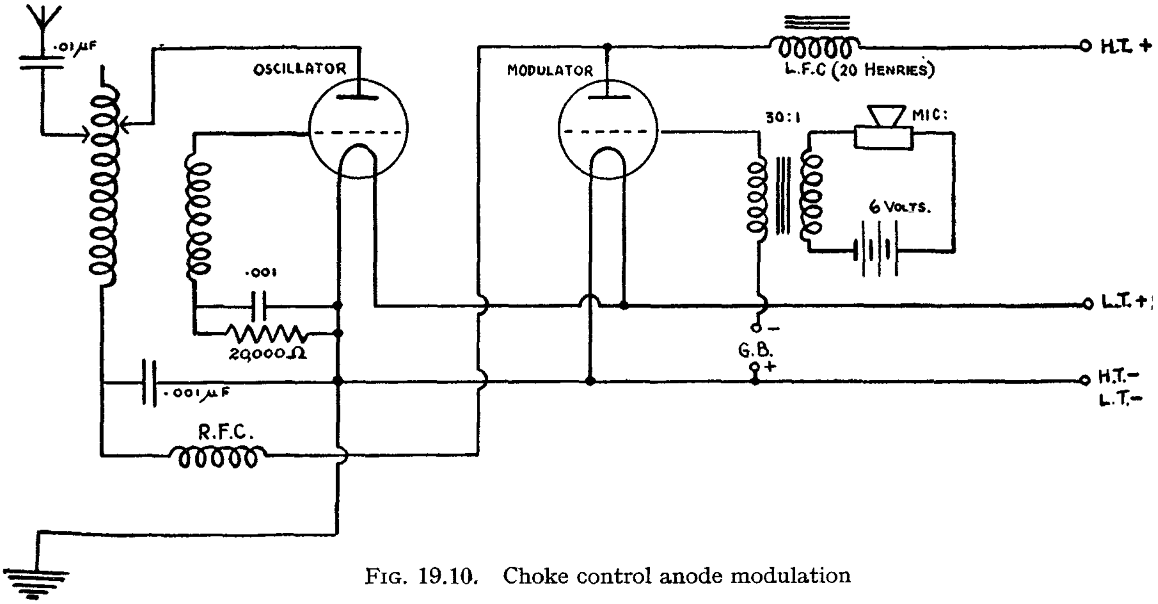


FIG. 19.10. Choke control anode modulation

The second of these methods is very convenient in practice because it lends itself easily to the employment of a modulating valve operating as an amplifier, magnifying the feeble audio-input from a microphone or a transformer.

Fig. 19.10 shows the arrangement for a radio-telephony transmitter, consisting of a simple Meissner oscillator with choke-control anode modulation.

The reader should attempt to modify this diagram using some other form of oscillator, by way of an exercise.

(*b*) Grid modulated valve oscillator

In this method of modulation the audio-voltage is applied in series with the grid coil so that it varies the grid bias and thus controls the output amplitude. In practice it is difficult to achieve deep modulation without serious distortion. Leak and condenser bias should not be used, fixed bias being preferable.

Fig. 20.10 shows the circuit of a simple grid modulated oscillator.

Sidebands

Let us return for a moment to the simple case in which an R.F. carrier is modulated at single acoustic frequency. The wave form of the modulated radiation will be something like that shown in fig. 17.10 depending, of course, on the depth of modulation. As before, we will let A represent the amplitude of the carrier, and B that of the audio-frequency. The resulting wave will then vary in amplitude at R.F. between A+B and A − B. Now the mathematicians have examined a modulated wave of this sort and have discovered a very striking thing about it. They find that it is exactly equivalent to *three radio-frequency* waves of constant amplitude, all co-existent. If we call the frequency of the original carrier p and that of the audio-wave q these three waves are :—

(i) A wave of frequency $p + q$ **and amplitude** $\frac{B}{2}$.

(ii) The original carrier of frequency p and amplitude A.

(iii) A wave of frequency $p - q$ and amplitude $\frac{B}{2}$.

Wave (i) is called the Upper Sideband wave.

Wave (iii) is called the Lower Sideband wave.

On this view the process of modulation can be regarded as the production of new sinusoidal R.F. oscillations in addition to the carrier, one above it and the other below it from the frequency standpoint.

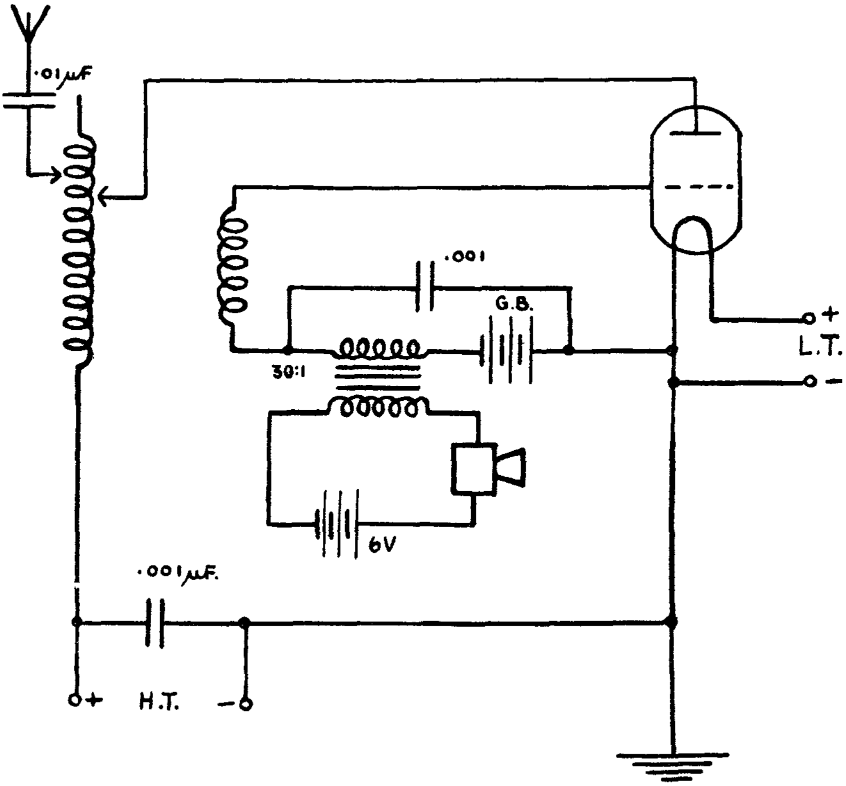


Fig. 20.10. Grid-modulated valve oscillator

The complex vibration of speech, we have already seen, consists of a whole range of simple frequencies; speech modulation will, therefore, result in the formation of an upper and lower sideband wave corresponding to every one of the simple waves going to make up the complex waves of speech. Suppose that the audio-frequencies range from 50 to 8,000 cycles per second, and that the carrier on which these frequencies are to be impressed has a frequency of 1,000 kc/s, then the frequency range of the upper sideband will extend from 1,000,050 to 1,008,000 c/s. and the lower from 999,950 to 992,000 c/s., so that the resultant R.F. oscillations now cover a whole band of frequencies, whose width is twice that of the highest modulating frequency.

This fact has a very important bearing on the allocation of frequencies to commercial broadcasting stations. Their carriers must be so far separated in frequency that there is room for each to develop its sidebands without overlapping. That means that they must have a separation of 20 kc/s if possible. Fig. 21.10 shows the frequency range required by a carrier of 1,000 kc/s, modulated with acoustic frequencies ranging from 50 to 8,000 cycles per sec.

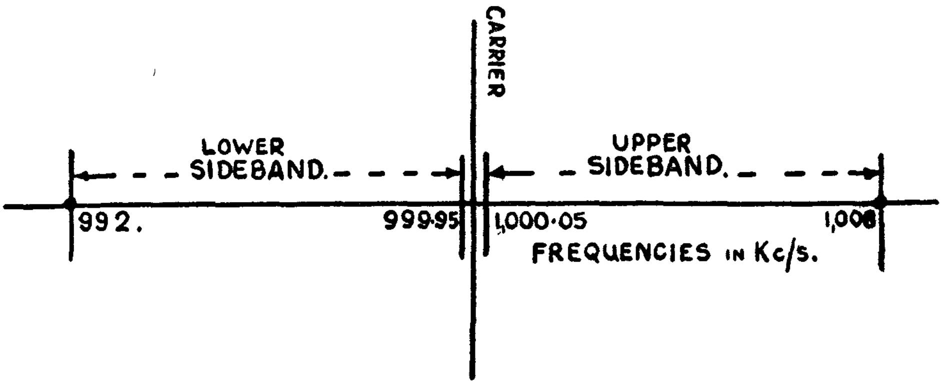


FIG. 21.10. Illustrating carrier and sidebands

We must bear this in mind when we come to discuss the reception of these modulated waves; our receiving circuit must be capable of responding to the frequency band as a whole. Its selectivity must be therefore a " band selectivity " and not a " line selectivity " if we want to ensure good reproduction of the original acoustic input. A receiver which is too selective produces an unpleasant effect called " side-band cutting ".

CHAPTER XI

DETECTION

The subject-matter of this chapter forms the complement of Chapter X; there we learnt how to radiate H.F. wireless waves into space and, further, how we could impress signals on these waves, either by the process of keying (radio-telegraphy) or by speech modulation (radio-telephony).

We now turn to the problem of how we can extract from the radiation the information imprinted on it at the transmitting station.

First, let us notice that the energy spreads out in all directions from the transmitting aerial, unless in some special cases we are able to concentrate it in a beam of some kind. At any distant point, therefore, the energy density of the waves is very small, and the collecting device (receiving aerial) must be as sensitive as possible; it must be high, well insulated, and with as little resistance as possible; the lead-in to the receiving apparatus must be as short and direct as convenient.

Aerial tuning circuit

We shall only consider simple types of aerial tuning; this tuning system, as previously explained on p. 82, is some kind of variable reactance whose adjustment brings the aerial into resonance with the incoming radio-frequency radiation.

Fig. 1.11(a) represents an aperiodic, or untuned aerial circuit, loosely coupled to a tuned circuit by an open transformer in a step-up ratio of about 1 : 2. Instead of this arrangement we can tap the lead-in to the coil of the tuned circuit, as shown in fig. 1.11(b). This is really another case of " matching "; the aerial with its incoming radiation has to match the tuned circuit and its associated apparatus and the matching is accomplished approximately in this way.

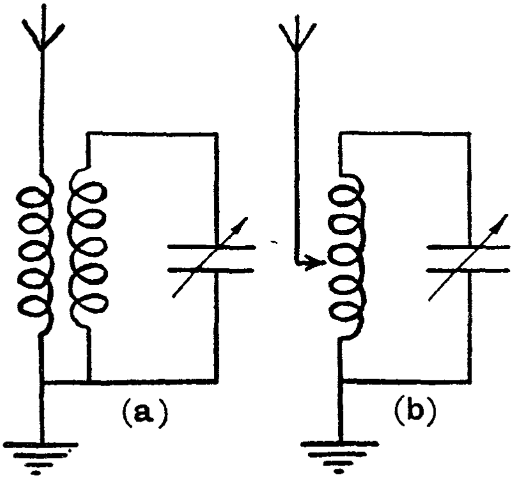


FIG. 1.11. Aerial tuning circuits

Other types of aerial tuning circuits are shown in fig. 2.11. In type (a) the tuning is accomplished by series condensers; in (b) the tuning is of the parallel type.

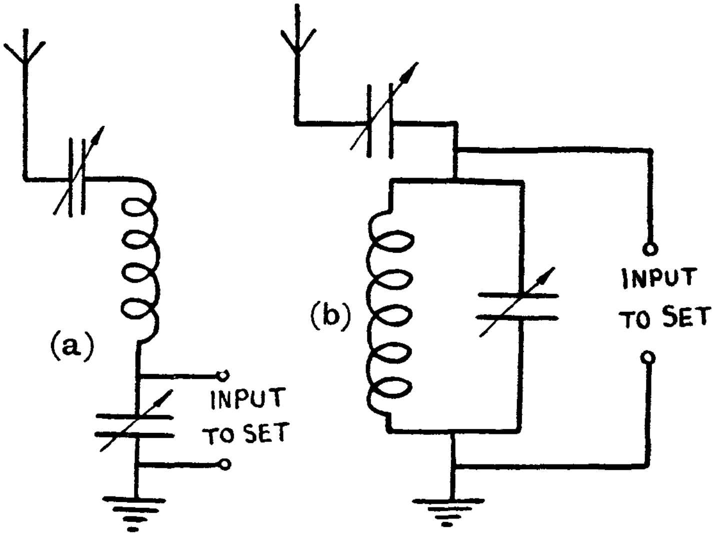


Fig. 2.11. Alternative tuning circuits

Whichever system we use, its object is to produce at the input terminals alternating voltages which are a replica of the received radiation; and further the circuit must be selective, so that it deals only with a certain band of frequencies and reduces interference from signals other than the ones wanted.

For high quality reproduction of broadcast signals, as was mentioned in the last chapter, the aerial tuning circuit must not be so selective that it results in side-band cutting; it must pass on to the receiving set the whole range of oscillations comprising a music-modulated carrier without altering their relative strength. A tuning circuit which will do this is called a " band-pass " circuit; the difference in behaviour at a range of frequencies between a simple line-pass tuning circuit and a band-pass circuit is shown in their respective frequency response curves :—

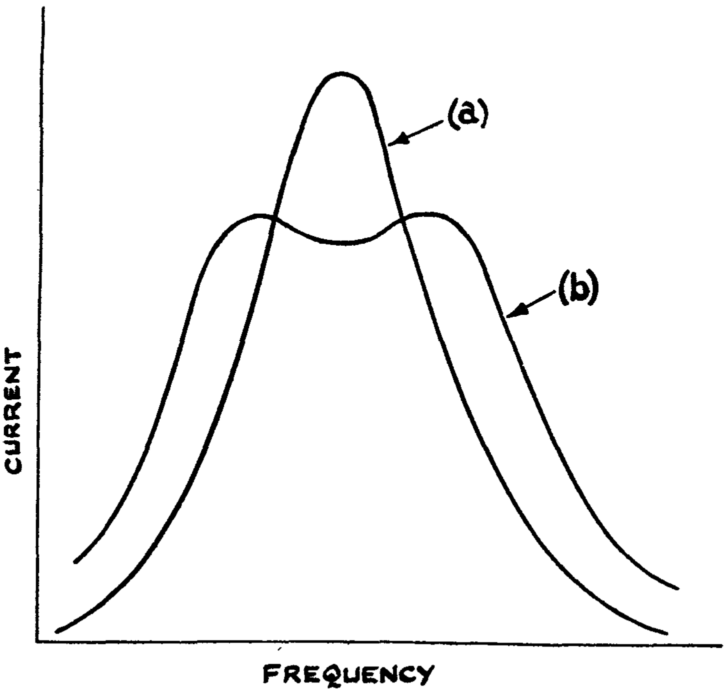


Fig. 3.11. Frequency response curves

(a) Line-pass circuit. (b) Band-pass circuit

The band-pass effect is obtained by using two tuned and coupled circuits ; the design of four such pairs is indicated in fig. 4.11 ; (a) consists of two circuits coupled by mutual inductance ; the latter is adjusted until the correct band-pass effect is obtained.

The two condensers of course could both be operated by the same control ; they would be " ganged " as we say. Arrangement (b) employs capacity coupling ; the coupling condenser is a small one of about 5 $\mu \mu$ F. Diagram (c) indicates another mutual coupling arrangement employing a " link " circuit.

The fourth circuit (d) is an important one from a practical point of view, and in it the coupling is brought about by means of a small reactance or resistance common to each circuit.

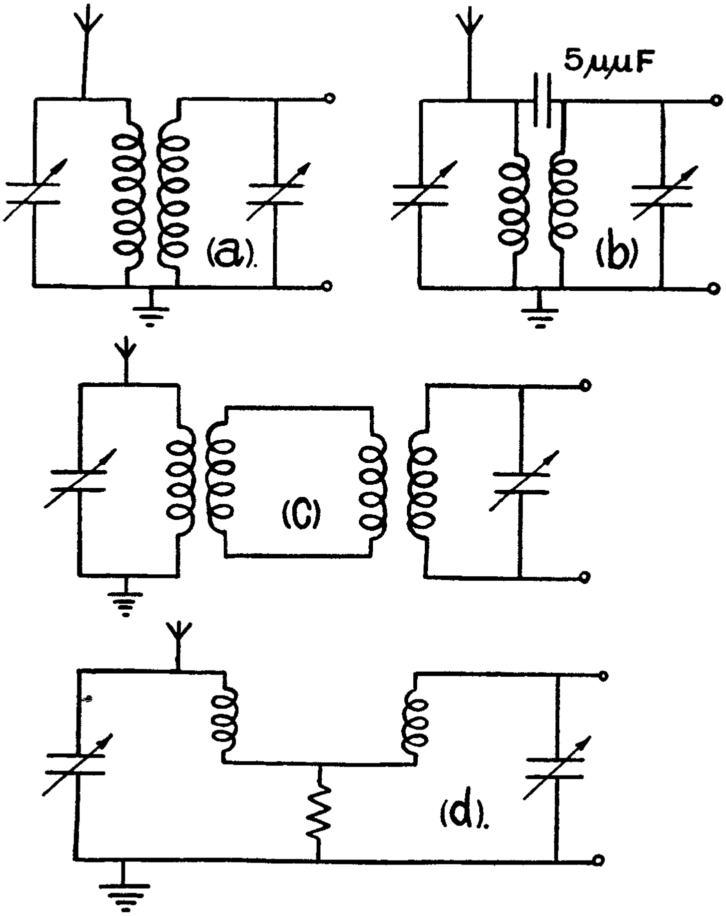


FIG. 4.11. Band-pass circuits

Due attention must be paid to stray coupling in circuits such as these ; coils and condensers must be carefully screened by being enclosed in earthed metal boxes so that their fields do not give rise to coupling other than that allowed for in the design ; otherwise the whole system defeats itself.

Detection

We are now confronted with the problem of extracting from the R.F. alternating voltages set up at the termination of our receiving aerial system, the acoustic frequency signals with which the carrier wave was modulated at the transmitter. We shall call this process *detection.* American engineers call it de-modulation and it is probable that this term will become more generally used as time goes on, because it does express so exactly the nature of the process in question.

First let us clear the ground a little ; three types of radiation have been mentioned, each of which has some special application to signalling ; these are (1) continuous wave (C.W.), (2) interrupted continuous wave (I.C.W.), and (3) speech modulated wave (R.T.). All of these require detection before they can deliver their message, and there are several methods by which it can be achieved. To avoid a long and tedious discussion, therefore, in which the essential points may become obscured, we will discuss only the detection of the speech modulated wave. Having done this with each type of detector, we shall be able to deduce, without further detailed examination, the corresponding action for C.W. and I.C.W.

The problem of detection is simple in theory, even though the practical technique leads to some formidable difficulties. A modulated wave as such can only be analysed by means of filters into a carrier and upper and lower side band, all of which are of radio-frequency and none of which therefore will operate our telephones and convey a message. We can actually separate out these components by electric filter circuits, but that is as far as we can go, and it does not help us at all. If, however, we distort the wave-form in some way, instead of merely analysing it with filters, we shall obtain frequency components other than the carrier and the side-bands, and one of these components will be at the audio frequency with which the carrier was modulated at the transmitter.

Unfortunately the process of distorting the wave-form, to which we must resort before we can extract the A.F. signal, introduces currents of frequencies which were in neither the carrier nor the modulating oscillations. They constitute distortion and it is one of the problems of the detector design to avoid such unwanted frequencies, or in any case to minimise their effects.

The most convenient way of distorting the modulated wave form is to use a rectifier, followed by suitable arrangements of resistances, chokes and condensers which pick out the wanted signal from among the products of rectification ; a rectifier used in this sense is called a detector.

(1) The diode detector

We have already seen in Chapter IX, p. 95, that if an alternating voltage of constant amplitude is applied between anode and filament of a diode, current will flow only during the half-cycles which make the anode positive, and that these current pulses, acting through a resistance and condenser load, will develop across it, in addition to an input frequency ripple, a direct voltage somewhat less than the peak of the input. As the result of this voltage the anode assumes a negative potential with respect to the filament so that only the extreme tips of the positive voltage half-cycles give rise to current—just enough in fact to maintain the charge on the condenser and keep the load volts-drop constant. Fig. 5.11 illustrates this process, with corresponding circuit diagram.

The first few cycles of input voltage drive current on the positive half-cycles and charge up the condenser, which, with the resistance in parallel, now acts as a bias. It is not quite a steady bias, but as is shown by the line OB, it fluctuates at input frequency, owing to the fact that the condenser partially discharges through the resistance on negative half-cycles and regains its lost charge on the positive half-cycles. The less the shunting resistance the more pronounced the input fluctuation will be (because the condenser will lose more charge on the negative swing and vice versa) ; if the shunting resistance is infinite (i.e. condenser only) the voltage would become quite steady equal to the peak of the input, and even the positive swings would produce no current because they would be equal to the negative bias provided by the condenser.

The mean voltage developed across the load consisting of the usual resistance and condenser combination, is proportional to the amplitude of the input (see p. 95) always

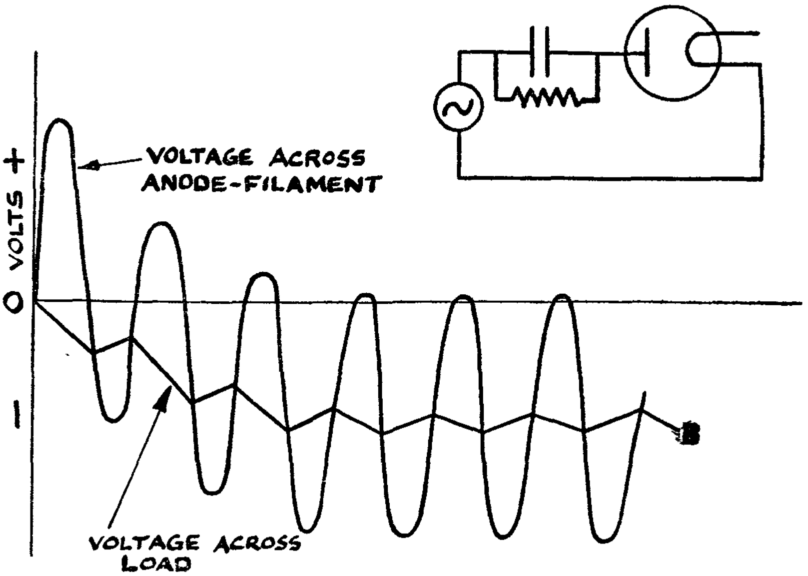


FIG. 5.11. Diode action with leaky condenser load

provided that the latter is big enough to overcome the initial curvature of the diode I*a*—V*a* characteristic, and here lies the solution to the problem of the detection of a modulated wave.

Let us suppose that the resistance of the load is so chosen that the mean voltage developed across it is 60 per cent. of the input amplitude ; this means of course that with a C.W. input of 1 volt peak, there will be developed across the load a ·6v. steady volts drop, which will act as a bias, making the anode ·6 volts negative with respect to the filament : superimposed on this will be a certain amount of input frequency ripple.

Now instead of a C.W., let us take as input an R.F. carrier modulated with a single acoustic frequency to a depth of say 50 per cent. As the amplitude of the wave varies so will the voltage across the load ; it will be 60 per cent. of the input amplitude, whatever changes the latter may undergo.

It follows, therefore, that the " steady " voltage across the load will rise and fall with the rising and falling amplitude of the input and its fluctuations will actually be of *the acoustic frequency*, with which the original carrier was modulated. Instead of a resistance as load, we can now substitute a pair of high resistance telephones, and in them we shall hear the notes corresponding to the modulation frequency, the D.C. component, the H.F. ripple, having no audible effect.

Let us illustrate the diode rectification of a modulated wave by means of a graph.

ABCD represents the envelope of a 50 per cent. modulated wave. OPQ is a line drawn at right angles to the axis XY ; OQ is the amplitude of the RF oscillations and OP (which is 60 per cent. of OQ) is therefore the mean voltage across the load.

The wavy curve RPS therefore represents the voltage across the load during the reception of the modulated wave ; it will act as a bias, driving the anode negative ; the operating point will fluctuate in consequence, at modulation frequency, with an R.F. ripple superimposed.

The whole action is illustrated in fig. 7.11, which also shows how the voltage across the load acts as a bias, making the anode negative with respect to the filament.

The modulated input sets up a varying voltage across the load which acts as a varying bias, shown by the wavy line PP^1 ; in the example we have taken the mean line PP^1 to the left of the zero voltage axis, by an amount of 60 per cent. of the peak of the input at any instant.

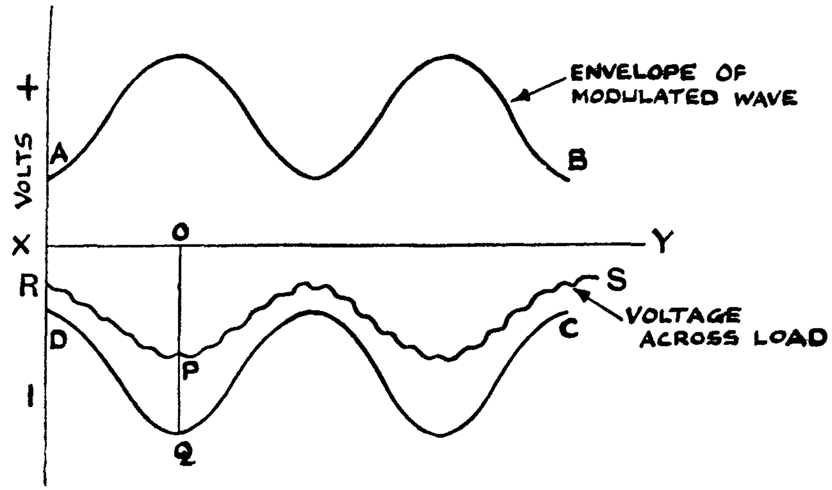


FIG. 6.11. Rectification of modulated wave

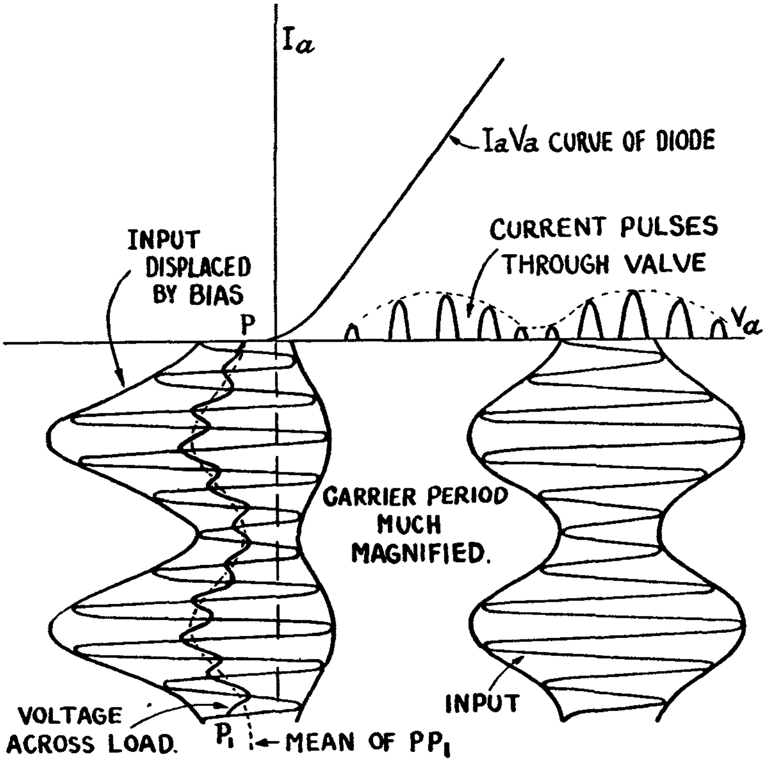


FIG. 7.11. Diode rectification of a modulated wave

If we examine the voltage represented by PP^1 (fig. 7.11) a little more closely, we shall see that it can be analysed into three components: the representative graph is re-drawn in the conventional way in fig. 8.11(a), and it will be seen that it is nothing more than a combination (by simple summation) of the three voltages represented in the graph of fig. 8.11(b). The R.F. ripple has been much magnified in these drawings for the sake of clearness.

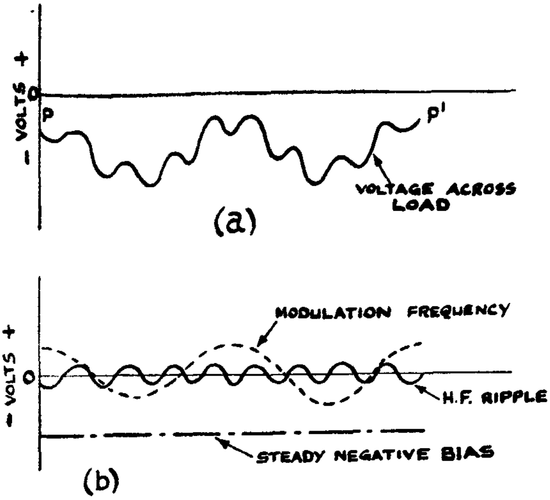


FIG. 8.11. Analysis of products of rectification

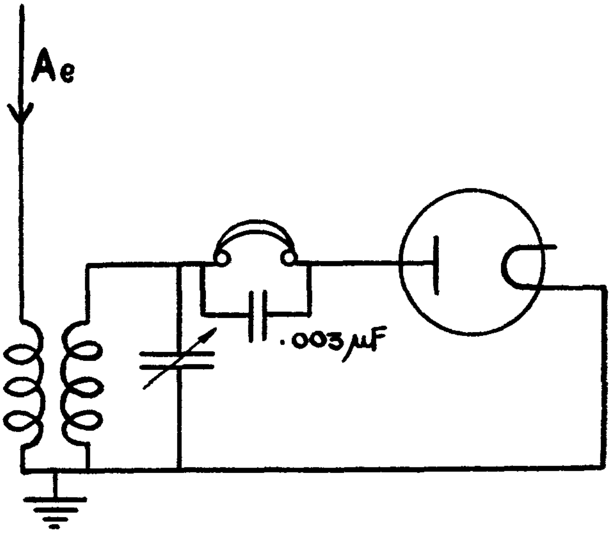


FIG. 9.11. Simple practical circuit for diode detection

These voltages are (1) A steady D.C. negative bias (chain line).
(2) A modulation frequency alternating voltage (dotted line).
(3) A ripple of input (carrier) frequency (full line).

These three components are the products of rectification, and it is only a matter of using suitable circuit arrangements to isolate the component at modulation frequency.

The simplest kind of circuit for this purpose is shown in fig. 9.11, where the leak resistance is replaced by a pair of high resistance telephones.

In practice, of course, this circuit would be extremely insensitive, and it would have to be followed by some form of amplifier to magnify the feeble signals produced, or alternatively it might be preceded by some kind of amplifier to boost up the modulated wave before detection. The latter course would carry the additional advantage of providing a bigger input to the diode, and would thus avoid, except on very deep modulations, the distortion arising from the initial curvature of the Ia—Va characteristic.

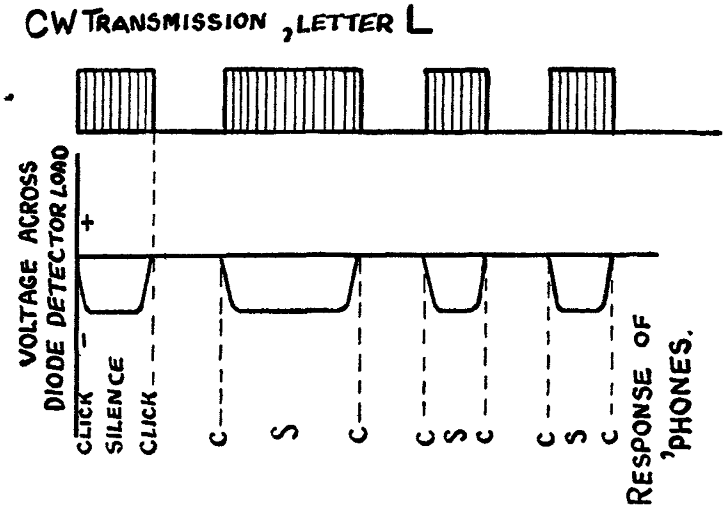


FIG. 10.11. Illustrating diode rectification with C.W.

Generally, some sort of device is employed to filter off the R.F. and the D.C. components. The reader might design such a filter on the lines of the scheme outlined in Chapter VIII, page 88.

Diode and C.W. detection

The reader will have no difficulty in seeing that a C.W. input will merely produce a steady direct voltage across the load; the telephones would register a click at the beginning and at the end; this is illustrated in the sketch of fig. 10.11.

Such an effect is not much use for signalling so that we can say that a straightforward diode detector is no use for C.W.

We shall describe how to deal with C.W. in a later section of this chapter.

Diode and I.C.W. detection

I.C.W. is of course only C.W. chopped into very short wave trains; if such radiation is detected by a diode system, every pulse (as in keyed C.W.) will produce two clicks in the phone, and as they may arrive at the rate of 1,000 per second, the phone clicks would become an audible note.

If this radiation is keyed at the transmitter the phone of the rectifier will reproduce the morse letter in sound whose pitch depends on the interrupter frequency at the transmitter end.

The anode detector

We next turn to the triode as a detecting device. In Chapter IX, page 100 *et seq.*, we have discussed the process called anode rectification; the application of an alternating grid voltage at an operating point near anode current cut-off (see p. 101) gave rise in the anode circuit to :—

(i) An increase in the mean anode current.
(ii) An anode current fluctuation of input frequency.

Further, owing to the curved nature of the I*a*—V*g* characteristic, we noted that the increase in mean anode current will be proportional to the square of the input amplitude, except for very strong signals, when it becomes linear.

These facts furnish sufficient material to see how an anode rectifier can function as a detector. Many books call this kind of detection "anode bend" detection; this is a term we shall not employ; we shall call it *anode detection*, meaning detection which depends on the properties of the anode current grid volts curve.

Fig. 11.11 represents the theoretical circuit on which we shall work.

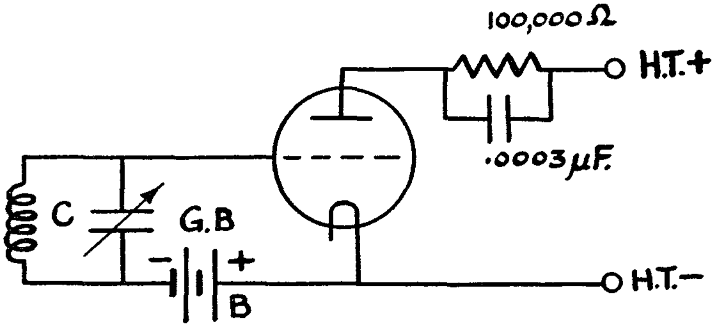


FIG. 11.11. Anode detection circuit

The grid is biased to anode current cut-off by means of the bias battery B; the tuned circuit C is tuned to a carrier with a simple single audio frequency modulation; the anode circuit has the resistance and condenser load as in the case of the diode. Notice first of all that the negative bias on the grid prevents current flowing in the grid circuit and thus we escape the disadvantage inseparable from diode detection, of loading the tuned circuit, and reducing its selectivity.

Fig. 12.11(a) shows what will happen when an amplitude modulated input is applied to the grid which, for the sake of generality, has not been biased quite to cut off; the positive swings of grid potential will produce an increase in anode current while the negative half swings will give a decrease, but as the former are always bigger than the latter owing to the curvature of the characteristic, there will be a fluctuation of mean anode current at modulation frequency.

Let us examine this action a little more carefully from the point of view of the voltage developed across our load; there will be a steady volts drop under pre-signal conditions; a rise of anode current will increase the volts drop and charge the condenser; the succeeding fall in anode current (which is less than the preceding rise) will decrease the volts drop, but the condenser will then discharge through the resistance and help to off-set this effect. The net result is that there is developed across the load a volts drop greater than that of pre-signal conditions, which rises and falls at modulation frequency, and which also has an R.F. ripple just as in the case of the diode, as shown in fig. 12.11(b). This voltage, of course, can be filtered to separate the D.C., audio- and radio-frequency components. In doing this we shall arrive at the following circuit: the pure resistance load is replaced by telephones, the load condenser is made to by-pass the H.T. battery, so achieving a desirable thing in all radio circuits, namely, the absence of R.F. in the high tension battery.

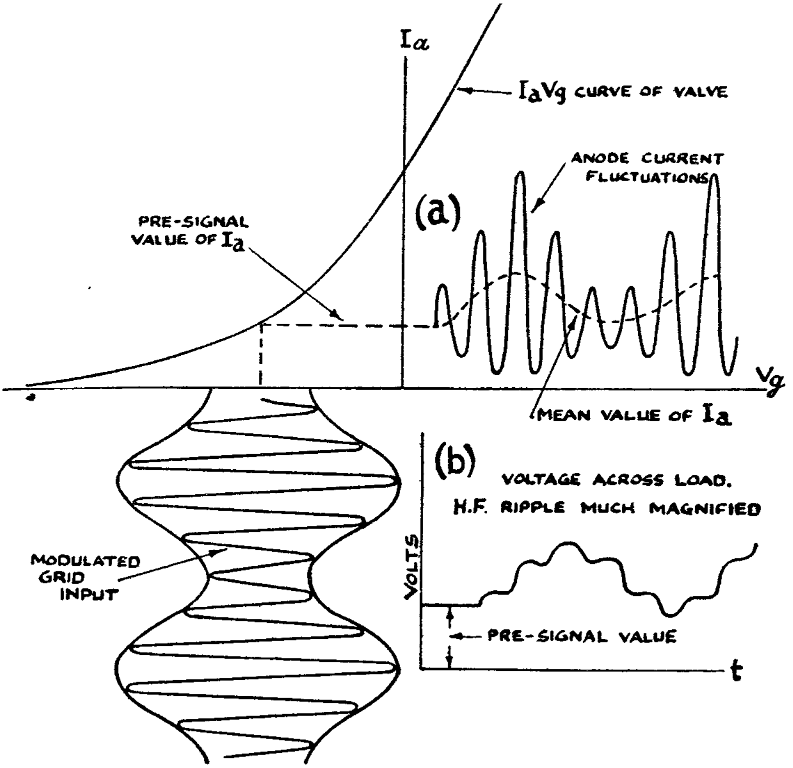


FIG. 12.11. Illustrating anode detection of modulated wave

The R.F. choke, of course, assists in this action (see fig. 13.11).

The combination of L.F. choke and 2μF condenser to separate the D.C. and modulation frequency currents, is a device called a *choke capacity output*. It eliminates D.C. from the phones (or the loudspeaker in the case of a large output) and improves their performance.

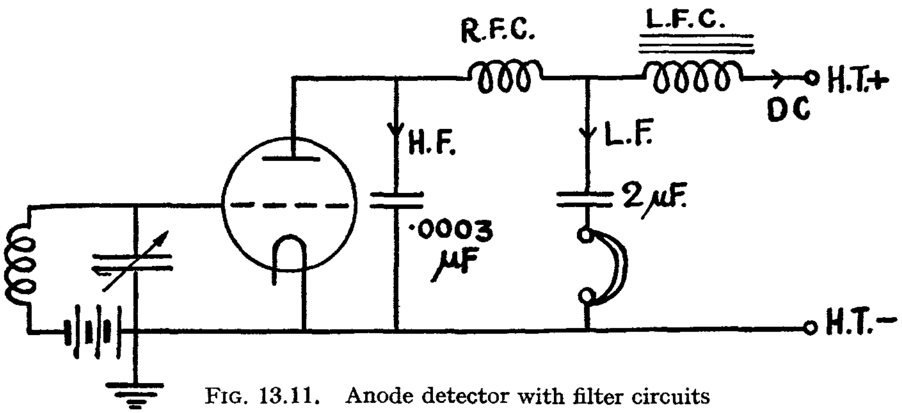


FIG. 13.11. Anode detector with filter circuits

Distortion

Owing to the marked curvature of the foot of the I*a*—V*g* curve, the anode current fluctuations are not proportional to the grid swings as they were in the case of the diode. The result of this is that the audio-frequency note developed in the telephones is not an exact copy of the note of the original modulator ; distortion is said to be introduced. It is equivalent in the case of the anode detector to the introduction of the octave of every note and is therefore called second harmonic distortion.

It is very marked in the detection of deeply modulated R.T. signals, but, of course, is of little consequence in the reception of morse signals. Even with R.T. the distortion is not very serious if the signals are not deeply modulated.

C.W. and I.C.W.

Anode detection is useless for C.W., but, like diode detection and for the same reason, it reproduces the interruption frequency note of I.C.W.

Adjustment

It is clear that we shall get the best results with the anode rectifier if the grid operating point is under that part of the I*a*—V*g* curve, where the curvature is changing most rapidly; further the grid swings should never take the grid positive and cause it to load the tuned circuit. We must, therefore, be able to adjust our anode-potential so that the foot of the I*a*—V*g* curve falls well to the left of the line V*g* = O, and also, by means of a potentiometer, we must be able to adjust our bias so that the operating point falls below that point on the I*a*—V*g* curve, where its curvature is changing most rapidly.

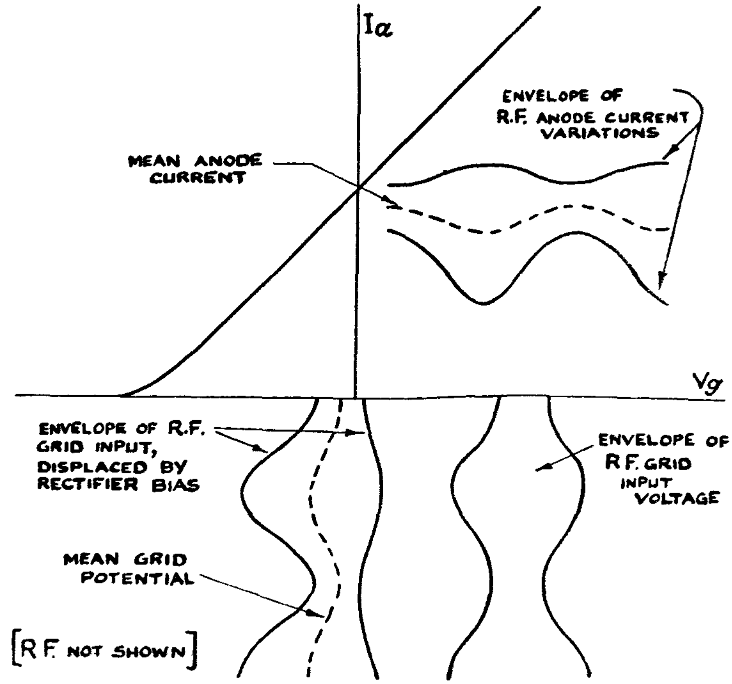


FIG. 14.11. Illustrating grid detection of a modulated wave

The grid detector

This type of detector, a practical form of which is shown in fig. 15.11, is really a combination of a diode detector with a triode acting as a class A amplifier.

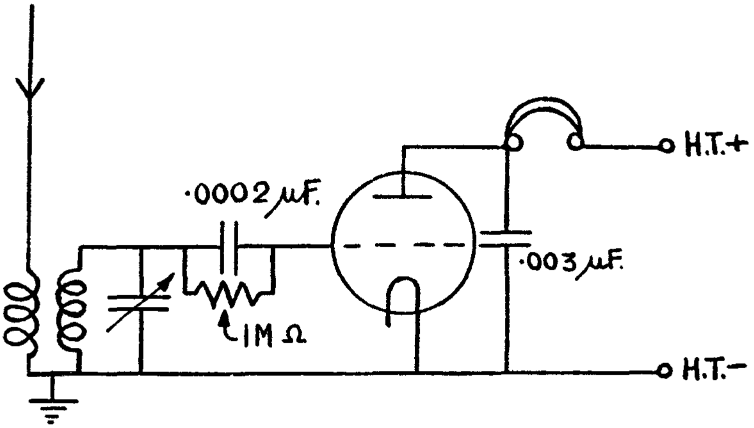


Fig. 15.11. Practical grid-detector circuit

The filament and grid of the grid-detector triode form a diode. If the arrangement is provided with the usual resistance and condenser load, a modulated input (fig. 14.11) will cause the mean grid potential to fluctuate with respect to the filament at modulation frequency.

By proper adjustment of the anode potential these grid fluctuations will, operating below the straight part of the I*a*—V*g* curve, give rise to similar undistorted fluctuations in the anode current which in turn will set up corresponding (magnified) voltage fluctuations across the anode load.

Fig. 14.11 illustrates this action. The grid behaves like the anode of the diode (fig. 9.11, p. 129), and its mean potential controls the anode current in the usual class A fashion.

Notice, however, that the grid potential, just like that of the diode anode, is still fluctuating at R.F., and these fluctuations (not shown in fig. 14.11), are also reproduced in the anode circuit.

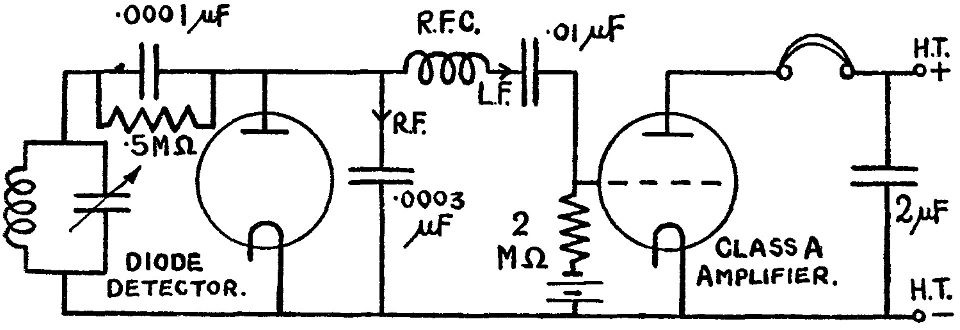


Fig. 16.11. Diode detector, followed by Class A amplifier

If the extreme negative tips of the R.F. grid swings come below the bend of the I*a*—V*g* curve, they will *cause anode detection* to take place and introduce serious distortion. To avoid this we must either raise the anode potential so that the foot of the I*a*—V*g* curve is at a safe distance away to the left as in power grid detection or limit our input to a safe value. This latter course may, with deeply modulated signals, bring in its train distortion due to the curvature of the foot of the grid characteristic. Thus, while the grid detector is sensitive (owing to the amplification it employs) it is very liable to serious distortions. These do not matter much in telegraphy, but are very objectionable in R.T.

detection. For these reasons modern practice is tending to substitute a separate diode followed by a triode amplifier for the combined detector amplifier described above.

Here the R.F. component can be eliminated from the amplifier altogether as in the circuit sketch of fig. 16.11.

H.F. By-pass condenser

In all these detector systems it will be observed that a fairly big condenser is employed to remove the R.F. product of rectification as soon as possible. It will be sufficient to remember that this must be done without going into the various reasons in detail. The two main reasons are :—

(i) The R.F. component will probably damp the preceding tuned circuit via the internal valve capacity.

(ii) It may set up all sorts of unwanted oscillations by " feed-back " through the various unavoidable small capacities in the circuits.

C.W. and I.C.W.—Grid detection

Again, the grid detector will function for I.C.W., but not for C.W. The reasons of course are the same as those given in connection with the diode detector.

Reaction

The triode detector circuits with which we have been dealing will not provide a very strong signal in the telephone unless the receiver is operated quite near to the transmitter. In the next chapter we shall deal more fully with this problem in discussing valve amplifiers in general ; at this point we shall describe a simple device called *reaction* by means of which a signal can be strengthened without the use of additional valve stages. We have already used this device without giving it this name, in describing the valve oscillator and the simple transmitter ; here a small portion of the output was fed back into the grid circuit and resulted in sustained oscillations in tuned circuit.

Now if a circuit will oscillate continuously, it follows that the effects of damping have somehow been overcome ; the effect of the feed-back then is to reduce the damping in the tuned circuit ; it will not give rise to oscillations until the feed-back is sufficient to overcome damping completely, but until this point is reached damping is reduced, and all the results of decreased damping make their appearance in the circuit.

Reaction is generally achieved by magnetically coupling a small coil in the anode circuit to the tuning coil in the grid circuit. The amount of feed-back can be controlled by altering the coupling of these two coils, or by altering the value of the R.F. current in the reaction coil, either by means of variable condensers or by actually altering the operating conditions of the valve, now functioning as a reacting detector.

The circuit of fig. 17.11 represents a simple grid detector with inductive feed-back. L_1 C_1 is the input tuned circuit ; the anode circuit will carry both R.F. and modula-

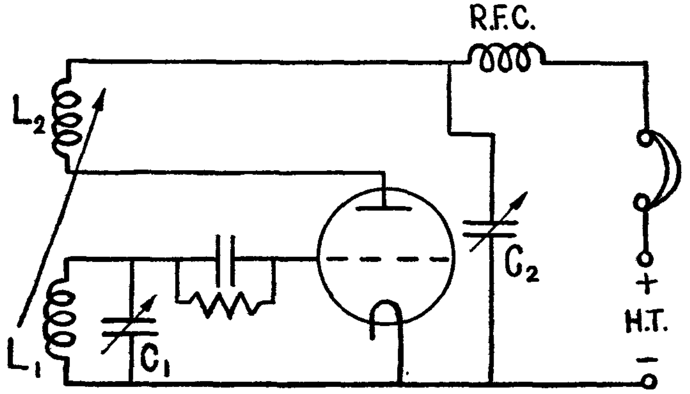


FIG. 17.11. Simple inductive reaction

tion frequency currents ; the former, flowing in L_2, will set up a magnetic flux which will couple with L_1 and either boost up or damp down the oscillations in L_1C_1 according to the sense of the coupling ; that is, there is a right and a wrong way about for coil L_2 and the correct way must be found by trial and error. Having got the correct sense for L_2, it is generally arranged that it can be brought nearer to or further from L_1, and so provide a control to the reaction effect. It will be found that if the coils are too close, the circuit will be set in oscillation, and although this oscillating condition has an important application in C.W. reception, it must be avoided in R.T. reception.

Reaction can also be controlled by varying the H.F. by-pass condenser C_2 ; the coils L_1 and L_2 can be fixed, and the feed-back will then depend on the setting of C_2 ; this makes a better mechanical job than swinging coil holders.

Another method of control is shown in fig. 18.11. The R.F. currents have now two paths to earth :—

(i) via the variable condenser C_2.
(ii) via the reaction coil and condenser C_3.

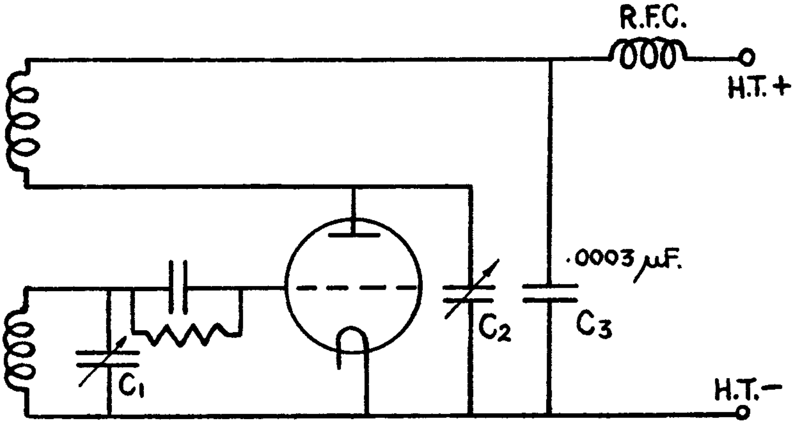


FIG. 18.11. Condenser control of reaction

When C_2 is adjusted to minimum capacity the R.F. current in the reaction coil (and therefore the amount of reaction) is a maximum and vice versa.

Reaction is often controlled by means of a differential condenser, as shown in fig. 19.11.

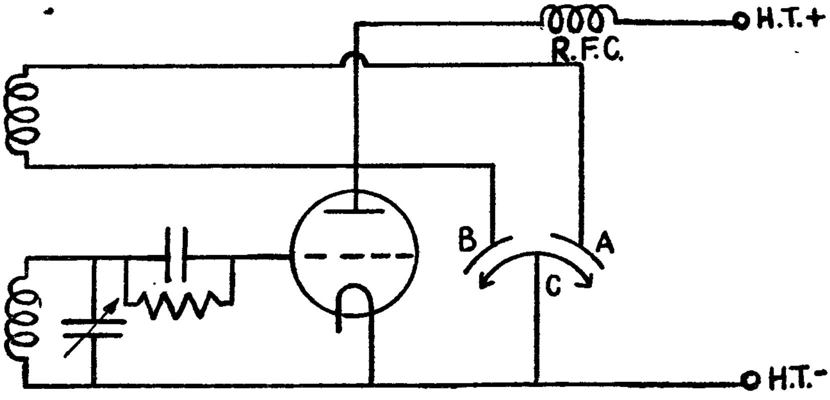


FIG. 19.11. Use of differential condenser for reaction control

The moving plates (C, fig. 19.11) can be turned towards A when a low impedance path is provided for R.F. currents via the reaction coil and reaction is increased. Turning C towards B short-circuits R.F. currents to earth direct and reaction is decreased. The advantage of such a system is that the total capacity between anode and earth is constant, and altering the reaction does not alter the operating conditions of the system.

Capacity Reaction

Instead of an inductive coupling between anode and grid circuits, we might have a condenser as indicated in fig. 20.11.

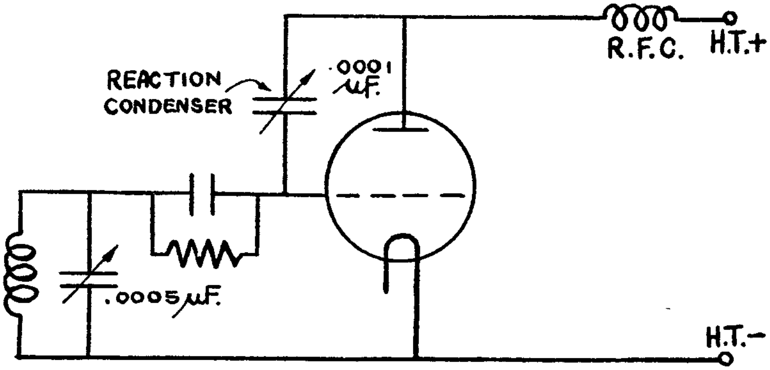


FIG. 20.11. Simple capacitative reaction

In fact, owing to internal anode-to-grid valve capacity, such a form of feed-back is always present in a triode and it is a potential source of trouble because it tends to throw a valve into oscillation. This form of reaction is not used ; it has been mentioned here to stress the possible effect of internal capacity of the valve ; indeed, generally it would act as a damping device rather than as a reaction effect, because an anode load of this nature is usually capacitive.*

Effects of reaction

The reduction of damping in a tuned circuit owing to the application of reaction will of course do two things :—

(i) Increase the gain ;
(ii) Increase the selectivity,

both of which are desirable ; but excessive reaction, while still below the oscillation level, will cause a serious falling off in the quality of R.T. reception ; speech becomes blurred and high frequency components become destroyed. As the oscillation point is reached, of course, the receiver begins to howl, for reasons we shall take up shortly, and we shall also see that this howl, while it is a hopeless nuisance in R.T., has a very important application in the reception of C.W. telegraphy.

Generally speaking the following are the points which should be looked for in a good reaction arrangement :—

(i) Smoothness of control ; the signal strength should rise gradually and progressively and the detector should glide gently into oscillation.
(ii) No effect on tuning ; with some arrangements it is necessary to re-tune the main control when reaction is varied.
(iii) No back-lash ; the set should go in and out of oscillation at the same setting of the dial.

Another way to control the reaction is to alter the potential of the anode of the valve ; in theory the Ia—Vg curves of a triode should be straight and parallel for different values of Va ; in practice this is not so ; they become steeper as the anode voltage rises. This is seen in the curves of fig. 21.11 (a).

Fig. 21.11 (b) shows the circuit in which this arrangement is applied.

The control in this case consists in altering the anode potential of the valve by means of the series resistance R in the anode circuit ; it has the advantage of being a very smooth form of control.

---

*NOTE : The inductance of an R.F. choke together with its stray capacity form a rejector circuit having a comparatively low resonant frequency. Since it is usually operated above this frequency it behaves as a capacitive and not an inductive load.

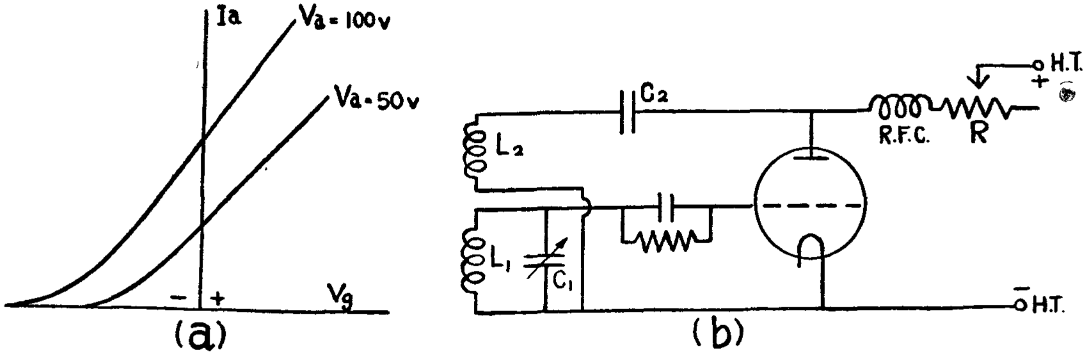


FIG. 21.11. Anode potential reaction control

L_2 and C_2 are chosen so that with moderate H.T. voltage the valve does not oscillate ; on raising the H.T. voltage the signal becomes progressively louder until oscillation sets in. Another advantage of this form of control is that it does not affect the tuning of the main circuit.

Class A, B and C operation

In general there are three ways in which a triode can operate ; when the grid input amplitude is not sufficiently large to take the swings into the curved regions of the I*a*—V*g* curve, the valve is under class A conditions. We have already met this type of operation in discussing the triode as a class A amplifier (see p. 98). The characteristic of class A operation is its linearity, i.e. the output is an exact and distortionless copy of the input. If the input is a modulated wave, i.e. a carrier with sidebands then the output will be a similar modulated wave, with no more and no fewer frequency components.

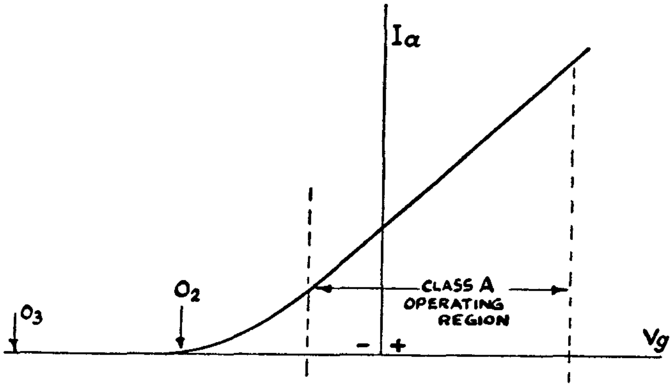


FIG. 22.11. Class A, B and C operation

In class B operation the grid is biased back to anode current cut-off (O_2) ; only the positive swings of the input will cause anode current to flow and the output will be distorted ; that is to say it will contain frequency components which were not present in the input. We have dealt with Class B operation in the section on anode detection ; it will be remembered that the H.F. modulated input is resolved, in this type of operation, into various frequencies in the anode circuit, one of which is the modulation frequency ; the carrier frequency and other H.F. components are present also, but beyond using the former for reaction we had no further use for them and by-passed them to earth.

Class C operation (biased to at least double cut-off) is an extreme form of Class B and is used in high frequency amplifiers for transmitters when a high power efficiency is sought. Like Class B it will give rise to distortion, and, of course, if the negative bias is very big indeed, the input may be insufficient to cause any current to flow in the valve at all.

This view of Class B and Class C operations as a method of introducing frequencies not present in the original input is one which merits a little study, because the same process enters into several items of radio practice, which might be thought to be unconnected. The particular frequencies introduced are harmonics and sum and difference frequencies. If the input is at a single frequency, then Class B or C operation will introduce a whole range of harmonics whose amplitudes depend on the valve characteristics, and the conditions of operation. If the input contains two or more distinct frequency components, the output will contain, in addition to a range of harmonics of each of them, frequencies equal to their sums and differences, taken in pairs. It is possible to design the output circuit to respond to whichever of this collection of frequencies we desire to preserve, while we can by-pass the unwanted components to earth and prevent their developing any power. As an example of this, take a simple grid modulator ; here we have a triode under Class B conditions, and the grid input consists of (1) the high frequency feed-back from the oscillating anode circuit (f) and (2) the modulation frequency (m). In the anode circuit therefore we shall find the following components :—

(1) f + a range of harmonics.
(2) m + its range of harmonics.
$$\left.\begin{array}{l}(3)\ f + m \\ (4)\ f - m\end{array}\right\}\text{sum and difference frequencies.}$$
(5) harmonics of $f + m$ and $f - m$.

The anode circuit being tuned to f will respond therefore to f, $f \pm m$ (because the two latter frequencies are so near in value to f), but the other components will disappear on account of the fact that the anode load offers them no appreciable impedance.

Again, take an anode detector ; the input is a modulated wave, consisting, with the same notation, of three components

$$f - m, f, f + m.$$

The characteristic Class B operation of the anode detector will therefore produce in its output circuit

(1) harmonics of f, $f + m$, $f - m$
(2) sum and difference frequencies :—

$$\begin{array}{ccc} 2f - m\,; & 2f\,; & 2f + m \\ m\;\;; & 2m\,; & m \end{array}$$

Examination will show that all these components fall into one of two classes, R.F. (f, $2f$, $2f - m$, etc.), and A.F. (m, $2m$) ; those in the first class are by-passed by a condenser to earth, and we are left with the modulated frequency (m) and its second harmonic ($2m$) to operate our telephones (see fig. 13.11). It is a matter of design and adjustment to make the effect of $2m$ as small as possible. It would take us far beyond the scope of this book to discuss the relative amplitudes of the frequency components introduced by Class B and C operation ; it is a matter of good fortune that the amplitudes of those components which we actually want to preserve are usually the greatest, while the amplitudes of the higher harmonics are very small indeed.

Similar considerations apply to diode operation ; the ordinary diode detector is of type B operation and the distortion so introduced is equivalent to the creation of sum and difference frequencies in the circuit ; the load is designed to select from the frequencies so produced the modulation frequency giving the signal. A grid-detector is a hybrid case; here the filament and grid act as a Class B diode detector and the filament, grid and anode, as a Class A amplifier.

C.W. reception—heterodyne

This leads us quite naturally to the reception of C.W. which, we have seen, cannot be received successfully by any of the standard methods.

If, however, an additional frequency (f_1) is injected into the tuned circuit of an anode detector resonating to the C.W. (f), the grid circuit will be driven at two frequencies, f and f_1, and the anode circuit will contain components at :—

$$(f+f_1) \text{ and } (f-f_1).$$

Now f_1 can be made as near f as we like (say 1,000 c/s less). Then in the anode circuit we shall have an audio oscillation of 1,000 c/s., the difference frequency, which will operate the telephones directly. Oscillations at $2f$ and $f+f_1$ will also be present; they are both at R.F. and can easily be by-passed to earth, but the important point is that we now have a method of C.W. reception of great flexibility.

There are two methods of applying this device :—

(i) Separate heterodyne oscillator.
(ii) Autodyne.

Fig. 23.11 shows a circuit using the first method. The valve V_1 is the local oscillator, and the valve V_2, acting as an anode detector, gives the heterodyne note in the telephones.

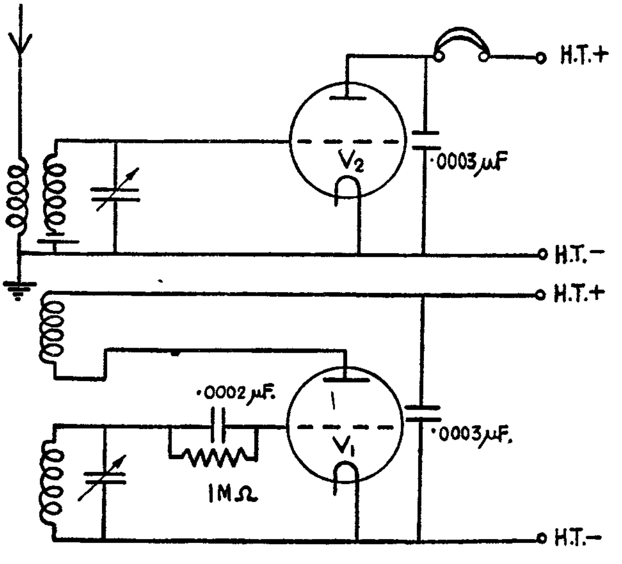


FIG. 23.11. Separate heterodyne circuit

A signal is "found" by adjusting the aerial condenser, and when the frequency of a received signal is such that the difference between it and that of the local oscillation is an audible frequency, the corresponding note is heard in the phones. As the local frequency is altered to become nearer to the signal frequency, the note in the telephones gets lower in pitch, and when the two become the same there is silence or a "dead spot."

There is no need to have a separate valve oscillator; a detector provided with reaction is sufficient; the reaction can be pushed to the point of oscillation and the valve is then in a condition to receive C.W. directly; it is in fact a small grid-modulated oscillator only; whereas a transmitter modulator produces a difference term of radio frequency, an autodyne produces a difference term of audible frequency, the pitch depending on the relative frequencies of signal and oscillation.

Fig. 24.11 gives the circuit of an autodyne arrangement.

The valve V_1 functions both as detector and oscillator; the input circuit now contains components at :—

(i) Input frequency f;
(ii) At the oscillation frequency f_1;

and therefore the anode circuit contains (among several R.F. components) an audio-frequency component (f-f_1).

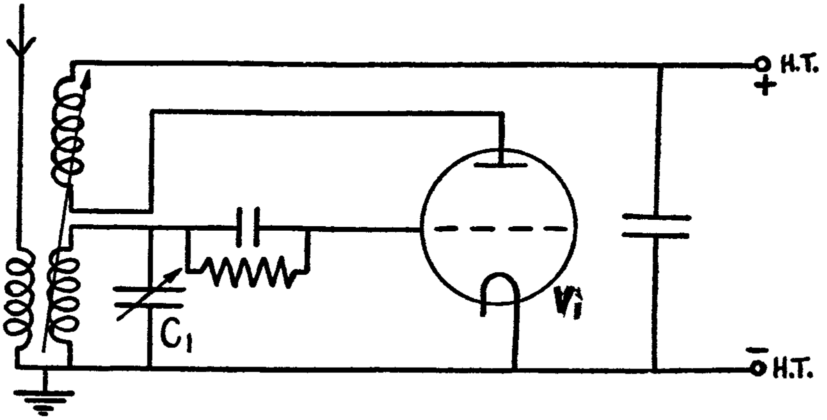


FIG. 24.11. Autodyne arrangement

The pitch of this note will depend on the setting of the condenser C_1; when this setting (which controls the oscillation frequency) makes f_1 higher (or lower) than f, we shall get a note of high pitch or vice versa.

It is interesting in connection with this discussion to note that the human ear, like a valve, is a device which makes sum and difference frequencies; if two whistles of slightly different pitch, p_1 and p_2, are sounded together the ear will become conscious of sounds of frequency :—

$$p_1, p_2, p_1 + p_2, p_1 - p_2,$$

probably $p_1 + p_2$ will be so high as to be outside the audible range, but $p_1 - p_2$ will be heard plainly and gives a piercing quality to the sound. Outside the ear, in the air, these sum and difference frequencies do not exist; the two sound waves interact to form a combined wave whose amplitude rises and falls at the difference frequency, but the new frequencies owe their objective existence to the characteristics of the human ear, just as the sum and difference frequencies in the anode circuit of a detector owe their existence to the characteristics of the thermionic valve.

Heterodyne wavemeter

An autodyne receiver can be made to function as a wavemeter; the tuned circuit is carefully calibrated, and for every setting of the condenser dial the oscillation frequency of the circuit can be found by reference to a calibration chart.

An incoming C.W. will produce a heterodyne note in the 'phones, and by adjusting the condenser the pitch of the note may be made lower and lower until it finally disappears. At this point the radiation frequency and the oscillation frequency are the same, and since the latter is known the former is determined.

Such an instrument is really a low-power C.W. transmitter; if it is set to a predetermined frequency its carrier can be picked up by heterodyne on a nearby receiver (provided the latter has some reaction device by which it can be thrown into oscillation) and the latter can be adjusted exactly to receive the desired frequency.

Sometimes a heterodyne wavemeter is provided with a modulating device, so that it sends out "*tonic train*", that is C.W. with a single low-frequency modulation. In that case its radiation can be picked up by a non-oscillating receiver and the latter can be set to any required frequency within the range of the wavemeter.

Fig. 25.11 gives the circuit of a typical heterodyne wavemeter.

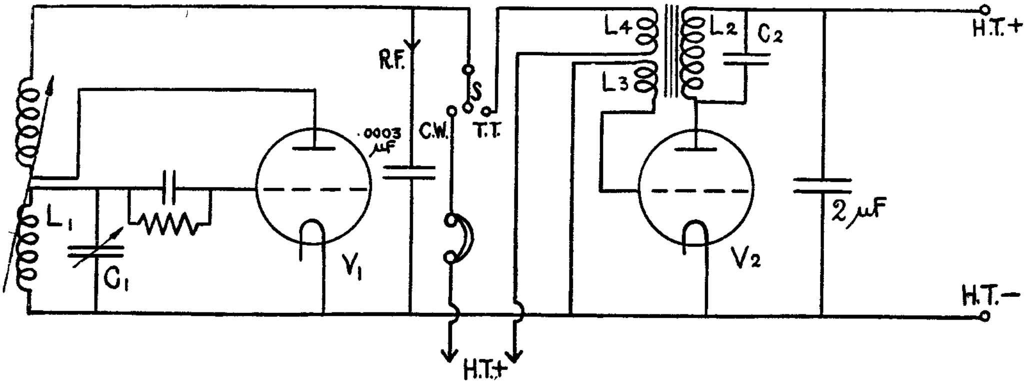


FIG. 25.11. Heterodyne wavemeter

V_1 is an oscillator whose tuned circuit L_1 C_1 is carefully calibrated ; with the switch S over to the left, the anode of this valve is connected via the telephones to H.T. and the instrument acts as a heterodyne wavemeter simply.

V_2 is a modulator valve ; the anode circuit L_2 (one winding of an iron cored transformer), together with condenser C_2, forms an audio-frequency tuned circuit of frequency a few hundred cycles per second. Reaction to this circuit is provided by the grid circuit L_3.

With the switch S over to the right the anode of the oscillator is connected to H.T. via another winding (L_4) on the transformer, and in consequence the voltage in the anode of V_1 rises and falls at the oscillation frequency of V_2 ; thus a modulated output is provided by V_1, which, as stated above, can be used to set a non-oscillating receiver to a given frequency.

Valve Voltmeter Wavemeter

The anode detector (p. 130) can be made to function as an A.C. voltmeter, and as such is often employed in place of the thermo-ammeter as a resonance indicator in a wavemeter circuit.

An alternating voltage applied between the grid and filament of a triode operating under class B conditions will, by rectifier action, give rise to a direct current in the anode circuit, and this current, passing through a suitably calibrated instrument gives directly a measure of the applied alternating grid voltage.

Further than this, the steady negative bias applied to the grid ensures that little or no current is drawn from the circuit operating it. In this way damping effects are avoided in the circuit of the wavemeter when using a valve voltmeter as a resonance indicator, and in consequence the determination of the resonant condition is sharp and accurate.

CHAPTER XII

AMPLIFICATION

In the previous chapter we saw how we could extract signals from our modulated R.F. waves by means of some form of detector.

The strength of a signal issuing from a detector is usually fairly low, and so far we have only mentioned reaction as a method of increasing it ; reaction, of course, has only a limited application, and we are therefore driven to adopt some device whereby the signal can be magnified or amplified, while at the same time it suffers no serious distortion. While such distortion is not serious in the case of telegraphy, it is a thing to be guarded against in telephony, even though first quality reproduction is not required.

In Chapter IX we discussed the amplifying property of a triode ; the alternating voltage applied to the grid reappears in amplified form across the load in the anode circuit, and if the grid-swings have been limited to the straight part of the Ia—Vg curve the magnified signal is distortionless.

Amplification which takes place after detection is called *audio-frequency amplification*, and an audio-frequency amplifier will consist of two sections :—

(i) one or more triodes operating as class A voltage magnifiers ;

(ii) output stage, in which the magnified A.F. signal voltage is required to deliver power into the 'phones or loudspeaker.

The circuits or components used to pass on the signal from detector to amplifier, or from the first to the second amplifier, are called *intervalve couplings* ; in A.F. amplifiers these couplings must be able to respond evenly at all the frequencies in the acoustic range without suppressing some or accentuating others ; they must obviously be circuits with very flat frequency response curves. The three common types of coupling are—

Resistance-capacity coupling ;

Choke-capacity coupling ;

Transformer coupling.

We shall discuss them in turn.

Another way of securing a strong signal from the detector stage is to magnify the R.F. voltage delivered by the aerial tuned circuit, in one or two steps, before it is applied to the detector for demodulation. This is *radio-frequency amplification* and its technique presents some serious difficulties ; the couplings in an R.F. amplifier will only have to deal with a narrow band of frequencies, represented by the carrier and its sidebands, and it will therefore be advantageous to make them selective ; for this reason such couplings are usually tuned circuits of some kind and, in addition to their coupling function, they afford a very considerable increase in the overall selectivity of a receiver.

Audio-frequency amplifiers—Section I

This section deals only with those valves and their associated couplings which form the A.F. voltage magnifiers. The output stage is discussed in the next section.

The following schematic diagram (fig. 1.12) will help us to visualise the layout of such an amplifier, which has two valves in this case. The audio-frequency voltage from the detector (1) is transferred by the coupling (2) to the first amplifier (3), which is usually

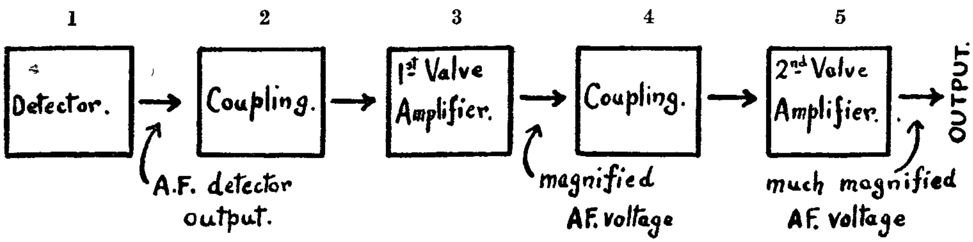


FIG. 1.12. Layout of 2-valve amplifier

a class A operated triode ; the magnified A.F. voltage then passes by way of coupling (4) to the second amplifier (5) (another class A triode), where it is magnified again, and is then passed on to the output stage where it is used to control the power supply to the loud speaker. The couplings of such an amplifier now claim our attention.

(1) *Resistance capacity coupling*

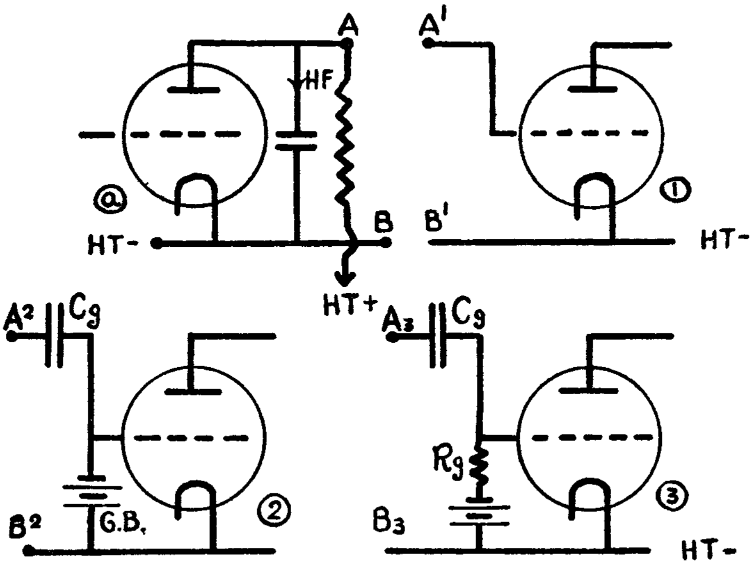


FIG. 2.12. Resistance capacity coupling

Fig. 2.12 (a) represents a detector stage (most of the detail being omitted for the sake of simplicity) with a resistance load R in the anode circuit. Across R will be produced the A.F. voltages which we have to amplify; let us proceed by trial and error. Try circuit (1) as a coupling, that is, join A to A_1, B to B_1 when the voltage across R will be acting across the grid-filament circuit of the amplifier ; *but so will the voltage of the H.T. battery* ; the grid of the amplifier will be something like 100 volts positive and probably the valve would be ruined by excessive grid and anode current—so this will not do.

Secondly, try putting a condenser C*g* (circuit 2) between A and the grid, and at the same time include a bias battery to maintain the amplifier grid at the correct operating point. C*g* will block the H.T. voltage entirely and thus protect the amplifier grid, and the bias will do the job assigned to it, but still we shall get no results for this reason :—

The audio signals across R are represented in fig. 3.12 by an alternating current generator G. The output of this generator will flow in a circuit consisting of condenser C*g* and a small resistance R*b*, the internal resistance of the G.B. battery ;

therefore most of the generator voltage will be developed across C*g* owing to its high reactance compared with R*b*, and very little will be applied between grid and filament of the amplifier.

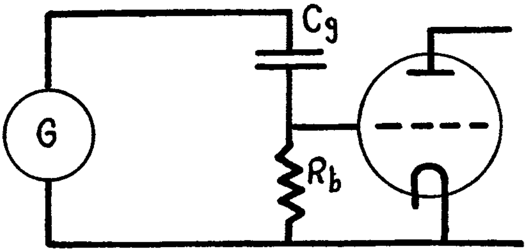


Fig. 3.12. Showing necessity for grid resistance

The next step, therefore, is represented by circuit 3 (fig. 2.12), where a high resistance (·25 to 2 megohms) is joined in series with the grid bias battery. In this circuit *most* of the output voltage is developed where it is wanted, i.e. across R*g* and not across C*g*. Thus we are led to the conventional circuit of a resistance capacity coupled amplifier (fig. 4.12).

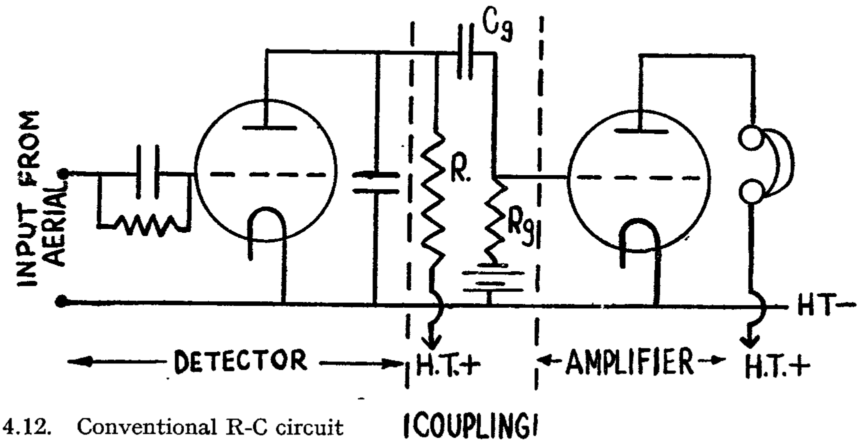


Fig. 4.12. Conventional R-C circuit

The amplifier valve must be supplied with sufficient high tension voltage so that a good length of the straight part of the I*a*—V*g* curve is within the negative-grid region ; this ensures that no grid current flows, and at the same time the amplification is linear (see fig. 5.12).

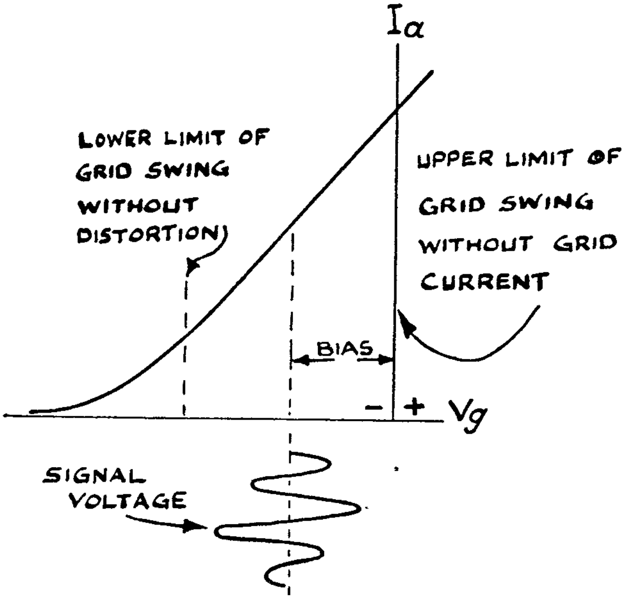


Fig. 5.12. Limit of permissible grid swing

The value of C*g* is of some importance ; if C*g* is too small its reactance to low audio-frequencies will be great and consequently too great a proportion of lower A.F. voltage will be wasted across C*g* ; the signals becoming attenuated, that is, not rich enough in bass notes. If C*g* is too big we shall get what is called " grid blocking " ; after an extra strong signal, the amplifier will distort badly for a second or so. Normal values of the condenser are 0·1 μF. to 0·002 μF. depending on the value of R*g* ; the product R*g*.C*g* should be about 0·005 for good all round performance. The anode resistance R should not be too big ; about three times the slope resistance of the valve in whose anode circuit it is connected is a good average rule for estimating its value.

(2) *Choke capacity coupling*

One disadvantage of a resistance capacity coupling is that an inconveniently high voltage must be used to operate the valve satisfactorily through the anode resistance; a big proportion of the H.T. voltage is lost as direct volts drop in this resistance. To avoid this, an iron cored choke (10 to 50 henries) can be used in place of the anode resistance. The choke is relatively low in D.C. resistance, but offers considerable reactance to alternating currents. Apart from this difference, choke capacity (C.C.) coupling behaves just like resistance capacity (R.C.) coupling; a typical C.C. circuit is shown in fig. 6.12.

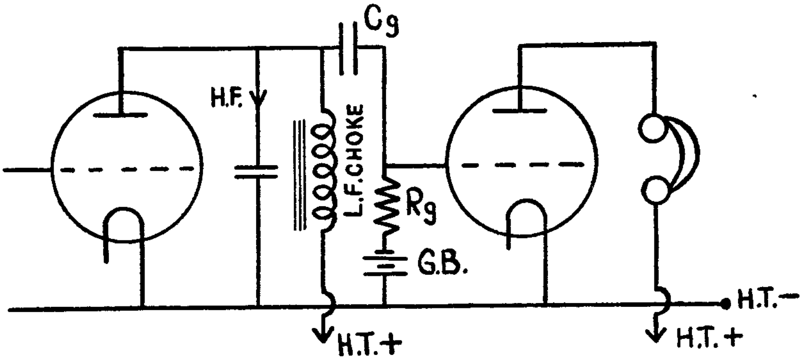


FIG. 6.12. Choke capacity coupling

Generally speaking R.C. coupling is to be preferred to C.C. coupling. In the latter case the A.C. voltage developed across the choke will depend on the frequency, hence some A.F. voltages will be amplified more than others; the low A.F. notes will be attenuated at the expense of the high ones because the reactance of a choke goes *up* with the frequency.

If we use a choke of extra high inductance with the object of avoiding low-note loss we shall be adding a lot of self capacity (see p. 77) and the choke will be equivalent to a low frequency rejector circuit. Notes whose frequencies are about the resonant frequency of this circuit will be unnaturally accentuated and notes above the resonant frequency will suffer serious cut-off since the corresponding A.F. currents will pass the choke via its self capacity and will produce no voltage worth speaking of to operate the succeeding amplifier.

(3) *Transformer coupling*

In R.C. and C.C. coupled amplifiers, the only gain is due to the magnification afforded by the valves themselves; in the type of coupling about to be described, another source of gain is present in the voltage step-up property of a transformer.

In the simplest type of transformer coupling, the intervalve transformer consists of two windings, primary and secondary, wound on a laminated core of a magnetic alloy. The primary winding forms the load of a detector valve, and the secondary forms the grid-filament input circuit to the amplifier.

If another stage is employed, the primary of the next intervalve transformer is in the anode circuit of the first amplifier, and the secondary is the input circuit to the second, and so on. The secondary is biased in order to maintain the grid at a suitable operating point. Fig. 7.12 shows a conventional transformer coupled detector and amplifier.

The signal voltages developed across the primary winding in the anode circuit of the detector are stepped up by transformer action into the secondary, whence they are applied to the grid-filament terminals of the amplifier. The amplifier load in fig. 7.12 is a pair of head phones, in which the signal is reproduced after magnification.

Unless the transformer is carefully designed, this type of coupling suffers from the same defects as choke capacity coupling, namely, low note loss, together with high note cut-off due to the necessity for fairly low primary inductance and to the self capacity of the windings respectively.

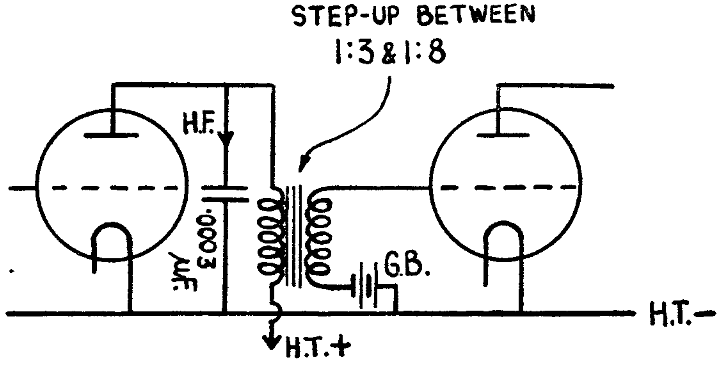


FIG. 7.12. Transformer coupling

The normal transformer coupling just described has one rather serious drawback ; the primary winding has to carry a direct current (the steady component of the anode current of the previous valve), which has all sorts of evil effects ; from the constructional point of view alone it renders a transformer core of relatively large bulk imperative if distortion due to magnetic saturation is to be avoided.

We can get out of the difficulty by using a *parallel-fed transformer.*

Fig. 8.12 shows a detector valve (V_1) following a parallel-fed transformer and a single amplifying valve (V_2).

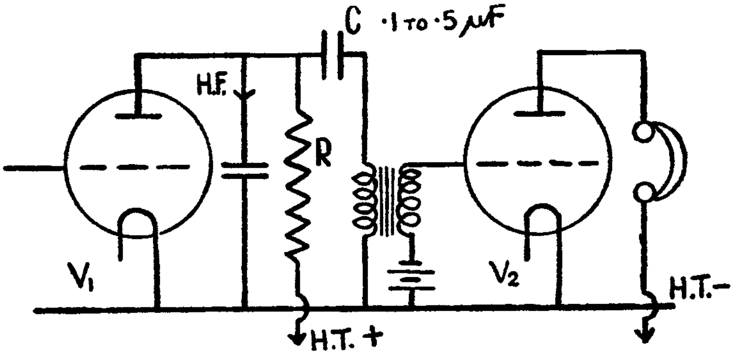


FIG. 8.12. Parallel-fed transformer

As in the case of R.C. and C.C. coupling, A.F. voltages are developed across the anode load R, which also conveys the steady anode current to earth. The coupling condenser C and the transformer primary are in parallel with R, and across them appear the A.F. signal voltages. The reactance of C is very low at such frequencies, and therefore they act almost entirely across the primary winding, into which no D.C. can get owing to the condenser C. Core saturation is thus avoided, and it is possible to make the core of high permeability material (such as mu-metal) and at the same time to reduce its bulk. The high permeability renders possible a reduction in the number of turns in the windings with a corresponding diminution of self-capacity and avoidance of high note loss. The overall magnification with a parallel-fed transformer is a little less than in a similar amplifier using a similar series-fed transformer.

In the example discussed above we have dealt only with a one-stage amplifier, following a detector valve. A second stage of amplification can be applied, using the appropriate coupling and another triode. It must be remembered, however, that the input to the second amplifying valve is now considerably bigger than that to the first, and additional care must be taken to avoid overloading. A two-stage transformer coupled amplifier is shown in fig. 9.12.

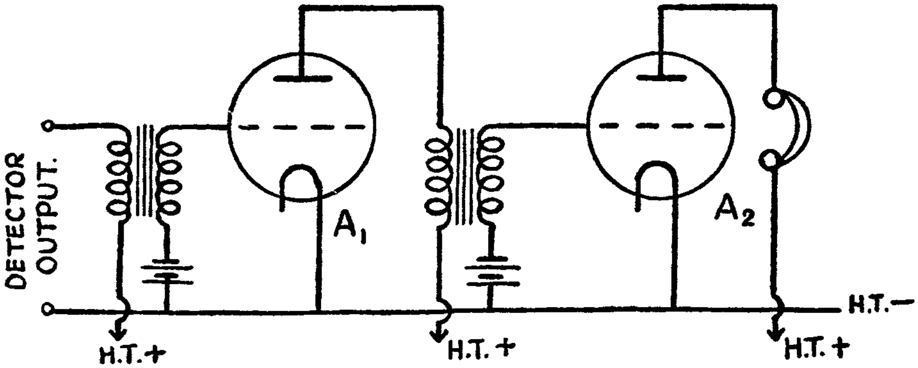


FIG. 9.12. Two-stage amplifier

A.F. Amplifiers—Section II

This section deals with the last of a series of A.F. amplifying valves, and with the coupling through which it delivers power to operate the telephones or loudspeaker. The grid-filament circuit of this valve is operated by the magnified voltage developed in the first section of the amplifier ; we are not concerned with voltage amplification any longer, we want to make this last valve deliver power, that is to say its load must develop a volts drop which is an undistorted replica of the input, and at the same time pass a considerable current. The power, of course, will come from the H.T. battery ; the amplified input voltage serves to " trigger " this power to operate the loudspeaker.

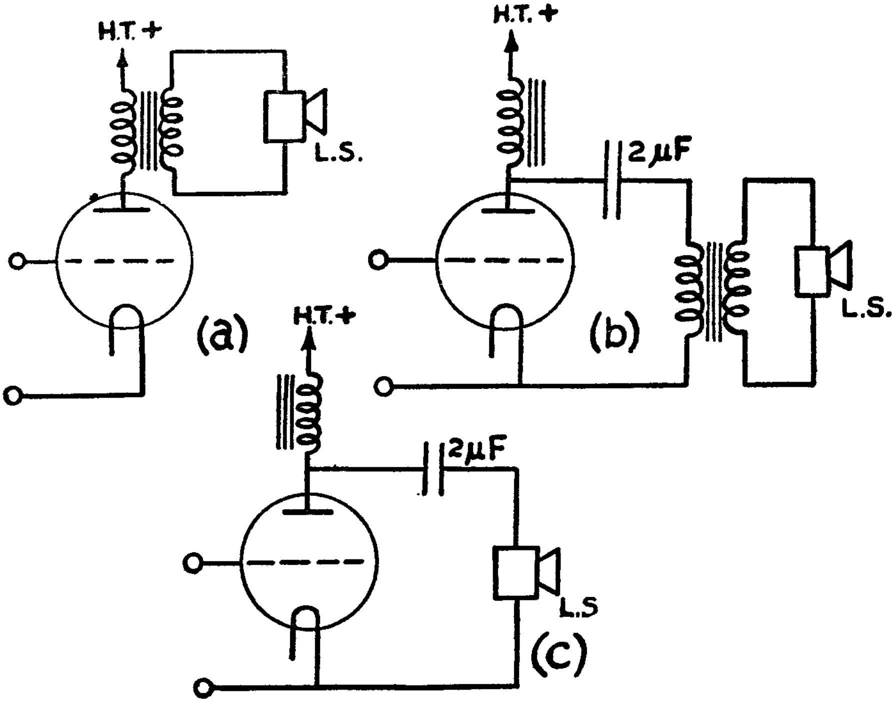


FIG. 10.12. Output circuits

In order to achieve this aim the valve must have a low A.C. resistance because, like any other generator, the amount of power it can deliver falls off as the internal resistance rises ; triodes designed for loudspeaker operation have an A.C. resistance between 500 and 3,000 ohms. The anode voltage must be high, and it must pass a fairly large steady anode current. Finally, it must be matched to the anode load so that it provides maximum power with minimum distortion.

The best load for a triode is about twice the A.C. resistance of the valve ; for a pentode the optimum load is about one-quarter of the valve slope resistance.

Of course the impedance of a speaker or of telephones varies with the frequency of the signal, so a compromise is made in which matching is adjusted to an audio frequency of about 400 c/s.

In order to fulfil the condition of high anode voltage the output coupling must be either of the transformer or choke capacity type. Resistance capacity coupling is undesirable because so much of the H.T. voltage would be wasted as volts drop in the anode resistance.

Three suitable output circuits are shown in fig. 10.12 ; (*a*) shows a transformer output ; the turns ratio is adjusted so that the reflected resistance in the primary is twice the A.C. resistance of the triode (see p. 77) ; (*b*) shows a choke-capacity output, with a transformer to give the necessary matching ; (*c*) is a choke-capacity output with a tapped choke which permits of matching on the auto-transformer principle.

Output valves in parallel

One way of achieving an output stage of low A.C. resistance is to put two small valves in parallel, grid to grid, anode to anode ; such a combination will certainly deliver more power, but it is just as prone to distortion as a single valve.

Output valves in push-pull

Fig. 11.12 shows a push-pull disposition. The input is divided between the two valves by means of an A.F. transformer with centre-tapped secondary. This is at once an advantage ; the complete input might overload a single valve ; half the input is not so likely to do so.

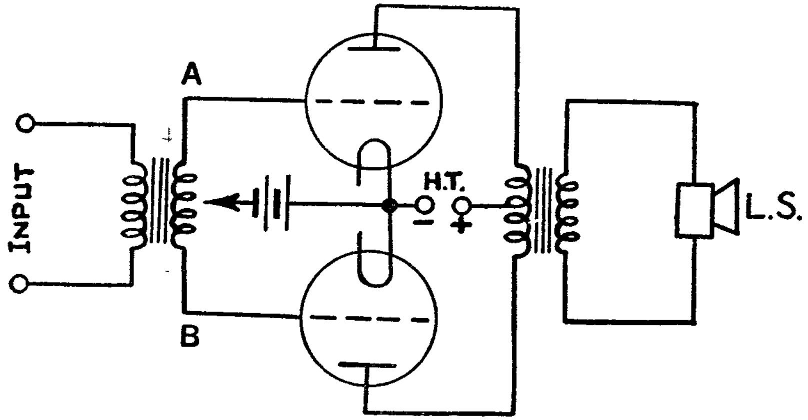


FIG. 11.12. Push-pull disposition

The centre point of the input secondary can be regarded as " earthy ". When end A is positive, end B must be negative, and thus the inputs to the two valves are 180° out of phase. So far as the output primary goes, the steady anode currents from the two valves will be flowing in opposite senses, and the effective current in this winding will be equal to their difference ; if the valves are accurately matched and draw the same anode current, the effective current will be zero and the core unpolarised.

Fig. 12.12 shows how, even if the inputs produce distortion in each valve, the combined output in the transformer is undistorted and, further, the steady component of anode current in the primary winding is zero. This last effect frees the amplifier from the possibility of distortion due to magnetic saturation of the core and permits the employment of relatively small transformers.

One disadvantage of a push-pull output stage is that it is a serious drain on the H.T. supply; it is inefficient; most of the power is wasted in heating up the valve anodes and only a relatively small amount is used up in operating the loudspeaker.

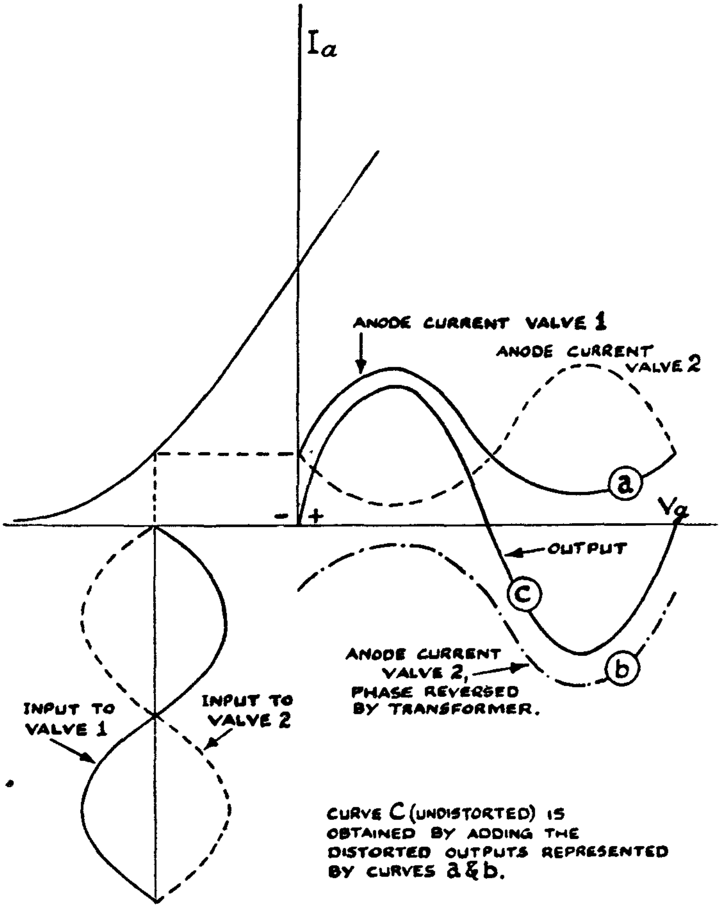


Fig. 12.12. Elimination of distortion in P.P.

But since by its very nature it cancels out distortion due to operation on the lower bend of the Ia–Vg curve, we can bias the valves to cut off if we so desire, and the current in the valve, when no signal is present, will be practically zero.

This system is what is called " quiescent push-pull " (Q.P.P.) and represents a considerable advance in the direction of efficiency and economy in H.T. supply; it is in fact a sort of class B push-pull system but it is also very insensitive. A much bigger input is required for full operation than in the case of normal class A push-pull.

Its development has led to another class of power amplifier called a Class B audio amplifier.

Class B amplification

In order to overcome the insensitivity of Q.P.P., and at the same time retain its efficiency, it seems a fairly obvious step to employ valves of higher magnification.

Fig. 13.12 shows the I_a–V_g curves for two valves, (*a*) being a normal valve of about $\mu = 4$, and (*b*) being a high magnification valve, say $\mu = 30$, both operated at the same anode voltage.

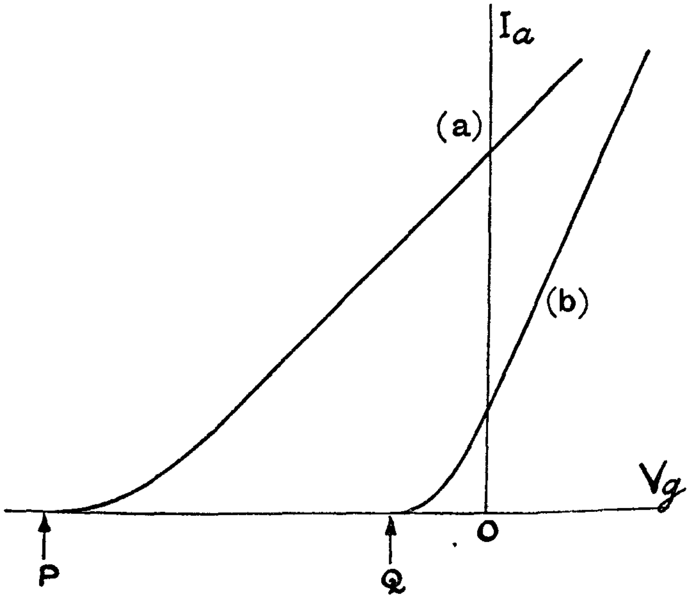


FIG. 13.12. Characteristics of high and low magnification valves

From the diagram it is clear that in the case of (*a*) used under Q.P.P. conditions there is room for a considerable input before grid current flows ; the input amplitude can rise to OP if necessary. In the case of (*b*) however, only a very small input (amplitude QO) is possible without grid current.

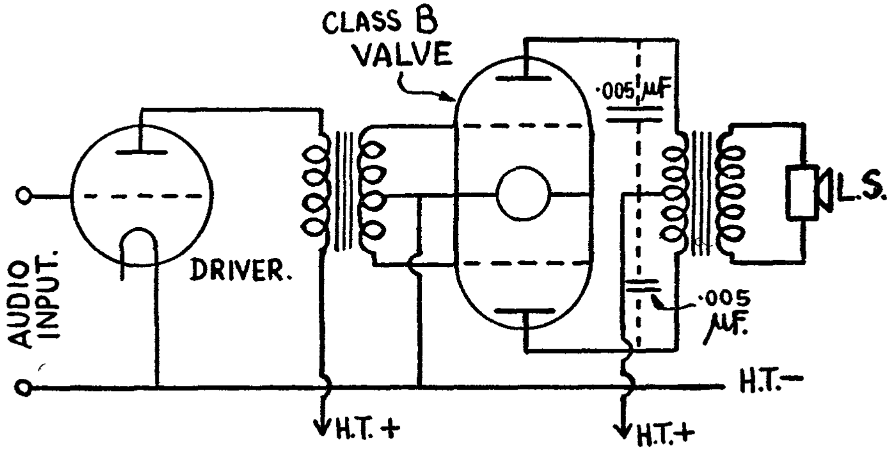


FIG. 14.12. Class B amplifier with driver

If, therefore, we are going to use the high μ valves in our amplifier we must face the possibility of grid current, and allow for it in the design ; this allowance and the high μ valves represent the difference between a class B audio amplifier and the ordinary Q.P.P. system.

The audio-frequency signals from the first section of the amplifier are applied to a small power valve called a driver ; the output from the driver is applied across the primary of the audio-frequency transformer, whose secondary is arranged in the normal push-pull fashion, with a centre tap.

The two high μ valves in push-pull are normally in a single glass envelope (Class B valve), the two grids being operated by the secondary of the input transformer ; the output follows ordinary push-pull practice. The design of the transformer is a matter of some importance : its secondary will have to carry the grid current of the Class B valve and to ensure that this *variable* load will not cause undue fluctuations in the voltage developed, the windings must have as little resistance as possible and the A.C. resistance of the driver valve should be reasonably small. The driver valve is acting as a power amplifier (since grid current is flowing in the output circuit) ; its output A.C. resistance must, therefore, be matched to the input A.C. resistance of the output system. This is done by a correct choice of driver transformer ratio. If the anode A.C. resistance of the driver is R_a ohms and the total input resistance to the class B valve is R_b ohms, then the correct ratio (n) for the driver transformer is given by

$$n = \sqrt{\frac{R_b}{R_a}}.$$

The ratio is usually about 1 : 1.

Class B amplifiers are rather apt to generate unwanted oscillations ; this tendency is checked by connecting small condensers (0·005 μF.) between anode lead and centre tap of the output transformer.

The pentode

Triodes designed for power amplification require large inputs for maximum power output and those capable of high audio outputs require a large D.C. power supply to the anode.

Greater sensitivity could be obtained by using a screen-grid valve, but such a valve would require the screen potential to be high if a large output is to be obtained, and there would be a great possibility that during at least parts of the output cycle, the anode potential would be less than that of the screen.

When this happens, the secondary electrons emitted from the anode by the impact of the normal anode current stream, will pass to the screen because its potential is higher than that of the anode and thus produce a fall in anode current. Under such conditions the valve is unstable and, therefore, owing to the effects of secondary emission, ordinary tetrodes cannot be used as output valves. By placing an earthed screen between the anode and the screen grid we prevent this action ; secondary electrons will return to the anode because its potential will always be higher than the earthed screen ; anode current does not fall and instability is overcome.

Such is the Pentode, which is a sensitive output valve. Pentodes are about 30 per cent. efficient ; output triodes are about 15 per cent. efficient.

Radio frequency amplification

We now turn to the pre-detector amplifier, which magnifies the R.F. voltages appearing across the aerial circuit before their application to the detector itself. For grid detectors the optimum input is about ·25 V. R.M.S. ; anode detectors require about 2·0 volts R.M.S. The diode detector operates satisfactorily over a fairly wide range, and if the input exceeds 5·0 volts R.M.S. it is practically distortionless.

The voltage across the aerial circuit is generally of the order of millivolts or even microvolts, so that to operate a detector satisfactorily it is necessary to resort to some method of magnification. We have already mentioned reaction, which is a simple and direct method of strengthening the signal, but it carries grave disadvantages in its effect on the purity of the final output of the receiver. The obvious course is to employ a valve amplifier between aerial circuit and detector. Such an amplifier consists of one or more valves, with their associated couplings, and, just as in the audio-frequency case, each valve and its coupling will contribute a certain stage gain.

The fundamental difference between an R.F. and an A.F. amplifier is that, while the latter has to deal with a wide band of frequencies, the former has to deal only with a very narrow band, and we may therefore expect to find " tuned " couplings in the R.F. amplifier. Such couplings not only increase the stage gain, but they also contribute materially to the overall selectivity of the receiver.

We proceed to discuss the various types of coupling.

(1) *Tuned anode coupling*

Resistance-capacity coupling is a simple and efficient one in A.F. amplifiers. It is practically useless in R.F. work on account of the stray valve capacities which offer so little reactance to radio-frequency current.

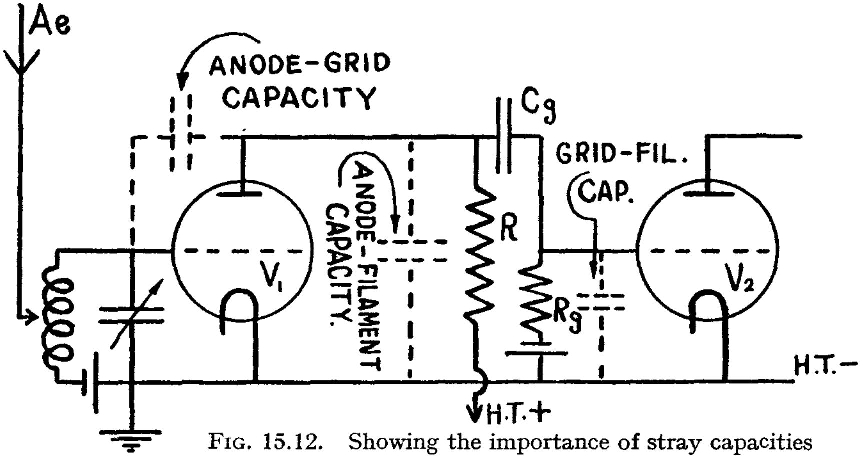


FIG. 15.12. Showing the importance of stray capacities

In fig. 15.12, which represents an R.F. resistance capacity amplifier, it will be seen that the anode resistance R has in parallel with itself :—

(1) The anode filament capacity of V_1.

(2) The grid filament capacity of V_2.

These capacities amount to only a few micro-microfarads, but at radio-frequencies their reactance is not more than a few thousand ohms, and in consequence the stage gain is small, and practically independent of the anode resistance R. Another important point is that the feed back through the anode-grid capacity acts as negative reaction (because the load is capacitative) and thus damps the aerial circuit badly. Fortunately, we are

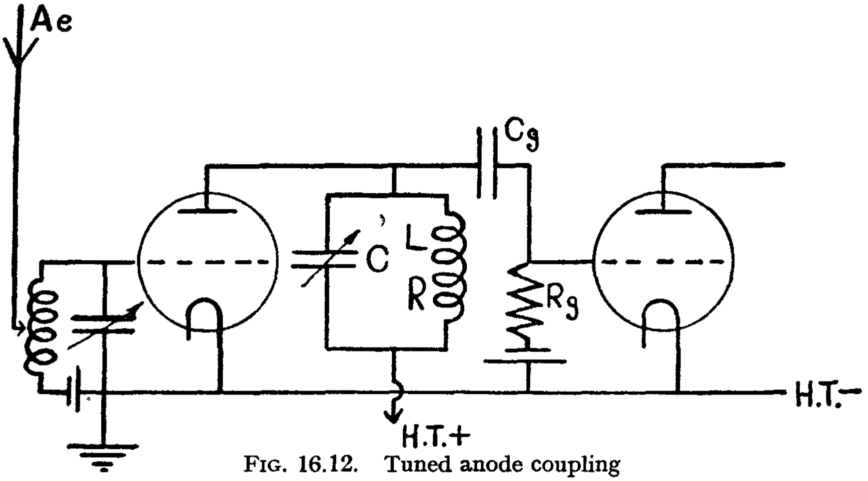


FIG. 16.12. Tuned anode coupling

able to meet most of these difficulties by using a tuned circuit in place of the anode resistance R.

In the circuit shown in fig. 16.12 the anode load consists of an inductive coil L in parallel with a variable condenser. The unwanted capacity can be considered as another small fixed condenser in parallel with the load. Condenser C can be adjusted so that with the coil and the stray capacities the whole circuit is in resonance with the input frequency, and the load behaves like a resistance of value $\frac{L}{CR}$ but without the effect of the stray capacities. These latter have, in effect, been now put to good use in contributing to the capacity required for resonance in the anode load.

This is the tuned anode circuit ; the condenser C will have to be readjusted whenever the aerial condenser is altered to tune the aerial system to a different frequency.

Notice that R*g*, the input resistance of the next amplifier, is in parallel with the tuned circuit which will therefore be damped unless R*g* is very big ; it should be not less than a megohm, and the capacity of C*g* should be such that the R.F. volts-drop across it is small compared with that across R*g*, because this latter volts-drop is the input to the next valve. C*g* is normally between ·0005 and ·00005 μ F. If the valve following the R.F. amplifier is a detector, then C*g* and R*g* have their detector valve values round about ·0003 μ F. and 2 M Ω respectively.

This seems a simple and effective coupling ; it has defeated the difficulty of stray capacities, but if the reader looks at the circuit carefully he will see that it is identical with our T.P.T.G. oscillator mentioned in Chapter X ; it is, in fact, unstable. It might be possible to operate one stage of such a coupling in an R.F. amplifier, but to operate two would be impossible unless we can overcome the feed-back due to the anode-grid capacities of the valve.

In the next chapter, in which we shall deal with R.F. amplifiers for transmitters, we have to face the same difficulty, and there will be found described what are called " neutralising circuits "; their function is to feed back, from anode circuit to grid circuit, alternating voltages equal to that fed back through the anode grid capacity, but *in the opposite phase*, so that the total feed back is zero and the system is stabilised. It would be quite feasible to apply such a circuit in the case under discussion, but we have a much better device for ensuring stability in the form of the tetrode or screen-grid valve.

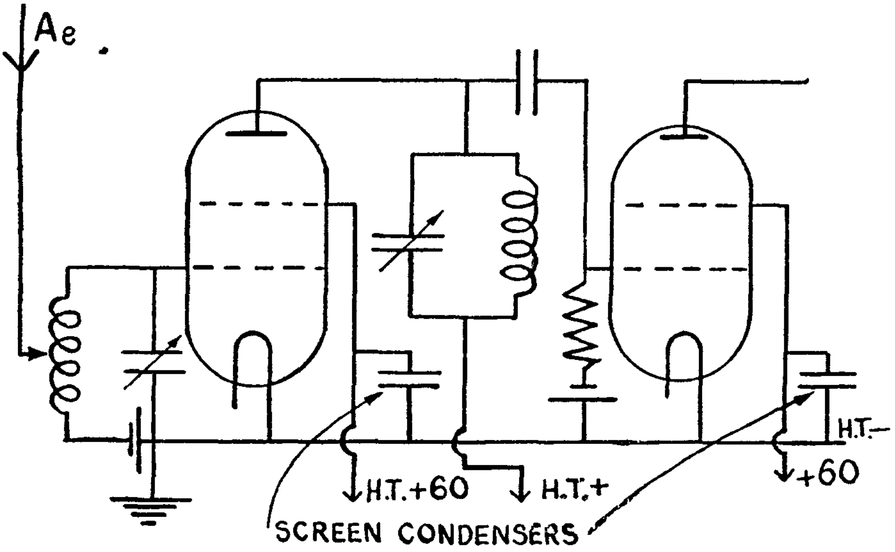


FIG. 17.12. Use of the 4-electrode valve

The tetrode was described briefly in Chapter IX ; the screen between control grid and anode reduces the anode-grid capacity to a few thousandths of a micro-microfarad, but in order to make the valve function the screen must be maintained at about 60 volts above the filament, and what is also vitally important, it must be joined through a fairly large condenser (·2 μ F.) to earth. This large capacity, with its very low reactance at R.F., can be considered to maintain the screen at a constant earth potential in respect of the R.F. operations going on in the valve. Apart from these two points (screen voltage and screen condenser) the tetrode behaves just like a triode whose anode grid capacity has been eliminated and we can, therefore, reconstruct our tuned anode H.F. amplifier as shown in fig. 17.12.

The screen condenser must be of a " non-inductive " type, otherwise it may constitute a tuned rejector circuit at some particular frequency, and so function as a high resistance and defeat the action of the screen. It is common practice to supply the 60 volts on the screen by means of a potentiometer arrangement across the H.T. supply.

As a matter of design it is found that a tetrode stage gives the highest gain when (*a*) the valve has a high A.C. resistance and (*b*) the tuned circuit has a high dynamic resistance.

Notice that in the circuit of fig. 17.12 the R.F. amplifying tetrode is in parallel with its tuned circuit. If, in our search for high gain, the latter is given a very high dynamic resistance, the voltage developed across it will be correspondingly large, and even the minute anode-grid capacity of a tetrode (0·003 micro-microfarad) may provide a path through which this high voltage may cause feed-back and set up oscillations.

Another result of this parallel disposition is that the valve damps the tuned circuit. These two difficulties can be overcome by what is effectively an auto-transformer device ; the anode lead is tapped into the coil of the tuned circuit about one-third of the way from its " earthy " end (fig. 18.12). As a result of this tapping only that part of the total voltage developed, which appears across OB, is applied to the tetrode, and the tendency to feed back is lessened. Further, since the valve is no longer in parallel with the whole of the tuned circuit, its damping effect on the latter is much reduced.

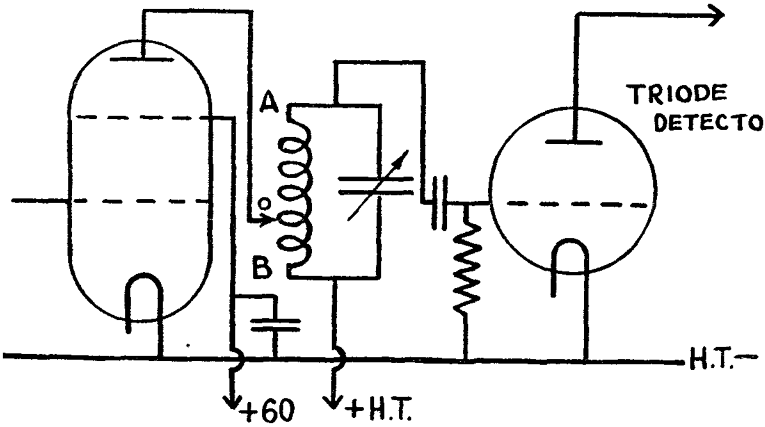


FIG. 18.12. Anode tap arrangement

It is a simple step to proceed from this auto-transformer coupling to transformer coupling where the coupling is a high-frequency transformer with a tuned secondary, as shown in fig. 19.12.

This method of coupling combines the advantages of the one previously mentioned, with one more, rather important from the point of view of safety of operation ; the condenser C is now isolated from the positive high-tension lead.

If it is desired to use more than one stage of R.F. amplification it is essential to employ either transformer or auto transformer coupling if both high gain and stability are to be achieved.

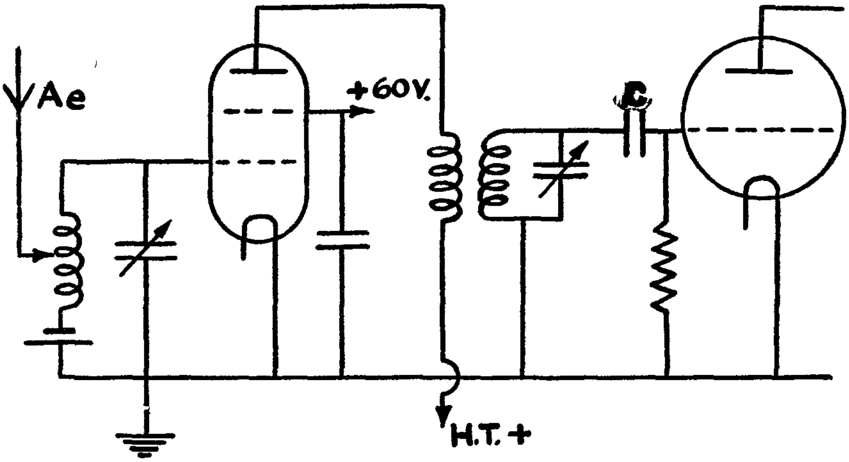


FIG. 19.12. Tuned H.F. transformer

Another type of R.F. coupling is the *tuned grid* circuit. As will be seen from an inspection of fig. 20.12 it is a modification of the simple tuned anode circuit, in which the tuning condenser of the load circuit is at earth potential.

The tetrode is coupled to its anode load circuit through an R.F. choke capacity circuit, in which the direct component of anode current completes its path through an R.F. choke of low resistance and the oscillating component is passed on to the next stage through a condenser.

In fig. 20.12 below we have a tuned grid amplifier followed by a grid detector. As before, the anode circuit of the amplifier can be tapped into the coil in the interests of stability and selectivity. In all types of R.F. amplifiers the detector grid circuit may be tapped down to reduce damping imposed by it on the tuned circuit.

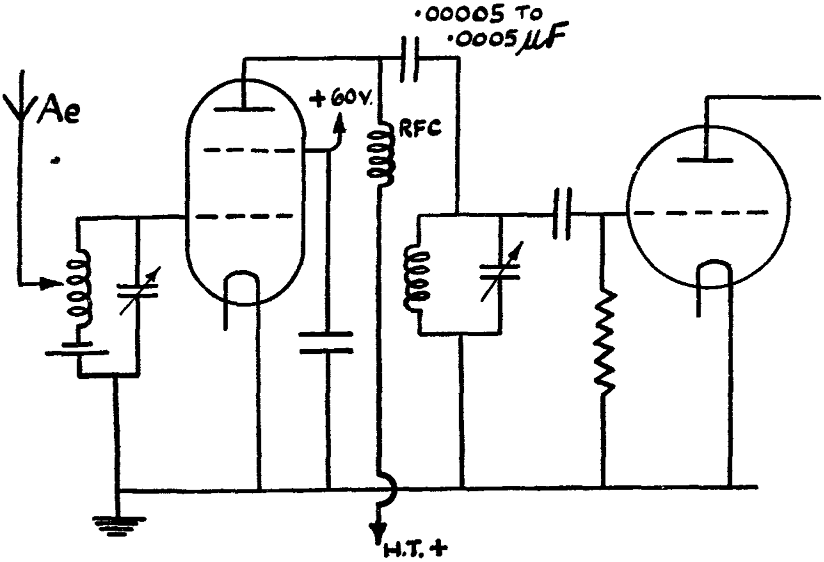


FIG. 20.12. Tuned grid coupling

Back-coupling and decoupling

In all the circuits which have been considered so far it has been assumed that the alternating components of the various anode currents flow through the power supply system and back to filament without opposition, i.e. that the power supply is without resistance. This is never the case, and in practice a valve will build up an alternating voltage of some value across the anode power supply. This voltage is also across the various valves and their anode loads, and as a result a certain proportion of it will appear in the different grid-filament circuits.

It will, in some stages, be in phase with the input signal and will cause instability, and in the other stages it will be in anti-phase with the input and loss of amplification will be the result. The instability usually takes the form of a series of " plops " (called " motor boating "), or an audio-frequency howl. It is essential to reduce to a minimum the chances of this back-coupling, or " feed-back ". The process by which the back-coupling is overcome is known as *decoupling.*

This is done by inserting, between the anode supply and the anode lead to each valve (except the last), a high-resistance R*d*, the junction of this resistance and the load being connected to earth via a high-capacity condenser C.

The behaviour of the system may best be understood by studying the circuit of fig. 21.12 (a), which is shown redrawn in fig. 21.12 (b), and its equivalent circuit is shown in fig. 21.12 (c).

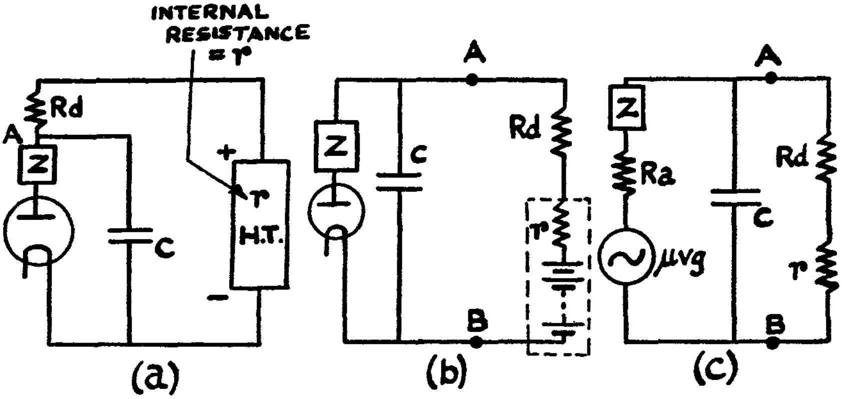


FIG. 21.12. Decoupling arrangements

The condenser C is of sufficiently large capacity to reduce to a small figure the total impedance of the circuit between A and B at the lowest frequency occurring in the apparatus. Therefore the current flowing round the circuit due to the imaginary generator μ V*g* develops only a small P.D. between A and B. This small P.D. is divided between the battery resistance *r* and the decoupling resistance R*d* in proportion to their resistances, and since R*d* is very much greater than *r* the voltage developed along *r* is negligibly small and exerts no appreciable influence on the grid-filament circuits of other valves in the equipment.

Decoupling condensers and resistances for R.F. amplifiers are smaller in value than those for A.F. amplifiers because of the higher frequencies occurring in the circuit.

It is essential that the decoupling condenser used at radio frequencies shall be non-inductive, otherwise, at one particular frequency, the condenser will behave as a high impedance rejector circuit instead of a low impedance by-pass, and decoupling at that frequency will not be effective.

It is not necessary to decouple the last valve of an amplifier, for the unwanted voltages are then, in any case, much smaller than the wanted ones, so that no harm results. Also, the last valve of a receiver is usually a power valve and requires the highest possible anode voltage if the maximum power is to be obtained.

The introduction of a decoupling resistance in such a stage is therefore avoided, but partial decoupling is usually achieved by shunting the anode supply system with a high capacity condenser.

Similar principles are applied to the grid circuits of an amplifier in which grid bias is derived from the common H.T. supply, or from the anode currents of the valves.

R.F. stoppers

It is essential in all receivers, to keep radio-frequency voltages from the audio-frequency stages. Their presence in an A.F. amplifier results in distortion and self oscillation. In spite of elaborate filtering in the detector anode circuit, R.F. voltages often appear in the grid-filament circuits of later valves. For this reason we have to devise a means whereby the full A.F. voltages and negligibly small R.F. voltages appear between the grid and filament of the first audio-amplifier. Fortunately, the valve itself provides half the solution.

Between the grid and filament, and inherent in the valve design is a capacity C_{gf} of a few micro-microfarads fig. 22.12.

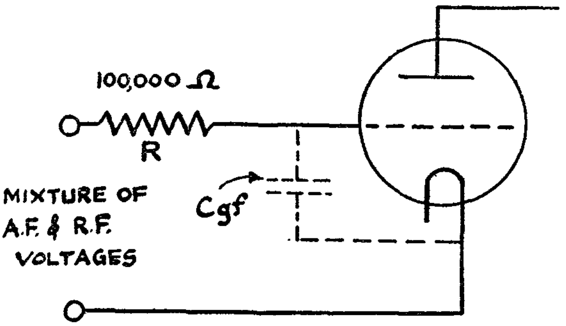


Fig. 22.12. R.F. stopper

Now, if a large resistance of about 0·1MΩ is connected in series with the grid, it forms, with the capacity C_{gf}, a potentiometer across the input. At high frequencies C_{gf} will have a low reactance compared with that of the series grid resistance R (the *stopper* resistance), so that the greater part of the unwanted R.F. voltage is dropped along R, and only a small portion will appear across C_{gf} (i.e. actually between grid and filament). This is just what is required.

The reverse is the case for audio-frequencies. The capacity C_{gf} has very high reactance at audio-frequencies so that a negligible amount of the audio-voltage is dropped along R, and practically the full audio-voltage across C_{gf} appears between grid and filament of the wave.

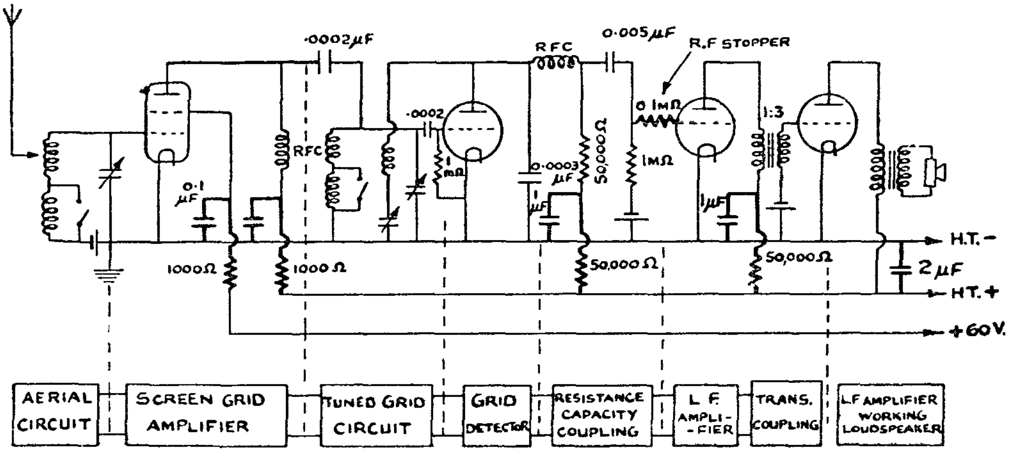


Fig. 23.12. Receiver showing decoupling circuits and R.F. stopper

Fig. 23.12 is a diagram of a receiver showing the decoupling arrangements and an R.F. stopper after the detector stage.

CHAPTER XIII

MASTER OSCILLATOR CONTROLLED TRANSMITTERS

In Chapter X we discussed a simple type of transmitter, which was, in effect, merely a straightforward development of the valve oscillator. Such a simple transmitter has an inherent drawback ; its frequency is unstable.

In theory, of course, we could design such a transmitter to give any desired output, but in practice we should find ourselves limited by lack of frequency stability. All valve oscillators consist of two main elements :—

(i) a frequency fixing circuit ;
(ii) an amplifier.

The first is usually a tuned circuit, containing inductance, capacity and resistance ; the second is usually an amplifying triode.

To secure stable frequency, the properties of (i) i.e. L, C and R, must be maintained constant ; we have already seen that with a simple transmitter this is not an easy matter ; the aerial is usually part of the oscillating circuit and its mechanical movements alter the properties of the latter. Temperature has a big effect on the coil and condenser spacings, as well as on the constants of the oscillating valve, and even more important than these effects is that due to " loading " the tuned circuit. The frequency of the circuit alters, that is, when it is called upon to emit radiation. If, for example, a simple transmitter is keyed, or modulated, rapid fluctuations of frequency (scintillation) take place which are extremely objectionable and may even give rise to serious distortion, especially on high-frequency transmissions.

To ensure a stable frequency in a valve oscillator it will be necessary to :—

(*a*) keep all the working parts at a constant temperature ; the most important of these are the tuning coil, condenser and valve.
(*b*) ensure that any part of the tuned circuit is not liable to mechanical motion or oscillation (e.g. if the aerial is part of the oscillating circuit it may sway in the wind).
(*c*) avoid damping losses in the tuned circuit, particularly those caused by valve resistance, grid current and aerial loading.

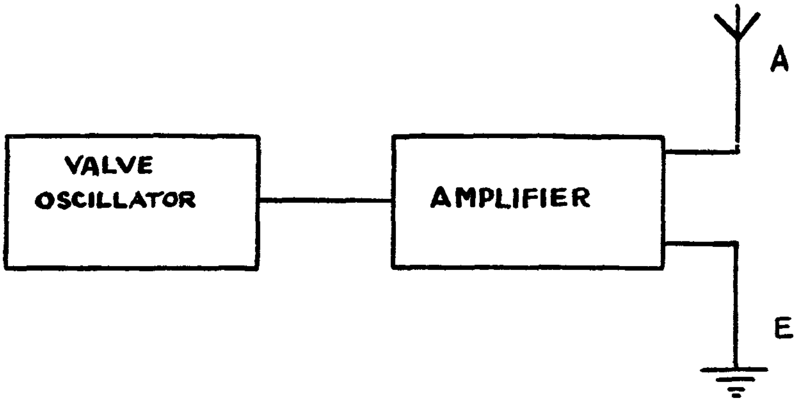


FIG. 1.13. Schematic master oscillator system

These conditions are secured by using a carefully screened *master oscillator* working into an amplifier which, in its turn, supplies power to the aerial. The general layout is indicated in fig. 1.13.

The amplifier in this system will provide a big output for the aerial, while at the same time imposing quite a small load on the oscillator, and so we shall achieve a much higher degree of frequency stability than if the oscillator itself supplied the load.

Design of amplifiers for transmitters

These amplifiers differ in several ways from those used in receivers. In the first place, much higher power is employed and great stress is laid on efficiency, which is the ratio of output power to high-tension input power. This is important because most of the energy loss appears as heat at the valve anode. The amplifiers of receivers are operated with sufficient negative grid bias to prevent the flow of grid current, but examination of any valve characteristic curve of I*a* against V*g* shows that more than half the characteristic is not used if we are subject to this restriction. In order to use the whole available curve, amplifiers for transmitters are operated with so large an input voltage that grid current flows for a large portion of each cycle. Again, R.F. amplifiers for receivers employ screen grid or pentode valves, but these are only used in low power transmitters at the present time, though there are a few high power S.G. valves in operation. Instead, triode valves are used. A simple circuit, suitable for a low-frequency transmitter is shown in fig. 2.13.

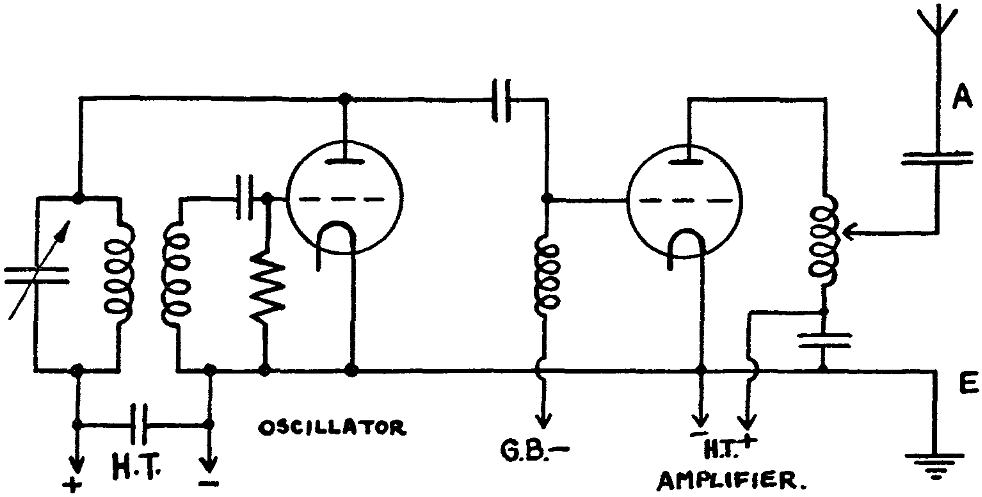


FIG. 2.13. Simple M.O. transmitter

A simple tuned anode oscillator is used to drive an amplifier valve in the anode circuit of which is a coil, coupled to the aerial, with the capacity of which it forms a tuned circuit. A proper choice of aerial tapping point brings the output circuit into resonance, while the anode tap matches the aerial load to the valve impedance so that high efficiency and large power output may be secured. The oscillator is coupled to the amplifier by a condenser of low reactance at the oscillator frequency. The amplifier grid bias can be produced by a variety of methods. In the present case a battery is used, while an R.F. choke prevents this battery forming a short circuit across the oscillator output. If the amplifier is driven so hard that grid current flows, a grid leak may be used.

It is possible to devise many other types of transmitter using the basic idea of a master oscillator followed by a power amplifier. Any oscillator circuit may be used, the most common being the Hartley, Colpitts, Meissner and tuned anode-tuned grid arrangements, and the selected oscillator may be coupled directly (via a condenser) or by mutual inductance to the succeeding amplifier. For a complete collection of methods of coupling an oscillator to an amplifier the reader should consult reference books or one of the handbooks prepared for the use of amateur radio operators. One interesting circuit is shown in fig. 3.13. Here a tuned anode-tuned grid oscillator is employed coupled by mutual inductance to its amplifier.

When the two tuned circuits of the first valve are properly adjusted it commences to oscillate, the necessary reaction or feedback being obtained by coupling output and input circuits through the interelectrode capacity C*ga*. The high frequency current in the tuned

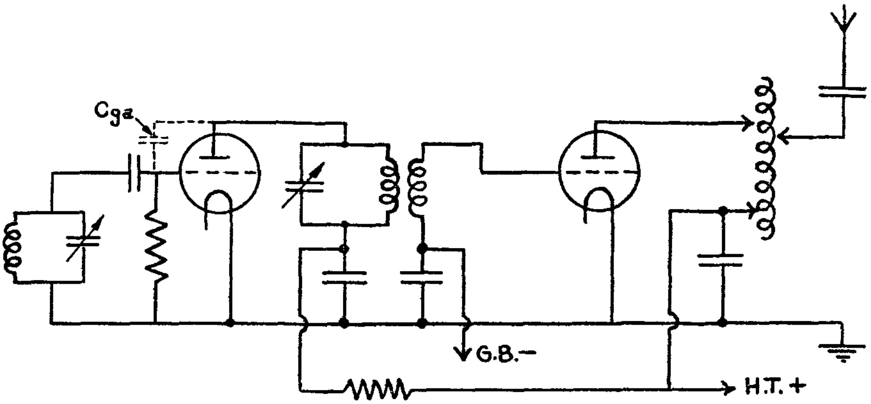


FIG. 3.13. M.O. system with T.A.T.G. oscillator

circuit in the oscillator anode supply induces a voltage in the grid coil of the amplifier, the output of which is delivered to the aerial-earth system.

Stability of amplifiers

Consideration of the action of the oscillator in the circuit just mentioned indicates a serious fault possessed by the simple triode amplifier.

Output and input circuits are coupled by the valve capacity *Cga*, and any valve in which the size of this capacity is sufficient will generate oscillations instead of amplifying with stability the power applied to it. The screen-grid valve was developed to avoid this effect in amplifiers used in receivers. The action of this valve has already been discussed, and circuits have been described in which it may be employed. Precisely the same theoretical circuits are used for transmitter or receiver amplifiers, except that in the latter case, voltages used are much lower than for transmitters.

There is, however, another method of dealing with the problem of instability or self-oscillation of power amplifiers. The general idea of this is to balance the feedback through valve-capacity by an equal amount in the opposite phase through another capacity, which may be varied until the two effects cancel, leaving no tendency for self-oscillation to occur. This method is generally known as *neutralising*.

Principles of neutralisation

Examination of the simple circuit (master oscillator and power amplifier) on p. 159 will show that the output circuit and the input circuit of the amplifier are coupled by means of the anode-grid capacity of the amplifier valve; in fact, the arrangement reduces to the one shown in fig. 4.13.

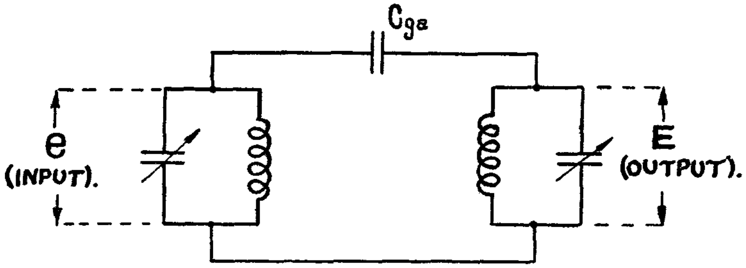


FIG 4.13. Showing coupling due to valve capacity

The amplified voltage E across the output drives a current round the circuit (overcoming the effect of the smaller input voltage) which sets up a P.D. across the input; this undergoes amplification and eventually self-oscillation ensues.

This circuit can be easily modified so that the feedback is in two parts, in opposite phase, the output voltage (so far as feedback is concerned) is split into two halves; one of these feeds back through C*ga* and the other through a small neutralising condenser, which can be adjusted until it is equal to the anode-grid capacity of the amplifier. Fig. 5.13 shows how this can be done.

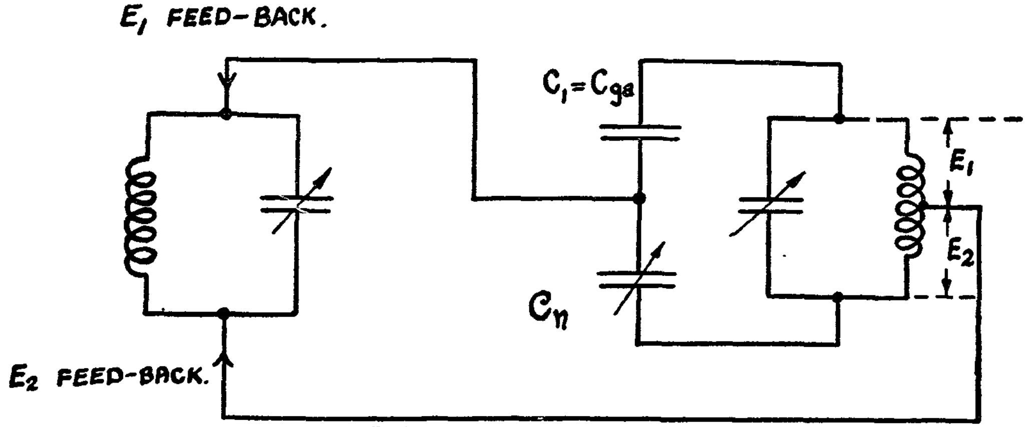


Fig. 5.13. Illustrating neutralisation

The voltage across the output coil is split by means of a centre tap into two parts E_1 and E_2; E_1 feeds back to the input circuit via the anode-grid capacity; E_2 feeds back via a small condenser C*n*. These reactions are in opposite phase and, therefore, if they are equal they neutralise each other. They are adjusted to equality by correct setting of the small condenser C*n*.

This method of neutralisation is due to Hazeltine and is shown applied on an actual amplifier in fig. 6.13.

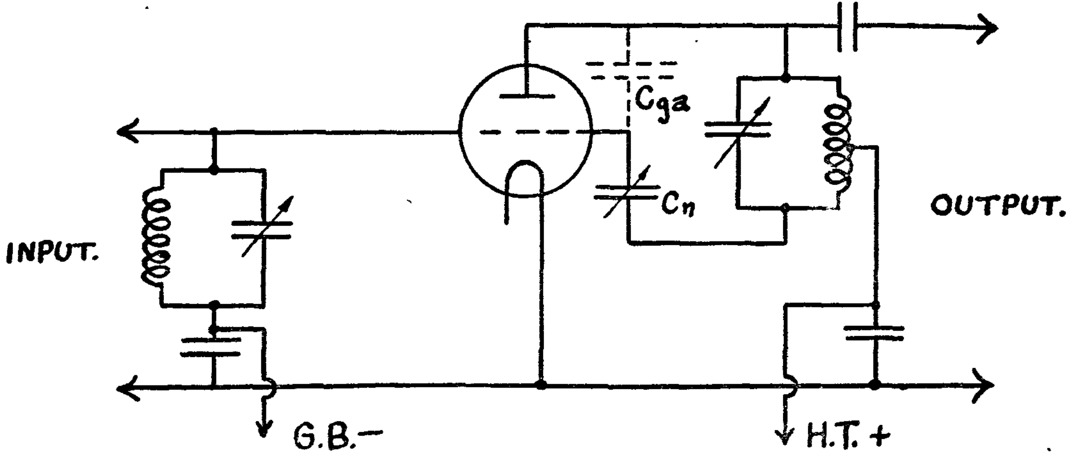


Fig. 6.13. Hazeltine neutralisation circuit

A serious disadvantage of the Hazeltine circuit is that centre-tapping the tuned circuit of the amplifier reduces its dynamic resistance to a fraction of its normal value, and so makes it difficult to get high efficiency operation with short-wave transmitters.

Another circuit, due to Rice, in which the input is centre-tapped is shown in fig. 7.13. Here only one half of the input voltage is applied to the grid-filament circuit amplifier, so that it means a considerable reduction in power output, due to the reduced drive.

Reverse reaction method of neutralisation

The Rice and Hazeltine methods when properly employed give a perfect balance which holds over a wide frequency range, and are used with transmitters of the highest quality and power. A small amount of unbalance or feedback is of value in small, light or portable equipment because the output is increased in just the same way as reaction increases the output of a receiver. The feedback must not be great enough to make the amplifier break into oscillation.

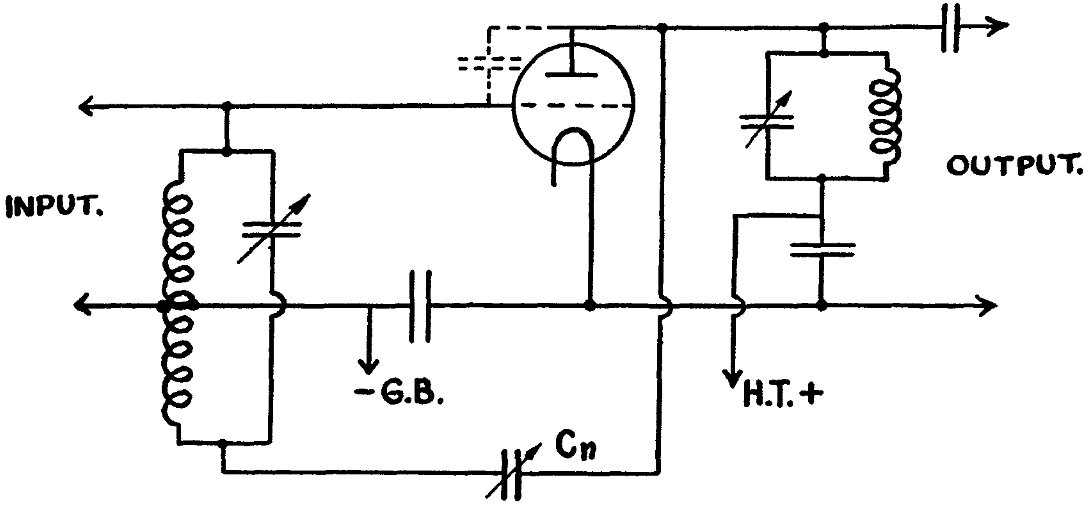


FIG. 7.13. Rice neutralisation circuit

To secure stability we may employ a reaction circuit in which the coil is arranged so as to give negative feedback, which damps any tendency to self-oscillation. Any of the arrangements previously discussed for receivers may be used. A typical example is shown in fig. 8.13.

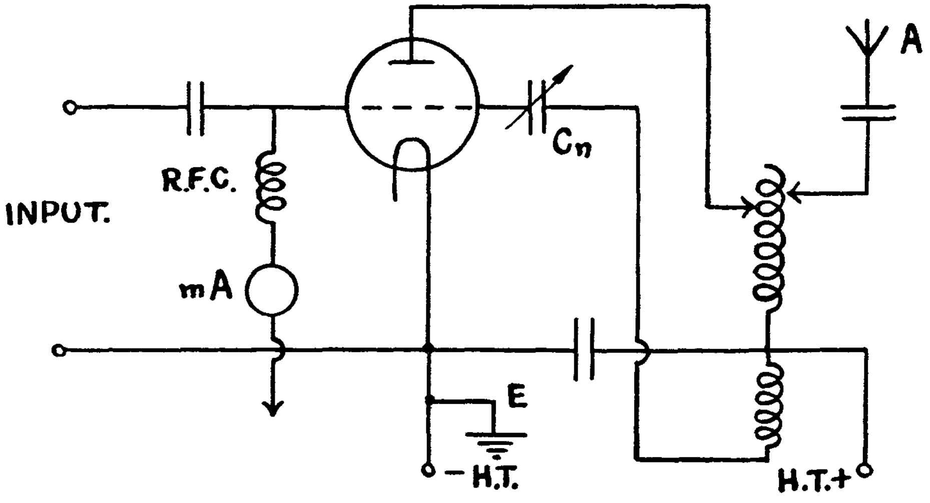


FIG. 8.13. Stabilisation by reverse reaction

Current in the main oscillatory circuit induces an e.m.f. in the neutralising winding which is so connected that feedback through the anode-grid capacity C_{ga} is opposed by an equal and opposite amount through the variable neutralising condenser.

Often a larger fixed condenser is joined in series with the latter to prevent damage if the variable condenser plates should touch.

Practical methods of adjusting for neutralisation

These all depend on applying an input to the amplifier with its high tension supply off, and then tuning its anode circuit to resonance. If there is any coupling between input and output, an e.m.f. will be induced in this tuned circuit. This may be great enough to operate a neon lamp as in the neon lamp wave-meter, or to light a flash lamp bulb in a loop of wire near the tuned anode circuit. Three practical methods of neutralising are described below.

(*a*) *Neon lamp method*

Turn the neutralising condenser to zero and switch on all power supplies except the amplifier H.T. Apply the drive or excitation from the oscillator to the amplifier input and tune its anode circuit to resonance, holding a neon lamp on the valve anode, or on that end of the tuned circuit joined to the anode. At a certain point the lamp will light. Now alter the neutralising condenser until the lamp goes out, showing that the original feedback has been cancelled. Retune the main circuit and then the neutralising condenser successively for a perfect balance. Some extra refinements are used in practice, but in any case this method is not very accurate due to stray capacity effects and to the peculiarities of the neon lamp.

(*b*) *Flash lamp method*

Here the procedure is exactly as above, but as an indicator use a small lamp bulb joined to a loop of a few turns of wire, coupled to the tuned circuit (by holding near the end of the tuning inductance). A thermo-ammeter of suitable range may be used in series with the tuned circuit instead of the lamp and wire loop.

(*c*) *Grid current method*

Connect a milliammeter in series with the grid-bias supply to the amplifier as in fig. 8.13. Set the neutralising condenser to zero. Tune the anode circuit through resonance. Here the grid current varies and gives a sudden kick on the meter. Now adjust the neutralising condenser until no kick is obtained on tuning through resonance. This method is much more sensitive and accurate than the others, but is not often used in the Service.

High efficiency amplifiers

We have seen how we can ensure frequency stability in a transmitter by carefully screening the oscillator and by ensuring that at no time is it called upon to deliver much

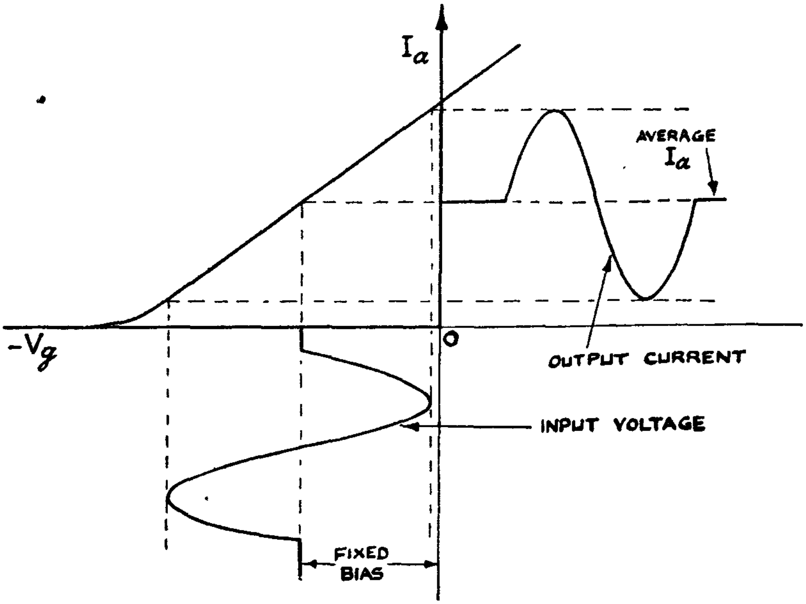


FIG. 9.13. Class A amplification

power. The next point which claims our attention is the efficient working of the amplifying valve which, operated by the output from the master oscillator, supplies energy to the aerial circuit. Some of the H.T. power input to this valve must inevitably be lost inside the valve—in heating the anode, among other things ; our problem is to see that such internal losses are reasonably small compared with the useful output in the form of aerial radiation.

We have already mentioned the three types of triode operation A, B and C.

In the class A amplifier anode current flows during the whole cycle of input voltage (fig. 9.13).

The output wave shape closely resembles that of the input, and would do so exactly if the characteristics were straight. This means that in class A operation the valve is a distortionless amplifier ; we saw its application as an audio-amplifier in the previous chapter.

Notice that when the signal voltage is zero there is still quite a large anode current ; even in the quiescent condition there is a considerable dissipation of H.T. energy, whose value is given by the product :—

Anode current × voltage between anode and filament.

Such an amplifier, then, is not very efficient.

In a class B amplifier the grid bias is adjusted so that the anode current is almost cut off when no signal is applied ; this alone means a considerable advance in efficiency from the class A system. Anode current flows only on the positive half-cycles of the input, but owing to the large negative bias the amplitude of the input has got to be much bigger than in class A if both types of amplifier are required to maintain an equal output.

Fig. 10.13 illustrates class B operation.

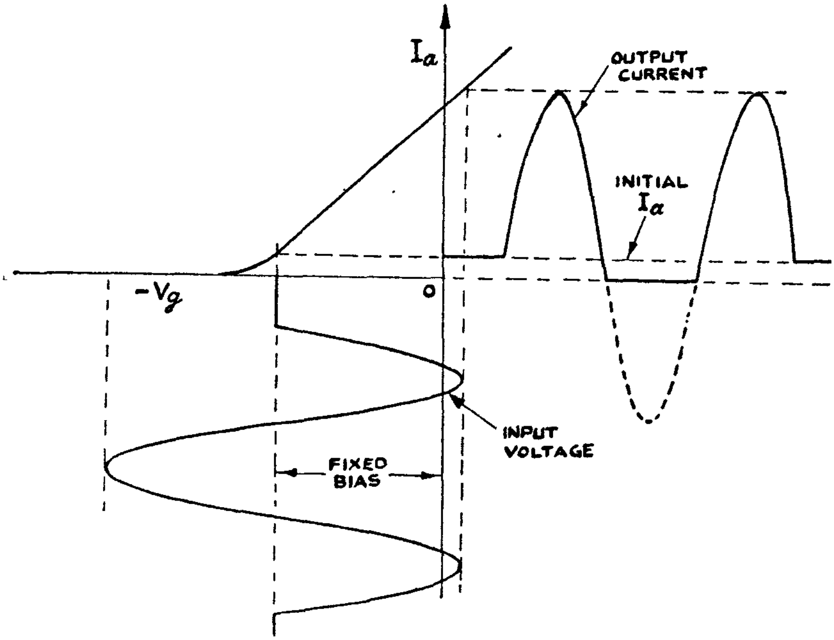


FIG. 10.13. Class B amplification

Class B output is not a facsimile of the input ; what we gain in efficiency we lose on the quality of the magnification, but we shall have more to say about this later. Note in passing, that we employed class B amplifiers in audio-frequency work, and managed to offset the distortion of one valve against that of another, in a push-pull arrangement.

We can go a step further ; the negative bias can be increased to double anode current cut off if necessary, and anode current will then flow for something less than half of each input cycle.

This is class C operation, indicated in fig. 11.13. It represents still higher efficiency than class B, but, like it, introduces serious distortion and cannot be used for audio-frequency work, even in a push-pull arrangement. Remember, however, that the problem which faces us is not that of audio-frequency amplification, where we have to amplify over a whole range of frequencies without discrimination.

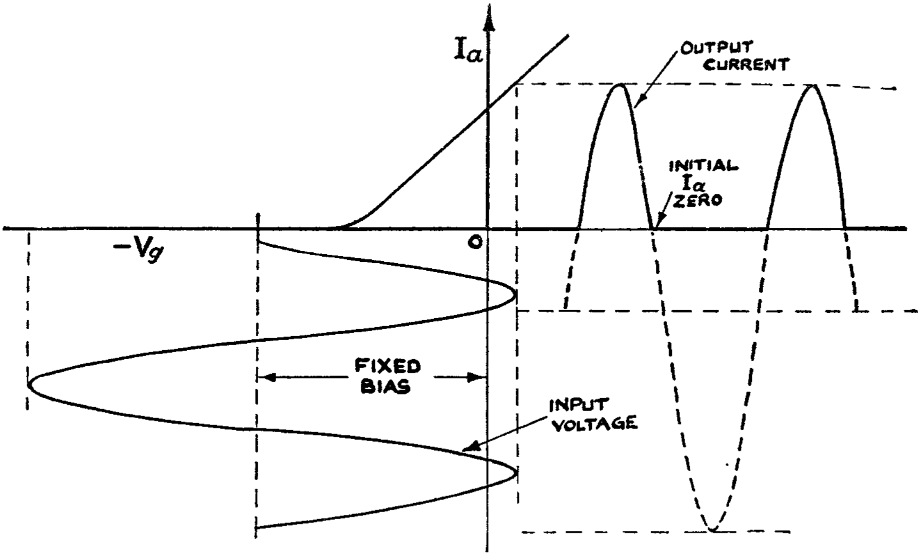


Fig. 11.13. Class C amplification

The amplifier of a master oscillator controlled transmitter has only one frequency to deal with, the R.F. wave generated by the master oscillator, so we can put a filter circuit in the output which will *only respond at the desired frequency,* and voltages of any other frequencies which have been introduced by the distorting amplifier will have no effect. So distortion ceases to worry us and we can take full advantage of the efficiency of Class C operation.

This point is worth a little more discussion ; we have already mentioned Fourier's theorem, in which it is shown that any periodic wave-form can be resolved, or analysed into a whole series of sine curves of different amplitude and phase, whose frequencies are in the ratio 1 (fundamental) ; 2 (2nd harmonic) ; 3 ; 4 ; 5 ; and so on.

It is quite a difficult piece of mathematics, but that need not prevent our making use of its results in any case that crops up. Take the output of a class B amplifier as an example ; the Fourier analysis shows that the half-sine pulses of anode current are equivalent to a combination consisting of :—

(i) A direct current.
(ii) An alternating current of input frequency.
(iii) An A.C. of smaller amplitude than (ii) of twice input frequency.
(iv) An A.C. of still smaller amplitude, of four times input frequency.

Now, if this output is made to pass into a circuit tuned to the fundamental (No. ii above) only the fundamental will be present in the circuit and the distortion represented by (i), (iii) and (iv) will be quite negligibly small. Just as when a receiver is tuned to a wanted station, the voltage induced in the aerial by that station is selected from dozens of others, perhaps equally strong, by the tuned circuit, so in an R.F. class B or C amplifier, the fundamental is selected from the distorted output by the tuned circuit in the anode.

Reverting then to the efficiency side of the question ; an amplifier with a tuned circuit as anode load will behave as if this circuit were just a high resistance, at the resonant frequency. When the current through valve and load is a maximum the volts drop across the load is maximum and the voltage between anode and filament is therefore a minimum.

In class C operation, the anode current will be concentrated in a series of maxima corresponding with the peaks of the input voltage ; during these bursts of anode current, the anode-to-filament voltage will be small and therefore the energy lost inside the valve itself will be small ; thus by limiting the anode current to pulses delivered at the right instant, which is what class C operation does for us, we can raise the efficiency of our amplifier to a figure in the neighbourhood of 70 per cent., and at the same time by using a tuned output circuit, escape the consequences of distortion.

To summarise ; to obtain high efficiency we must :—

(*a*) use a large negative bias to cut off anode current for most of the cycle, i.e. use class C operation ;

(*b*) have a high input voltage, whose peak value must be large enough to cause a pulse of anode current.

This has the following consequences :—

(i) the amplifier " sensitivity " is low—because it wants such a big input to operate it ;

(ii) power amplification is low because grid current flows which absorbs energy from the input circuit. The usual value of the ratio :—$\frac{\text{output R.F. power}}{\text{input R.F. power}}$ is about 8 to 10.

The disadvantages of the high efficiency system are :—

(i) rather large driving power is necessary ;

(ii) excessive instantaneous currents (sometimes ten times as large as the D.C. feed) occur, and damage the filament ;

(iii) in spite of the filtering properties of the anode load, excessive harmonic radiation occurs, especially on high frequencies, or if the aerial is tightly coupled to the final tuned circuit.

Frequency multiplication

So far we have considered a transmitter consisting of a master oscillator followed by one stage of R.F. amplification ; more stages can be added if necessary and even in simple apparatus two stages are generally employed. The intermediate or buffer amplifier serves to isolate the oscillator more completely from the output circuit.

When the aerial is required to radiate considerable power, there may have to be a whole series of amplifier stages, and if the operating frequency is high there will be a considerable difficulty owing to feed back, even when screen grid valves, or neutralising circuits are used.

Now a class B or C amplifier generates harmonics and we can pick out by means of a tuned circuit in the anode, any one of them which we desire to use ; thus with an input frequency (f) we can develop $2f$ or $3f$ in the anode circuit, by the simple process of tuning. We can start then with a *low frequency* (f) low power oscillator, amplify its output in class B or C, and from the anode of the amplifier select say $2f$; we can repeat this process and obtain $4f$ and so on till we reach our desired frequency and *there will be no feedback* because the outputs are not at the same frequency as the inputs, and the system is very stable.

Using this principle it has been found possible to start with an audio-frequency oscillator controlled by a 1000 c/s tuning fork, amplify the output, and increase the frequency to 100 Kw. and 1 Mc/s respectively.

The next problem which confronts us is that of impressing signals on the radiation from our transmitter; how shall we (*a*) key the transmitter for telegraphic signals or (*b*) modulate its output for radio telephony ?

1. Keying a transmitter-amplifier

When fairly low powers are involved, keying the H.T. circuit is a satisfactory method; a modification which gives good results is shown in fig. 12.13. Here the grid circuit and the H.T. negative connection to the amplifier filament are opened; a large negative bias is thrown on to the grid at the moment the H.T. circuit is broken, and a sharp cessation of radiation occurs. This is probably the best all round method of keying. It should be noted that a practical amplifier would require neutralisation. This is not shown on the diagram.

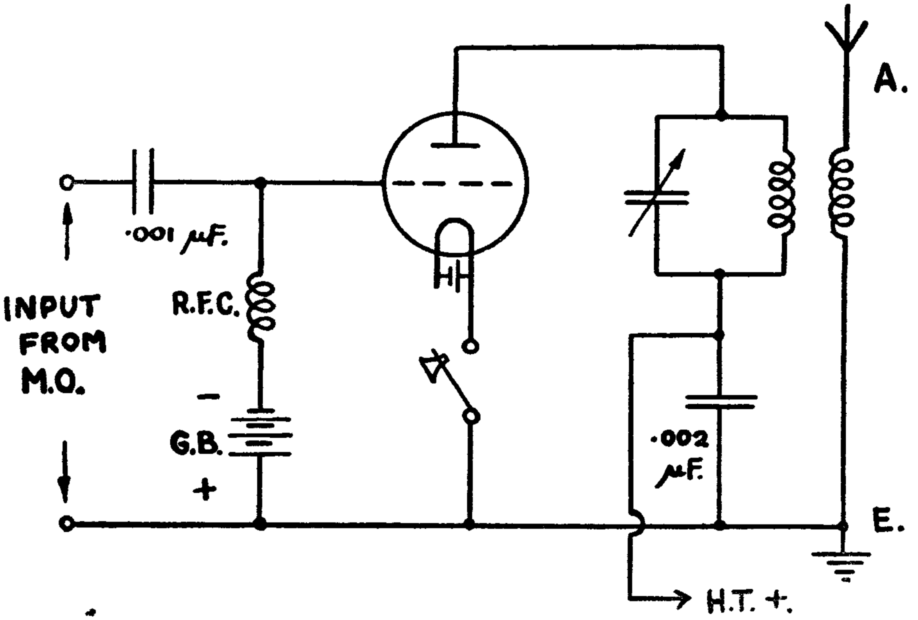


FIG. 12.13. Keying a transmitter-amplifier

2. Keying a frequency doubler stage

Transmitters employing the frequency doubling principle described above are usually arranged according to this scheme :—M.O.—Amplifier—Doubler—Amplifier—Doubler—Amplifier. In such cases it is possible to key one of the doubler stages, just as we did in the single amplifier stage above.

There is a serious drawback to this method of H.T. keying, especially with high-power sets. When the drive is removed from a class C amplifier the anode current sinks to zero; the sudden makes and breaks of the key therefore causes rapid current changes in the high tension supply and the induced voltages which are thereby set up may cause considerable damage, particularly in any D.C. machines, which may be employed.

This disadvantage may be overcome by using a device called an " absorber valve "; this valve is so arranged that when the key is down it draws only a small current and directly the key is up its current rises; in this way it keeps the H.T. load fairly constant. Absorber keying is satisfactory, but the extra equipment employed makes it rather costly.

3. Keying transformer primary

When a transmitter is operated from an A.C. supply it is possible to remove the H.T.

by opening the primary circuit of the main rectifier transformer. This needs a large key and only permits slow signalling. It can, however, be combined with methods 1 and 2 above by using a synchronised key which (*a*) operates in the amplifier H.T. (as in 1 or 2) and (*b*) also opens the primary of the main transformer and thus obviates large induced voltages.

When the load on the high-tension supply is removed by keying the valves, the H.T. voltage rises. The possibility of damage to components or insulation can be prevented by keying the primary of the transformer, and as this is on no-load at the time, only a moderately large key is required.

Induced effects can be stopped by using an artificial load across the transformer primary in small sets.

Modulation of output

As in the case of the simple transmitter there are two main ways of modulating the output from an M.O. controlled R.F. amplifier. These are :—(*a*) grid modulation, in which the modulation voltages are introduced into the grid circuit of the amplifier ; (*b*) anode modulation, in which the modulation voltages are introduced into the anode circuit of the amplifier.

It is important to observe that it is not possible to achieve modulation when an amplifier is operating under class A conditions. This has already been pointed out in Chapter XI, p. 131 ; in order to produce the sum and difference frequencies in its anode circuit (which is one way of regarding modulation) a valve must be operating under class B or class C conditions. It is found in practice that a class C operated amplifier lends itself best to modulation ; class B operation gives rise to distortion.

The circuits shown in fig. 13.13 illustrate the general principles of grid and anode modulation of an R.F. amplifier. Fig. 13.13 (a) shows a grid modulated amplifier, the A.F. voltages being introduced in series with the grid bias, which therefore varies at modulation frequency.

In fig. 13.13 (b) the modulation voltages are introduced in the anode circuit ; this causes the anode voltage on the amplifier to fluctuate at the modulation frequency. The audio signal introduced in the anode circuit is an alternating voltage which periodically adds to or subtracts from the fixed H.T. voltage. When the total anode voltage is high the aerial current is also large ; in fact, when the correct adjustments are made the aerial current is proportional to the anode voltage. The variations in anode potential thus result in the production of a modulated aerial current.

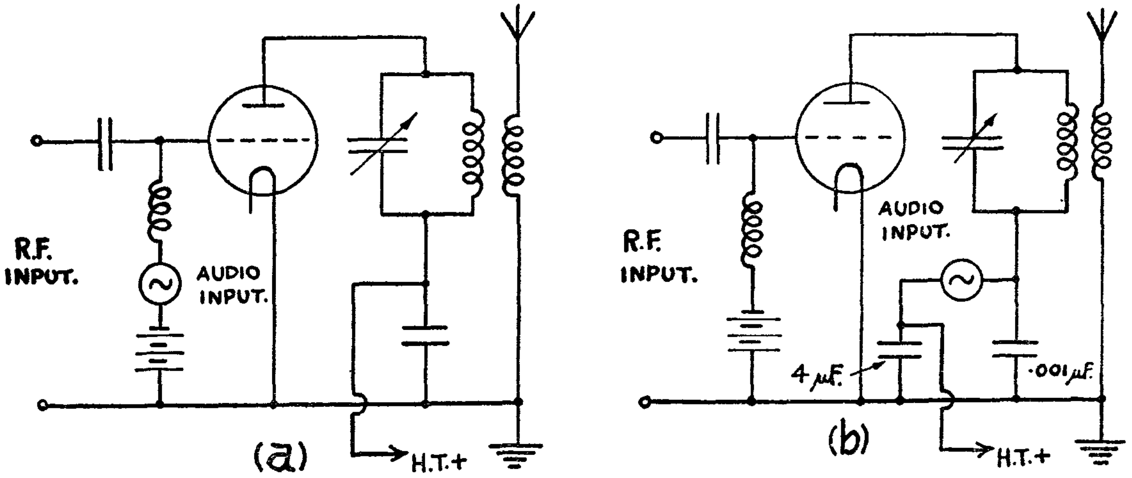


FIG. 13.13. Modulation of a transmitter-amplifier

In both cases the alternating current in the output circuit will be varying in amplitude at modulation frequency, or looked at from the side-band point of view, the anode circuits will contain currents at the carrier frequency f and at two other frequencies $f + m$ and $f - m$, where m is the modulation frequency.

Practical modulation circuits

The main principles of grid and anode modulation have been explained, and it now remains to describe the actual circuits used to couple the modulator valve, or low-frequency generator to the radio-frequency amplifier. The methods used are exactly the same as those used to couple an ordinary output valve to a loudspeaker. These are three in number :—

- (i) direct coupling, the loudspeaker being connected in the anode circuit of the valve.
- (ii) choke capacity coupling.
- (iii) transformer coupling.

A minor point of difference is that, unlike the average loudspeaker, the R.F. valve, which is really the audio load of the modulator, needs a D.C. supply as well as the audio power output of the modulator valve. Typical circuits illustrating the various methods are given below. In all cases it is desirable to keep radio-frequency current out of the audio stages, which is done by the liberal use of R.F. chokes and small by-pass condensers (of the order of ·001 μF.). These have no appreciable effect on the audio output of the modulator.

Transformer coupled grid modulated amplifier

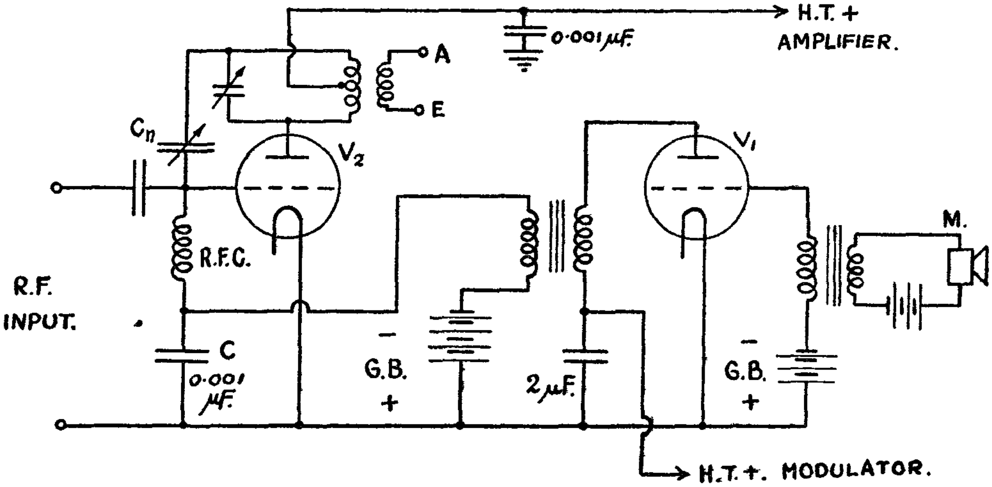


FIG. 14.13. Transformer coupled grid modulated amplifier

As shown in fig. 14.13, a microphone M is coupled to a transformer and, on speaking into it, an audio-voltage appears between grid and filament of the modulator valve V_1. The amplified output is taken from the secondary of a transformer in the modulator anode circuit, and varies the grid bias of a neutralised amplifier valve V_2. The output of this valve is then the desired modulated wave. The choke R.F.C. and by-pass condenser C keep R.F. current out of the modulator transformer. Anode by-pass condensers are also shown for the two valves, ·001μF. for the amplifier and 2μF. for the modulator. The aerial would be joined to such a set by plugging into a socket in the anode coil (through a blocking condenser) to insulate the aerial from the D.C. H.T. supply, or by inductive coupling through a coil situated near the anode coil ; this latter method is preferable, and is shown on the diagram.

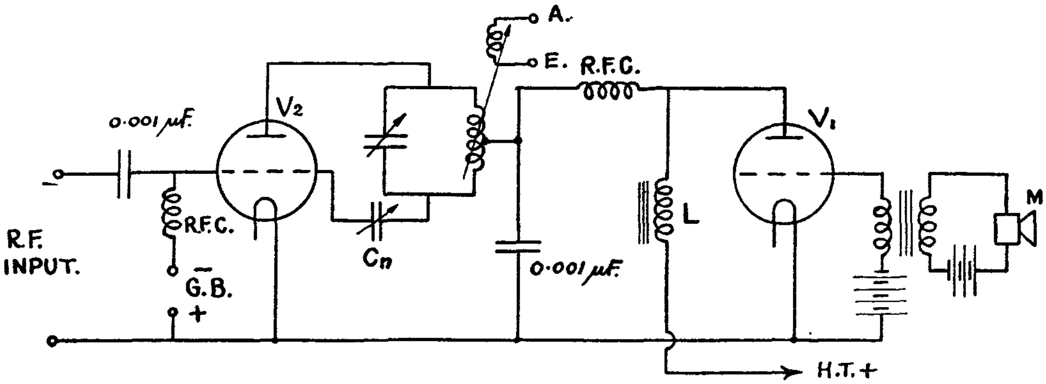


FIG. 15.13. Choke coupled anode modulated amplifier

In the case illustrated in fig. 15.13 variations of the modulator grid voltage cause variations in its anode current which passes through the modulation choke L (about 20 henries inductance). The current variations cause potential changes across the anode choke, and therefore the anode voltage of V_1 also changes.

But V_1 is coupled to the amplifier V_2 through the choke R.F.C. and so the output of V_2 varies according to the instantaneous anode voltage, and a modulated wave is produced. With the arrangement shown, full modulation is impossible, because the anode voltage of the valves cannot be reduced to zero. To secure full modulation, a resistance is joined in the anode lead of V_2 only. This drops the anode voltage of the amplifier, so that less audio output is required for full modulation. The resistance is shunted by a large condenser (1 to 10μF.) so that the audio-voltages generated by the modulator can still be applied to the amplifier.

Transformer coupled anode modulation

The difficulty of securing 100 per cent. modulation with the choke coupled modulator can be avoided by the use of a transformer as shown in fig. 16.13 which may be used to step up the audio-voltage to any desired level, or which may be used to supply the amplifier and modulator valves with independent H.T. voltages so that full modulation can be secured.

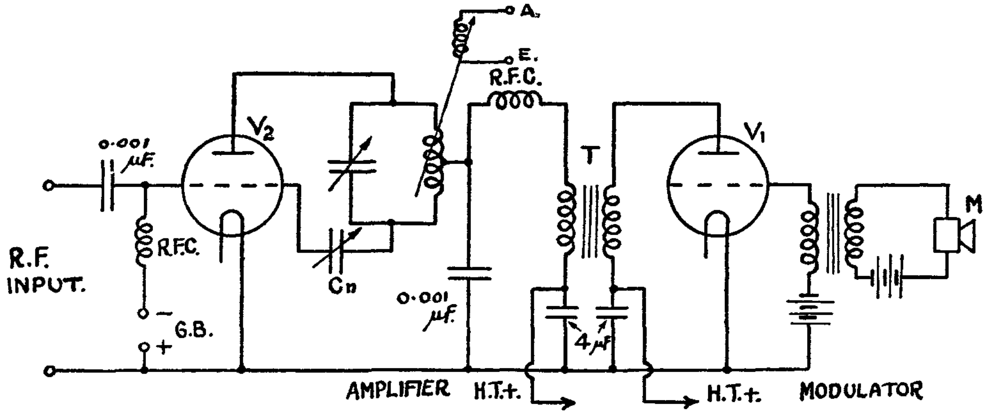


FIG. 16.13. Transformer coupled anode modulated amplifier

In this case the microphone output causes audio-frequency variations of the current in the primary winding of the modulation transformer T. These induce a voltage in the secondary which is in series with the H.T. supply to the amplifier whose output varies with the supply voltage. This again results in a modulated wave. Separate H.T. supplies are shown, but a single supply can be used by connecting the two H.T. leads together.

An important advantage of the transformer coupling is that it enables the modulator load to be adjusted correctly.

Suppose a particular modulator requires an anode load of 2,500 ohms, and that the amplifier takes 100mA. at 1,000 volts, representing a load of 10,000 ohms ; the transformer ratio required is given by $\sqrt{\frac{10,000}{2,500}} = \sqrt{4} = 2$

That is, a 2 : 1 step up is required between modulator and amplifier.

Quartz controlled transmitters

The frequency-fixing element of all the transmitters discussed in this chapter has been a tuned circuit associated with a valve oscillator. Unless carefully designed, the inductance and capacity of the circuit are liable to vary because of temperature changes, humidity or mechanical vibration. Even when sufficient precautions are taken the apparatus becomes costly, delicate or unwieldy so that it is only suitable for fixed ground station transmitters operated by skilled personnel. The properties of quartz crystals give an alternative method of securing frequency stability, even under adverse conditions such as are found with small portable equipment, while, under good conditions with temperature control, a frequency stability of one part in a million can be maintained over long periods.

Some of the properties of quartz crystals have already been mentioned and it will be sufficient here to quote some operating particulars and to show practical circuits in which they may be used.

(*a*) *Range of frequency of quartz oscillators*

This is normally 25 to 4,000 kc/s, but recently crystals have been produced up to 7 Mc/s and even to 14 Mc/s but these latter are very fragile, give low output and are not of very high frequency stability.

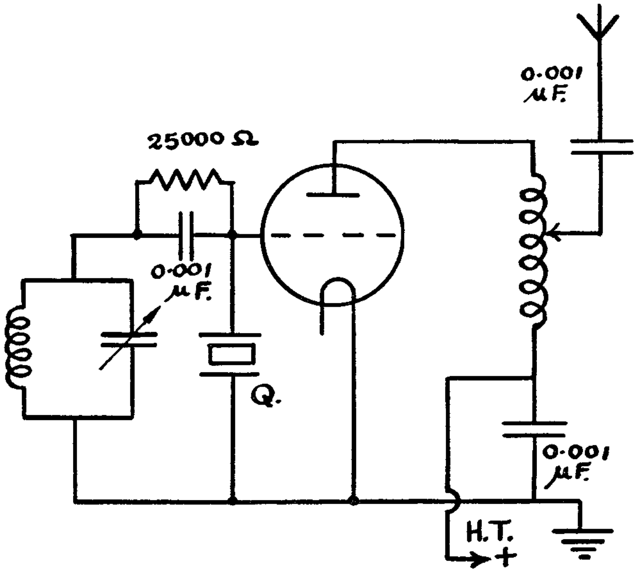


FIG. 17.13. T.A.T.G. oscillator with quartz control

(*b*) A good crystal will fully drive a valve having an anode dissipation of 10 to 50 watts, depending on the frequency of operation. The output is greatest at high frequencies, where the limit is set by crystal heating. On low frequencies the limit is set by the risk of crystal fracture due to the large amplitude of vibration.

(*c*) *Crystal mounting*

The object is to restrain the vibrations as little as possible, and the crystal is usually mounted between flat metal plates (sometimes in a vacuum), a small air gap being left. Adjustment of this gap gives a slight control over the frequency.

(*d*) *Crystal oscillator circuits*

There are two main methods of employing quartz crystals in oscillators. The first is to fit up any oscillator using reaction to obtain oscillation, and then to place a crystal across the frequency fixing circuit. The frequency of the generated oscillation then becomes very stable. This method is rather crude, and is only used in cases where apparatus must be small, light in weight or simple in construction and operation, such as in aircraft transmitters. The circuit of fig. 17.13 shows a tuned plate-tuned grid oscillator modified by the addition of a quartz crystal.

Circuits such as this can only be used with low-power valves on account of the risk of damaging the crystal. The circuits shown in fig. 18.13 (a) and fig. 18.13 (b) are extensively used in modern practice, the first being more usual.

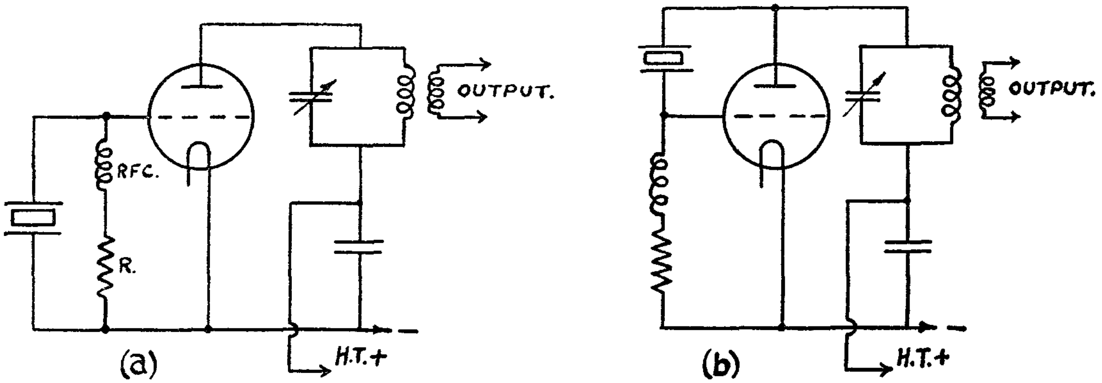


FIG. 18.13. Quartz controlled oscillator circuits

Screen grid or pentode valves may be used in either of the above circuits, but sometimes a small capacity is needed between grid and anode or between grid and filament, because it is feedback through this capacity which maintains the crystal in continuous oscillation.

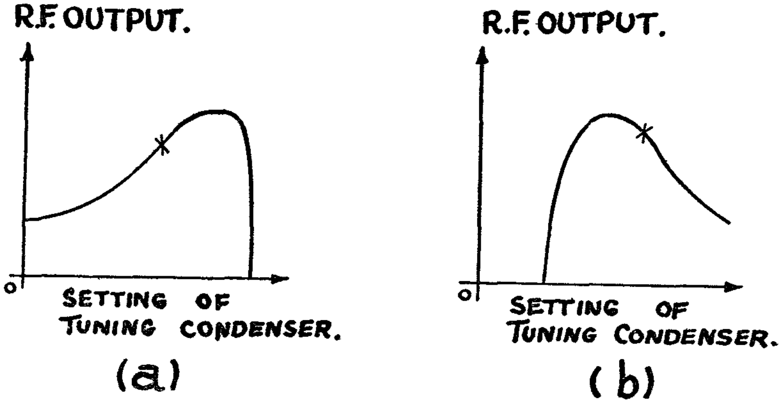


FIG. 19.13. Output from quartz controlled oscillator

If the tuning condenser is varied the R.F. output varies according to the graphs shown in fig. 19.13 (a) and fig. 19.13 (b), which correspond to the circuits shown in fig. 18.13 (a) and fig. 18.13 (b) respectively.

The output drops abruptly at a particular setting of the condenser. The best operating points are marked on the curves.

Circuit diagram of a typical transmitter

To illustrate the principles discussed in this chapter, a circuit diagram is shown of a crystal controlled transmitter in which the final class C amplifier is anode-modulated (fig. 20.13). Starting with a crystal oscillator stage, we have next a frequency doubler (neutralising not strictly necessary but sometimes used). This feeds the final neutralised amplifier, though in practice an intermediate stage would probably be used. The audio side consists of a modulator supplied from the microphone via an audio-amplifier (sometimes called a sub-modulator).

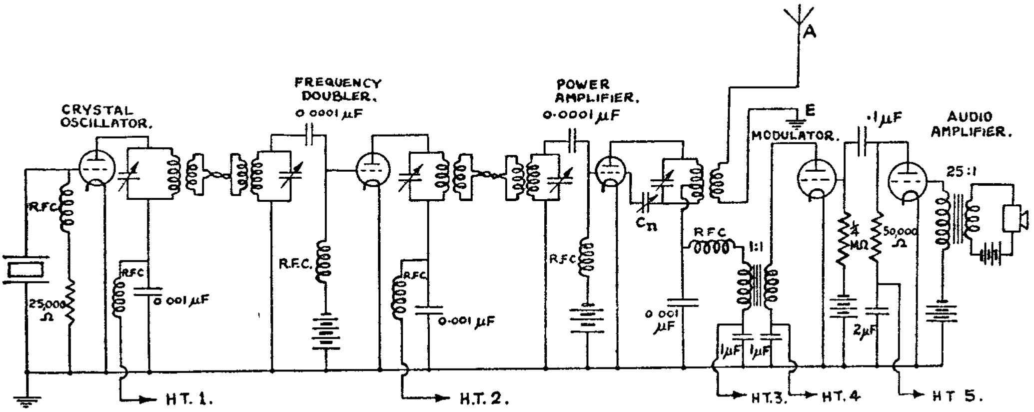


FIG. 20.13. Typical circuit diagram of high-power radio-telephony transmitter (crystal controlled).

Many details are omitted, but the figure gives a general idea of the layout required for the radiation of good quality programmes of high-frequency stability.

A note should be made of the method of coupling the various stages together. Each anode circuit has a small coil coupled to it in which a voltage is induced. This coil is coupled by a twin cable or transmission line to the next stage (which may be some distance away). The anode inductance and coupling coil behave like a step-down R.F. transformer. At the input to the next stage the voltage is stepped up again by another tuned R.F. transformer. The power is thus passed on from stage to stage in the form of a large R.F. current at low voltage. Dielectric and other losses are thus kept small and the influence of stray capacity much reduced. The system is known as *link coupling* and is much used in modern practice.

CHAPTER XIV

THE SUPERHETERODYNE RECEIVER

The essential requirements of a communications receiver are few in number, but difficult to satisfy. High sensitivity is necessary so that weak signals may be received, whilst, to prevent interference from undesired stations, a receiver of great selectivity is needed, and it must be extremely stable in operation, with no tendency to go into uncontrolled oscillation. Fidelity and power output are questions which concern chiefly the user of a broadcast set operated for entertainment purposes.

The usual receiver consists of a high-frequency amplifier, a detector and a low-frequency amplifier, and in practice is not easily designed to give extreme selectivity or sensitivity. The selectivity depends to a large extent on the number of tuned circuits employed, and on their individual design. To secure single dial control all the tuning condensers are ganged together, and this raises difficult mechanical problems if more than four sections are used. The ease of separating two signals depends not on their actual difference in frequency but on the percentage difference. Thus, it is very much easier to separate stations transmitting on 99 and 100 kc/s respectively than if their frequencies were 999 and 1,000 kc/s.

In the first case the difference of 1 kc/s is 1 per cent. of the higher frequency. In the latter case the same difference of 1 kc/s is only one-tenth of 1 per cent. of the higher frequency signal. The selectivity problem is therefore much more acute at high than at low radio frequencies.

The same remark applies to the sensitivity and stability. It is easy to design good tuned circuits for use at low radio frequencies, because resistance and dielectric losses are small.

Again, stray capacity couplings are not effective at low frequencies, because their reactance is large and thus the feed back is kept below the level at which it might cause self-oscillation of the amplifier valves. At high frequencies of the order of 20 Mc/s the internal screening of a screen grid valve may not be good enough to secure stable operation. The difficulties of high-frequency amplification may be appreciated when it is noted that a single stage of radio-frequency amplification can easily give a gain of 200 times at 200 kc/s (1,500 metres). At 20 Mc/s (15 metres) the gain might be only five times and the selectivity proportionately bad.

The superheterodyne principle

The problem of securing high gain and selectivity has been solved by an extended application of the principle of heterodyne reception described at the end of Chapter XI. When two alternating voltages are together applied to a detector valve, then the output contains a whole range of alternating voltages of different frequencies. If a triode valve is used as a grid or anode circuit rectifier, it is immaterial whether the two voltages are applied to the grid circuit or to the anode circuit, or even if one voltage is applied to each, the output will still contain the frequencies just mentioned. The practice of applying one voltage to the grid circuit is clearly allowable, for elementary valve theory shows that a given voltage in the grid circuit exercises the same control over the anode current as a voltage μ times as large in the anode circuit. We may repeat these remarks in an amplified form as follows :—If the grid or anode circuit of a valve operating under class B conditions is excited simultaneously by voltages of different frequency, f_1 and f_2, then, in the anode circuit of that valve there will be in addition to f_1 and f_2 their harmonics

$2f_1$, $2f_2$, $3f_1$, $3f_2$, etc., and, more important still, sum and difference frequencies $f_1 - f_2$ and $f_1 + f_2$. By making the anode load of the valve a tuned circuit any desired component of this collection can be picked out, just as in the case of the class C operated power amplifier valves described in the last chapter. In C.W. heterodyne reception we arrange that the difference frequency $f_1 - f_2$ shall be within the audible range, 400—1,000 c/s, and the simple insertion of a pair of telephones with an H.F. by-pass condenser in the anode circuit of the valve is sufficient to render this frequency audible and to by-pass all the others to earth. But clearly it is not necessary to have $f_1 - f_2$ within the audible range. Suppose f_1 is a radio frequency, say 1,000 kc/s, then if f_2 is 900 kc/s, our difference frequency will be 100 kc/s which, whilst it is low from the radio-frequency standpoint, is well outside the upper limit of audibility. Such a frequency is called a *supersonic frequency**, and it can easily be selected from the other frequencies in the anode circuit of the valve producing it by the usual tuned circuit arrangements.

If now we had a C.W. transmitter radiating at a frequency f_1, we could receive the signal with an ordinary aerial system, mix it with the output of a local valve oscillator of frequency f_2 and then apply the two signals into a class B operated valve whose anode circuit is tuned to $f_1 - f_2$: this device would therefore in effect, *change the frequency of the received signal* from radio frequency f_1 in the aerial to a supersonic frequency (S.F.) $f_1 - f_2$ in the anode circuit of the class B valve, and this is the essence of the superheterodyne principle. The immediate advantage of a frequency change (R.F. to S.F.) is that it becomes a comparatively simple matter to amplify S.F. signals and to secure high selectivity in the circuits which have to deal with such signals.

The stray capacities which render R.F. amplification so difficult cease to cause trouble at S.F. because at this lower frequency their reactances become comparatively high. For a similar reason feedback troubles leading to instability in R.F. amplifiers are absent in S.F. amplifiers.

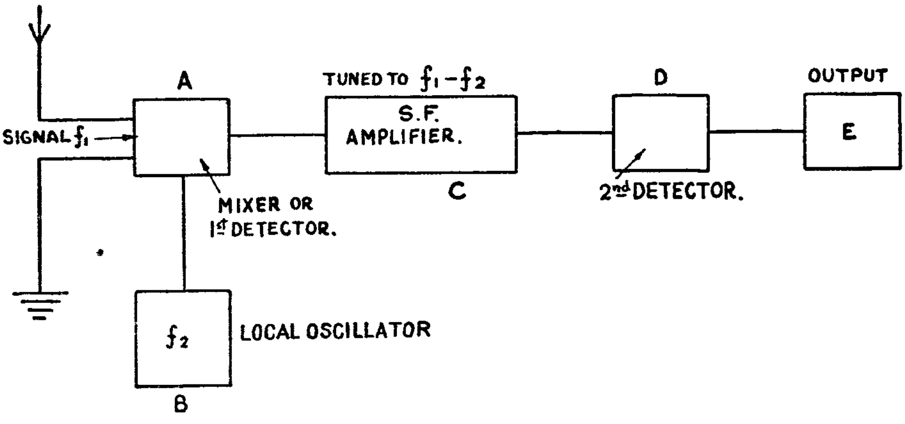


FIG. 1.14. Block schematic diagram of superheterodyne receiver

Reference to the schematic diagram above (fig. 1.14) will help to make these ideas more clear.

" A " is an anode circuit rectifier, although under special conditions a grid detector may be employed and this valve is known as the " first detector ".

" B " is a local oscillator generating oscillations of frequency f_2. " A " and " B " constitute what is known as a frequency changer.

In the grid circuit of " A " the signal oscillation (f_1) and the local oscillation (f_2) give rise to an oscillation of frequency $f_1 - f_2$ in the anode circuit.

---

* Note: The term *intermediate frequency* is often used in place of *supersonic frequency*.

This difference frequency is normally of the order of 100 kc/s or 450 kc/s, although in certain specialised receivers it may be as low as 50 kc/s or as high as 1,600 kc/s.

" C " is the S.F. amplifier : it may consist of several class A operated valves whose couplings are tuned once and for all to the supersonic frequency. Leaving " C " we shall have quite a strong signal of S.F. but it is still radio-frequency in the sense that it is much too high to operate telephones.

It must be " detected " just as an ordinary R.F. signal is detected, to make it deliver its message in a pair of phones. This is effected by the stage " D ".

" D " is the second detector, whose design follows ordinary detector practice.

" E " is a normal output stage such as described in Chapter XII. A signal is " found " in such a set by adjusting the frequency of the local oscillator till, in combination with the wanted signal, a difference frequency equal to that to which the S.F. stages respond is produced. It should be realised that these S.F. stages are very selective and in practice will only respond to signals of their designed frequency. A circuit designed on the block diagram of fig. 1.14 is given in fig. 2.14.

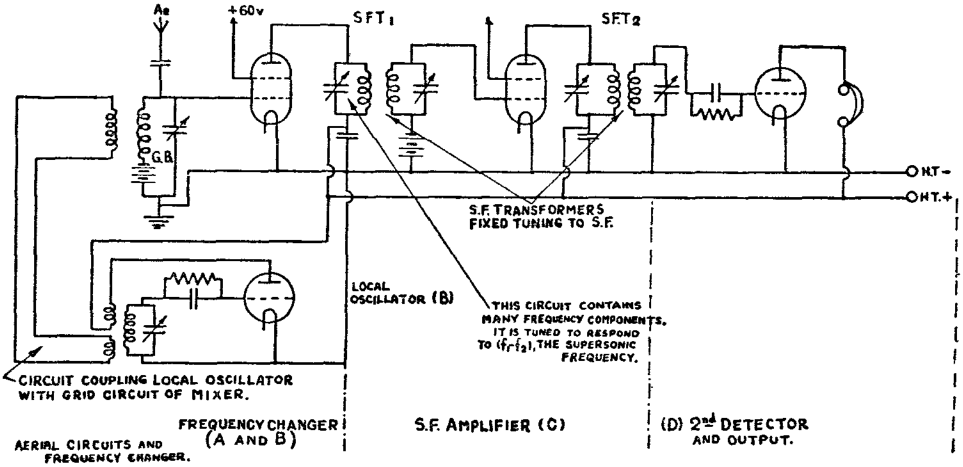


FIG. 2.14. Theoretical diagram of superheterodyne receiver

Examination of the circuit shows a tuned grid oscillator which injects a voltage into the grid circuit of an anode circuit rectifier by means of suitable coupling coils. This voltage is combined with a signal from the desired transmitter. The action of rectification produces a current in the anode circuit of the required difference frequency. A supersonic frequency transformer S.F.T.$_1$ is tuned to the difference frequency and the output voltage across the secondary of this transformer is applied between grid and filament of a screen grid amplifier valve, in the anode circuit of which is a second supersonic frequency transformer S.F.T.$_2$. This is followed by a grid detector, in the anode circuit of which are placed the telephones.

Objections to the circuit shown

The circuit of fig. 2.14, although quite practicable, has two very serious disadvantages. The oscillator, which is really a miniature C.W. transmitter, is directly coupled to the aerial. It therefore radiates and causes serious interference in neighbouring receivers. A different method of coupling the local oscillator and first detector or frequency changer must be devised. If this is not possible a high frequency amplifier between aerial and frequency changer will prevent radiation provided it includes a screen grid valve.

A second serious disadvantage of the circuit given arises as follows :—

Consider a local oscillator generating a frequency of 1,000 kc/s, and suppose the S.F. is 100 kc/s.

Let a signal of 900 kc/s be received ; then the beat frequency (S.F.) is 1,000 kc/s—900 kc/s = 100 kc/s and would be amplified by the S.F. stages.

Now suppose a signal of 1,100 kc/s is also received. The oscillator frequency is still 1,000 kc/s and the difference frequency is 1,100 — 1,000 = 100 kc/s as before ; this signal would be amplified just as well as the first. These signals of 1,100 and 900 kc/s are equally capable of mixing with an oscillator of 1,000 kc/s to give the correct S.F. of 100 kc/s. It will be noted that the two signals 1,100 and 900 kc/s differ by 200 kc/s, i.e. by twice the S.F. ; this effect is called " second channel interference ". It can best be removed by the use of selective input circuits, which when tuned to say 900 kc/s entirely exclude the interfering signals on 1,100 kc/s.

The use of an R.F. amplifier in front of the frequency changer will prevent the effect, because such an amplifier will include two or three tuned circuits capable of rejecting the interfering signal.

Oscillator interaction

When two oscillators supply a rectifier in such a way as to make a heterodyne whistle we notice that as the frequency of one is made to approach that of the other, the beat note falls in pitch and suddenly disappears.

The two oscillators have now pulled into step. The effect is worst on the highest frequencies corresponding to short waves. This is because any given difference of frequency is a smaller percentage of the signal frequency than at low frequencies.

This effect becomes important in the superheterodyne arrangement, where the frequency of the local oscillator tends to pull into step with that of the incoming signal. Multiple valves have been designed to reduce the trouble and will be discussed later. In some cases a radio-frequency amplifier is used between the oscillator and frequency changer in an attempt to isolate them from one another and so prevent this locking effect.

Superheterodyne receiver for C.W. reception

The receiver so far described will be useless for C.W. reception. The first detector or " frequency changer " would in effect change the incoming C.W. signal from R.F. to S.F. and allow us to achieve high selectivity and stable amplification in the S.F. stages, but at the output end of the S.F. stages we shall still have C.W. signals, which will not deliver up their message in any of the ordinary detectors. The heterodyne method must be employed and the set must be provided with another oscillator (second heterodyne oscillator) whose output mixed with the C.W. output from the S.F. amplifier will, by the usual heterodyne action, give an audible note of the difference frequency in the anode circuit of the second detector.

A receiver will now be described in which due consideration has been given to all the points just discussed.

The purposes of the various sections or units indicated in fig. 3.14 may be summarised as :—

" A ". R.F. amplifier

Here the incoming signal is amplified as much as possible at the original radio frequency. The amplifier contains selective circuits which keep out interfering signals spaced from the desired signal by twice the supersonic frequency. The amplifier, in other words, prevents second channel interference by its high selectivity. At the same time any radiation from the oscillator is prevented from reaching the aerial, since the amplifier will contain screen grid valves in which there is practically no coupling between output and input.

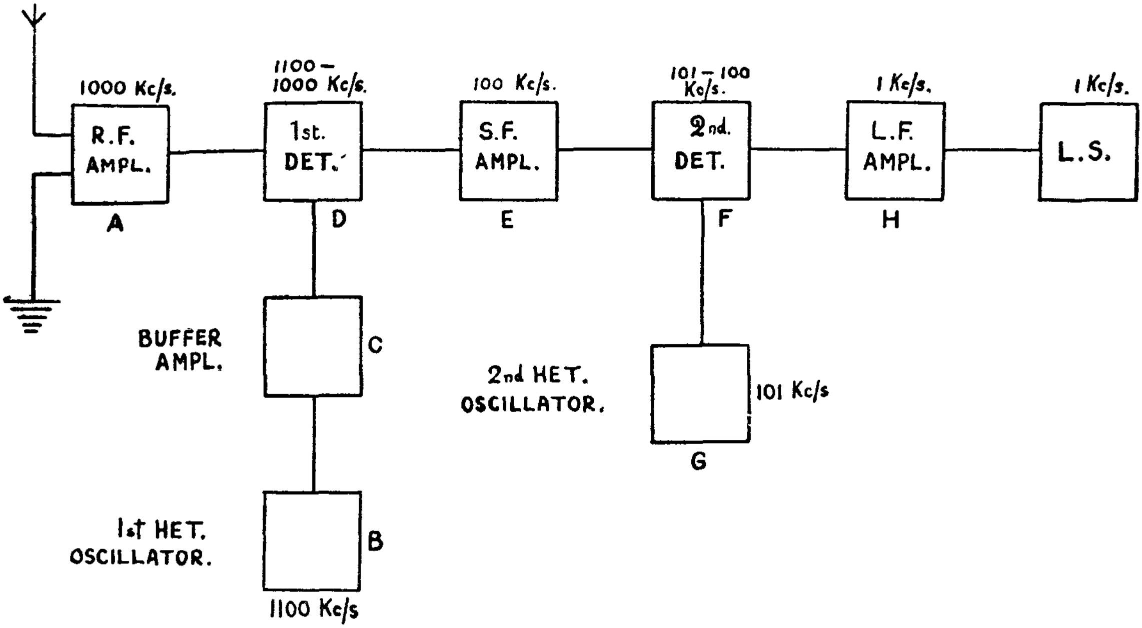


FIG. 3.14. Block schematic diagram of superheterodyne receiver for C.W. reception

Any amplification which can be secured before the frequency changer lessens the amount needed from later stages in the receiver, and it is found that the ratio of signal to noise is much better. Much of the hiss and noise heard in a badly designed superheterodyne receiver is due to the oscillator, so that the less amplification used after it, the quieter is the receiver.

"*B*". *First heterodyne oscillator*

In this stage, which we have previously referred to as the "local oscillator", we generate a radio frequency voltage which is ultimately combined or mixed with the original signal, giving rise to the supersonic frequency output. This oscillator must be capable of tuning over a range of frequencies so that any desired signal may be combined with it to produce the S.F. oscillation.

"*C*". *Buffer amplifier*

The function of this is to isolate the first heterodyne oscillator output from the original signal, so that interaction, pulling, or locking is prevented. It can also be made to give any desired output level if fitted with some form of volume control, as by the use of a variable-mu valve.

"*D*". *First detector*

The difference frequency or S.F. is produced by mixing the signal with the output from the first heterodyne oscillator in the grid circuit of a suitably operated valve. The anode circuit of this valve is tuned to respond to the difference frequency provided.

"*B*", "*C*" and "*D*" together constitute the frequency changer.

"*E*". *Supersonic frequency amplifier*

A straightforward radio frequency amplifier is employed to raise the signal level to a suitable value for application to the second detector. Screen grid or R.F. pentode valves are employed, and usually have variable-mu characteristics. It is usual, though not essential, to use transformer coupling between stages, the degree of closeness of coupling being adjusted to give a compromise between selectivity and quality of reproduction when the set is used for broadcast reception. In Service sets tuned anode coupling is used. It has the advantage of giving greater stage gain, but at the expense of selectivity which is better when a high-frequency transformer is employed. Often a choice of coupling is provided so that high selectivity may be used on distant stations, subject to interference, while a band pass characteristic gives good quality reproduction from local stations, where high selectivity is not needed.

" F ". Second detector

If I.C.W. or R.T. is being received, any form of detector may be employed, either diode or triode using grid or anode rectification. All these methods give satisfactory results.

" G ". Second heterodyne oscillator

If C.W. is being received, then the S.F. amplifier is delivering C.W. at supersonic frequency to the second detector.

We have already seen that an ordinary detector (diode, anode or grid circuit) will not detect C.W. and we have to make use of the heterodyne principle again. The output of the second heterodyne oscillator is of such a frequency that, mixed with the S.F. amplifier output in the second detector valve, it provides an audio difference-frequency component in the valve anode circuit.

The output stages consist of audio-frequency amplifying circuits operating usually a loudspeaker. Their design and function are precisely the same as described in Chapter XII.

Reception of telephony using the superheterodyne principle

If the original signal is modulated, it may be regarded as a carrier and two sidebands covering a range of frequencies differing from the carrier frequency by a very small percentage.

Thus a 1,000 kc/s carrier modulated with a 500 cycle/sec audio note can be regarded as :—

(i) The unmodulated carrier of 1,000 kc/s.
(ii) An upper sideband oscillation of 1,000 kc/s + 500 c/s.
(iii) A lower sideband oscillation of 1,000 kc/s — 500 c/s.

All these are mixed with the first heterodyne oscillation, say of 1,100 kc/s. Three intermediate frequency outputs are produced :—

(*a*) Due to carrier 1,100 — 1,000 = 100 kc/s.
(*b*) Due to upper sideband oscillation 1,100 — (1,000 kc/s + 500 c/s) = 100 kc/s — 500 c/s.
(*c*) Due to lower sideband oscillation 1,100 kc/s — (100 kc/s — 500 c/s) = 100 kc/s + 500 c/s.

Thus we still have a carrier and two sidebands, but the new carrier is 100 kc/s instead of the original 1,100 kc/s. The S.F. amplifier is not selective enough in the ordinary way to distinguish between 100 kc/s, 100 kc/s + 500 c/s, and 100 kc/s — 500 c/s since the percentage difference of frequency is very small (only $\frac{1}{2}$ of 1 per cent.)

Voltages of all these frequencies are amplified in the S.F. stages, and reach the second detector. There they are rectified just as an ordinary R.F. modulated wave is rectified by the detector in a straight set and an audio-output corresponding to the original modulation is obtained. It must be noted that to receive R.T. and I.C.W. no second heterodyne oscillator is required and must not be used; in general purpose receivers a switch is provided to put it out of action except when C.W. is being received. A circuit diagram of a practical receiver will be given later when the design of the frequency changer has been dealt with a little more fully.

Miscellaneous features of the superheterodyne receiver

Volume control

It is convenient to control the amplitude of a received signal both before and after rectification at the second detector. This is because grid and anode circuit detectors operate best when a particular signal input is employed.

Larger signals overload them, while small signals involve operating on curved valve characteristics, giving rise to distortion on telephony. Post-detector control is also useful, for the required audio-output depends on whether telephones, loudspeakers or recorders are in use.

In modern receivers, pre-detector volume control is secured by the use of variable mu valves, the stage gain of which is altered by varying the grid bias. This variation may be obtained manually, using bias potentiometers, or automatically by rectifying the received carrier, and feeding back the resulting steady voltage as additional grid bias. Post-detector volume controls usually take the form of a potentiometer across the second detector output. This may supply audio-frequency amplifiers giving any desired final power level. Further reference to volume control will be made at the end of this chapter.

Aids to high selectivity

In spite of the very high selectivity provided by a normal superheterodyne receiver, cases of interference sometimes occur which need special treatment, especially when C.W. signals are being received.

Suppose two C.W. transmitters differ in frequency by 500 c/s, and that one of them is caused to give a final audio-output at 1,000 c/s. The other will give a note of either 500 or 1,500 c/s, depending on the particular heterodyne frequency used.

The desired tone of 1,000 c/s will be heard with a strong interfering background on 500 or 1,500 c/s. This may be removed by the use of a tuned rejector circuit, resonating at 1,000 c/s, which may be wired in the anode circuit of an audio-amplifier as in the diagram fig. 4.14. It will present a low reactance to currents above or below the resonant frequency, but a high resistance at the resonant frequency of 1,000 c/s. By suitably adjusting the second heterodyne oscillator, any desired C.W. signal may be made to give rise to the 1,000 c/s beat note, and others will be excluded because their final heterodyne note differs from the required 1,000 c/s, and in consequence they are by-passed by the rejector circuit or " note filter ".

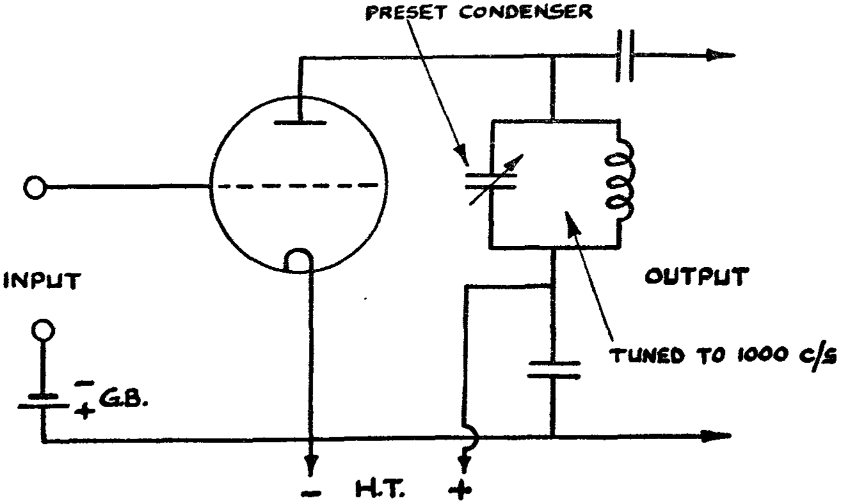


FIG. 4.14. Audio note-filter

Extremely high selectivity is frequently obtained by the use of filter circuits or intervalve couplings in which quartz crystals are employed. Full information on these can be obtained from current periodicals dealing with receiver construction and design.

Frequency changers for superheterodyne receivers

The design of a satisfactory frequency changer, is a most complicated problem and the difficulties increase rapidly with the frequency of the signals it is desired to receive.

In the first place, an oscillator is required, the performance of which is satisfactory in the following respects :—

(i) *Frequency stability.*—This must be high because any drift will cause the S.F. to differ from the required value, with consequent low gain and strong possibility of interference. Stabilisation of the oscillator frequency can be secured, but only at the expense of low output, limited tuning range and complicated construction. Some modern receivers incorporate automatic arrangements for stabilising the oscillator frequency.

(ii) *Constancy of output.*—The output of an oscillator whose frequency may be varied is not constant over the whole range unless special precautions are taken, e.g. automatic volume control. A buffer amplifier is of use in this connection. The output must not merely be constant, but also of the correct amplitude for application to the first detector.

(iii) *Isolation from the signal frequency circuits.*—This is necessary to avoid the effect known as pulling or locking, whereby the oscillator becomes synchronised with the signal. This is a serious trouble at high frequencies but can be avoided by the use of a buffer amplifier or by the use of special multiple valves.

(iv) *Freedom from harmonics.*—These are apt to cause interference in the form of whistles.

Next, the oscillator output must be combined with the signal by the first detector or mixer, which may be a simple rectifier, or a multiple valve of a type to be discussed later. A good frequency changer should possess the following features :—

(i) no radiation.
(ii) no locking or interaction between signal and oscillator circuits.
(iii) should not be critical in respect of applied voltage, from either signal or oscillator.
(iv) should operate well at all frequencies.
(v) should not generate harmonics of the applied voltages, because these are a fruitful cause of whistles and interference.
(vi) should give a high S.F. output.

Modern frequency changers are divisible into two main classes which may be described as (*a*) detector-type mixers and (*b*) electron-coupled mixers. In the first case the output of the oscillator is coupled to the detector portion of the frequency changer by the use of ordinary circuit components such as coils, condensers and resistances. The valves used may be either separate, or combined in the same envelope.

In electron-coupled mixers the heterodyne action is secured by applying the oscillator voltage to a grid placed somewhere in the cathode-anode electron stream of the first detector. Here, again, separate or multiple valves may be used.

As examples of the first class of mixer we may quote the anode or grid-circuit rectifier in which the oscillator voltage is injected either in the grid circuit along with the desired signal, or in the anode circuit.

In the second class we have as a simple example the high-frequency pentode, in which the oscillator voltage is applied to the suppressor grid, where it exercises a control over the electron stream in a manner to be described later.

The detector type frequency changer

The diagram (fig. 5.14) shows the mixer stage of a certain receiver, some of the details having been simplified. The action is as follows :—

The output from a local oscillator is amplified by a variable-mu valve whose anode-voltage variations are applied, together with the signal-frequency output of the R.F. amplifier, to the grid-filament circuit of the first detector. The output of this detector contains the desired difference frequency voltage which is selected by means of a tuned circuit and passed on to the S.F. amplifier. The first detector is shown as a grid rectifier, but this practice is unusual. The anode rectifier is more commonly chosen for this position because the grid condenser and resistance tend to smooth out the desired supersonic-frequency voltages unless both leak and condenser are made very small, which reduces sensitivity. A simplified diagram of a complete superheterodyne receiver using this type of frequency changer is given in fig. 6.14.

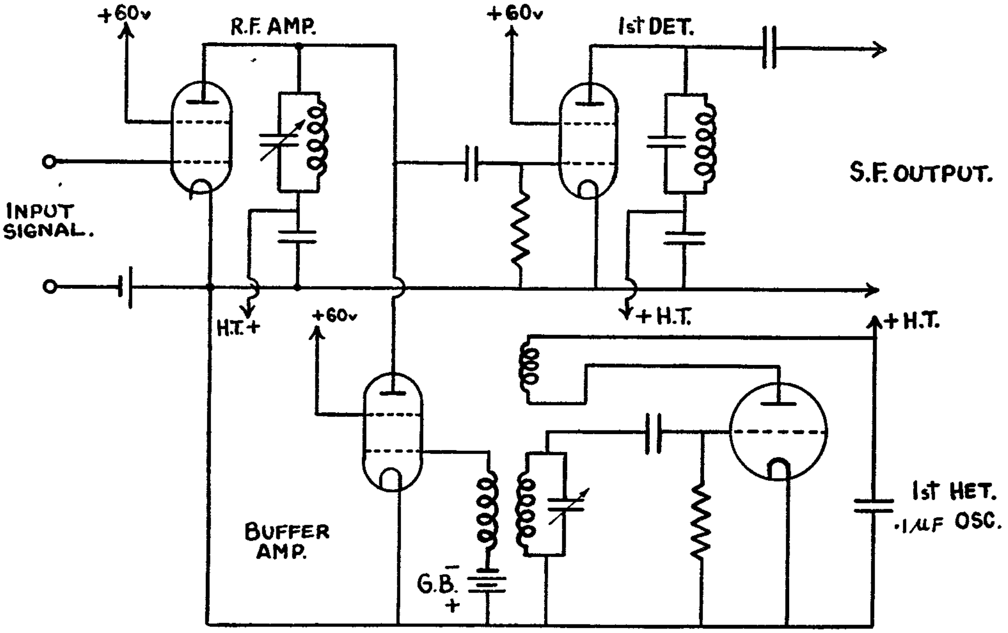


Fig. 5.14. Detector-type frequency changer

Up to the present the terms first detector and mixer have been applied to that stage in the receiver in which the difference frequency is produced by rectification of the signal oscillation and local oscillation. When superheterodynes were first produced this result was always secured by rectification, and the term " first detector " was quite appropriate. Modern multiple-valve frequency changers do not depend on rectification for the mixing process. Present-day practice is to use the term frequency-changer or mixer in such cases and to reserve the name " detector " for the stage in which signal rectification occurs.

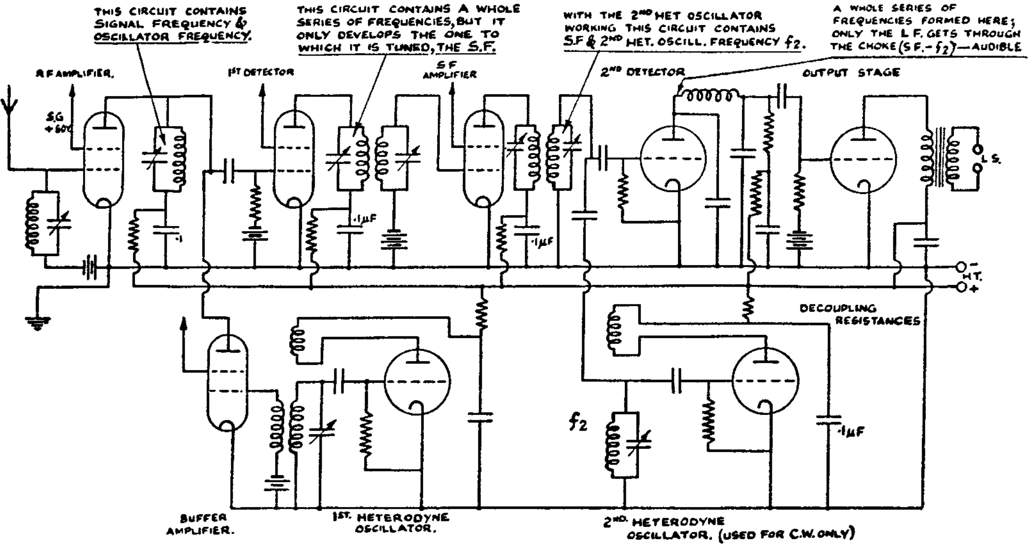


Fig. 6.14. Circuit diagram of a superheterodyne receiver

Multiple valve frequency changers

The electron stream of a valve may be controlled by any number of grid or mesh electrodes situated in the region between anode and cathode. We have already studied the effect of placing one, two, or three grids in this space when dealing with the triode, tetrode and pentode respectively.

To accelerate the electrons we maintain grids at positive potentials, and to retard them, negatively biased grids are used. Thus electrons approaching negative grids are slowed down, and in this case we have the effect of a cloud of slow moving electrons forming a space charge near the grid, rather like the space charge between the filament and control grid of an ordinary triode. This cloud of electrons is sometimes called a " virtual cathode ".

Consider the pentode valve shown in fig. 7.14.

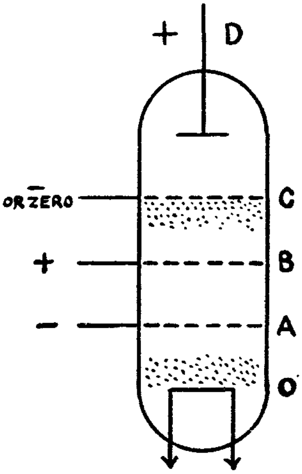


Fig. 7.14. Virtual cathode of pentode

Starting at O, an electron leaves the filament, and experiences two forces :—

(i) Repulsion due to the first negative grid A.

(ii) Attraction due to the positive grid B.

If the latter force is great enough, the electron moves rapidly across the space AB, and then enters a retarding region between, B and C, due to the grid C, but is finally attracted to the positive anode D. The electron density is greatest in the spaces OA and BC. The density in the region BC is affected by the potential of the grid C, being greatest when this potential is most negative.

Just as the electron stream from the original cathode is controlled by the first grid A, so the stream between B and D is controlled by the potential of the grid C.

The whole action may be summed up by saying that the final electron current reaching the anode is influenced by the potentials applied to any controlling electrodes in the anode-filament space, and if these potentials are alternating and of different frequency, theory and experiment indicate that the anode current contains an alternating component whose frequency is the difference between the frequencies of the original input voltages.

The pentode valve, therefore, can be used as a frequency changer, and a practical circuit of a pentode frequency changer with separate oscillator is shown in fig. 8.14.

When the tuned anode oscillator (fig. 8.14) is working, an alternating voltage is induced in the coupling coil in its grid circuit. This voltage is applied between the suppressor grid (third grid) of the pentode and earth.

The input signal is applied between the first grid and earth so that a suitable transformer S.F.T. in the anode circuit can be tuned to the different frequency and give an output which may be further amplified by later valve stages.

This converter is very good in most respects, but is rather expensive in valves and components.

It has been superseded by a multiple valve, known as the triode-pentode, which is merely a combination of the two valves in one envelope.

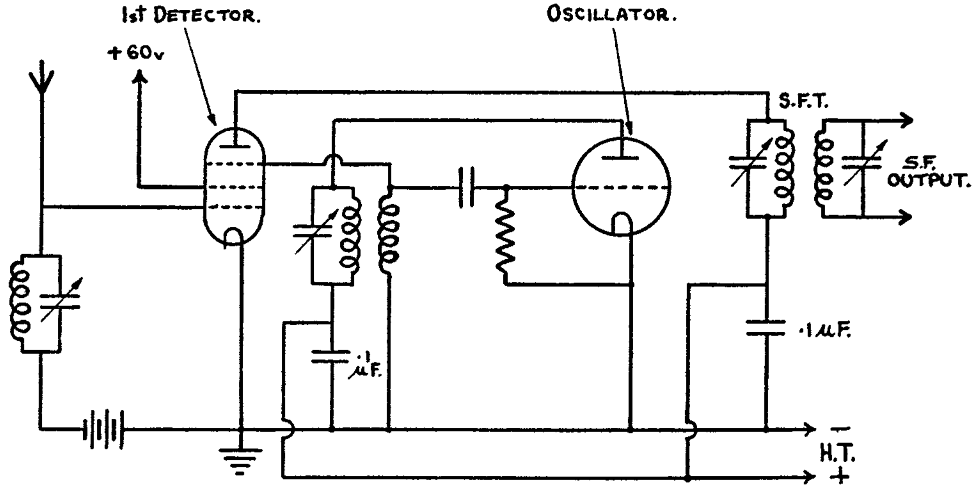


FIG. 8.14. Pentode frequency changer

The hexode

An objection to the pentode frequency-changer is the fact that there is no screening between suppressor grid and anode. If another screen is added, we obtain the hexode valve which may be used with a separate oscillator in the circuit shown in fig. 9.14.

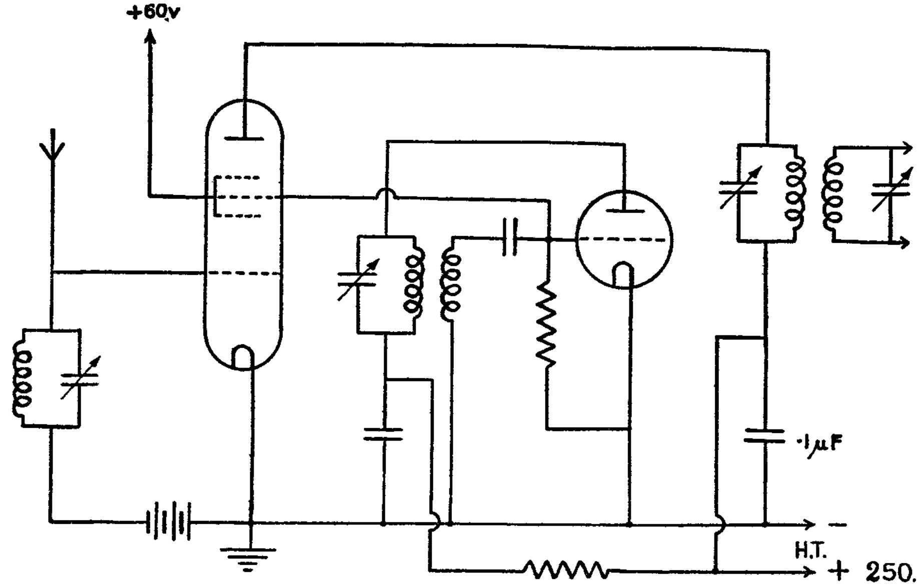


Fig. 9.14. Hexode frequency changer

The triode-hexode

If the two valves shown in fig. 9.14 are placed in one envelope we obtain a triode-hexode (fig. 10.14). The circuit used is the same as if the valves were separate.

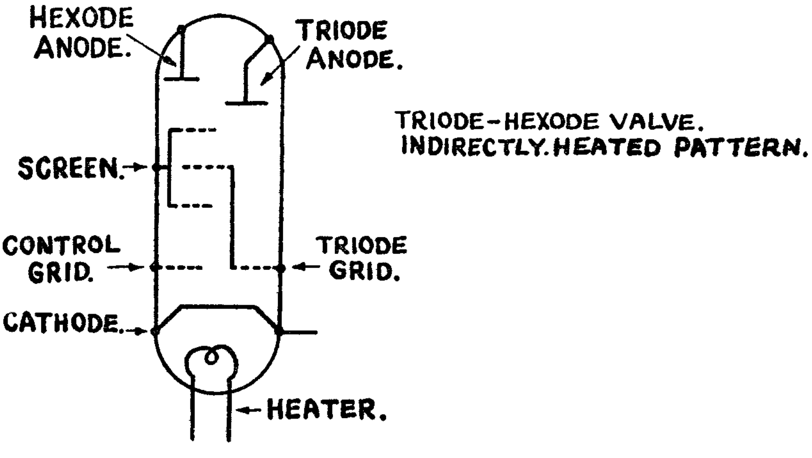


Fig. 10.14. Triode-hexode valve

The heptode

The addition of one more grid to a hexode gives a single valve frequency changer also not requiring any additional oscillator valve. Starting from the filament the electrodes are, in order, oscillator grid, oscillator anode, screen, signal grid or modulator grid, screen and anode. Such a valve is termed a heptode or pentagrid.

The three inner electrodes, i.e. filament, grid and oscillator anode are connected as an ordinary reaction oscillator. The resulting variations of the electron stream are mixed with those due to the signal input, and a supersonic frequency output is finally delivered from the transformer S.F.T. (fig. 11.14). Although extensively used for broadcast reception up to 1,500kc/s, the heptode is not satisfactory for high-frequency working.

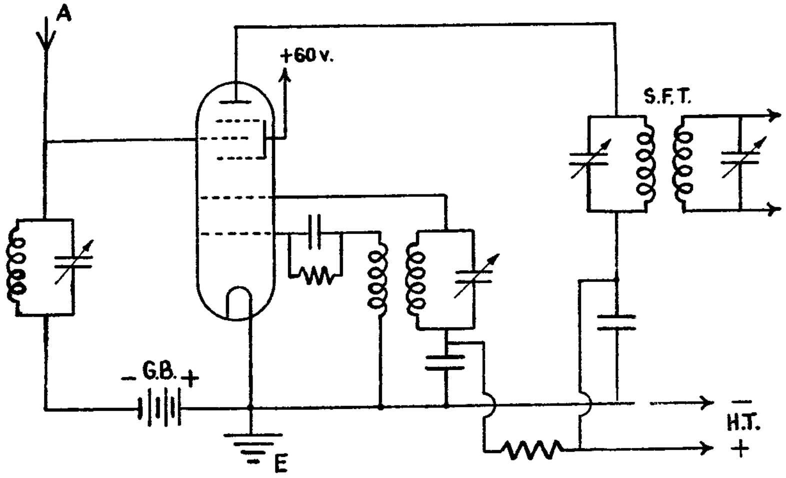


FIG. 11.14. Heptode frequency changer

In the first place, the screening between the oscillator anode and signal grid is not perfect, and there is some pulling and other interaction. The oscillator mutual conductance is rather low, and oscillation is difficult to maintain at high frequencies. This may be overcome by placing an external triode oscillator in parallel with that normally used. Harmonics of the oscillator can also cause interference or whistles.

The most effective frequency changer is probably the triode-hexode and this is extensively used in commercial receivers.

Interference problems

By reason of multiple detection and the presence of oscillators, the superheterodyne receiver is particularly subject to interference troubles which do not affect straight sets.

Second channel interference

This has already been discussed, but may be mentioned again for completeness. If a signal can reach the first detector having a frequency removed from the desired signal by twice the supersonic frequency, the oscillator frequency lying between these, then both signals will give the desired supersonic frequency. The resulting interference cannot be removed by any subsequent selective amplification.

The remedy is to use enough selectivity prior to the frequency changer to reduce the unwanted signal to a harmless level.

A band-pass filter H.F. stage, or rejector circuit, will secure this effect. Image-suppression circuits have been devised to deal with special forms of this interference, but these are not in use in any existing Service sets.

Beat interference

If a strong local station differs in frequency from the wanted signal by an amount equal to the intermediate frequency, there is a possibility of voltages due to both these sources reaching the first detector. Here they will give rise to a beat-frequency equal to the intermediate frequency, causing interference. Here again the remedy is sufficient pre-selection or selective amplification preceding the first detector.

Intermediate frequency harmonics

The second detector inevitably generates harmonics of the S.F. If these are coupled to the input circuit they may cause interference. Suppose, for example, a signal of 550 kc/s is received, the S.F. being 110 kc/s. The fourth harmonic is 440 kc/s which can beat with 550 kc/s to give the correct S.F. of 110 kc/s. Thorough screening and good circuit lay-out will prevent this effect.

Oscillator harmonics

These may beat with undesired signals reaching the first detector. Here, again, the best cure is the use of adequate pre-selection. Good oscillator tuned circuits help, because they offer very small reactance to harmonics of the oscillator frequency, and thus harmonics only develop small voltages across the tuned circuits.

Extensions of the superheterodyne principle

The most highly developed superheterodyne receivers are used to receive weak signals from distant short-wave transmitters. These signals may have an amplitude of a few microvolts, and this, combined with their high-frequency, sets a very difficult reception problem. In an effort to overcome this, two frequency changes are made, with two S.F. amplifiers. The first S.F. is much higher than the second and is employed principally to secure immunity from second channel interference. It may be 1,600 kc/s. The output from the first S.F. amplifier is mixed with that from a second heterodyne oscillator giving an S.F. of, say, 110 kc/s.

After still further amplification, the signal is rectified and passed into the audio-amplifier stages, or, if C.W. is being received, a third heterodyne oscillator is caused to beat with it to give a final audio tone of from 400 to 1,000 c/s in the telephones.

The use of a low second S.F. gives extremely high selectivity. The schematic diagram of such a receiver is shown in fig. 12.14.

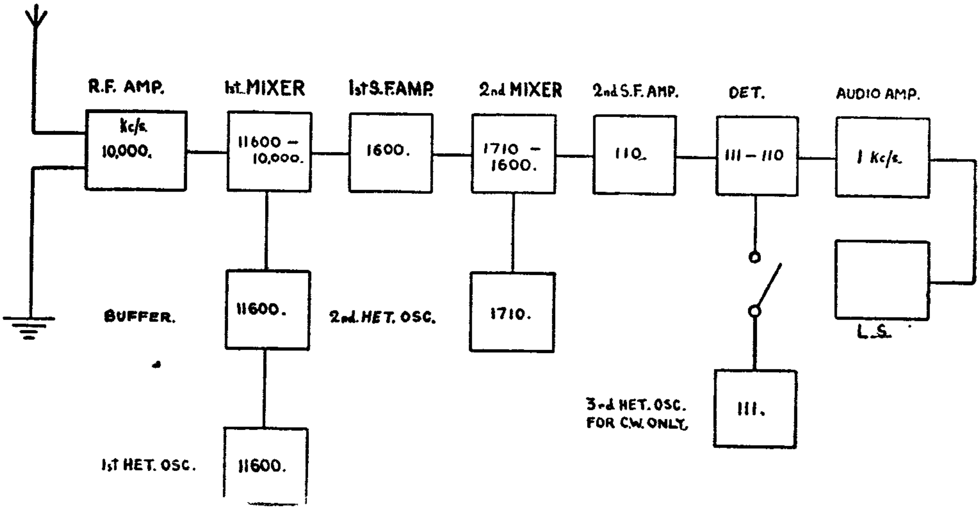


FIG. 12.14. Block diagram of double superheterodyne receiver

The operating frequency (kc/s) is shown in each unit of the receiver. Extreme care has to be taken in the design and construction of such receivers to avoid the production of unwanted whistles and interference, and to secure a good ratio of signal to noise.

In spite of the apparent complexity, such receivers can be constructed in portable form and are so used in the Service. It should be noted that the only tuning elements which need be made variable are the condensers of the R.F. amplifier and of the first heterodyne oscillator.

Choice of the supersonic frequency

Now that most of the problems of superheterodyne reception have been discussed, we may refer to the considerations underlying the choice of the supersonic frequency.

In the first place this cannot lie in any part of the frequency bands occupied by desired signals, and this immediately restricts the choice to the regions of 100 kc/s, 450 kc/s or some high value like 1,500 kc/s.

If the first figure is chosen, care is needed to avoid second-channel interference, especially if high-frequency signals are to be received. High adjacent channel selectivity* is secured with ease, and receivers of high gain and stability are readily constructed. Finally it is difficult to gang the tuning condenser sections if this frequency is used and we must employ condensers having specially shaped vanes.

If 450 kc/s is chosen, second-channel interference is not troublesome, controls may be ganged, and selectivity is fairly good as regards separating stations on adjacent frequency channels. Padding condensers must be employed to preserve a constant frequency difference between signal and oscillator-tuned circuits.

The third possible choice, 1,500 kc/s, is only used in specialised receivers, because of the poor gain and selectivity obtained by the choice of a S.F. of this value. The first S.F. amplifier of a double superheterodyne receiver is commonly operated at this frequency.

The choice of a high S.F. also secures immunity from whistles due to oscillator harmonics and beat interference, because the interfering signals lie outside the tuning range of the receiver or are too weak to become obtrusive.

Gain or Volume control

Receiving sets are required to receive signals of all strengths. For instance, a nearby transmitter may set up a large voltage across the first tuned circuit, whereas another station which is some thousands of miles away will, at best, set up a signal of only a few microvolts in the receiver aerial system. Therefore, unless some form of control is employed at the receiver, a wide range in the output strength will occur at the telephones or loud speaker. Some signals will cause severe overloading of all the valves in the set. others being barely audible after the output stage.

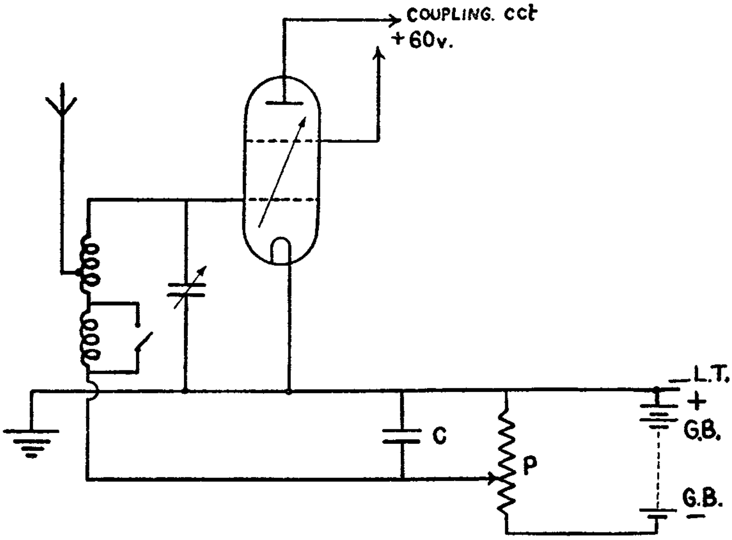


Fig. 13.14. Volume control using potentiometer

Besides this, there are other reasons why a control of signal strength is desirable at the receiver. First, the detector stage operates best at one particular value of the input so that there should be a " gain control " or " volume control " before the detector, and, second, the audio output required depends on what type of device is being driven by the

---

* NOTE—Adjacent channel selectivity refers to the ability of a receiver to receive signals from a desired station without interference from stations using carrier frequencies near that of the desired station.

output valve. Telephones are very sensitive and are placed close to the ear so that only small power is necessary to operate them. Loudspeakers are not so sensitive, and recorders require a power of several watts if the recordings are to be successful.

One method of volume control—reaction—has already been described. By the use of reaction we are able to bring up the strength of a very weak signal, but we cannot conveniently employ reaction to decrease the strength of a particularly strong signal. We have to use other means.

The most common method of pre-detector volume control is by the use of variable-mu valves. The principle of this special type of valve is explained on p. 106, Chapter IX ; a circuit employing this kind of control is given in fig. 13.14.

A potentiometer P is connected across the grid bias battery, and its moving arm is connected to the " earthy " ends of the grid circuits of the variable-mu valves in the receiver. If several amplifier stages are used, the grid circuits must be decoupled from each other, but if there is only one, it is only necessary to by-pass the resistance between potentiometer arm and earth by a condenser C (fig. 13.14) of about 0·1μF. This must be done, for part of the potentiometer resistance is in series with the tuning coil and would spoil its selectivity and gain unless it were short-circuited as far as signal-frequency currents are concerned.

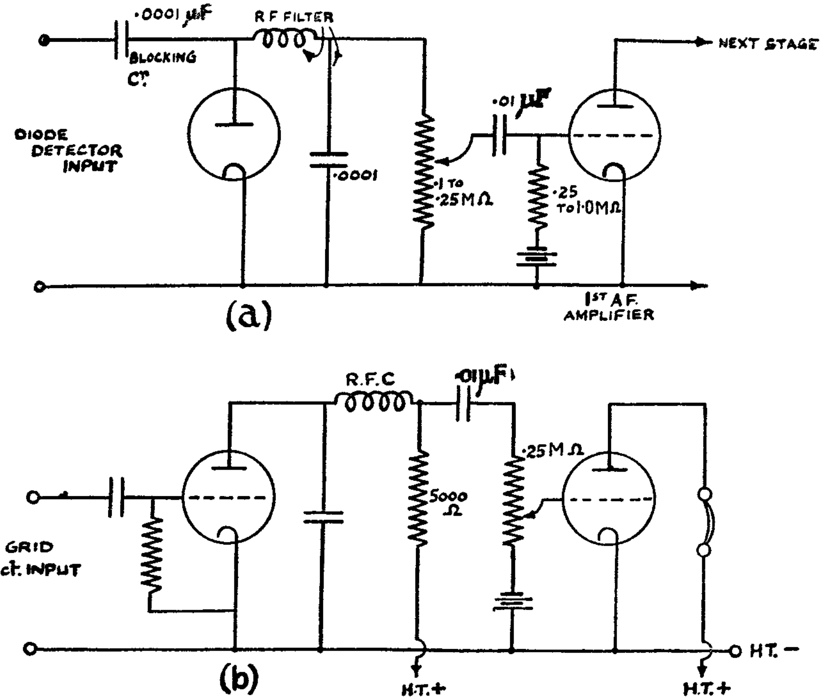


FIG. 14.14. (a) Post-detector control, diode detector
(b) Post-detector control, triode detector

Post-detector volume control is best obtained by using a potentiometer as the detector output load and tapping off from it whatever voltage is required to give the desired output level. Two circuits illustrating this method are shown in fig. 14.14 (a) and fig. 14.14 (b).

In the circuit fig. 14.14 (a) it is necessary for the amplifier grid leak to be much larger in value than the diode load, or distortion occurs with deeply modulated signals.

Automatic volume control

Transmissions of high frequency (greater than about 600 kc/s) often suffer from fading, which can prove very troublesome. At one instant the telephones or loudspeaker of the set receiving the transmission may be violently overloaded, and shortly afterwards the signals may disappear entirely. This is a serious disadvantage of such transmissions, and it is usual to compensate for the fading out by judicious use of the manual volume control. A better way of overcoming signal variations is by *automatic volume control* (A.V.C.). In this arrangement the amplitude of the received carrier controls the sensitivity of the receiver in such a way that the audio-output is rendered practically constant in intensity for wide variations in input signal strength.

One simple method of securing this control is to rectify the received signal, and to use the resulting D.C. voltages as grid bias on variable-mu valves. The stage gain of a variable -mu valve decreases as the negative grid bias is increased. Now after rectification strong signals will produce the largest bias voltage, and so will be amplified less than weaker signals, and a compensating action is obtained whereby the final audio-frequency output does not depend greatly on the amplitude of the received carrier.

It is obvious that the control voltages must not depend on the modulation used, or the loud and soft passages of music would tend to be equalised, and so in practice the carrier only is arranged to provide the required grid bias. A more detailed explanation of this method of volume control follows :—

It is well known that, in the anode circuit of the detector valve of a straight set, and in the second detector of a superhet, there are current components of radio-frequency in addition to the A.F. currents. The amplitudes of the R.F. currents present depend directly on the amplitude of the signal in the aerial. To obtain A.V.C. the radio-frequency components are rectified by a diode and caused to develop a steady voltage across a load resistance. The polarity of this voltage is arranged to be suitable for providing negative bias on the controlled stages. The voltage developed across the load resistance, which is usually of the order of ·5 to 1 megohm, is in such a direction that the end remote from the filament is negative with respect to the filament. If this point is connected to the grid circuits of the preceding variable-mu valves, these grids will acquire potentials negative compared with the filament, the magnitude of which will depend on the aerial input. The larger the incoming signal, the greater will be negative bias on the grids of the valves connected to the A.V.C. system. The magnification afforded by a valve depends upon the working slope, which in turn depends upon the grid bias. At large values of grid bias the working slope is low and hence the stage gain is small, so that for large inputs the sensitivity of the receiver is reduced and the opposite is the case for small signals. The resultant audio-output is in consequence nearly constant for a wide range of variation in the aerial input.

The diagram of fig. 15.14 gives in a simplified form the general layout of an A.V.C. system ; the diode B is the second detector of a superhet, and it produces across its load resistance R the usual products of rectification, namely, a direct voltage, a modulation frequency voltage and some supersonic frequency voltages. The first of these is used, via the heavy line, to provide the A.V.C. voltage to the grids of the previous variable-mu valves ; the modulation frequency and the S.F. components are prevented from reaching the grids of these valves by means of the resistance-capacity filters R_2, C_2, R_3, C_3 ; this is a very necessary precaution because it renders the control bias independent of the programme modulation.

The disadvantage of this system of simple A.V.C. is that an A.V.C. bias becomes operative even on weak signals where the full sensitivity of the receiver is required. By using a separate diode for A.V.C. purposes and applying to its anode a potential which is negative with respect to its filament, the A.V.C. system is rendered inoperative to all signals of peak value less than this. A further improvement in control may be obtained by extra amplification of the R.F. or S.F. before rectification for A.V.C. purposes, which gives a large control voltage.

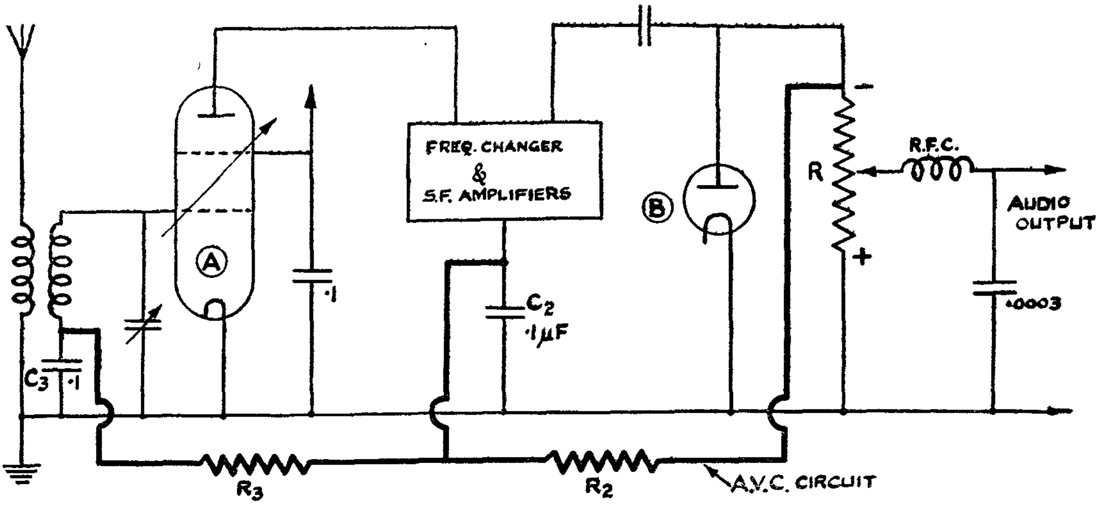


FIG. 15.14. Simple A.V.C.

Fig. 16.14 below is a diagram of the R.F. and S.F. stages of a superhet receiver with simple A.V.C. The heavy lines of the diagram indicate the components involved in the A.V.C. process.

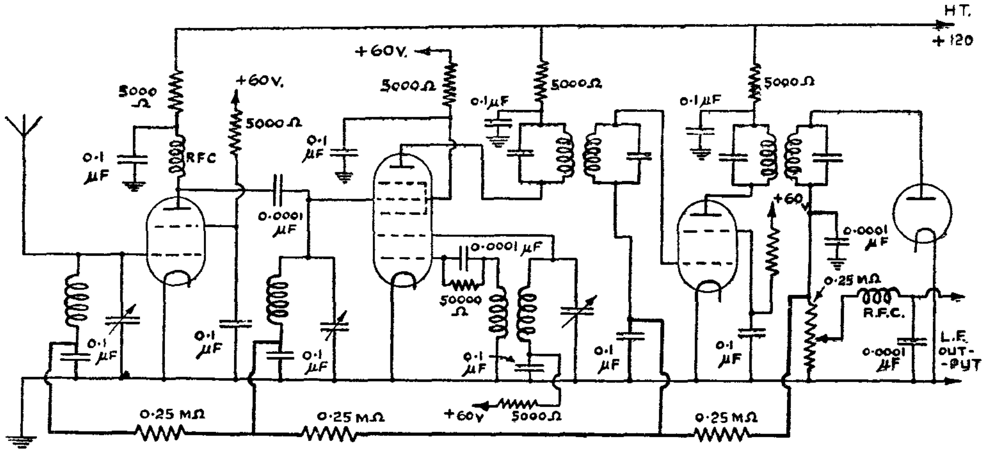


FIG. 16.14. Typical R.F., S.F., and second detector stages of a superheterodyne receiver provided with simple A.V.C.

CHAPTER XV

AERIALS—DIRECTION FINDING

The aerial or antenna is that part of a transmitting station which radiates wireless waves, or that part of a receiving station which picks up the radiation sent out by a distant transmitter. In aircraft the same aerial is used for both purposes, but at ground stations it is usual to have two aerials.

Frequently the transmitter together with its aerial is situated at some distance from the signals office and it is " remote controlled " ; the receiver on the other hand cannot be so easily operated at a distance, and it is therefore housed in the signals office and its aerial leads in directly.

While a receiving aerial is usually small and lightly constructed, a transmitting aerial is a much bigger and more expensive affair.

Doubling the height of an aerial may easily increase fourfold the amount of energy it can radiate : a 1 kw. transmitter would become as effective as a 4 kw. installation simply by doubling the height of the aerial, so that although an efficient aerial involves a high first cost, it becomes an economy in the long run. In the case of reception it is less expensive to provide an extra stage of amplification in the receiver than to build a costly aerial, although even then a good aerial is desirable because it increases the signal to noise ratio ; several stages of amplification tend to introduce noise.

Transmitting and receiving aerials possess the reciprocal property common to all physical absorbers and radiators ; a good transmitter is also a good receiver, and in what follows, any reference to an aerial applies equally well to both transmitter and receiver aerials.

Aerials are of three main types :—

(1) The Hertz aerial—used at high frequencies.
(2) The Marconi aerial—can be used at all frequencies but is specially useful at low frequencies.
(3) The loop aerial—used in direction finding.

The Hertz aerial

In Chapter VI we discussed at some length the properties of the oscillatory circuit ; fundamentally such a circuit must possess inductance and capacity and, further, the resistance must be less than a certain value. In such a circuit, as we have seen, the free electricity will surge to and fro, first charging the condenser, then creating a magnetic field around the inductive coil, then returning to the condenser and so on, until the energy is eventually lost and the electricity comes to rest. In Chapter IX we saw how these oscillations could be maintained by means of a valve and how, by opening out the circuit so that its condenser field became very widely spread, it could be made to radiate wireless waves.

The method in which the circuit is modified to ensure a large spread of the condenser field underlies the design of the aerial. So far as we have considered it, all we have done is to open the condenser of the oscillatory circuit into the conventional " aerial and earth " system, but that is not the only way to proceed.

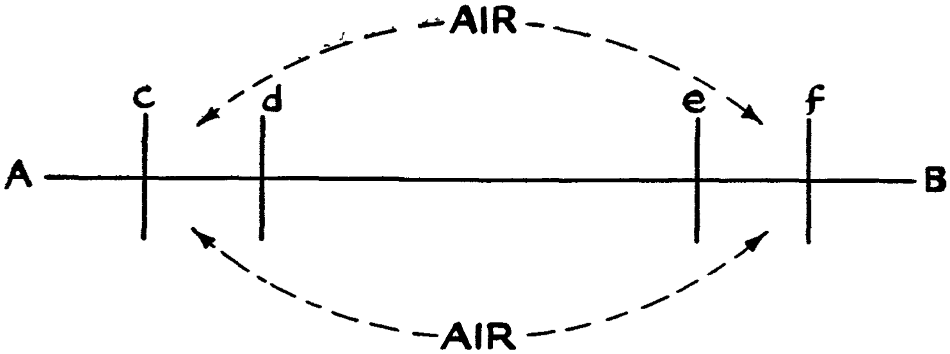


FIG. 1.15. Distributed inductance and capacity

Consider a single horizontal rod made, say, of copper (fig. 1.15) ; then :—

(*a*) the system possesses inductance ; if a current were passed in it magnetic lines would be set up around it ; any variation of current would cause a variation in the magnetic field which, in its turn, brings about an induced E.M.F. tending to oppose the current changes. It is this sequence of events which is characteristic of an inductive system.

(*b*) the system possesses capacity. How a *single* conductor can have capacity is a little difficult to see, but take any two sections (*cd, ef*) : these sections are two conductors and they are separated (externally) by air ; capacity, therefore, exists between them, and so from any pair of such elements there will be a contribution to the general distributed capacity of the system. The problem is something like that of the self-capacity of a coil mentioned in Chapter VIII.

(*c*) if the system possesses inductance and capacity (provided the resistance is low enough) it must be capable of oscillations.

Imagine for a moment that we could drive all the free electrons to the end A and then let them go ; the electricity would then surge backwards and forwards with a frequency determined by the effective inductance and capacity of the rod. Let us defer discussion as to how the oscillations can be maintained and examine the results of such a state of affairs in the rod itself.

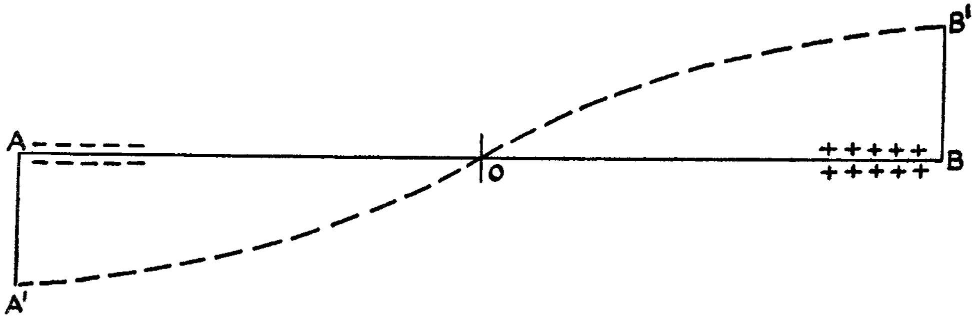


FIG. 2.15. Voltage distribution (1)

With the electrons driven up to the end A (fig. 2.15) we shall have a negative potential at A, with a positive potential at B. As we proceed along the rod from A the negative potential will gradually diminish till at the middle of the rod we reach zero potential and pass into a region of positive potential ; this is indicated by the dotted curve A^1 O B^1.

When the negative charge is released it will surge to the end B, and we shall have the state of affairs indicated in fig. 3.15, after which it will return to the conditions represented by fig. 2.15.

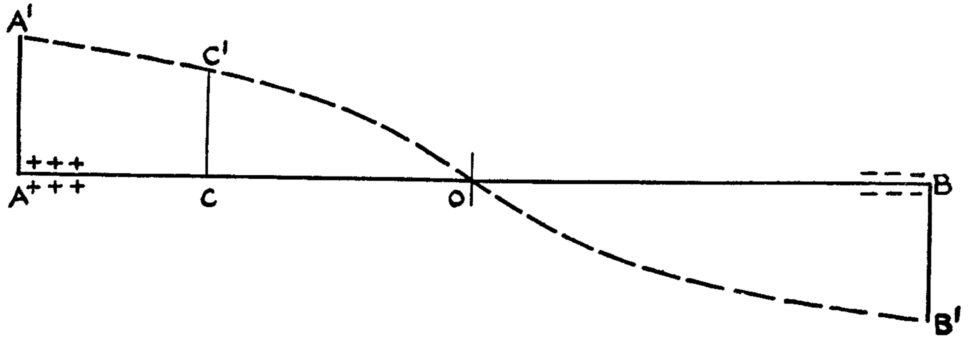


FIG. 3.15. Voltage distribution (2)

At the ends of the rod therefore we shall have voltage variations of amplitude A A^1 or B B^1, while as we move towards the centre O the amplitude of the voltage swing falls to zero ; at C, for example (fig. 3.15), the amplitude of the voltage variations during oscillation is equal to C C^1.

On the other hand, since the points A and B are open circuited, there can be no current amplitude at A and B, but the current will reach its maximum at the centre. Fig. 4.15 indicates the current distribution in such a system.

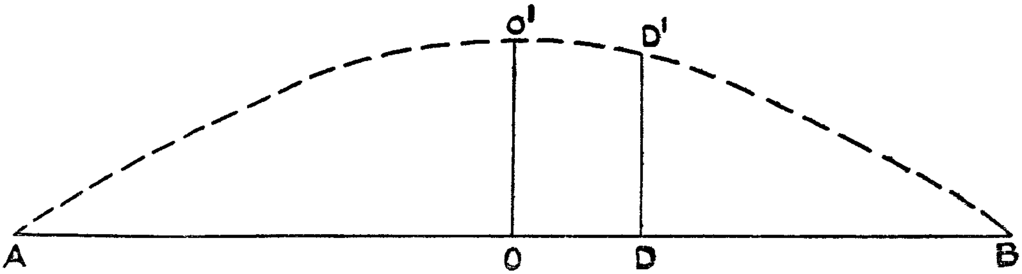


FIG. 4.15. Current distribution

At A and B the amplitude of the oscillatory current is zero ; at O it is O O^1, while at D it is D D^1.

The rod therefore behaves as an oscillatory circuit, and at the same time its widely distributed capacity enables it to radiate freely ; it constitutes what we call a Hertz aerial.

The wavelength of the radiation emitted during oscillation is nearly *twice the length* of the rod, to be exact :—

Wave length in metres = length of rod in metres × 2·1.

For this reason this type of aerial is known as a " half-wave aerial " ; it will only radiate appreciably at its own frequency or on harmonics of that frequency.

Contrary to the practice with which we are familiar, no part of a half-wave aerial is earthed, but as we shall see shortly, the presence of the earth will exert a very profound effect on the character of the radiation it transmits.

A half-wave aerial is maintained in oscillation by means of an independent oscillator, operating at either the fundamental or harmonic aerial frequency.

This oscillator or " transmitter " is coupled to the aerial which may be either *centre* or *current* fed, or *end* or *voltage* fed. These terms simply mean that the energy is introduced

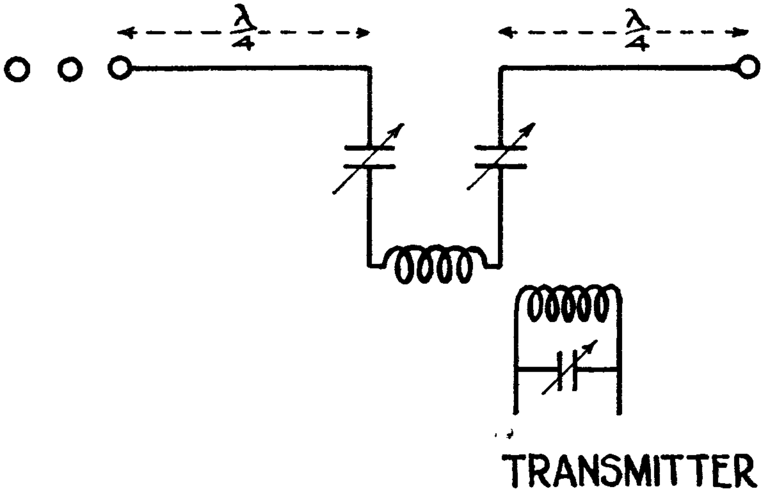


FIG. 5.15. Current fed Hertz aerial

either at the centre (where the current is maximum) or at the end (where voltage is a maximum.)

In the case of centre feed, we must use a coupling circuit appropriate to ***maximum current***, i.e. an acceptor circuit.

Fig. 5.15 indicates this: the condenser of the circuit is usually in two sections to ensure a complete balance.

For end feed we must use a circuit appropriate to maximum voltage, namely a " rejector circuit ", to couple the transmitter to the aerial (fig. 6.15).

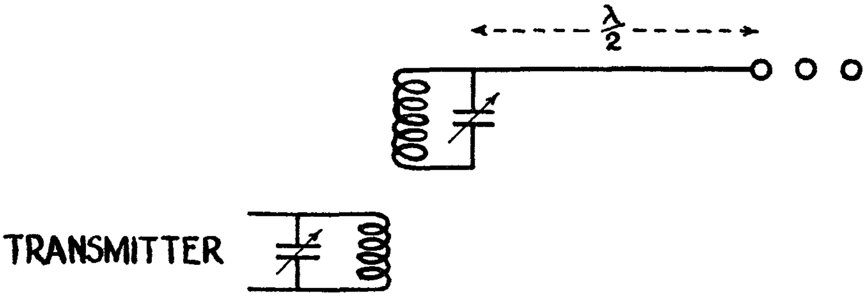


FIG. 6.15. Voltage fed Hertz aerial

The reader will realise that a serious difficulty arises here ; it will be impracticable to bring the transmitter close to the feeding point, and the latter must be connected to the aerial by means of feeder lines. These lines will carry R.F. currents which will tend to radiate independently of the aerial, and destroy the simplicity of the system.

To avoid feeder line radiation very careful design is necessary, the general idea being to ensure that at any instant the effect due to one feeder line is exactly neutralised by that due to the other.

There is one aspect of the behaviour of a horizontal half-wave aerial which is of great importance : it does not send out any radiation along the plane of the earth unless it is situated at a height equal to several wavelengths above the earth's surface. Normally it throws its radiation upwards into the sky. Although this fact has important application in certain branches of radio communication it is clear that it limits the use of such an aerial for ordinary point to point communication or for broadcasting.

The Marconi aerial

The half-wave aerial described above is limited in its application ; it can only be used

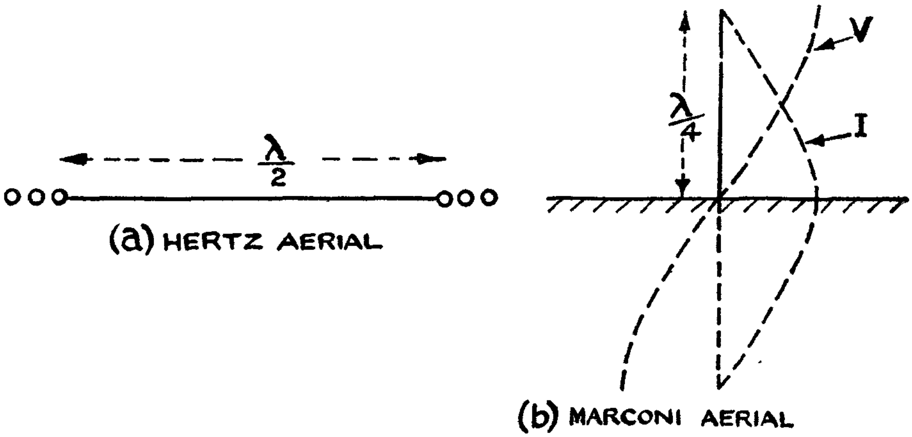


FIG. 7.15. Development of the ¼-wave aerial

for high-frequency work. It does not give any radiation along the horizontal plane. It was Marconi who suggested the use of one half of the Hertz aerial in a vertical position, with the lower half earthed. With such an arrangement we shall get general radiation in the horizontal plane. Fig. 7.15 shows this development; the dotted line of fig. 7.15 (b) corresponds to the missing half of the original aerial. The earth's surface behaves like an electrical mirror and we may regard the Marconi aerial as a vertical Hertz aerial with the lower half as the " image " of the upper half.

The dotted curves in fig. 7.15 (b) indicate the voltage and current distributions in the corresponding half-wave aerial. It is only the top halves of these curves which have any practical reality, so that for the vertical quarter wave aerial we shall have maximum current and minimum voltage at the bottom, with maximum voltage and minimum current at the top. This will mean that if such an aerial is used as a transmitter it must be " current fed " at the earthed end.

Now it is clear that if we are proposing to use a quarter-wave aerial to transmit at low frequency it might have to be inconveniently high; at 300 kc/s for example it would have to be about 800 ft. high. If we want to retain the efficient radiating properties of the quarter-wave aerial there is no way out of it; it must be a quarter-wavelength high, but if we are prepared to sacrifice some efficiency we can use two devices to overcome the difficulty of height on low-frequency work.

(i) The aerial can be bent over, into an L shape (or T shape). The " top " will not contribute appreciably to the horizontal radiation for reasons already discussed, but it will serve to make the current distribution in the vertical limb more uniform. This is indicated in fig. 8.15.

Of the original length OC, it is now only OA which functions as a radiator (or receiver as the case may be), but the effect of the top is to make the current amplitude at the top of the aerial equal to AA^1.

(ii) The aerial can be loaded with an inductive coil at the earthed end; such a coil, of course, will increase the effective length, but it will not contribute a corresponding increase in the radiation properties of the aerial. Even so, such a loading coil has other advantages.

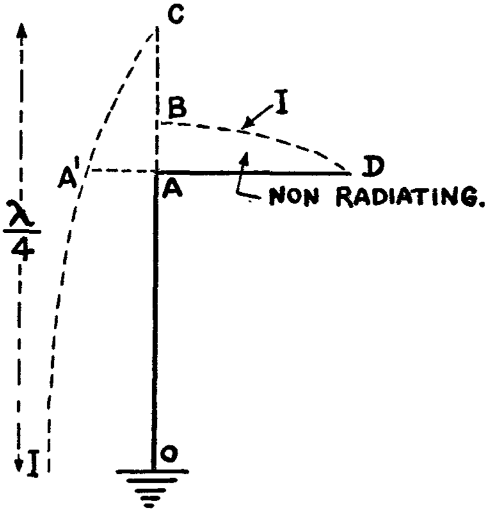


FIG. 8.15. The L-shaped aerial

(*a*) by varying its inductance (tappings or variometer), we can readily alter the natural frequency of the aerial.

(*b*) it is a convenient arrangement for coupling the aerial to the transmitter or receiver as the case may be.

It is interesting to note that in some modern transmitting systems the problem of height has been accepted, and the aerial is actually tuned by having an adjustable (vertical) top which can be raised or lowered in order to tune the aerial, in order that is to make the height exactly a quarter-wavelength and to retain the radiation efficiency of the aerial without recourse to loading.

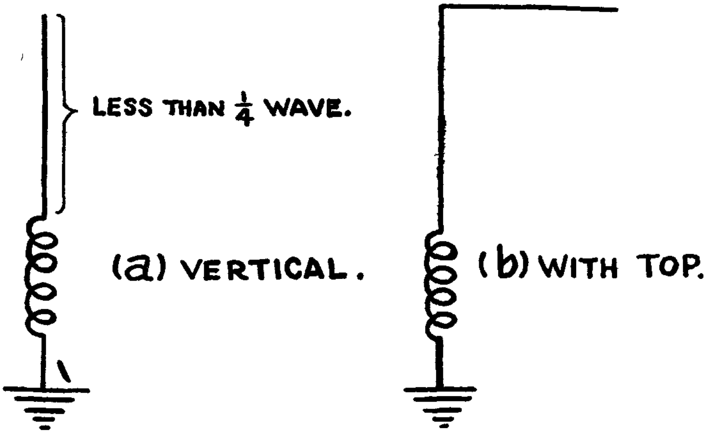


FIG. 9.15. Loading an aerial with inductance

To return to the loaded aerial, the reader will realise that in the consideration of such an aerial we have returned to something already quite familiar. This is the aerial which we developed in an earlier chapter, by opening out the condenser of an oscillatory circuit into the familiar " aerial and earth " arrangement. In such a system the inductance was concentrated in the coil, while the capacity was concentrated in the aerial, and that is effectively the case in our quarter-wave loaded aerial. The aerial tap in a transmitter is the device whereby the effective length of the aerial is made a quarter-wavelength simply by altering the loading inductance, a process we have previously called " tuning the aerial ".

Loading with inductance will increase the effective length, that is to say, it will make a physically short aerial capable of resonating to long waves : if we wanted to make a physically long aerial resonate to short waves we should have to load the aerial with a condenser.

A quarter-wave aerial can be made to function on harmonic frequencies : supplied with energy at twice its natural frequency it functions as a half-wave aerial, or with three times its natural frequency as a three-quarter-wave aerial.

Fig. 10.15 indicates the voltage and current distribution under these conditions : it is left to the reader to decide what type of feed would have to be used in each case.

While it is desirable in the interests of efficiency to design a transmitter aerial to be resonant with its input, that is to say to make the aerial the correct length, or to load it so that its effective length is correct, we do not go to this trouble in the case of a receiving aerial unless very special circumstances are involved. All we need to guide us in putting up a receiving aerial (apart from questions of atmospherics, insulation and of mechanical stability) is the general statement " the higher the better ".

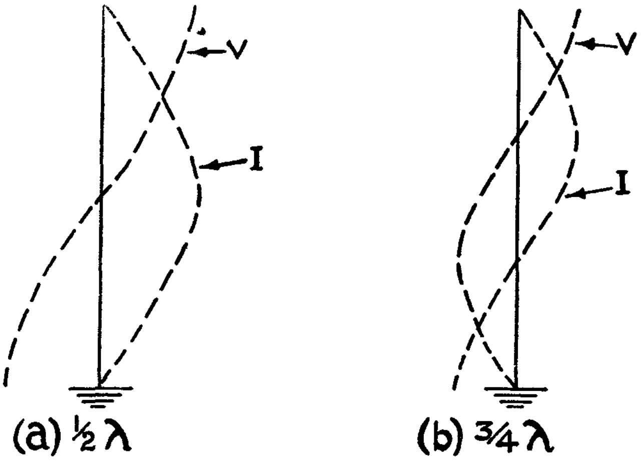


FIG. 10.15. Harmonic excitation

Polar diagrams

A motor car headlamp is designed to throw forward as much light as possible, and to send none at all backwards.

We could represent this state of affairs by drawing lines from a point representing the headlamp and making the length of the line represent the amount of light thrown out along the path corresponding to the line. Thus (fig. 11.15), the centre forward line OA will be longest in this case, and the line, corresponding to light which just escapes the edge of the lamp (OC) will be shortest. Since there is no light escaping to the left of OC or OC^1, the lines in that region will have zero length.

The curve joining the ends of those lines CBAB^{1}C^1 is called the *horizontal polar diagram* of the lamp.

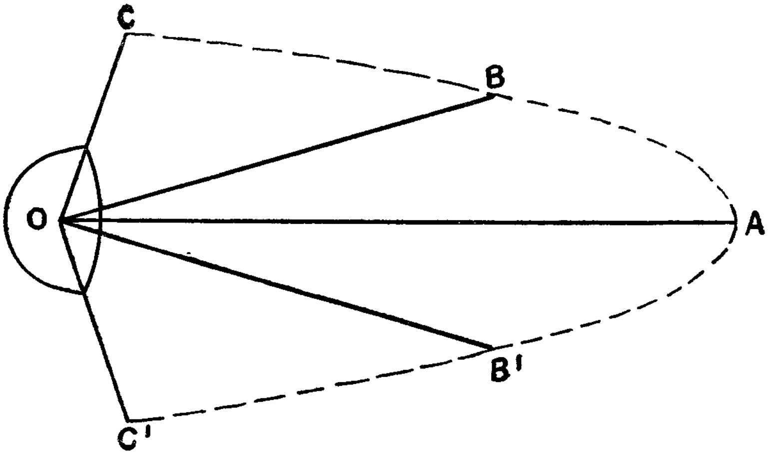


FIG. 11.15. Illustrating a polar diagram

We could, of course, carry out a similar process in a vertical plane passing through O and get the vertical polar diagram, which in this case would be about the same as the horizontal diagram.

A point of light in free space unhindered by mirrors or any other kind of obstructions would emit light uniformly in all directions: both the horizontal and vertical diagrams would be circles. An ordinary incandescent lamp might be expected to give a circular horizontal diagram, but owing to unavoidable irregularities in the filament, the curve departs from the circular form somewhat as shown in fig. 12.15.

The illustration gives the diagram for a lamp with a straight tungsten filament.

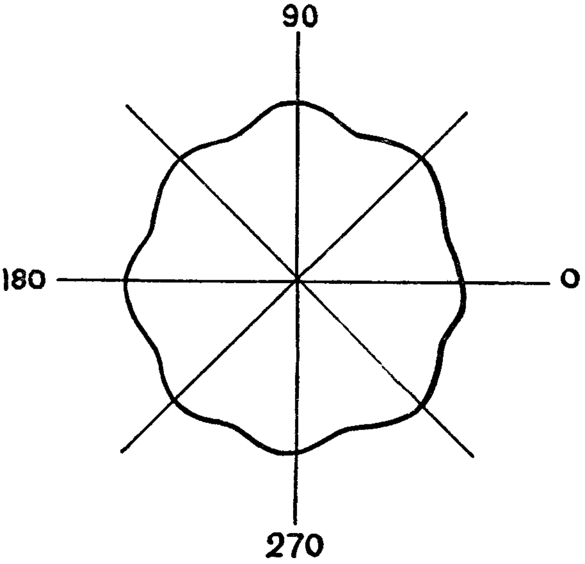


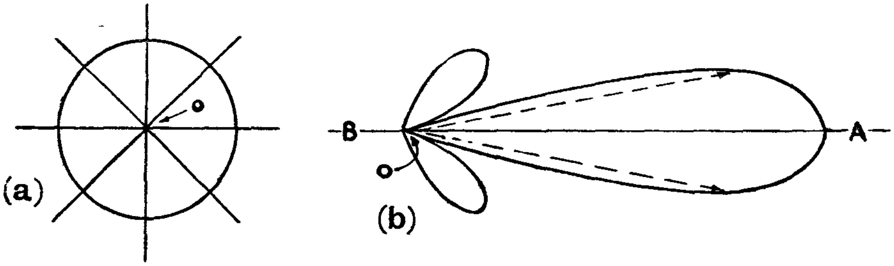
FIG. 12.15. Horizontal polar diagram of a filament lamp

Now a transmitting aerial gives out radiation—not luminous radiation like a lamp, but the invisible radiation constituting wireless waves—and we can represent on a polar diagram the amount of radiation streaming away in any direction.

A simple vertical quarter-wave aerial radiating freely (i.e. unhampered by trees, buildings, and so on) will radiate equally well in all directions along the earth's surface : its horizontal polar diagram is a circle : it is like a lamp, shining, unhindered by mirrors or opaque objects, evenly in all directions.

The general radiation from a quarter-wave aerial has obvious advantages, but if we desire to establish wireless communication between two points A and B on the surface of the earth, it is only the radiation going from A to B or vice versa which is of any use, and for that purpose we make use of a complicated aerial system called an *aerial array* which throws most of the radiation in the required direction, just as our motor car headlamp throws its light where it is wanted. The horizontal polar diagram of an aerial array will not be a circle, but it will show a predominance of radiation in one direction with a certain number of secondary maxima.

Fig. 13.15 (a) shows the horizontal polar diagram of a free quarter wave aerial, and fig. 13.15 (b) that of a certain type of aerial array. In each case the aerial is understood to be at O, perpendicular to the plane of the paper.



(a) ¼ WAVE AERIAL. (b) AERIAL ARRAY.

FIG. 13.15. Examples of polar diagrams

The polar diagram of an aerial tells us not only in which direction it radiates its energy most freely when acting as a transmitter but also *from which direction* it will most readily receive wireless waves when acting as a receiving aerial. Thus, in fig. 13 15 (a) the aerial at O will receive equally well radiation arriving from all directions ; that of fig. 13.15 (b) will receive signals coming from the direction of A quite well, but it will respond hardly at all to signals arriving from the direction of B.

The loop aerial

The most important polar diagram from our point of view is that of a loop aerial.

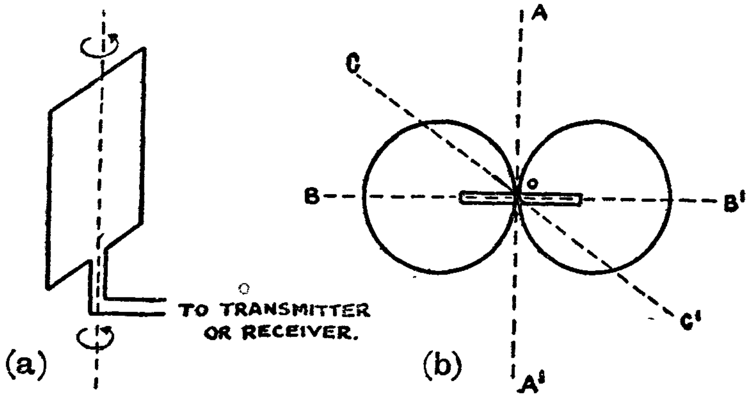


FIG. 14.15. The loop aerial

In its simplest form a loop aerial would consist of a rectangle of wire supported with its plane vertical and capable of rotation (for reasons we shall see presently) about a vertical axis. Its horizontal polar diagram is a " figure of eight ", two circles touching at the centre of the loop.

Fig. 14.15 (a) illustrates this : (*a*) represents the loop aerial and (*b*) shows a plan of the aerial together with its polar diagram.

The polar diagram tells us that the loop will neither transmit nor receive along the line A A^1, while its maximum effectiveness will be along B B^1; along any other intermediate direction C C^1 the effectiveness will be somewhere between the maximum and zero values.

As our chief interest in the loop concerns its receiving properties, let us investigate its behaviour in this respect from first principles.

Fig. 15.15 shows in plan a vertical loop and T represents a transmitter sending out general radiation. At some distance away the radiation will consist of sets of vertical electrolines, accompanied by horizontal magnetic lines ; the figure shows three such pairs.

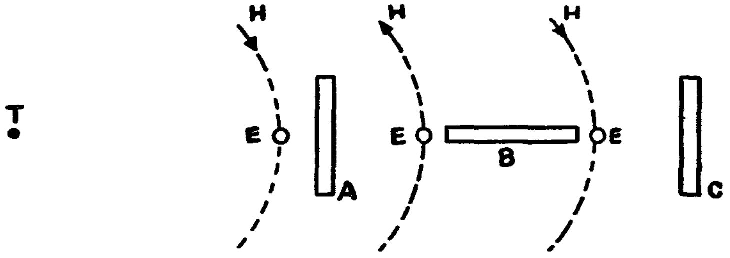


FIG. 15.15. Directive property of a loop aerial

With the receiving loop in the position A there will be no change in the flux threading the loop as the radiation sweeps past, simply because, on account of the fact that the magnetic lines are parallel to the plane of the loop, there is no flux through it at any time, so that no E.M.F. will be set up in the loop at all if it faces the transmitter at A. This, of course, corresponds to radiation coming along A A^1 in fig. 14.15 (b).

In position B, however, we shall have as much flux as possible (because the plane of the coil is at right angles to the magnetic lines), and as the flux changes due to the passing radiation we shall have E.M.Fs. induced ; this corresponds to radiation coming along B B^1 [fig. 14.15 (b)].

It seems reasonable to expect that as the loop wire turns from position A to B the strength of the signal set up in it will vary from zero to a maximum, and as the loop passes from B to C it will fall from maximum to zero again. As the loop is turned around it will receive maximum strength signals when its plane " points to " the transmitter, and zero strength signals when its plane faces the transmitter. It is this property which enables us to use the loop aerial as a basis of direction finding.

The manner in which the response of a loop aerial to the signals of a distant transmitter varies as the loop is turned round a vertical axis can be investigated by means of the polar diagram.

In fig. 16.15 A B represents a plan of the loop and T a distant transmitter. The distance OP then represents the strength of the signals in the aerial, with the loop in that particular position. Suppose now we turned the loop about the vertical axis through O and took a series of readings of the angle θ, and of the intercept corresponding to OP, we could plot a graph which would show how the signal strength varied (assuming the transmitter to be working steadily) throughout a complete revolution of the loop.

Such a graph is shown in fig. 17.15 and the corresponding loop positions are indicated below.

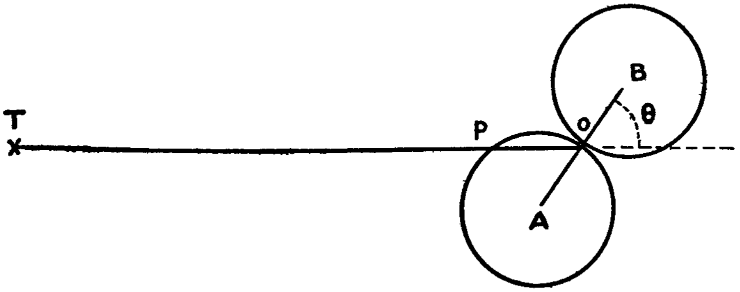


FIG. 16.15. Illustrating reception by a loop aerial

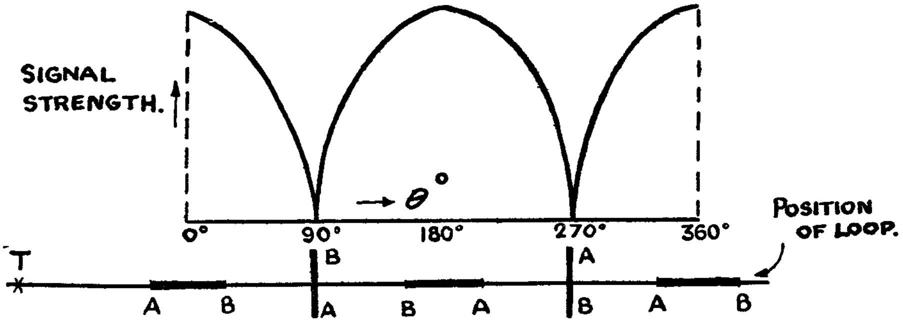


FIG. 17.15. Variation of signal strength during a revolution of the loop

In a complete rotation there will be two maxima and two minima, and there are two important points to notice :—

(i) the minima are sharper than the maxima ; the maxima persist at about the same level while the loop is turned through nearly 30°, while the minima only persist over a few degrees of rotation. For this reason, when a loop is used to find direction it is always adjusted for minimum signal. Its plane will then be facing the transmitter.

(ii) while these properties of a loop aerial enable us to determine the line in which the origin of the radiation must be, we cannot say on which side of the loop the transmitter is situated. Fig. 18.15 will make this clear.

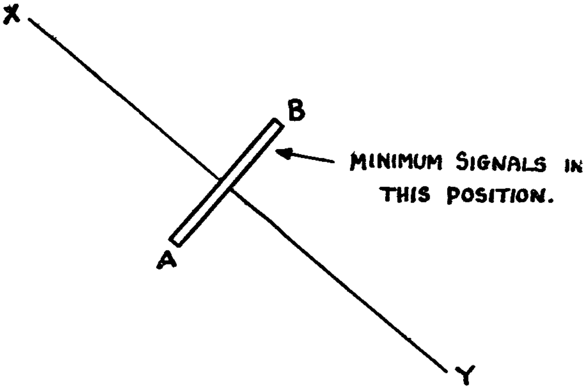


FIG. 18.15. The 180° uncertainty

The loop AB has been turned to give a minimum, then we know that the transmitter must be somewhere on the line XY, but whether it is on the X side of the loop or on the Y side we have at the moment no way of telling. This indefiniteness is called *the 180° uncertainty ;* the loop is able, we say, to determine *direction* but not *sense.*

In practice a simple loop such as we have in mind will not give sharp minima for two main reasons (i) antenna effect and (ii) night effect. The first of these is due to the vertical sides of the coil acting as vertical aerials or antennæ, deriving voltages from the electric component of the incoming wave and having different impedances to earth. The net result is an out-of-balance current round the loop which does not vary with the orientation.

Night effect is due to the fact that during darkness much of the received radiation does not arrive along the horizontal plane, but downwards from the sky. For a further discussion of these effects the reader should consult Admiralty Handbook of Wireless Telegraphy, Vol. II, (1938 edition), Sections T8–25. It will be enough for our purpose to observe that in order to eliminate vertical effect a D.F. loop usually consists of many turns of insulated wire, enclosed in an earthed metal tube, provided with an insulated section as shown in fig. 19.15.

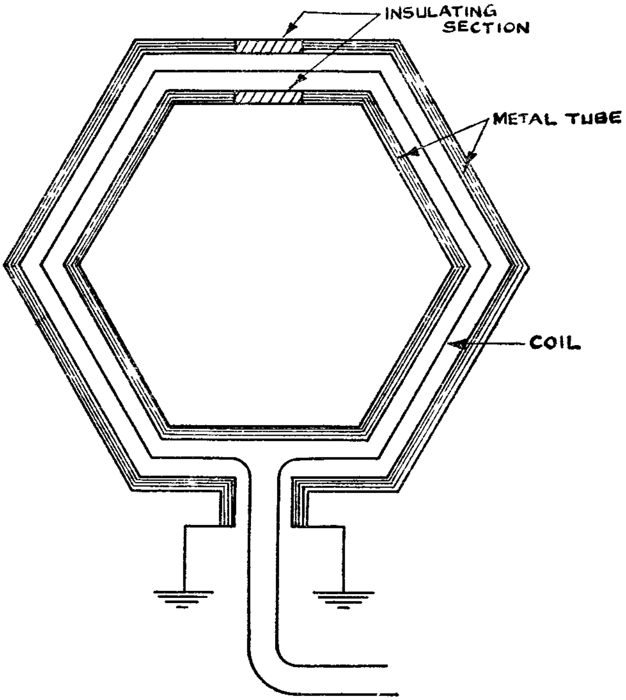


FIG. 19.15. D.F. Loop with screening

The loop is mounted on a vertical axis about which it can be rotated by means of a worm and wheel mechanism over a scale of degrees fixed to the body of the aircraft as shown in fig. 20.15.

The coil is provided with a pointer, which, as the former is turned, indicates on the scale of degrees the angle between the normal to the coil and the fore and aft axis of the aircraft.

To take a simple example: suppose that a certain transmitting station gives a minimum signal when the arrow points to 45°, then the transmitter is somewhere along the line AB [fig. 21.15 (a)] making 45° with the axis of the machine. The sense of the station can be found by taking a second bearing after the aircraft has flown for a few minutes on the same course.

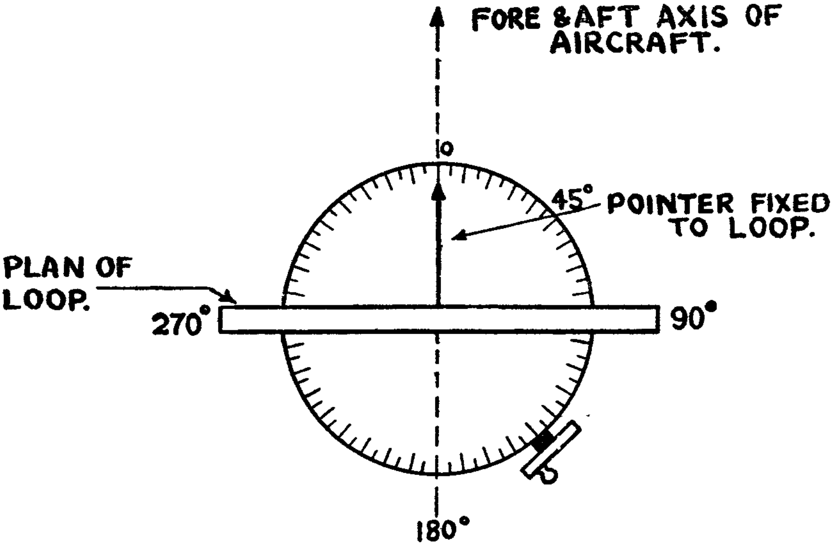


FIG. 20.15. Mounting of D.F. loop

Suppose this new bearing is 90° [fig. 21.15 (b)], then the transmitter must be along the line CD. It follows that it must be at the intersection T of the two bearings, and the 180° uncertainty is resolved.

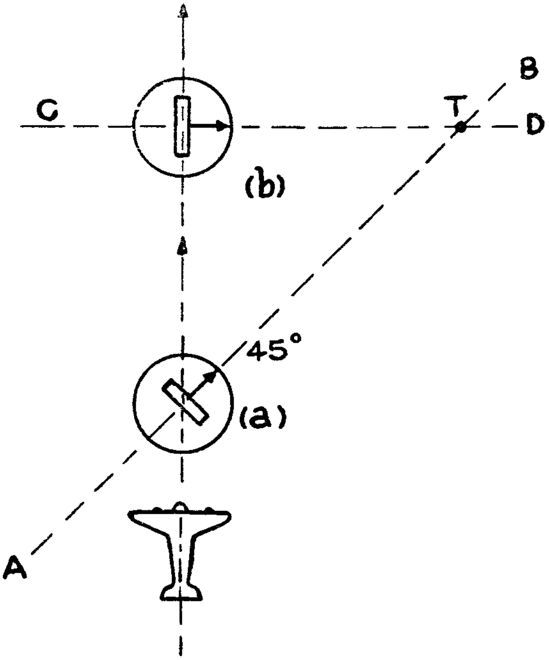


FIG. 21.15. Sense finding by second observation

Where such a procedure as that outlined above is impracticable, a sense finding unit must be employed. Consider the circuit shown in fig. 22.15.

A B is the loop aerial capable of rotation about its vertical axis: it is shown lined up on a distant transmitter, to give maximum response. C is an ordinary open aerial, and the resistance R is adjusted so that the signal in the telephones of the receiver unit is equal in strength to that due to the loop in its " maximum " position.

With both aerials receiving at the same time, the input to the receiver will be a combination of two equal alternating voltages, and its magnitude will therefore depend on their relative phases.

This raises difficult considerations beyond the scope of this book, so let us be content with the statement that in the circuit shown above, the two component voltages are either :—(i) in phase or (ii) 180° out of phase *according to the orientation of the loop*. Let us assume the former alternative : then the input signal is the *sum* of the two, and the response is a maximum. If, however, we turn the loop round so that the points A and B are interchanged, we shall reverse the phase of the input due to the loop and the total input will be the *difference* between the two components and the response will be a minimum. It should be zero if the preliminary adjustment of R has been carried out correctly. This allows us to discriminate between the positions of the coil, which, uncombined with the signal from the open aerial would have given identical maxima.

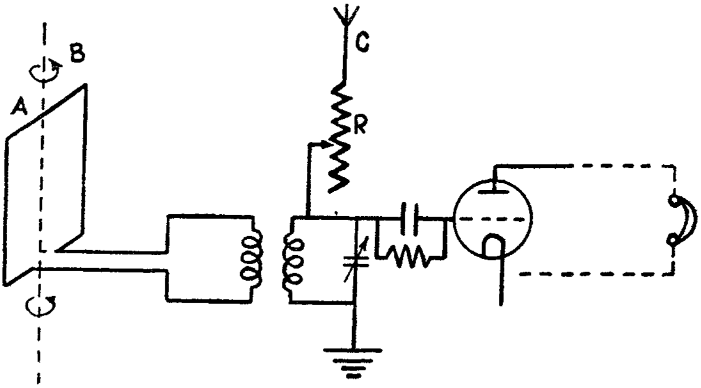


FIG. 22.15. Sense finding unit

To make sure of this discrimination, we now adjust the loop with respect to a transmitter whose position is known, to give minimum response in combination with the open aerial, and then inscribe an arrow on the casing of the loop, pointing to the transmitter.

In any subsequent adjustment this arrow will point in the direction of the unknown transmitter when, in combination with the open aerial, it is adjusted for minimum signals.

Notice that, when the bearing of a station is being found, we work with the loop alone and the adjustment for minimum signals causes the loop to face the transmitter [see fig. 23.15 (a)].

With the open aerial switched in we shall get a minimum with the coil " plain in line with " the transmitter (fig. 23.15 (b)).

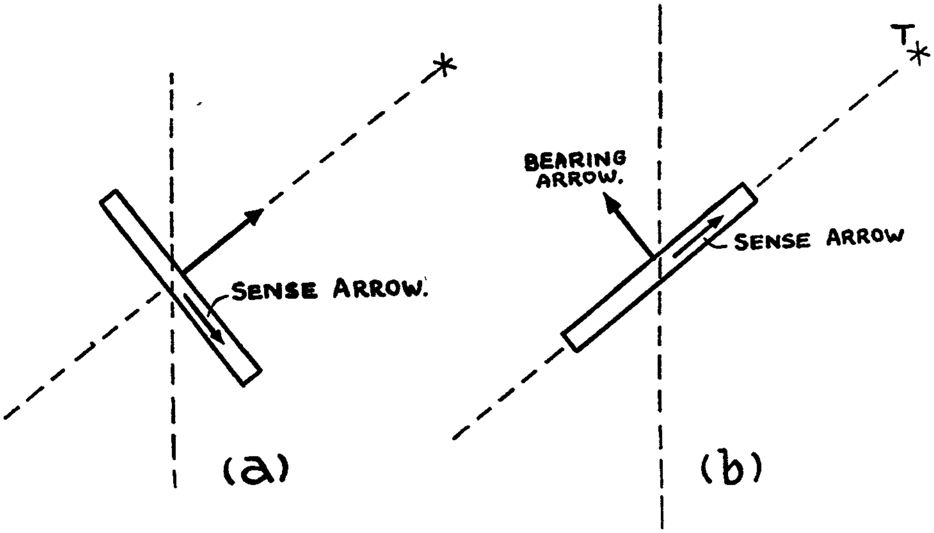


FIG. 23.15. Sense and bearing positions of a D.F. loop

It is quite possible to determine bearing and sense in one operation, with both aerials working: it is not desirable because the minimum is much less sharp than with the loop alone.

Another way in which an aircraft can obtain bearings, is to transmit signals whose direction can be determined by loops on the ground. From the aircraft operator's point of view this is a much simpler process. All he has to do is to call up a ground D.F. station and ask for a bearing. The ground station operator then adjusts a loop for zero signal on the radiation from the aircraft and then gets the bearing which is duly communicated. In practice a loop is not used on ground D.F. work, but a special kind of aerial called an " Adcock " aerial is employed.

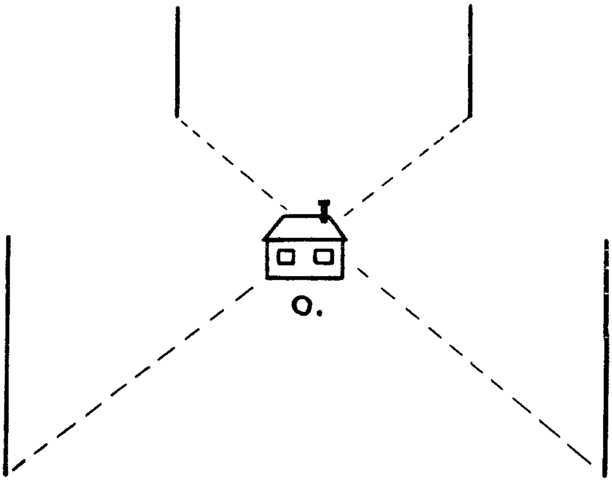


FIG. 24.15. Adcock system

Four vertical aerials are erected at the corners of a square (fig. 24.15) and the lower ends are connected by cables carefully screened in metal conduits, and buried in the earth, to a signal office O in the centre. In this office is an instrument called a Goniometer which, by employing the four vertical aerials at once, is able to indicate the direction from which a signal is arriving.

With three such installations as this, well separated, it is possible to tell from the signals sent out by an aircraft its location on the map. The importance of this in time of bad visibility cannot be exaggerated.

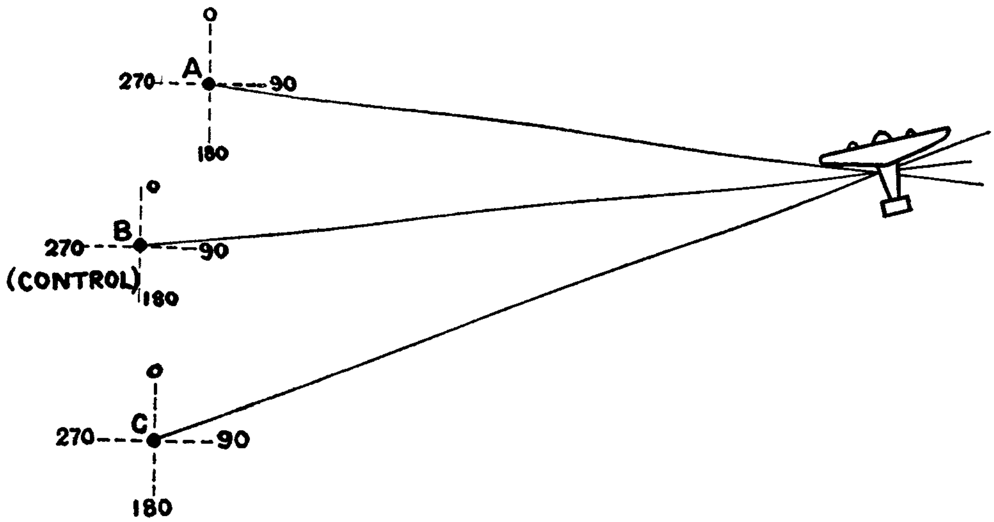


FIG. 25.15. Intersection by three bearings

A, B and C are the Adcock stations, B being the control station. A lost aircraft sends out a message for position, and all three stations take a bearing. The results are telephoned to the control station where they are plotted on a large-scale map, and the intersection of the rays shows the position of the aircraft. This position is communicated to the aircraft, whose pilot is then able to set his course accordingly.

The Adcock D.F. system is not subject to errors from Night Effect.

CHAPTER XVI

POWER SUPPLIES FOR RADIO TRANSMITTERS AND RECEIVERS

Radio equipment using thermionic valves requires the provision of three distinct sources of electric energy. These are respectively for filament heating, for grid bias and for high tension supply. In the early days of wireless it was the custom to employ batteries of dry cells or accumulators. These are readily available, completely portable and do not give rise to electrical interference problems. They are still extensively used to operate receivers, especially when a public supply is not available. Dry batteries may be used to supply H.T., grid bias and filament power, but when possible accumulators are used for low tension purposes because they give a relatively constant voltage during discharge and are easily re-charged when necessary. Grid bias batteries are not normally required to deliver any output power, and their life is thus the normal shelf-life of the type. They fail by drying-out of the paste electrolyte or by corrosion of the zinc element. Receivers or small transmitters requiring a H.T. consumption below 10 mA can be supplied economically from dry batteries, but above this figure H.T. accumulators of the lead-acid or nickel-alkaline type should be employed.

The lead-acid type has a higher voltage and lower resistance per cell, but needs more maintenance than the other type. Neither type offers the advantages of portability possessed by the dry battery. Receivers of higher power and most transmitters must be supplied from large and heavy accumulators or from public mains. Various methods of supply will be described in turn and their chief properties discussed.

In all cases where power is obtained from other sources than batteries it is necessary to smooth the output, that is to render it free from irregular variations or ripples which would affect adversely the performance of transmitters or receivers. Such ripples or irregularities are caused by the presence of alternating voltages in addition to the desired steady voltages, and if applied to a receiver will cause a loud hum in the telephones or loudspeaker. Similarly, the output of a transmitter will be modulated by the hum voltages, and its signals, even when received on a battery set, will be marred by a background of hum of sufficient level to cause annoyance. Fig. 1.16 shows how a voltage having a pronounced ripple can be split up into a steady and an alternating part.

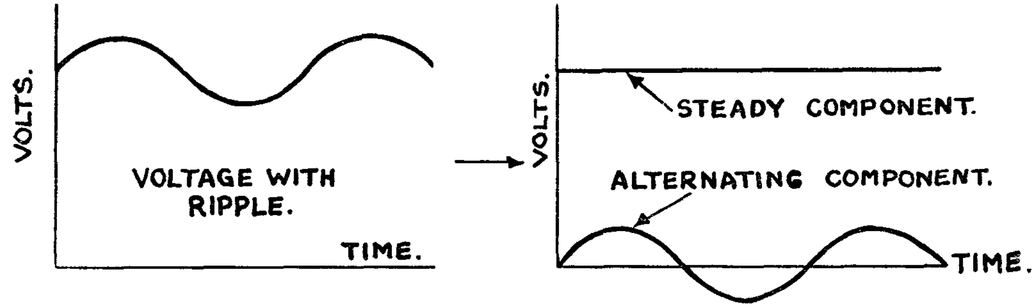


FIG. 1.16. Analysis of pulsating voltage

The possibility of separating voltages into steady and alternating components suggests a method of smoothing the supply, using the properties of inductances and condensers. The current from the generator or other source is passed to the receiver or transmitter through a large iron-cored inductance, which offers a large reactance to alternating currents but which passes freely the steady component. The terminals of the radio set are also shunted by a condenser of several microfarads capacity which serves to by-pass any remaining trace of alternating current. Such arrangements of inductances and condensers are known as filters or smoothing circuits and detailed descriptions of suitable types will be given later. Any power supply unit will therefore consist of two main parts, the supply unit proper and the filter circuit.

Motor-generator sets

One of the most obvious methods of securing power supplies for transmitters is by the use of motor generators or rotary converters. A motor is provided capable of working off any available supply. This is caused to drive a dynamo, which is wound so as to deliver the desired output. This output may be at a high voltage for H.T. purposes, or at a low voltage for filament heating or grid bias supply. In some cases the dynamo armature has two distinct windings brought out to separate commutators, one for H.T. and one for L.T. purposes. As an example, consider a transmitter which requires a high tension supply of 4,000V 250mA and a filament supply of 18V 10A and suppose that the supply mains available are 220V D.C. The machines required will be two generators, each equipped with its own driving motor, or a single motor with a coupling on each end of the armature shaft for driving the two generators. The advantages of the separate machines are (i) greater flexibility in that the speed, output and position of the machines are separately under control and (ii) greater reliability, e.g. if the L.T. machine breaks down the transmitter filaments may be run from accumulators. The efficiency is, however, less than if a single motor is used. The L.T. machine would be shunt wound, with its own field resistance for voltage regulation. The H.T. generator cannot be shunt wound because of the high voltage and is usually separately excited from the 220V mains and provided with a field regulator.

Motor generator sets are available in all sizes giving an output from 120V 20mA to 15,000V 6A, and have the advantage of great reliability and low internal resistance. This latter requirement is necessary so that the terminal voltage does not fall excessively as the load current is increased. We may sum up this property of the motor generator by saying that the voltage regulation is very good. The disadvantage of this form of supply is that skilled attendance is required and rotary machines need careful maintenance. Concrete foundations must be provided for large machines, and the problems of noise and vibration must be overcome. The insulation of high-voltage machines is difficult and the design of commutators and brushes is a complicated matter. In the event of a fault developing the short-circuit current is very large, due to the low internal resistance of the generator, and may do extensive damage to the transmitter.

The output from a D.C. machine can be used for filament heating without any smoothing, but the H.T. supply must be filtered to remove the voltage ripple. Reference to any text-book on dynamo design will show how this ripple originates, for each armature coil generates alternating voltages as it rotates and these are rectified by the action of the commutator. Another objection to the use of motor generators arises from the sparking which often occurs at the brushes. These sparks generate radio-frequency voltages which can cause interference in nearby receivers unless chokes and condensers are arranged to prevent the radiation of energy from the machine.

Utilisation of A.C. supplies

Now that the grid system is in operation in this country it is usual to find that a D.C. supply is not available. It is still possible to use motor generators, the motor being designed to operate from the A.C. supply, and this is the type of apparatus used to operate broadcast transmitters. Low-power transmitters and most broadcast receivers are now operated from the A.C. mains.

The great advantage of A.C. is that its voltage may be raised or lowered to any desired value by means of transformers of low cost and great reliability, which have no moving parts, and which need no maintenance.

The filaments of transmitting valves may be heated directly from 50-cycle mains without trouble from hum because the filament is so thick and heavy that it does not vary in temperature by an appreciable amount during the time of a single cycle, just as an electric soldering iron works equally well on A.C. or D.C. mains, and for the same reason.

The filaments of receiving valves cannot be heated from A.C. without causing hum. This led to the development of the indirectly heated cathode valve, which was described in Chapter IX. The filament of the output valve of a receiver may be operated on A.C. because there is no subsequent amplification. For the high-tension supply it is necessary to convert the A.C. to D.C. and this object may be achieved by many methods. All of these utilise the properties of rectifiers.

Properties of rectifiers

A rectifier may be loosely defined as a one-way conductor, i.e. it has a low resistance in one direction and a high resistance in the opposite direction. The diode valve is a practical example, but it is only one of a large number of devices which exhibit this property. Most of the others are only of scientific interest and have no great practical value at the present time, but several have been developed to a stage of great commercial importance. These include the copper-oxide rectifier, selenium rectifier, thermionic diode, mercury vapour diode, and the mercury arc. In addition, mechanical rectifiers or commutators are well known and take the form of rotary or vibratory switches either self-acting or motor-driven.

The copper-oxide rectifier

If a copper disc or plate is oxidised on one side by suitable heat treatment or chemical processing it is found to exhibit the one-way conductivity described above. It has a low resistance to current flow from oxide to metal, but its resistance is high in the metal to oxide direction. The maximum current that the rectifier will pass in the forward direction is settled by the surface area of the plate and oxide. The voltage which may be applied to a single element is rather low, and if increased is liable to cause breakdown of the layer of oxide. If a high voltage must be applied then a number of rectifying elements must be joined in series. To obtain a low resistance in the forward or conducting direction large plates must be used, and a good contact made to the oxide layer. This is often done by using plates of soft metal like lead between the rectifying elements. The rectifying effect is lost at high temperatures, and so the apparatus must be kept cool by the use of radiator fins or discs. The practical copper-oxide rectifier has taken two forms, the disc type and the plate type.

In the former the oxide coated elements are discs having a central hole. They are assembled on an insulating sleeve, with lead washers, metal spacers and cooling fins as shown in the diagram, the whole being bolted together on a long spindle passing inside the insulating sleeve. Spring washers are fitted to maintain a relatively constant pressure on the discs. Pressure is not necessary to effect rectification, but the good contact secured in this way reduces the rectifier resistance to a minimum.

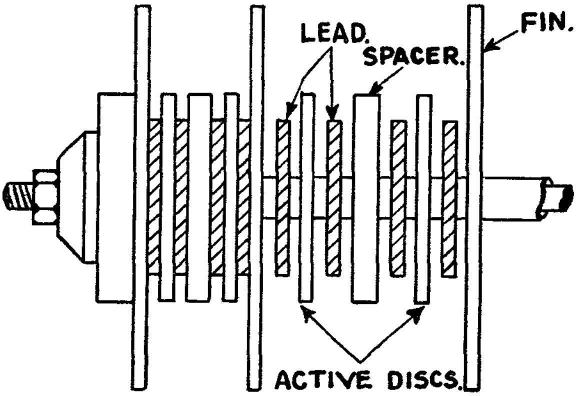


FIG. 2.16. Copper oxide rectifier

The current which may be passed through a rectifying element of the disc type is limited, and it is often necessary to connect rectifiers in parallel. This may be done for currents up to about 10A. Above this figure another form of construction is adopted, known as the plate type. Copper plates are oxidised on both surfaces and then sprayed with metal to make a good contact with the oxide layer. Plates treated in this way are provided with suitable connectors, varnished and bolted together in series or parallel batches according to the output required. Such assemblies are usually fan-cooled as normal fins cannot be provided.

The selenium rectifier

An alternative form of metal rectifier has been in use for some years on the continent and has now been introduced into England. In this case the rectifying properties of selenium are employed. Nickel-plated iron discs are coated with metallic selenium and subjected to heat treatment. As the selenium is a bad conductor it is sprayed with metal alloy to distribute the current over the surface area, and to make intimate contact with the rectifying element. There is a low resistance offered to current in the direction iron to selenium, but a high resistance in the opposite direction.

The rectifying discs are assembled on an insulated spindle, being separated by spring washers in contact with the metal layer with which the discs are faced. Both copper-oxide and selenium rectifiers appear to have an indefinitely long life, and only deteriorate by a very small amount with age. This takes the form of a slight increase in resistance in the conducting direction of the order of 5 per cent.

They are reliable, can stand heavy overloads for short periods and are not affected by vibration, dust, oil or chemical fumes except as regards corrosion of the metals employed in their construction.

In the present state of development the selenium rectifier shows some advantages over the copper-oxide type as regards weight and bulk. This is because each selenium element can withstand a greater reverse voltage than the corresponding copper-oxide element and also the permissible working temperature is higher. This statement may be put into the form that weight for weight the selenium rectifier shows the better voltage regulation characteristic. Both types are extensively employed in situations where extreme reliability and long life are essential, and where skilled maintenance is not required.

Thermionic diodes

Details have already been given of the construction and operation of diode valves. When used as H.T. rectifiers, the filament is usually heated from a low-tension transformer.

These rectifying valves are extensively used in broadcast receivers which require relatively low power for their operation. Standard sizes give respectively outputs of 250V 60mA, 350V 120mA and 500V 120mA. A wide range is also available, suitable for operating all types of transmitters up to the largest sizes.

These valves give the best performance when used to supply small currents at very high voltages. During that half-cycle of the applied alternating voltage in which the valve conducts, there is a relatively small volts drop between anode and cathode, but during the idle half-cycle a very large voltage is applied making the cathode positive with respect to the anode. High-vacuum valves are able to withstand enormous reverse voltages, and this is one of their great advantages. For this reason they can be used to give high voltage D.C. supplies for transmitters, X-ray tubes, cathode ray oscillographs and similar equipment. Their efficiency is not very high because of power lost in filament heating and in losses at the anode. This anode loss arises as a result of the internal resistance of the valve, due to the effect of the negative space charge between filament and anode. The losses due to this effect appear as heat at the anode and in large equipment water cooling is essential to prevent the destruction of the valves.

The filaments of high-power valves are always of pure tungsten, and the anodes of nickel or molybdenum. When water cooling is used, steel or copper anodes are employed. Valves giving up to 3,000V output have oxide coated or thoriated-tungsten filaments, which have a higher electron emission per watt of filament power.

Mercury vapour diodes

The main objection to the thermionic diode is its relatively high resistance, due chiefly to the space charge effect. Since this is a negative charge, it would appear possible to neutralise it by means of a properly placed positive charge. This has been done by the production of positive ions in the neighbourhood of the filament. When an atom of any substance has an outer electron removed it is left positively charged and is known as a positive ion. Atoms may be ionised in various ways, one of the most important being by collision with other atoms, or by the impact of high speed electrons.

If a gas or vapour is introduced into a diode valve, the region between filament and anode is filled with atoms. Electrons which leave the filament collide with these atoms on their way to the anode and provided the speed is high enough, the collision results in the ejection of an electron and the formation of a positive ion. The electrons continue on their way to the anode, the positive ion being attracted to the filament. Since current flows through a valve from anode to filament, the filament must be negative with respect to the anode, and this accounts for the attraction of the positive ions. The violence of the collision depends upon the speed of the primary electron, and this is decided by the anode voltage. The potential required to cause a collision of such violence as to eject a secondary electron is known as the ionisation potential, and for the substances employed in these valves is of the order of 20 volts. Higher potentials than these may eject two secondary electrons, giving a positive ion having twice the charge of a singly ionised atom. Mercury vapour is commonly employed as a source of ions, and is produced by introducing a few globules of mercury into the valve during manufacture.

The detailed action inside a mercury vapour diode is as follows :—When the filament is heated, electrons are emitted from it and start on their journey to the anode. In a short distance they have reached a speed high enough to eject an electron from the mercury atoms with which they soon collide. The electrons continue on their way to the anode, their speed reduced by the collision. The positive ions drift towards the filament. Now the ions are much more massive (several thousand times) than the electrons, and so move very slowly, unless a very high voltage is applied. The collection of positive ions near the filament neutralises the negative space charge of the electrons. This space charge in the high-vacuum diode prevents the easy transport of electrons from filament to anode by its repulsion of those newly emitted form the filament. The net result is that when the space-charge is neutralised, only a low voltage is required to draw electrons to the anode.

The volts-drop across the valve must not in fact exceed about 15 volts, or another effect comes into operation. Higher voltages cause double ionisation of the mercury atoms, and these ions, having twice the charge of the singly ionised atoms are attracted to the filament with twice the force. Since the ions are very heavy, the resulting bombardment of the filament causes the destruction of the emitting surface, particularly if this is formed by an oxide coating. Whatever the size of the mercury-vapour rectifier valve, the volts-drop across it should never exceed 15–20V, and thus the losses are very small and air cooling is sufficient even in the largest sizes.

As an example, suppose a valve of the type in question gives an H.T. output of 2,500V 0·8A and that its filament takes 5V 10A. Assuming the volts-drop across the valve is 15V the losses are—

(i) Due to filament, 5V 10A = 50 W.
(ii) Anode dissipation, 15V 0·8 A = 12 W.
The power output = 2,500V 0·8A = 2,000 W.
Total power = output + losses
= 2,000 + 62 = 2,062

$$\text{Efficiency} = \frac{2{,}000}{2{,}062} \times 100 = \underline{97 \text{ per cent. approx.}}$$

The efficiency of a high-vacuum diode of this size would not exceed about 60 per cent.

Attention should be drawn to several special characteristics of this type of rectifier, most of which arise from the vapour filling. In the first place, only low filament voltages are permissible or the electron emission is not uniformly distributed along the filament length, but most comes from the positive end. If a filament voltage in excess of 20V is employed then ionisation may set in and a discharge occur in the valve from the positive to the negative end of the filament. The pressure inside a vessel containing gas or vapour depends on the temperature. If the temperature is low the pressure is also low, and there are few atoms of mercury between the anode and filament. The space charge is then not capable of being neutralised and a high volts-drop occurs across the tube and the ions fall into the cathode with sufficient force to destroy its emitting surface.

If the valve is run too hot the gas pressure is high and the valve will only stand a limited voltage of polarity which makes the filament positive with respect to the anode.

It will be remembered that when used for rectification purposes a half-wave rectifier must stand a high reverse voltage during the non-conducting phase or half-cycle which in some cases is about three times the D.C. output voltage. Under these conditions, mercury vapour rectifiers are liable to break down and an arc discharge occurs inside the valve which rapidly destroys it. To ensure a long life for this type of rectifier a number of practical precautions must be observed.

After transit, or prolonged disuse, the mercury tends to collect in undesirable situations in the tube, and globules may even short-circuit the filament. To avoid this, and to secure a correct distribution of mercury, it is necessary to apply the full voltage to the filament for half an hour or so with no anode voltage. At all times it is essential that the filament voltage is kept to the rated level. Less damage is done by overheating than by under-running the filament. A cold filament implies a poor electron emission and consequently there is insufficient ionisation to keep the volts-drop down to a value below 20 volts. Consequently, there is greater risk of double-ionisation occurring and the filament is liable to suffer damage from bombardment.

Momentary short-circuits must be avoided for the same reason that they cause damage to the filament by a temporary increase in the anode-to-cathode potential difference. If maximum power is to be taken from a mercury-vapour rectifier, it is necessary to use a correctly designed filter which will be discussed later. Special rectifier circuits are often designed with the object of reducing the peak reverse voltage to which the valve is subjected.

In transmitter circuits it is essential to arrange some system of delayed switching whereby the filaments may be heated to the working temperature before the anode voltage is applied. This requires the use of a time switch or one which uses a thermal delay.

Practical rectifier circuits

Both thermionic diodes and metal rectifiers possess the property of one-way conductivity. The circuits of apparatus using the former always appear more complicated than metal rectifier circuits because of the provision which must be made for filament heating.

Fig. 3.16 shows the conventional representation of a metal rectifier. Beside it appears for comparison a diode valve. The arrow indicates the conducting direction in both cases, i.e. the direction in which current flows through the device.

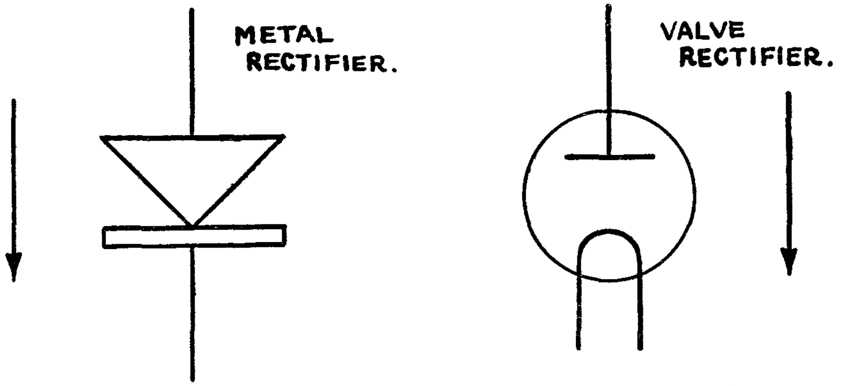


FIG. 3.16. Rectifier symbols

We shall first draw circuits using metal rectifiers. If we wish to substitute valves, then the triangle in the metal rectifier becomes the anode and the rectangle becomes the cathode or filament, and we must arrange for a power supply for filament heating.

Half-wave rectifiers

In this case a single rectifier is joined in series with the load, which in all subsequent circuits is shown as a resistance, but which in practice would be a receiver or transmitter. Current is passed in alternate half-cycles and suppressed in the others. The diagram of the wave-form of current and voltage obtained shows that there are idle periods in which no current flows in the load. For some purposes like battery charging this is of no consequence, but in others it is necessary to devise means whereby the current may be made to remain constant during the whole time.

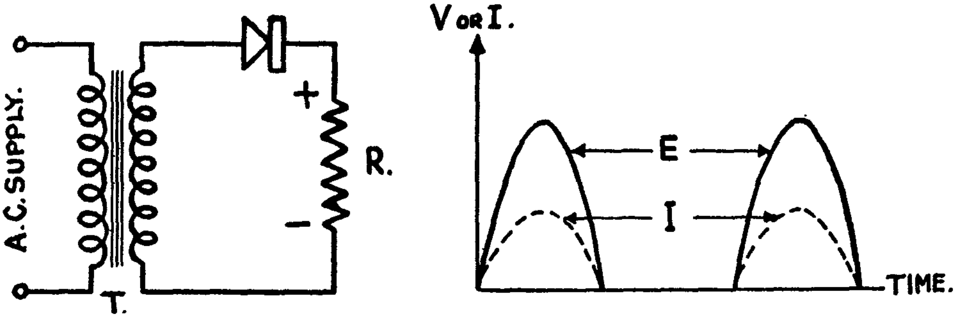


Fig. 4.16. Half-wave metal rectifier

Referring to the diagram, T is a transformer of which the primary is connected to the A.C. supply. The ratio of turns on the winding is arranged to give the desired output voltage. During that half-cycle of the secondary voltage in which the rectifier conducts, current flows through the rectifier, the load resistance and the transformer. The rectifier is non-conducting in the reverse direction, and so no power is delivered to the load. The load current is thus a succession of pulses having the shape of half-sine curves and similarly the load voltage has the same wave-shape as shown on the diagram.

The output voltage and current may be measured using moving coil meters. These of course cannot indicate the instantaneous values of current and voltage during the cycle, but they settle down and indicate the average values. For a pure resistance load the average is about 30 per cent. of the peak value of the applied alternating voltage.

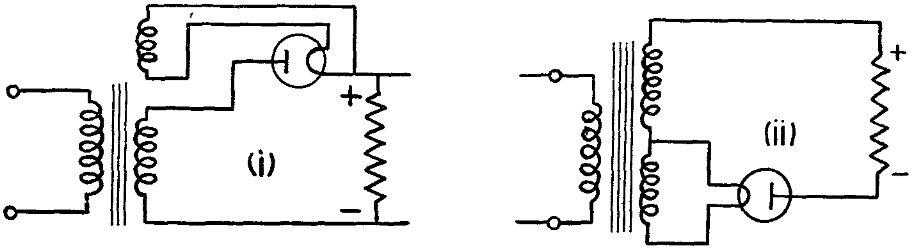


Fig. 5.16. Half-wave valve rectifiers

The same fundamental circuit may be used with a valve rectifier. The transformer is now provided with a low tension winding in addition to the main H.T. secondary. This is used instead of a battery, and the use of A.C. for filament heating is possible because the relatively thick filament maintains a constant temperature and thus gives ample electron emission during the whole cycle.

Fig. 5.16 shows the two most common circuits, the first being the exact equivalent of the metal-rectifier circuit. The second has some practical advantages. During normal operation, the voltage-drop across the valve is small compared with the voltage across the load. Generally the negative end of the supply is earthed, and so the potential difference between the valve filament and earth is very low. The insulation required is thus very slight. In the other case the rectifier filament is at a potential above earth of more than the total H.T. supply, and must therefore be exceedingly well insulated; this adds considerably to the cost of the transformer. In addition four wires must be taken from the transformer secondaries, as against three in the second diagram. The wave-shapes of current and voltage are exactly the same as in the case of the metal rectifier.

Full-wave rectifiers

The half-wave circuit has one obvious disadvantage in that there is no load current or voltage during alternate half-cycles, and these interruptions may not be desirable, especially in radio work where steady and well smoothed power supplies are required. The load current also flows through the secondary of the transformer and as it is always in one direction the core becomes magnetised, and may even be saturated. This increases the transformer losses and decreases the output, besides distorting the waveform of the supply voltage.

There is a further disadvantage of the half-wave supply which arises as follows. A pulsating current causes more heating than a steady current equal to the average of the varying current. This is because the heat generated is proportional to the square of the current. As an example, let us compare the heat produced by a steady current of 5A and then that produced by a current of 10A flowing in the same resistance for half the time. The figure representing the heat produced in the first case is 5^2 or 25, and in the second case is half of 10^2 or 50. The factor 1/2 is introduced because the current of 10A only flows for half the time. The heat produced by 10A is thus twice that produced by 5A, although the average values are equal in the two cases. The only way out of this difficulty is to spread the power more evenly over the whole cycle. Considerations like these led to the design of rectifier circuits which utilise both half-waves of the alternating input voltages.

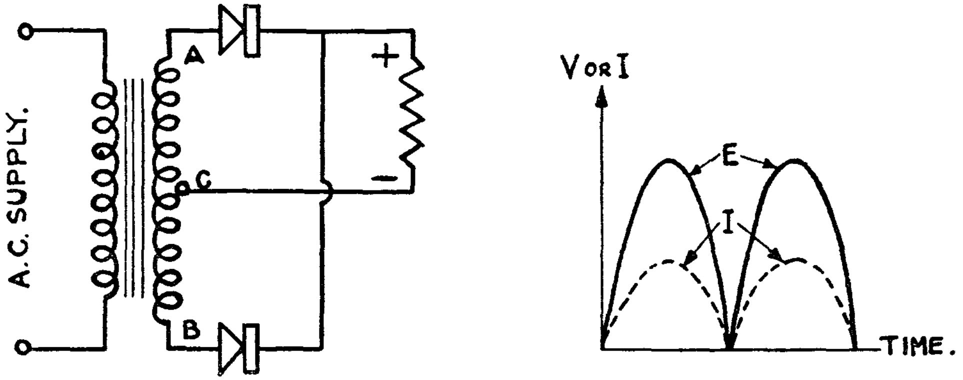


FIG. 6.16. Full-wave metal rectifier

Fig. 6.16 shows the simplest method of obtaining full-wave rectification, using metal rectifiers; it also shows the principle of a valve rectifier without the complication of filament wiring.

There are now two secondary windings joined in series, each connected to a half-wave rectifier. To understand the action it is best to cover up on the diagram one rectifier and half-secondary. Then, regarding the remainder as a half-wave rectifier, trace out the path taken by the current. Next, cover up the other half-secondary and trace the current

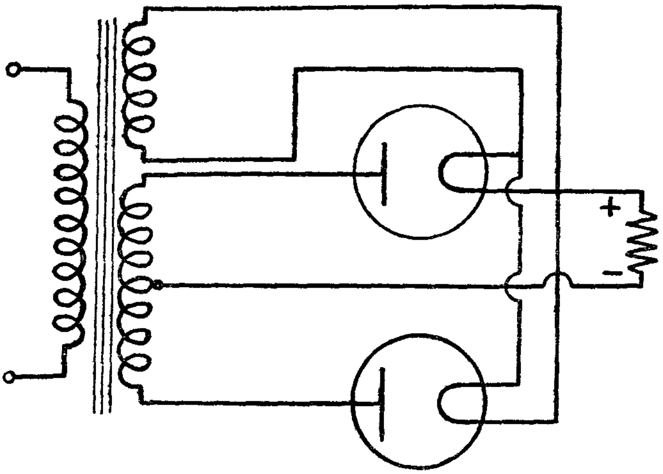


FIG. 7.16. Full-wave valve rectifier

from the first. Now imagine that voltages are measured from the centre-tap C to the ends of the secondary. During the half-cycle in which the end A is positive, and B negative, the upper rectifier conducts and the lower one is idle. The polarity is reversed in the next half-cycle and B is positive while A is negative. The lower rectifier now operates alone. The half-secondaries thus each give a half-wave output, but they are interlaced so that as regards the load current, full-wave rectification is secured, the wave-form being as shown on the diagram. As the rectifiers are in series across the secondary it might be thought that they would short-circuit the winding, but they are seen to be back-to-back or connected in opposition. In actual practice this circuit is seldom used with metal rectifiers, but is practically standard for valve-rectifiers when the circuit becomes as shown in fig. 7.16.

Instead of using two separate half-wave rectifiers, it is usual to mount them in a single envelope, especially in small sizes. The diagram fig. 8.16-shows a valve of this type. Note the conventional drawing of this circuit, which is the simplest method of representing the complete layout.

In this diagram a further refinement has been indicated. The rectifier filament-winding has been centre-tapped for the purpose of taking out the positive H.T. lead, instead of taking it, as before, from one side of the filament. This is so that the H.T. current divides equally down each filament limb instead of passing chiefly down that side from which it was led out in the first case. With modern valves the H.T. current is so large in proportion to the filament current that if it all passes down one side of the filament, the life of that half is reduced. This precaution is not necessary if rectifiers having an indirectly heated cathode are employed.

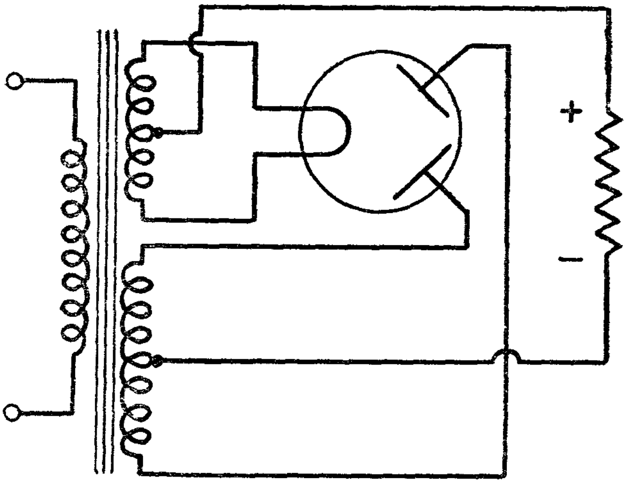


FIG. 8.16. Full-wave single-valve rectifier

The bridge-circuit rectifier

The full-wave rectifier just described has a number of disadvantages. First, the transformer secondary must be wound for a total voltage at least twice the maximum D.C. output voltage. This may be expensive and dangerous. Also, since the half-secondaries are only used alternately the heating effect is large compared with the heating which would be produced by the average output current, because of the half-sine waves of current followed by idle half-cycles. This means that the transformer material is not being used economically. These disadvantages are so great that the circuit of fig. 9.16 was devised to overcome them.

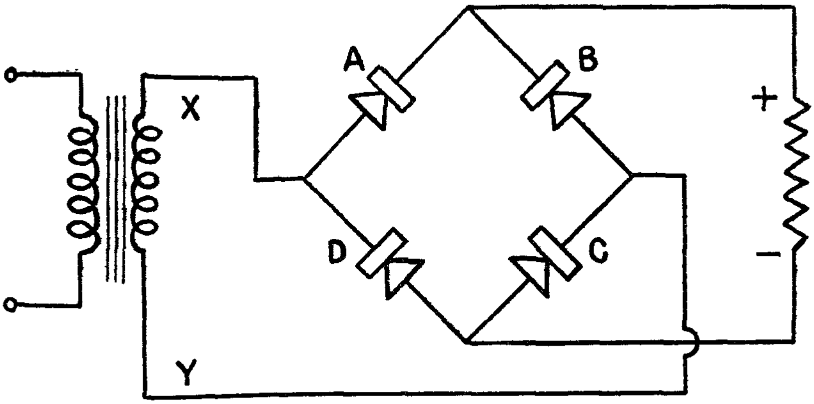


Fig. 9.16. Full-wave bridge rectifier

It employs four rectifier elements and during that half-cycle of the transformer secondary voltage which makes the terminal X positive current passes from X, through the rectifier A and the load, back to the transformer through the rectifier C, the elements B and D being out of action. During the next half-cycle, current passes from the terminal Y through B and the load back to the terminal X through the element D. This time the rectifiers A and C are idle. Practically the only disadvantage of this circuit is that it requires four rectifiers. This is of no consequence when metal rectifiers are used because they are built up of units, and it is merely necessary to connect these up in the proper order and bring out the required connecting leads. With valve rectifiers the situation is complicated by the provision which must be made for filament heating. If the circuit is examined, and an attempt made to substitute valve rectifiers, four tubes will be required. Of these, A and B can be supplied in parallel from a single winding, but C and D require separate windings, i.e. three filament windings in all. This makes the tranformer complicated and expensive, and for that reason the circuit is seldom used with valves, and in cases where it is so used the reason is that the rectifiers are only subjected to a low reverse voltage, and mercury vapour valves of high efficiency can be employed with safety. The outstanding advantages are that the peak currents are low in relation to the average D.C. output current compared with the corresponding figure for other rectifier circuits, and thus the heating is correspondingly low. Also the secondary voltage of the transformer is not much greater than the required rectified output. On both these counts the transformer is cheap, small and simple.

The voltage-doubler circuit

All the circuits previously described give a D.C. output voltage which is less than the

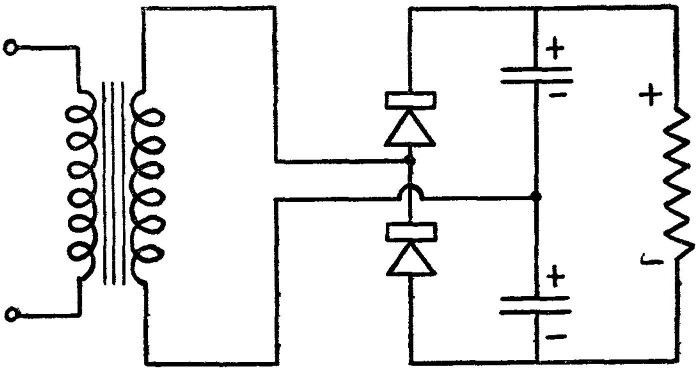


Fig. 10.16. Voltage doubler rectifier

peak value of the A.C. input. For some purposes where very high voltages are required, e.g. X-ray work, cable testing or for cathode-ray oscillographs, it is convenient to have available a method giving a greater output than the peak input. Fig. 10.16 shows one way of attaining this object.

The principle of this method is that we use rectifiers to charge up two condensers and then join these in series. Referring to the figure, and ignoring for a moment the lower rectifier, condenser and load resistance, it can be seen that the upper condenser will become charged with the polarity shown. Next consider the action of the lower condenser and rectifier. During the next half-cycle this condenser also becomes charged, but now the mode of connection of the rectifier is arranged so that the polarity of the condenser is the same as the upper one. The two are in fact now joined in series and the load resistance is connected across them both. The no-load voltage is twice the secondary peak alternating voltage, but on load this drops somewhat because the condensers become discharged. The use of this circuit is limited to low current outputs. Modifications have been made whereby the final output can be raised to practically any level, e.g. 700,000V by the extension of the doubling principle.

Although simplest with metal rectifiers the circuit is extensively used with valve rectifiers where very high voltages are required.

It must not be thought that the present list of rectifying methods and circuits is exhaustive, because only single-phase circuits have been mentioned, and if a three-phase supply of A.C. is available, there is practically no limit to the possible arrangements. These are chiefly of interest where high-power output and efficiency are of importance, the additional complexity and cost more than counterbalancing any gain in efficiency in the smaller sizes.

Vibratory rectifiers

Up to modern times practically all small portable wireless equipment has been operated from batteries. These are heavy, expensive and short-lived. Recently, efforts have been made to obtain all supplies from 6V or 12V accumulator, and to eliminate all other batteries. The accumulator can be recharged, and if the apparatus is mounted on a car or in an aircraft, the battery may " float " constantly across a charging dynamo.

The induction coil is a simple means of obtaining a high output voltage from a low-voltage D.C. supply. If the output is rectified and smoothed, the unit may form a replacement for the H.T. battery. Recent developments of this apparatus have led to the production of cheap and compact units capable of delivering 250V 60mA from a 6V battery.

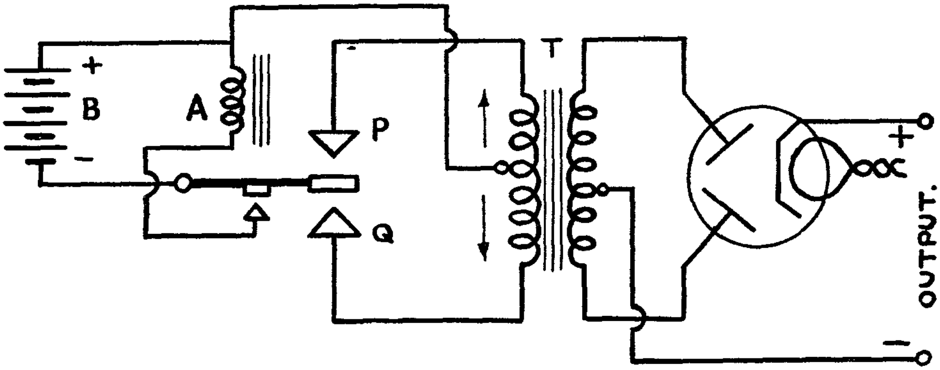


FIG. 11.16. Vibratory converter

Referring to fig. 11.16 a battery B energises a small electromagnet A in series with which is a contact breaker. The spring of this is extended and at the end carries two contacts. In operation these vibrate against two fixed contacts P and Q. When the moving contact touches P, current passes from the battery through the upper half of the transformer primary to the contact, along the spring, back to the battery. Later, when in contact with Q, current travels in the lower half-primary in the opposite direction. The effect is to induce an alternating voltage in the secondary which is rectified in the usual way using a valve or metal rectifier. If a valve is used it must be of the indirectly heated cathode type since its potential is that of the H.T. positive and the same battery may also heat the filaments of receiver valves which are at the potential of the negative H.T. supply. The use of a rectifier valve may be avoided by introducing a second pair of contacts which form a mechanical commutator, and which serve to reverse the secondary connections so as to obtain a D.C. output. Fig. 12.16 shows one way of attaining this object.

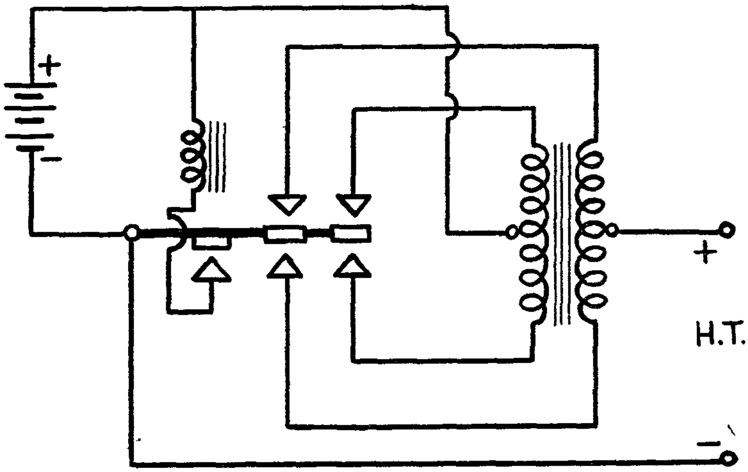


FIG. 12.16. Self-rectifying vibratory converter

Special precautions must be taken to suppress electrical interference caused by sparking at the contacts, and this is done by the use of chokes and condensers, and by good screening.

Filter circuits

For most radio purposes power supplies are required to be of constant voltage, whereas the output from common rectifier circuits is of a fluctuating or varying nature. Now any pulsating voltage of a periodic nature can be split up into a steady voltage together with a series of alternating voltages of different amplitudes and frequencies. The steady component is the only part in which we are interested, and a filter circuit is designed to pass this freely, and to suppress all other components, using the known properties of inductances and condensers. The two commonest circuits are shown in fig. 13.16.

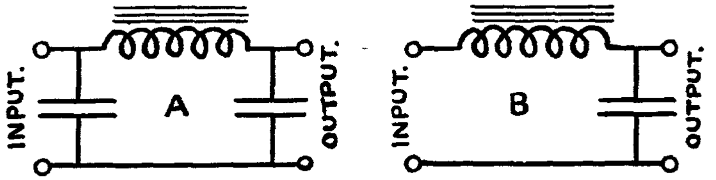


FIG. 13.16. Filter circuits

The input-condenser filter

The circuit A starts with a condenser, and is known as an input condenser filter, while the other (B) is known as a choke input filter. There appears only a slight difference, but the performance and operation are totally different. Consider first the input condenser type and imagine a half-sine wave of voltage applied to it. The first condenser charges up to the voltage of the supply, but later, when the applied voltage falls, it discharges into the load, and this discharge continues until the next half-cycle when it is once more charged up to the peak voltage of the supply. By reason of its action the first condenser was formerly known as a reservoir condenser. We may summarise the process by saying that at instants when the applied voltage from the rectifier is above the average, the condenser charges up, and discharges into the load when the rectifier voltage is below average, or when the rectifier is non-conducting. The choke tends to oppose any changes in the load current which passes through it, and any slight ripple present in the output is removed by the last condenser. Fig. 14.16 shows graphically the action of this type of smoothing circuit, (*a*) referring to a half-wave rectifier and (*b*) to a full-wave rectifier.

One very pronounced effect of the condenser must be noted. The average output voltage is much increased. Provided a large condenser is used, the output of a half-wave rectifier is increased three times, and that of a full-wave rectifier is increased by 50 per cent. compared with the average or mean value as indicated on moving coil meters with no condenser in circuit. The disadvantage of the input condenser is that it causes very large peak currents to be drawn from the rectifiers, which cause excessive heating, and if valve rectifiers are used, may spoil the filament emission. In consequence it is not used in high-power equipment. It is obvious that if the condenser remains charged nearly to the peak A.C. voltage, then the rectifier cannot possibly deliver current until its voltage exceeds the condenser back E.M.F. When therefore it does deliver current it is in the form of a brief pulse of high maximum value which causes excessive heating of transformer and rectifiers.

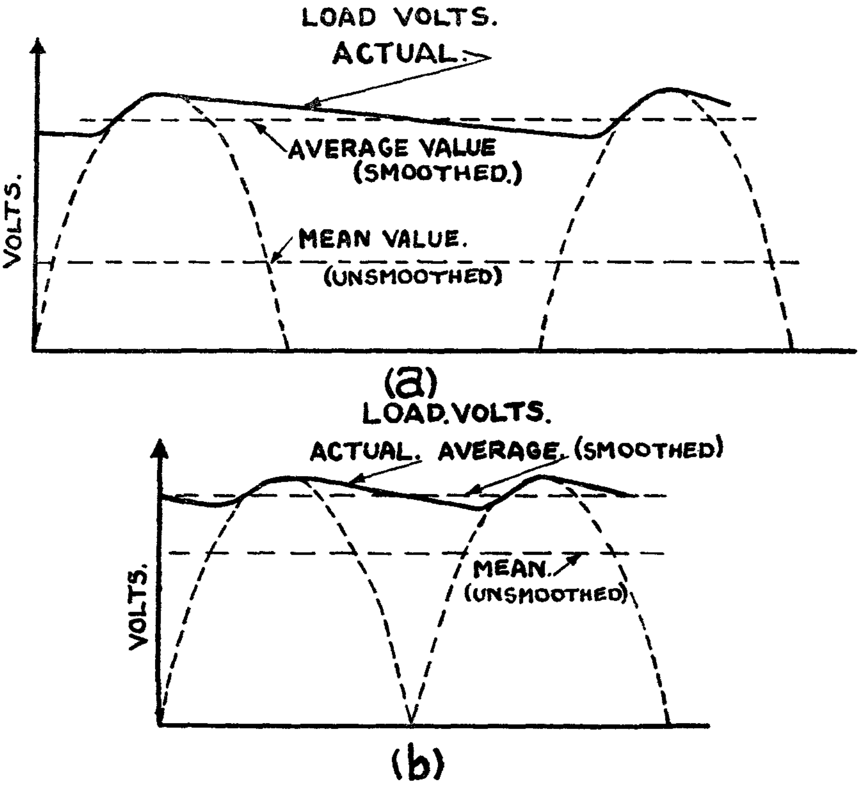


FIG. 14.16. Input condenser filter

The most important difference between the half and full-wave rectifier is that in the first case the fundamental frequency of the ripple voltage is the supply frequency, while in the second case it is twice this value, and therefore easier to smooth.

The choke-input filter

This type of filter is now used in all high-power rectifier equipment. When we attempt to alter the current in an inductance, it resists the change by generating a back E.M.F. If we apply to a choke a complex voltage consisting of a steady and various alternating components, the steady part will drive current through the choke and load resistance, but the alternating component will only drive a small alternating current round this circuit because of the high impedance presented by the series combination of choke and load resistance. The higher the frequency the more effectively the alternating components are suppressed.

If the wave form of the choke current is examined it will be found to pass in pulses each lasting half a cycle, and approximately rectangular in form. This is provided the inductance of the choke is large enough in comparison with the load resistance and this is one of the design requirements of a successful filter.

Fig. 15.16 shows graphically the wave form of the load-current and of the voltage applied to the input choke.

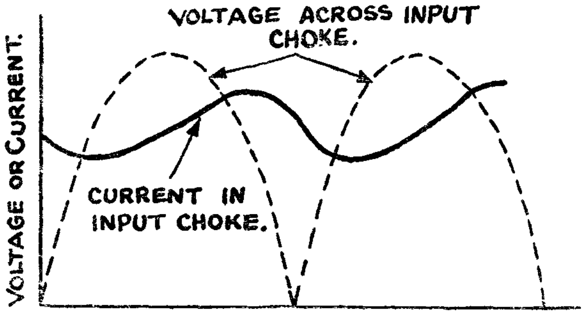


FIG. 15.16. Voltage and current relations

A disadvantage of this type of filter is that less smoothing is secured than in the condenser-input type, and it is usually necessary to follow it by a second choke and condenser. The action of the condenser following the choke has not been discussed, but as in other cases it serves to provide a low reactance path to earth for any ripple current which may pass the choke. Another disadvantage of the circuit is that it cannot be used with a half-wave rectifier since its action depends on the maintenance of a steady current from the rectifier, and in the half-wave case, this current is not available during the non-conducting phase.

A minor practical disadvantage is that because a very large alternating voltage is applied to the choke, a certain amount of mechanical vibration of the stalloy core laminations occurs. This leads to the production of a fairly loud hum of twice the supply frequency.

In spite of these disadvantages, it offers such advantages of efficiency, good voltage regulation and high output that it is universally employed in high power sets.

A difficulty common to all supply units is that of voltage-drop on load. If any source or generator has internal resistance, then there must be an internal voltage drop when current is supplied to an external circuit. Now rectifiers of the high-vacuum valve or metal type have a fairly high resistance, and transformers and chokes also have considerable resistance, so that these supply units show a volts drop on full load of between 15 per cent. and 50 per cent. This indicates poor efficiency. Mercury vapour valves give a much better performance, often more like 2 per cent. or 3 per cent. The rectifier using the condenser input filter has a very poor regulation, since its very action depends on the discharge of a condenser, the voltage of which must naturally fall in the process.

The voltage doubler circuit is exceptionally bad in this respect, although in high-voltage testing the effect is rather valuable, for the damage done by an accidental short circuit is not serious because the large fall in voltage limits the short-circuit current.

Nothing has been said as to the sizes of chokes and condensers necessary to obtain adequate filtering. The chokes are wound on laminated stalloy cores, and may have an inductance between 5 and 40 henries. The condensers are of the waxed-paper or oil-filled paper type and range in capacity between 1 and 4 μF. The higher the voltage and power output, the smaller the components used. Often in receivers, electrolytic condensers are employed. These are cheap and obtainable in large capacities up to about 32 μF for 400 volt working, but cannot be made for high voltage equipment unless several elements are connected in series.

PRINTED AND PUBLISHED BY HIS MAJESTY'S STATIONERY OFFICE
To be purchased from
York House, Kingsway, LONDON, W.C.2 429 Oxford Street, LONDON, W.1
P.O. Box 569, LONDON, S.E.1
13a Castle Street, EDINBURGH, 2 1 St. Andrew's Crescent, CARDIFF
39 King Street, MANCHESTER, 2 Tower Lane, BRISTOL, 1
2 Edmund Street, BIRMINGHAM, 3 80 Chichester Street, BELFAST
or from any Bookseller
1940
(Reprinted 1951)
Price 3*s* 6*d* net
PRINTED IN GREAT BRITAIN

(98566) Wt. 2185 K.76 6/51 Hw. (T.S.22130)
S.O. Code No. 22-154

E.M.I. INSTITUTES

The only College which is part of a great Industry

- ★ **Three Year Course on Telecommunications Engineering** (including one year's works training in E.M.I. Factories)—leading to C & G Full Technological Certificate.
- ★ **One Year Course** in the Principles and Practice of Radio and Television Servicing.
- ★ **Three Week Practical Courses** in Radio and Television.

Also Home Study Courses in all Branches of Radio and Television, Electrical, Production, and Aeronautical Engineering.

- ★ **Write for FREE BROCHURE giving full details of courses**

E.M.I. INSTITUTES
Dept. 116
10, PEMBRIDGE SQUARE, NOTTING HILL GATE,
LONDON, W.2. Telephone: BAYSWATER 5131/2

I. 31

METEOROLOGICAL MAGAZINE

The Magazine contains authoritative articles on a wide variety of topics, including the first accounts of much recent research of great importance, international collaboration in meteorology and aviation, reports of scientific meetings, notices and reviews of new books, descriptions of new instruments, notes on interesting weather and optical phenomena in this country and abroad, letters from correspondents. Figures are also given every month, summarising the weather experienced at many places in the British Commonwealth

PUBLISHED BY HIS MAJESTY'S STATIONERY OFFICE

YORK HOUSE, KINGSWAY, LONDON, W.C.2; 429 OXFORD STREET, LONDON, W.1 (POST ORDERS: P.O. BOX 569, LONDON, S.E.1); 13a CASTLE STREET, EDINBURGH, 2; 39 KING STREET, MANCHESTER, 2; 2 EDMUND STREET, BIRMINGHAM, 3; 1 ST. ANDREW'S CRESCENT, CARDIFF; TOWER LANE, BRISTOL, 1; 80 CHICHESTER STREET, BELFAST
or through any bookseller.

Price **1***s.* **6***d.* *By post* **1***s.* **7***d.* *Annual Subscription* **19***s. including postage*

37.25.