ELEMENTS OF ELECTRONICS

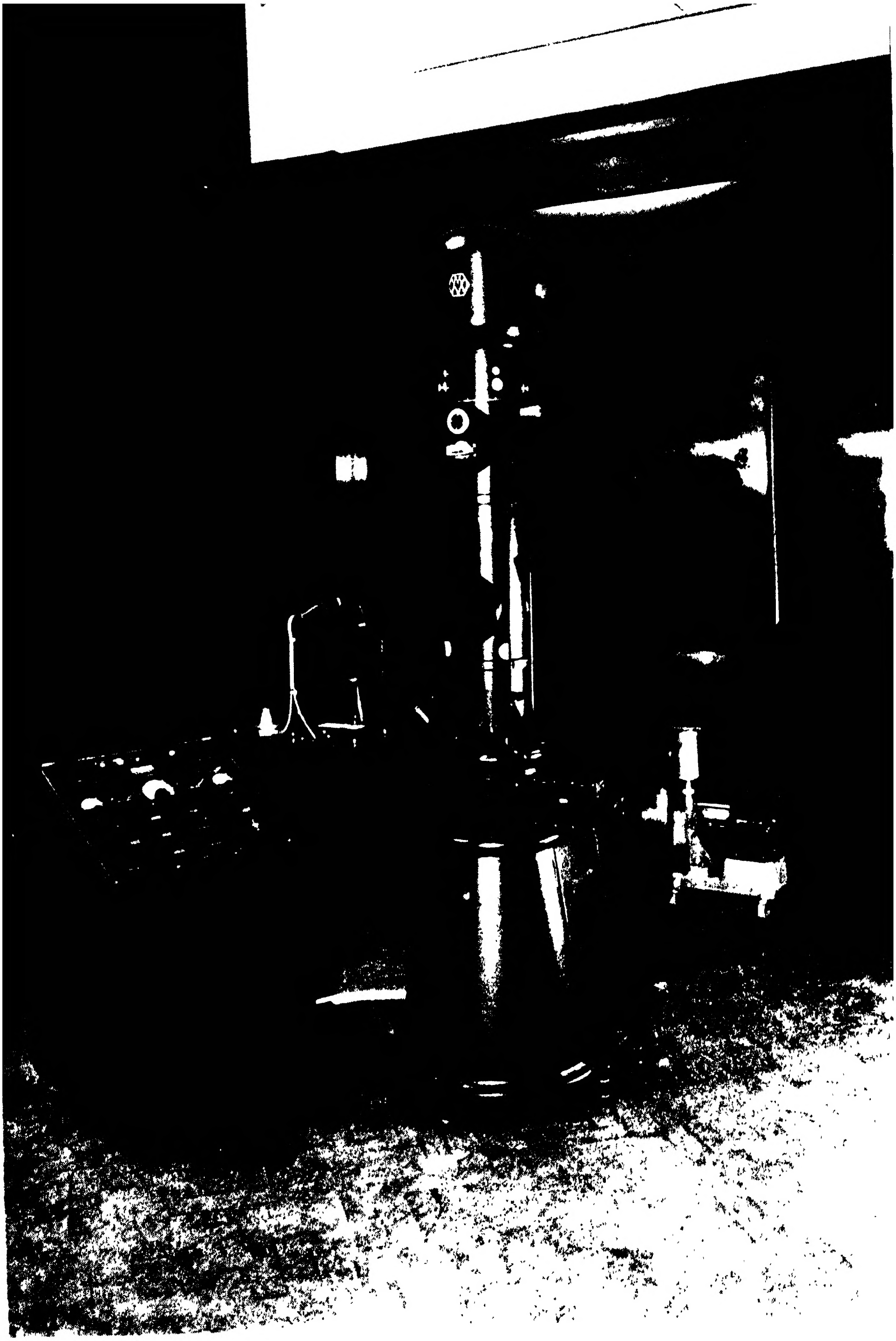


PLATE I

50 KV. Electron Microscope made in Britain by the Metropolitan-Vickers Company. The magnified image on the fluorescent screen is viewed through the hoods above the base of the instrument. At the rear is the vacuum pump equipment, and the control

ELEMENTS OF
ELECTRONICS

by

G. WINDRED
A.M.I.E.E.

LONDON
CHAPMAN & HALL LTD.
37 ESSEX STREET W.C.2
1949

First published 1949

BOOK
PRODUCTION
WAR ECONOMY
STANDARD

~~THIS BOOK IS PRODUCED~~
~~IN COMPLETE CONFORMITY~~
~~WITH THE AUTHORIZED~~
~~ECONOMY STANDARDS~~

Catalogue No. 384/4
Printed in Great Britain by W. & J. Mackay & Co. Ltd., Chatham
Bound by G. & J. Kitcat Ltd., London
Flexiback Binding

PUBLISHER'S PREFACE

THE science of Electronics, although well established, is still obscure to many who are fully qualified in other branches of physics and engineering. The increasing complexity of the whole subject of electrical engineering makes it difficult to include an adequate survey of electronics in the curriculum of university and technical college courses, and graduates are often introduced to electronic apparatus without a study of the background against which it has been developed.

In this book, Mr. Windred, himself an electrical and electronic engineer of many years' experience, has undertaken to provide the background by surveying the growth of electronics from the original discovery of the electron to the modern applications in radar and electro-medicine. The text, while descriptive more than mathematical, assumes a knowledge of physics and engineering, but the research worker in fields other than these will find the information concise and easily assimilated.

It is greatly to be regretted that Mr. Windred's untimely death prevented his final revision of the MS. and the reading of the proofs. This task was undertaken by Mr. G. Parr to whom we extend our grateful thanks.

CHAPMAN & HALL LTD.

CONTENTS

LIST OF PLATES

CHAPTER I

THE DOMAIN OF ELECTRONICS

The science of electronics may be defined as dealing with the production, control and utilisation of electrons for the attainment of particular results. Strictly speaking, we should recognise two distinct branches of the subject; one dealing with the physical phenomena attaching to the movements and interactions of electrons, which we may call pure electronics or electron physics, and the other dealing with the applications of these principles. This second branch is properly referred to as applied electronics or electronic engineering, and has its foundation in electron physics as represented by established experimental and theoretical work in this field.

And what exactly is this electron, with which we are so much concerned; what does it look like, what are its properties and how is it recognised? Science can give only a partial answer to these questions. It cannot tell us exactly what an electron is or what it is like, for the good reason that we have no means delicate enough for observing an electron directly; we can only observe the effects it produces under given conditions and infer from this its general properties. Much of our information has been derived from the study of large groups or streams of electrons; from electrons in bulk, as it were, and has thus applied to the behaviour of crowds rather than the character of individuals. A study of the effects due to individual electrons has, however, been undertaken by advanced research in electron physics, and from this work much has been learned about the electron itself, as we shall see later.

It may be pointed out that electronics is not so much a new subject as a new way of looking at electrical phenomena, all of which may be explained by the principles of electronics. All electrical effects are basically electronic because all electric currents are due to the movement of electrons and all electric charges are due to the accumulation of electrons. In the case of electrodynamics or current electricity we deal with electrons in motion; in the case of electrostatics with electrons at rest.

These are the two branches into which the broad subject of electricity is usually divided.

How then is it that the term electronics has come to be applied to a particular group or range of electrical phenomena, and what exactly does it comprise ? A close definition is impossible, but it may be said that in the generally accepted sense electronics is concerned with electrons which are *free*, in so far as they are not limited to the confines of solid conductors. The chief concern of electronics is with the passage of electrons through evacuated or gasfilled spaces as contained within the familiar electron tubes of electrical engineering practice, of which the radio vacuum valve is the outstanding example. We shall see later, incidentally, that the passage of electricity through gases involves other particles, known as ions, as well as electrons.

Applications of Electronics

The range of uses to which electronic apparatus has been put during recent years gives the most striking evidence of the utility of such methods in relation to ordinary electrical engineering practice. In many cases, such as the familiar example of television, electronic methods have created entirely new industries and provided means for achieving results which would have been unattainable in any other way. The photoelectric cell, for example, in its various forms has provided industry with a means of measurement and control sensitive to changes of light and offering untold possibilities in practice. In this case, as in practically all other applications of electronics, the thermionic valve, originally developed purely for radio purposes, forms an indispensable supplement for attaining practical results ; for increasing the inherently small effects produced by devices such as photocell to the level required for satisfactory industrial applications. Some idea of the immense range of electronic apparatus may be gained by contrasting, at one end of the scale, the modest photocell, about the size of a radio valve and producing a few millionths of an ampere of current, and at the other extreme the large power rectifier in its steel tank standing perhaps ten feet high and dealing with several thousands of amperes for driving electric trains or providing the power for heavy industries.

Somewhat outside the industrial sphere there is another contrasting development ; the electron microscope, with its immense magnification, opening up a prospect of great promise in the study of particles, structures and organisms of a smallness far beyond the range of the most powerful optical microscopes. In this field alone electronics offers perhaps the greatest of all services to man as a weapon for studying his age-old enemies to health in the sub-microscopic world of the virus and bacillus. Great progress has also been made recently in the development of electronic instruments for the study of brain disorders and for the treatment of diseased tissues.

Another great aid to medical progress, the X-ray tube, comes from a much earlier period of electronic development and is now taken very much for granted, although it would be impossible to assess the direct benefit to humanity which has resulted from its use in assisting the diagnosis of ailments and as an aid to surgery. During recent years X-rays have also been widely used for studying the internal structure and detecting hidden faults and flaws in all kinds of materials. By employing new and special electronic methods the research physicist has been able to gain a remarkable insight into the ultimate structure of matter and even to achieve in part the alchemist's dream of transmutation.

All this, and very much more, lies to the credit of electronic development during the relatively short time of its existence and may serve to give an idea of the contribution which electronics has made in practically every field of technical activity and endeavour. The future possibilities are quite untold, but obviously very great.

Advantages of Electronics

It is interesting to consider some of the reasons for the rapid development and progress of electronics. At the outset it may be pointed out that the number of cases where it had to compete with established methods is relatively small, and that the great majority of its important applications are new in the sense that they are feasible only with the aid of electronics ; the results could not have been achieved in any other way.

The unique advantages of electronic apparatus have two main sources ; the possibility of controlling large amounts of power with an almost negligible expenditure of energy, and the extreme rapidity of this control. By suitable amplifying arrangements employing the thermionic valve in its various forms it is readily possible to control currents many millions of times greater than the initial value required to effect the control. There is, moreover, no appreciable time-lag ; in spite of the numerous changes and magnifications to which the original energy may be subjected in the course of amplification the response is to all ordinary intents and purposes instantaneous. Secondary, but none the less important, considerations are that electronic apparatus is clean, compact, highly efficient and silent in operation.

Effect on Electrical Theory

The study of electronic principles has given a new insight into basic theories of electricity and cleared away much of the mystery which previously centred around fundamental ideas. The whole imposing edifice of electrical theory, built upon a mass of observations and experiments extending over more than a century, had hitherto rested, strangely enough, on altogether vague and unsatisfactory notions of the nature of electricity itself. In the early days electricity was regarded as having the nature of a subtle fluid, an idea which could not long be maintained without repeated modifications in the face of continual new discoveries.

These early ideas were gradually abandoned, and electrical theory developed for some time independently of any clear idea as to the nature of the effects with which it dealt. With the advent of the experimental research giving rise to the idea of the electron new light was cast upon many problems, and eventually it became possible, among other things, to give a satisfying if not complete answer to the question : what is electricity? Great progress was also made in our knowledge of the relationships between electricity, chemistry and physics, notably in connection with the constitution of matter.

Consideration of this experimental work, comprising what we may call the story of the electron, forms perhaps the best introduction to the real foundations of electronics.

FOR FURTHER READING

The following books deal with the general subject of electronics from an elementary standpoint.

1. D. Grimes. *Meet the Electron.* (New York, Pitman), 1944. A very elementary and non-mathematical introduction for

the general reader, outlining the properties of electrons and the chief applications of electronics.

2. J. Mills. *Electronics Today and Tomorrow.* (New York, Van Nostrand), 1944. Described by the author as " a quick introduction to the new things electronics is producing." Entirely descriptive, excluding all unnecessary details.
3. D. P. Caverly. *A Primer of Electronics.* (New York, McGraw-Hill), 1943. Seeks to explain the basic principles to those associated with electronics but lacking the training to follow technical details. Text covers elements of physics and electricity and magnetism as well as electronics.

CHAPTER II

THE STORY OF THE ELECTRON

The idea that an electric current actually consists of a stream of very small particles of some kind, each carrying an electric charge, had occurred to several scientists long before there was any thought of trying to detect such particles or measure their charge. The chief source of the idea was Faraday's laws of electrolysis, which showed that when a current was passed through a liquid solution, called the electrolyte, the particular salts in the solution are split up into their constituents and deposited at the respective electrodes through which the current is conveyed. The conditions are shown in Fig. 1. If, for example, the electrolyte consists of a solution of potassium chloride in water, the passage of current from the anode (positive) to the cathode (negative) will result in the deposition of chlorine at the anode and potassium at the cathode. The material is conveyed to the respective electrodes by the movement of negatively charged chlorine atoms towards the anode and positively charged potassium atoms towards the cathode. Similar results are obtained with other solutions, and it is found that the

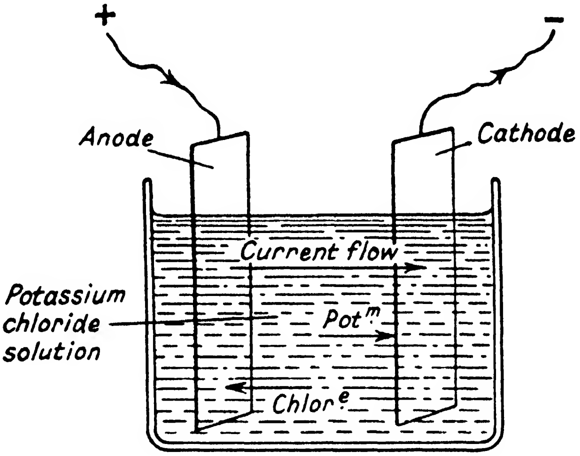


FIG. 1.—Separation and depositing of salts by current passing through chemical solution.

amount of material transferred by these charged atoms, or *ions* as they are called, is always proportional to the quantity of electricity passed through the electrolyte, that is, to the strength of the current multiplied by the time of its flow while producing the deposit. There is thus strong evidence for believing that the current itself is conveyed by the moving ions, each of which is the means of transporting from one electrode to the other a small quantity of electricity, representing its particular share of the total amount. The greater the strength of the current, the greater the number of ions in transit and the greater the quantity of electricity conveyed in a given time.

On this basis we may picture the flow of current through an electrolyte as being due to the movement from one electrode to the other of a swarm of ions, each carrying in some way a small but definite electric charge or quantity of electricity. The movement of the swarm commences as soon as voltage is applied between the electrodes, each ion delivering up its small contribution to the total current upon arrival at the electrode. When the voltage is removed the motion of these carriers ceases and the current flow is ended. For the present purpose we do not need to complicate this simple picture by considering the chemical processes involved in the action.

From experiments in electrolysis it was possible to determine quite accurately the relationship between E, the electric charge carried by each ion, and M, the mass of each ion. The ratio E/M was always the same for a given ion, that is, for a particular element deposited by the electrolytic action, and had its highest value in the case of hydrogen ions. This means that hydrogen atoms or ions in electrolysis carry a greater electric charge *in relation to their mass* than the atoms of any other elements, but it does not follow that the charges they carry are actually greater than for other ions. The same result would be observed if the hydrogen atoms were proportionately lighter ; the charges being the same, and this is now known to be the case.

Cathode Ray Experiments

In 1897 the late Sir Joseph J. Thomson carried out a series of experiments on cathode rays which were highly instructive in showing the nature of these rays and afforded interesting comparisons with the nature of electricity as revealed by the earlier experiments on electrolysis.

The name cathode rays was given to the rays which start from the cathode, or negative electrode, in a highly exhausted tube when a suitable voltage is applied to the electrodes. Vivid phosphorescence is produced by the rays when they strike the glass walls of the tube, and it was found that their path was deflected by the action of magnetic and electric fields. When no such fields are present the rays travel in straight lines from the cathode. Most of the facts about cathode rays were established by Sir William Crookes.

The form of cathode ray tube used by Thomson in his experiments is shown in Fig. 2. The cathode *C* takes the form of a small metal disk carried on a connecting stem sealed into the end of the glass tube. In the narrow neck of the tube immediately opposite the cathode are two metal plugs *A* and *B*, of which the first forms the anode. Both plugs have a small central aperture to allow the rays to pass from the cathode into the larger end of the tube in the form of a thin straight beam. The point *P* where the beam strikes the bulbous end of the tube is brightly luminous and its position is located by means of the scale *S*. The beam passes between the flat parallel metal plates *D* and *E*, to which a voltage may be applied when desired by means of the terminals carried on their extensions through the walls of the tube. The electric field produced in this way between plates *D* and *E* causes a deflection of the beam upwards or downwards, according to the respective polarities of the plates, the beam being always attracted towards the positive plate. From the laws of electrostatics, which state that like charges repel each other and unlike charges attract, it thus appears that the rays have a negative charge.

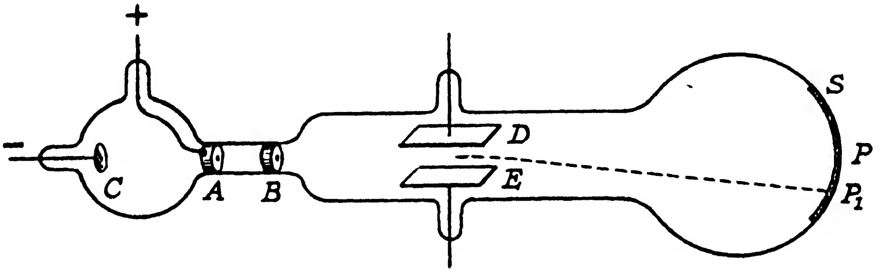


FIG. 2.—Sir J. J. Thomson's cathode ray tube of 1897.

The amount of deflection depends upon the voltage applied to the plates, and is measured by the distance traversed by the light spot on the scale *S*. If, for example, the beam is deflected into the position shown by the dotted line in Fig. 2, then the distance, or number of scale divisions, between P and P_1 is a measure of the deflection. The effects of a magnetic field can be observed in a similar way by arranging a magnet with its poles on either side of the neck of the tube so that the lines of force are parallel with the plates *D* and *E*. Since the deflection of the beam is always at right angles to the direction of the lines of magnetic force, a magnetic field arranged in this way gives deflections of the beam in the same plane as those produced by the electric field, the direction of deflection being dependent upon the direction of the magnetic field.

By employing an electromagnet consisting of a pair of coils carrying current, one on each side of the tube adjacent to the deflecting plates, Thomson applied a magnetic field of known strength as determined by the current flowing and the number of turns in the coils, and from the observed deflection of the beam worked out mathematically the relationship between :

(*a*) the charge and mass of each particle comprising the beam, this ratio being denoted by e/m, and
(*b*) the velocity of the particles, denoted by v.

He then applied a suitable voltage to plates *D* and *E*, arranging the polarity so that the deflecting action of the electric field was opposed to that of the magnetic field. The respective field strengths were then adjusted until the opposing forces on the beam just balanced each other, i.e., so that the luminous spot returned to position *P* corresponding to its undisturbed state. It can be shown mathematically that under this condition the velocity of the particles forming the beam is equal simply to the ratio of the electric and magnetic field strengths. As these field strengths could be accurately estimated, so also could the beam velocity v, and this in turn allowed the calculation of the ratio of charge to mass e/m of the particles from the relationships (*a*) and (*b*) above already established when using the magnetic field alone. As we shall see, this ratio e/m, the ratio of electric charge to mass of the cathode ray particles, became an important factor in tracking down the electron itself, though at the time of Thomson's experiments there was not sufficient evidence for regarding the cathode rays simply as a stream of electrons.

It was observed at the time that the value of e/m as determined by the cathode ray experiments was something like 2,000 times as large as the corresponding ratio E/M for the hydrogen ion in electrolysis. This meant that if the respective masses of the hydrogen ion and the cathode ray particle were at all similar, then the charge carried by the latter must be some 2,000 times the value of the charge carried by the hydrogen ion in electrolysis. Alternatively, if the respective charges were similar, then the hydrogen ion must have a mass about 2,000 times as great as the mass of a cathode ray particle. Recent research has established this latter alternative the actual mass of the hydrogen atom being 1,834 times that of the cathode ray particle, now identified as an electron.

This was the extent of the important step represented by Thomson's experiment. He had found a way of measuring the ratio of charge to mass of the individual particles forming the cathode rays. To find out more about their actual dimensions it was necessary to measure or determine in some way either the mass m or the charge e. This seemingly impossible task was tackled successfully by several investigators, practically all of whom were members of that remarkable group of scientists working in the Cavendish Laboratory at Cambridge towards the close of last century under the leadership of Sir Joseph Thomson.

Townsend's Experiments

The first attempt to estimate the electric charge carried by a single ion was described by J. S. Townsend in a paper read before the Cambridge Philosophical Society on 8th February, 1897. This work, which formed the basis of a great deal of subsequent research on the subject by various experimenters, may be regarded as a classical example of experimental method.

It had been known for many years that the hydrogen produced when a metal is dissolved in an acid is electrically charged, but the effect had not been studied to any extent. Townsend discovered that when current was passed through sulphuric acid the oxygen and hydrogen appearing at the respective electrodes carried a positive charge, whereas if caustic potash is used for the electrolyte both gases are negatively charged. He found moreover that if these gases are bubbled through water they form a cloud, owing to the condensation of moisture on the ions forming the gas. This moisture can be eliminated again by a drying process, as by bubbling the gas through sulphuric acid, without losing more than about 25 per cent of the original charge of the gas. In these facts Townsend saw a means of arriving at the amount of electric charge carried by each ion.

His apparatus is shown in simplified form in Fig. 3. The gas produced by electrolysis in vessel *A* is collected as shown and bubbled through potassium iodide in flask *B* to remove ozone, and then through the water in flask *C* to form a cloud around the ions. The cloud passes through the flasks *D*, *E* and *F* containing sulphuric acid and the interconnecting drying tubes which remove all moisture, the weight of which can be found by weighing the tubes. The charged gas is collected in flask *G*, which has an outer tin foil coating. By the well known laws of electrostatics the charge appears at this coating, and its amount is measured by the very sensitive quadrant electrometer *H*.

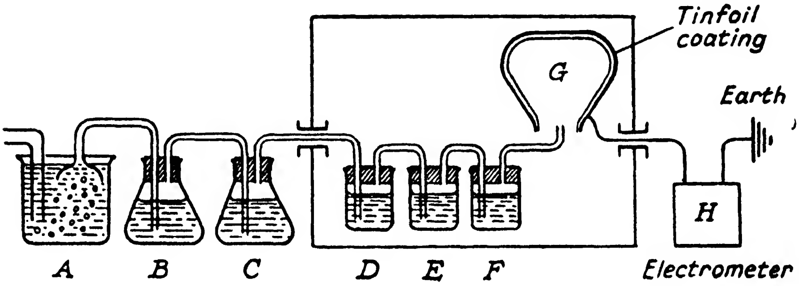


FIG. 3.—Townsend's apparatus for measuring the charge on an ion.

The experimental procedure was as follows :—

1. In the first place, only negatively charged gas was used, so that only charges of the same sign had to be considered, and it was assumed that each and every ion acquired condensation in the process of cloud formation, so that the number of droplets comprising the cloud was the same as the number of ions.

2. The charge on the gas was measured by the electrometer as already described. If now the number of droplets in the cloud, which we will call n, could be estimated and if Q was the charge as measured by the electrometer, then obviously the amount of charge on each droplet (or ion) would be given by $e = Q/n$. The value of n was found indirectly in the following steps.

3. The weight of the cloud produced on a given volume of gas was found by bubbling it through the flasks *D*, *E*, *F*, and then measuring the increase of weight of the drying tubes.

4. An estimate of the average weight of each droplet in the cloud was then made by means of a purely theoretical formula due to Sir George Stokes. This had been developed some time before, and quite independently of any application to experimental work. The equation gave the velocity of fall of spherical bodies of a given size and density through a gas of known viscosity under the action of gravity. Townsend observed the velocity of fall of his cloud in a flask connected to flask *C* and by applying Stokes' formula determined first of all the size of the droplets and then their mass, assuming them to be spherical.

5. After this it was a simple matter to find the number of droplets in the cloud by dividing the weight of the cloud (Step 3) by the mass of each droplet.

6. As a final step, the charge on each ion was arrived at by dividing the measured charge on the gas (Step 2) by the number of droplets.

There are several sources of error in this investigation: notably the assumption that the number of ions forming the charge is equal to the number of droplets, and the possibility that Stokes' equation might not apply in the case of such small droplets. It was in fact found later on that Stokes' equation required modification under the conditions applying to the experiments. As a result of his work, Townsend came to the conclusion that the average charge on each ion was approximately 3×10^{-10} electrostatic units (e.s.u.), that is to say about a *three thousand millionth part* of an electrostatic unit. This figure may not convey much to the reader, especially if he is not well acquainted with units of electrical measurement,* but it may be noted in passing that to maintain a current of one ampere in a circuit it would be necessary for 10^{18} (ten million million million, or ten billion million) such charges to flow past a given point in the circuit every second. We shall see later that Townsend's result was quite as accurate as could reasonably be expected in view of the experimental difficulties.

---

* The electrostatic unit of quantity or charge is $\frac{1}{3 \times 109}$ or 1/3,000,000,000th of the practical unit, or coulomb, representing the flow of 1 ampere for 1 second.

Thomson's Method

Another method, which is really a variant of Townsend's, was used by J. J. Thomson in 1898 for measuring the charge on gaseous ions. In this case the ions were produced in a closed vessel by means of an adjoining X-ray tube, the rays from which caused the gas in the vessel to become ionised, i.e., filled with electrically charged atoms or ions. The resulting ionised space was subjected to a small difference of potential and the corresponding current was measured. This current is dependent upon several factors, being proportional to :—

(*a*) The number of ions present.
(*b*) The amount of electric charge carried by each ion.
(*c*) The difference of potential causing the movement of the ions, i.e., the flow of current.
(*d*) The ionic mobility of the ions ; that is to say the velocity imparted to them by the action of unit difference of potential.

The mobilities of ions in gases at atmospheric pressure and temperature had already been determined by Rutherford, and the current and potential difference could be measured directly, so that only the number of ions and the ionic charge remained to be determined. If the number of ions could be arrived at, then the ionic charge could be calculated from the known relationship between the various quantities.

Thomson proceeded to find the number of ions by making a cloud of droplets form on them and then using more or less the same technique as did Townsend ; but for the formation of the cloud he employed a discovery made by C. T. R. Wilson during the years 1895 to 1897. This discovery was that the cooling effect produced by sudden expansion of the volume of a chamber containing an ionised gas caused the ions to acquire condensed moisture and thus produce a cloud. Thomson calculated the weight of the cloud from the theoretical amount of cooling resulting from the expansion and the difference between the densities of saturated water vapour at the respective temperatures. The number of ions was finally calculated from Stokes' formula as in Townsend's experiment.

The experimental apparatus is shown in simplified form in Fig. 4. The ionisation chamber takes the form of a glass bulb *A* covered at the top by an aluminium plate *D* and connected by the aperture *B* to the chamber containing the X-ray tube shown at *X*. The X-ray chamber has an opening at the bottom to admit the rays to the ionisation chamber *A*, the intensity of the rays being adjusted by inserting sheets of tinfoil or aluminium at the opening. A voltage applied between the aluminium plate *D* and the surface of the pool of water *W* by means of wires *E* and *F* produces an electric field acting on the ionised gas in chamber *A*. This field causes the positive ions to move towards the negative electrode and the negative ions towards the positive electrode in accordance with electrostatic principles. The resulting current through the gas caused by this movement of the ions was measured and compared with theory so as to yield a value for the charge on each ion.

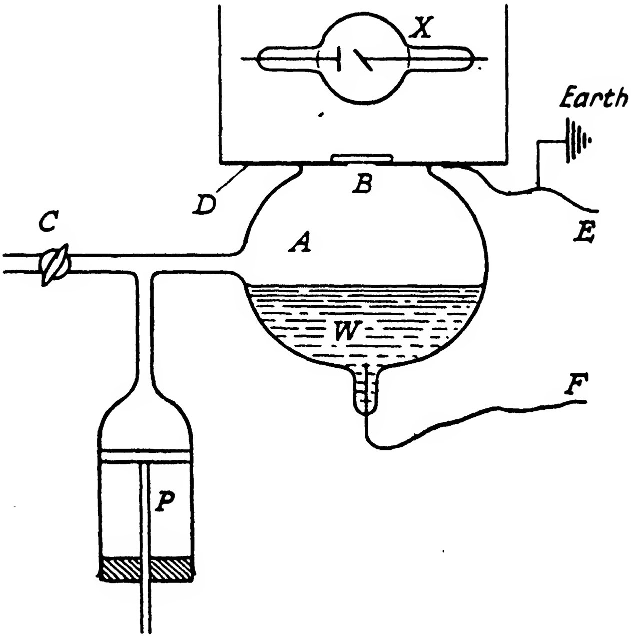


FIG. 4.—Thomson's cloud apparatus for measuring ionic charge.

The expansion required to produce the cloud in chamber *A* was caused by suddenly jerking down the piston *P*, the cock at *C* being closed. This is not the precise method actually used, but illustrates the principle. The uncertainties in this determination are considerable, and probably greater than in Townsend's work. The results gave the ionic charge as 6.5×10^{-10} e.s.u., or more than twice Townsend's value. Later on, Thomson repeated his experiments, using the radiation from radium to ionise the chamber instead of X-rays and making some other modifications to obtain more accurate results. The new determination gave a figure of 3.4×10^{-10} e.s.u., for the ionic charge ; a value in close agreement with Townsend's results and somewhat closer to the correct figure as determined many years later.

Wilson's Method

An important modification of Thomson's arrangement was used in 1903 by H. A. Wilson, who arranged his apparatus so that the charged cloud in the ionisation chamber was produced between two horizontal metal plates 35 mm. in diameter and from 4 to 10 mm. apart, connected to a 2,000 volt battery. The resulting electric field acted upon the charged droplets in accordance with the known laws of attraction and repulsion of electric charges, giving a vertical force on each droplet as well as the normal gravitational force due to their weight. The action of the field, according to its direction, was either to assist or oppose the gravitational force and thus to accelerate or retard the normal downward drift of the cloud caused by gravity alone.

With no electric field, a negatively charged cloud was formed by sudden expansion of the ionisation chamber and the rate of fall of the cloud between the plates was observed. This observation was then repeated with the field on and assisting gravity to drive the droplets downwards : that is, the bottom plate was positive. This gave two conditions for each droplet in the respective clouds : firstly the velocity due to gravity alone, and secondly the velocity due to gravity *plus* the electrical force acting on the droplet. From the known relations between the respective forces and velocities and with the aid of Stokes' formula it was possible to calculate the value of the ionic charge. Wilson obtained several different values in the course of his experiments ; the average of eleven measurements being 3.1×10^{-10} e.s.u., a value surprisingly close to those obtained by precious experimenters.

There are still several sources of error. Apart from relying on Stokes' formula being valid under the conditions applying in the experiments there is the consideration that the calculations for each determination are based on observation of two *different* clouds.

Millikan's Work

About the year 1906, R. A. Millikan of America conceived the idea of holding the uppermost part of the cloud stationary, with the object of observing rates of evaporation and thus making allowances for this disturbing factor in cloud methods of measurement. This idea led to his important series of oil-drop experiments, commenced about 1909.

In the year 1910, or thereabouts, Millikan began to suspect that Stokes' formula might not be valid for the tiny droplets used in the cloud experiments, the diameter of which was of the order of 0.0004 cm. It so happened that the validity of the formula had also been questioned on purely theoretical grounds at about this time. Accurate determinations of the viscosity of air had also become available, so that the possibility of more accurate results became very much improved.

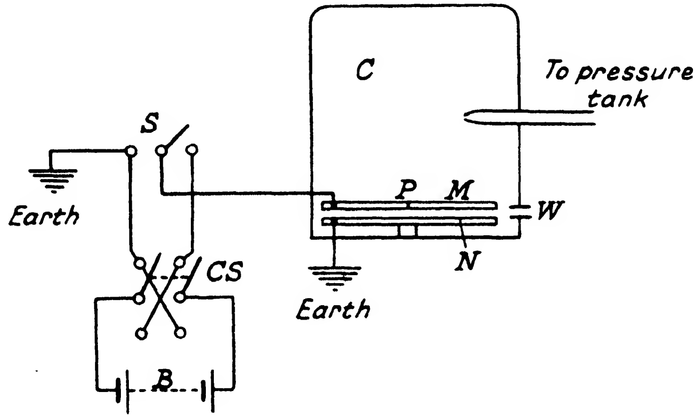


FIG. 5.—Basic arrangement of Millikan's apparatus for measuring charge on cloud droplets.

As the work we are now to describe resulted in the most accurate determinations of elemental charges which have ever been made, it may be advisable to go into rather more detail. The general theory is not at all difficult and the method has obvious affinities with that of H. A. Wilson.

The basic principle of Millikan's experiments involves the measurement of the forces on an electrified droplet of vaporised liquid maintained in suspension between two horizontal charged condenser plates. A very elementary representation of the apparatus used is given in Fig. 5. The oil, or other liquid employed, enters the chamber *C* in the form of a fine spray or cloud of very small droplets, each of which will have a radius of the order of two thousandths of a millimetre. The forces brought into action by the spraying operation give rise to frictional electrification of the droplets, so that each bears a definite charge of electricity. As the cloud drifts downwards under the action of gravity, one or more droplets will find access through the small hole *P* in plate *M*, and will then be under the action of the electrostatic field between the condenser plates *M* and *N*, which in this apparatus were located 16 mm. apart. With the switch *S* in the right-hand position, condenser plate *M* is connected to either the positive or negative terminal of the source of potential *B*, usually of about 6,000 volts, according to the position of change-over switch *CS*.

From the laws of electrostatics it follows that under these conditions the droplet, or more strictly its electric charge, will be acted upon by a force proportional to the charge on the droplet and the voltage applied to the plates, i.e., the electric field strength between the plates. The behaviour of a droplet under the action of this force is observed through the window *W* by means of a microscope. According to whether the plate *M* is at positive or negative potential with respect to plate *N*, the direction of the electric force on the droplet will be upward or downward respectively, so that it will either oppose or assist the downward gravitational force on the droplet representing its weight. In the experiments it was found possible, by suitably varying the applied voltage, to maintain the droplet quite stationary for half an hour or more at a time. Under these conditions the electric force must exactly balance the weight of the droplet. The electric field strength was known, so that if the weight of the droplet could be found it would be possible to calculate its charge directly from the simple relations connecting the quantities. In practice, the weight of a droplet is, of course, far too small to be measured by direct means but it can be determined from Stokes' formula, as had been done by previous experimenters.

In later refinements of this experiment it was found preferable, instead of keeping the drop stationary, to measure its upward movement due to the field and its downward movement or drift due to gravity only, i.e., with the field off, corresponding to the left-hand position of switch *S*, and to allow the drop to pick up radiation from a convenient source of some radioactive substance, or from X-rays. Experiments of this kind, with numerous modifications and different arrangements, showed that the electric charge acquired by the drop was always an integral multiple of a certain value of charge, so that for all observations the amount of charge arrived at was always once, twice, or some other whole number of times a definite elemental value of electric charge, which has consequently been taken to be the charge carried by a single electron.

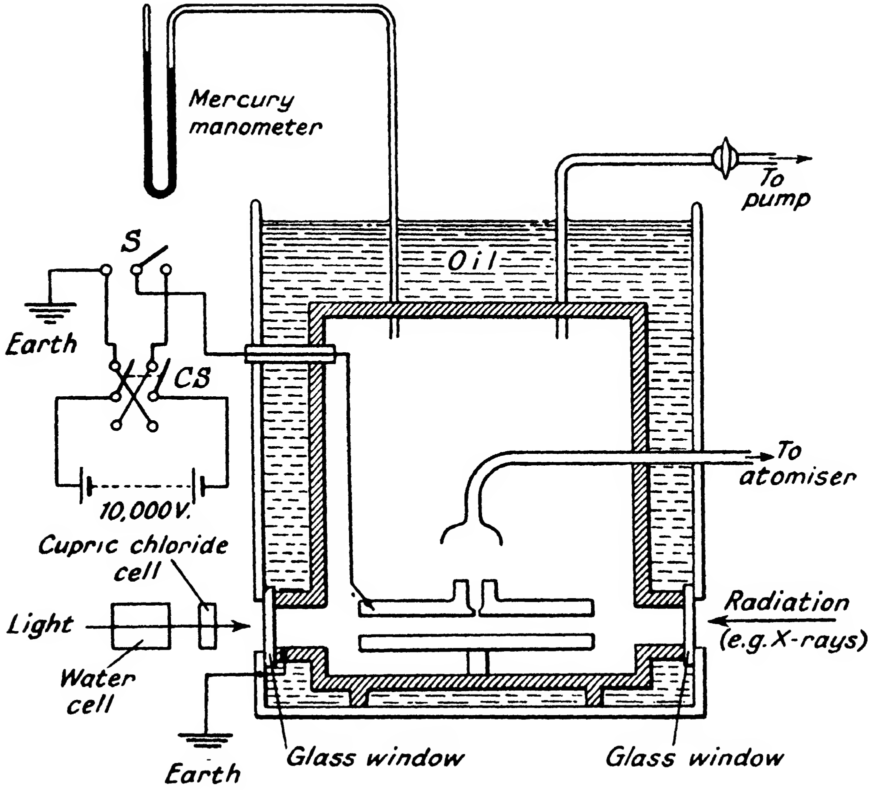


FIG. 6.—Final form of Millikan's apparatus.

Work with this apparatus revealed various sources of error, and improvements were carried out from time to time as indicated by experience in actual measurements as well as theoretical considerations. The final apparatus is shown in simplified form in Fig. 6. The inner vessel is of brass, and designed for working at pressures up to 15 atmospheres, so as to observe the effects of pressure on the test results. The absence of disturbing convection currents in the air between the condenser plates is ensured by filtering the heat rays from the light source by a water cell 80 cm. in length and a cupric chloride cell. Thermal stability is further ensured by immersing the whole vessel in an oil bath of 40 litres capacity. During observations the temperature remained constant within about 0.20 C. The atomiser for producing the cloud is blown by a puff of carefully dried and dust-free air. Only two windows are shown in the drawing, but there are three in the apparatus : one for admitting light, one for the entry of radiation when required and one for observation of the droplets by means of a short-focus telescope having a scale on the eyepiece for the measurement of their speeds.

After a total of some ten years of research, Millikan obtained his final result which, after considerable speculation, correction and comparison with other atomic data, was given as $(4.770 \pm 0.005)\ 10^{-10}$ e.s.u.

Other Experiments

In the meantime, entirely different and somewhat more direct evaluations of the elemental charge were made. In 1908, Rutherford and Geiger utilised the so-called alpha-particles radiating from a speck of radium to cause an impulse on the needle of a sensitive electrometer every time one of the particles entered an ionisation chamber. The total charge accumulated in the chamber in a given time was evidently the sum of the individual charges carried by the particles entering the chamber and could readily be measured, so that all that was necessary to find the average value of each individual charge was to divide the measured total charge accumulated in a given time by the number of impulses registered by the electrometer in the same time. It must be pointed out here that the alpha-particles are known to carry a charge of twice that of an electron. As the measurements in question gave a figure of 9.3×10^{-10} e.s.u., division by two gave as the electronic charge 4.65×10^{-10} e.s.u.

A similar method was used in 1909 by Regener, but he counted the number of particles emitted from a speck of radium by observing the number of scintillations they produced by striking a fluorescent screen in a given time. The corresponding total amount of charge was then found by measuring the charge delivered to a condenser by the radiation in the same period. After suitable corrections to take care of the different relative sizes of the screen and condenser the individual charge was found to be 9.58×10^{-10} e.s.u. Halving this value as before gave 4.79×10^{-10} e.s.u. as the electronic charge. These values are seen to be strikingly similar to Millikan's figure.

Other independent work has established similar results, and slightly differing values due to different authorities are quoted in reference books. The ratio of charge to mass of the electron has also been determined in a variety of ways by different experimenters, the generally accepted value being $e/m = 1.76 \times 10^7$ in *electromagnetic units* (e.m.u.).* Taking this value, and Millikan's figure of $e = 4.770 \times 10^{-10}$ e.s.u., it may readily be calculated that the mass of an electron is approximately 9×10^{-28} gram.

In practical engineering units the charge we have quoted is equivalent to 1.59×10^{-19} coulomb (ampere-second). Thus it is seen, by dividing this figure into unity, that a current of one ampere flowing for one second represents the movement of no less than 6¼ million million million electrons.

The Structure of Matter

For the study of electronics it is necessary to have a general knowledge of the part played by the electron and other elementary particles in the structure of matter. Some of these particles are of importance in connection with the conduction of electricity through gases.

Most of our present knowledge, hypothetical and otherwise, of the structure of matter is derived from experiments with the so-called radioactive substances. These have the curious property of emitting continuously and without stimulation of any kind three different sorts of radiation, which are known respectively as alpha-, beta- and gamma-rays. When passed through a magnetic field the behaviour of these rays is entirely different, as shown in Fig. 7. The alpha-rays are deflected as if they consisted of positively charged particles, the beta-rays in the opposite direction as if they were negatively charged particles, while the gamma-rays pass straight on as if they were not charged

---

* The electromagnetic unit of quantity or charge is ten times the practical unit (coulomb) and 3×10^{10} or 30,000,000,000 times the electrostatic unit.

at all. Research showed that the alpha-particles each have a charge equal to twice the charge on an electron but *positive*, whereas the beta-particles were found to be electrons moving at a speed approaching that of light. The gamma-rays proved to be very high frequency radiation having the same general nature as light.

About the year 1911 Rutherford made the experiment of passing alpha-particles through thin metal foil and found that a small proportion of the particles were strongly deflected in passing through, whereas the majority went straight on. The number of areas which caused deflection was found to be the same as the number of atoms in the foil and the amount of deflection was the same as if each area was the seat of a particle which repelled the alpha-particle according to the known laws of gravitation. Work of this kind enabled Rutherford to formulate his nuclear theory of the atom, according to which it consists of a positively charged and relatively massive central part, called the *nucleus*, surrounded by a number of electrons whose aggregate charge is equal and opposite to that of the nucleus, so that the atom is electrically neutral, i.e., without a surplus charge of either sign. The positive charge of the nucleus is ascribed to one or more particles called protons, each carrying the same electric charge as the electron but of opposite sign.

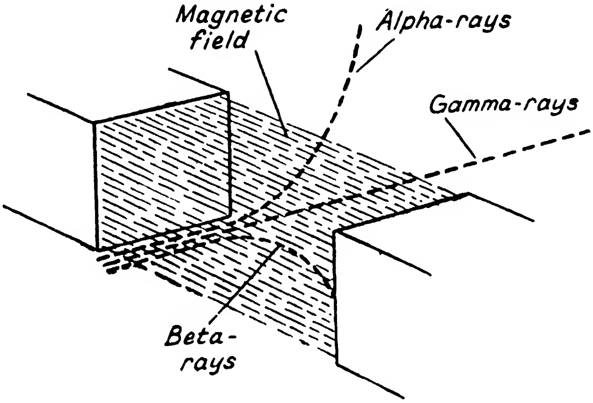


FIG. 7.—Behaviour of rays from radioactive substance in passing through magnetic field.

The simplest, and lightest atom is that of hydrogen, having a nucleus of one proton and a single accompanying electron.

Next in weight and complexity comes the helium atom, with two protons and two electrons. Next is the lithium atom, with three of each kinds of particle, and the sequence goes on with a few gaps up to the heaviest element, uranium, having a nuclear charge 92 times that of hydrogen. The arrangements of the atoms for the first three elements in the scale are shown in Fig. 8. It is assumed that the electrons must rotate around the nucleus at a speed such that their outward centrifugal force exactly balances the inward force due to electrostatic attraction. This is the same law as applies to the sun and planets, except that in this case the inward force tending to bring the bodies together is a gravitational one.

By various means it is possible for one or more of the planetary electrons to be detached from an atom. In this case the atom ceases to be electrically neutral and has a positive net charge. It thus becomes subject to influence by electric and magnetic fields and in this condition is called an *ion*. The original neutral state is usually soon restored by the acquisition of new electrons. The nucleus is much more stable, and can be disrupted only by radioactive changes or bombardment with high speed alpha-particles. Such action causes a radical change in the atom, which is virtually transformed into the atom of a different element. There is this such a thing as transmutation of the elements, and the change is enduring, as the nucleus shows no tendency to revert to its original state.

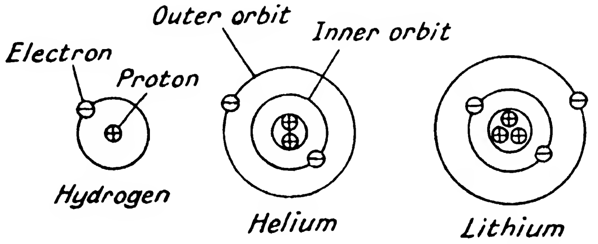


FIG. 8.—Atomic structure of simple elements.

In addition to atoms we also have to consider *molecules*, which may be defined as the smallest amounts of any substance, whether element or compound, capable of independent existence and retaining the properties of the original substance. They are elementary constituents of matter which cannot be further subdivided by mechanical means, but which by chemical and electrical means can be resolved into their constituent atoms. In solid bodies the molecules are bound tightly together, and strongly resist any forces tending to separate them ; in liquids they offer limited resistance to relative movement, and in gases there is a minimum of resistance to change of position. The movement of molecules in a gas is apparently random and chaotic, although of course it is actually subject to definite, and highly complex, physical laws.

Modern Electrical Theory

The conclusions of electrical science which are of chief interest and importance for our present purpose may be summarised as follows in terms of electrons :—

1. In the unelectrified or neutral state a piece of matter contains a definite number of negative electrons, such that two pieces of matter in this condition will experience no mutual electric forces.
2. If the number of electrons in the material is greater than the foregoing number corresponding to the uncharged state, the material appears to have a negative charge. If the number of electrons present is less than that corresponding to the uncharged state, the material appears to have a positive charge.
3. From the foregoing considerations it follows that when a body is positively charged, the process is actually the removal of electrons. Conversely, the process of imparting a negative charge is the addition of electrons, the number of electrons removed or added determining the amount of the charge. This principle is in accordance with the established hypothesis of electrostatics, that when any body is charged, equal amounts of positive and negative charge are produced.
4. In good conductors there is a relatively large number of the so-called " free-electrons," which move about inside the material with comparative freedom. In poor conductors or insulators the movement of electrons is very severely restricted.
5. The free electrons in a conductor may be regarded as wandering about apparently at random in all directions between the atoms of the substance when no external

electrical forces act upon the conductor. If an electrical force is applied, each free electron will have superimposed upon its previous chaotic motion a definite translatory motion in a definite direction, which depends upon the direction in which the applied electrical force is acting. This unidirectional translation of electrons constitutes the electric current.

FOR FURTHER READING

Much of the history of the early experiments is to be found in the following books :—

1. Sir J. J. THOMSON and G. P. THOMSON. *Conduction of Electricity Through Gases.* (Cambridge), 2 vols., 1928–33. This work is a classic of electrical literature, and reviews experimental research in all its branches, including the determinations of charge of electrons and ions. Although the work is advanced, much of the text can be followed without specialised training.
2. R. A. MILLIKAN. *Electrons* (+ and —), *Protons, Photons, Neutrons and Cosmic Rays.* (Univ. of Chicago Press), 1935. A good and easily readable account of experimental work leading up to the author's determinations of electronic charge, of which a full account is given. Also deals with atomic structure, the nature of radiant energy and cosmic rays. Largely non-mathematical.

CHAPTER III

THE PRODUCTION OF ELECTRONS

A source of electrons is an evident necessity for any form of electronic device, and the various ways in which electrons may be produced are consequently of primary importance in applied electronics. For the present purpose we are concerned with four methods :—(1) Thermionic emission, in which the electrons are driven out of a conductor by the action of heat. (2) Secondary emission, in which the electrons are ejected by the bombardment of high-speed particles such as electrons and ions impinging upon a conductor. (3) Field emission, also known as cold emission or auto-electronic emission, in which the electrons are drawn from a conductor by the application of an intense electric field. (4) Photoelectric emission, in which the electrons are ejected by the action of light falling on a conductor. This form of emission is dealt with in Chapter V along with other photoelectric effects which do not rely upon emission.

THERMIONIC EMISSION

This method of producing electrons is very widely used in practice. It is employed not only for all kinds of valves but also for a wide range of other electronic devices. In consequence of their practical importance, thermionic effects have been widely studied and extensively developed during recent years, so that the thermionic cathode in its many forms has become a highly efficient and adaptable arrangement.

Basic Researches

In 1880, the physicists Elster and Geitel made the first definite experimental inquiry into the electrical effects produced by the increase of temperature of a conductor. They used an evacuated glass vessel containing a filament, either of metal or carbon, near which was located an exploring electrode in the form of a small metal plate, externally connected to an electroscope. The filament was heated by passing current through it, and under these conditions it was observed that the electroscope indicated the presence of a negative charge on the exploring plate electrode. This experiment showed conclusively that the hot filament was able in some way to project across the evacuated space a negative charge of electricity; some of which was collected by the plate or *anode*.

The next step was made in 1883, during experiments carried out by Edison in connection with his early carbon filament lamps. On this occasion he was trying to discover the cause of filament failure, which he had noticed was associated with a progressive darkening of the inside of the glass bulb. In an attempt to determine the nature of this deposit, caused evidently by the hot filament, Edison sealed into one of his lamps a metal plate with an external terminal, to which was connected a galvanometer, as shown in Fig. 9. It was found that when the free end of this instrument was connected to the positive side of the filament supply there was indication of a feeble current, but no current flowed if the connection was made to the negative side of the filament supply. This result showed that the plate had the ability to withdraw a current from the region of the hot filament only when the plate was positive with respect to the filament, and thus implied that the current was composed of negative charges.

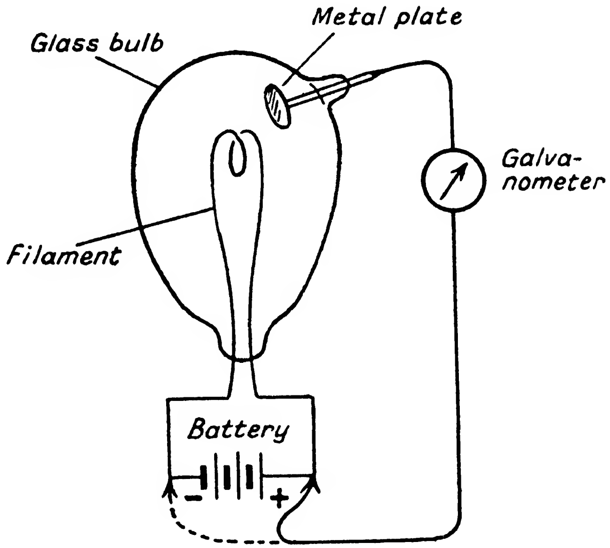


FIG. 9.—Edison's experiment with exploring electrode inside lamp

The practical possibilities of these discoveries were not realised at the time: in fact the state of electrical science was not suggestive of any application, and for many years the subject remained in the stage of academic experiment. It was brought from this stage to the very forefront of potential application mainly as a result of the series of painstaking researches carried out by the late Sir J. A. Fleming between the years 1889 and 1896, culminating in the development of the thermionic valve.

The evolution of the radio valve and the almost immediate application of this device in communications gave an immense impetus to theoretical and experimental study of thermionics. Systematic study from the theoretical viewpoint, supported by experiment, dates from 1901, when O. W. Richardson published a paper on the negative radiation from hot platinum. In 1903 the same author published a paper on the electrical conductivity imparted to a vacuum by hot conductors. This paper is one of the classics of the subject, and firmly established the basic principles.

General Theory

The problem of finding out just how the electrons are ejected from a heated metal into the surrounding space has naturally received a great deal of attention, and various theories have been proposed. The modern view is that in a metal at normal temperature the free electrons inside the metal, which may be imagined as darting about in all directions apparently at random, experience powerful retaining forces whenever they find themselves at or near the boundaries of the metal, in that indefinite region which separates the space and the metal. In this so-called boundary region there is a distribution of positive charges due to surface atoms which are assumed to be in the continuous process of losing and acquiring again their requisite number of accompanying electrons. These charges attract any escaping electron which may find itself tottering as it were on the brink of the boundary as a result of arriving at too high a speed, and draw it back before it can escape. At normal temperatures very few electrons escape, and those that do are pulled back after an excursion which would be far too small for measurement by ordinary means.

As the temperature of the metal is increased, the thermal agitation of the electrons and their speeds also increase, and so does the chance of their getting clear of the metal. The conditions can be represented by the analogy of a bowl or basin containing some dried peas. If the bowl is mildly agitated, the conditions correspond to normal temperature, and the peas, representing the electrons, do not escape, although a few may roll a little way up the sides. As the bowl is agitated more briskly the peas climb higher up the sides, and with sufficient agitation, corresponding to a sufficiently high temperature, they begin to escape from the bowl altogether ; thermionic emission has started. This idea, clothed in mathematical garb, forms the basis of the modern theory of thermionic emission, as presented in the specialised textbooks on the subject.

The amount of work required to remove an electron from a material depends upon the retaining forces to which it is subject, and these are different for different materials. For a given material the work required is known as its *work function*, and the term *electron affinity* is also used with the same meaning. The work function, representing the ratio of work to unit charge, is expressed in volts and is a measure of the difficulty of extracting electrons. Its value ranges from about one to six volts for the various emitter materials.

Experimental Work

Many of the earlier experiments were carried out under conditions which would now be regarded as very crude, owing to accumulated knowledge on various aspects of the subject. Foremost among the recognised essentials is the necessity of precise information concerning the physical state of any conductor under investigation, and particularly the state of its surface, which has a profound effect upon the thermionic emission. From this point of view a metal is only " clean " when its surface is as free as possible from contamination due to oxides, sulphides, gases or other metals. Many surface impurities emanate from the interior of the metal, which must be specially prepared and treated to eliminate the possibility of contamination of the surface after preparation. It may be said that no metal prepared in air can be really clean, and any measurements of emission from a contaminated surface will be representative of this contamination rather than of the metal itself.

In most investigations of thermionic emission a collecting plate or anode is used to receive the emitted electrons, and the current which they produce can be measured as shown in Fig. 9. The anode is maintained at positive potential with respect to the emitting surface or cathode, and thus forms a source of attraction to electrons. In the course of his investigations Richardson developed a formula which gave the electron current in terms of the temperature for a given kind of cathode material.

The relationship between emission current and temperature according to Richardson's formula is shown in Fig. 10. The point *A* at which emission commences depends upon the cathode material. With rising temperature there is at first a gradual and then a sharp increase in the emission current. According to Richardson's theory the current is independent of the applied anode voltage, so that however strong the field produced by this voltage, the number of electrons emitted in a given time, i.e., the measured anode current, depends only upon the temperature of the cathode. It was shown by Schottky, however, that emission increased with anode voltage as well as temperature. This increase, which became known as the *Schottky effect*, necessitated modification of Richardson's formula, and also of ideas concerning emission.

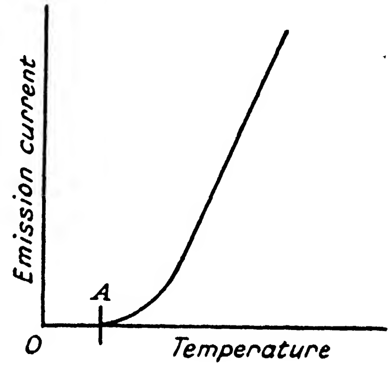


FIG. 10.—Relation between thermionic emission and temperature.

Another interesting phenomenon is the *shot effect*, representing a fluctuation of emission current due to the fundamentally random nature of thermionic emission. The escape of electrons is a random event, so that the emission may be compared with a stream of shot. The number of electrons emitted in a given time, and therefore the average current, may be constant under steady conditions but there are instantaneous fluctuations because of the random escape of individual electrons.

The *flicker effect* also represents a current variation, particularly noticeable in the case of composite cathodes, and is thought to be due to the arrival or loss of individual atoms at the emitting surface, causing changes in the emission at the corresponding points of the surface.

Thermionic Cathodes

As long ago as 1904, A. Wehnelt published the results of his work on emission from metallic oxides, leading to the development of the oxide-coated, or Wehnelt, cathode with an emission much greater than that of a corresponding metal filament. The technique of coating, usually with barium or strontium oxide, is somewhat special, and is followed by an activating heat treatment which considerably enhances the emission. The thickness of the coating has no appreciable effect on the emission, and it has been found that a mono-molecular layer, i.e., a layer only one molecule thick, emits as well as a coating several thousands of molecules in thickness. It may be noted, however, that according to recent work the emission from a given coating is influenced to some extent by the core material.

The greatly increased emission obtained from tungsten filaments to which a trace of thorium oxide (thoria) had been added during manufacture was reported by Langmuir in 1914. Cathodes of this type give a much greater emission than plain tungsten, but there is not much to choose in this respect between thoriated cathodes and the oxide-coated type. Both have advantages and disadvantages in particular applications.

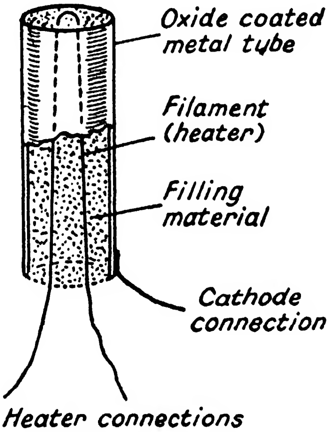


FIG. 11.—Form of indirectly-heated cathode.

The directly heated cathode in the form of wire or strip, coated or otherwise, is not satisfactory under some conditions of operation. For example, in the case of valves used in A.C. mains operated radio receivers the cyclic variations of voltage along the length of the filament produced a mains hum which became amplified at each stage. This difficulty was overcome by the use of an indirectly-heated cathode, a simple form of which is shown in Fig. 11. In this case the filament is of hairpin form, usually of tungsten, and is generally called the heater. The cathode proper, consisting of an oxide-coated metal tube, is brought to its operating temperature by conduction from the heater through the filling material, usually silica, porcelain or magnesia, which tightly fits the tube and insulates the heater from it.

Another objection to direct heating arises from considerations of current distribution. In order to preserve a uniform temperature along a directly heated cathode it is necessary that the heating current should be several times the emission current. Since the emission current flows through only a part of the cathode in obtaining egress from it under normal conditions, different parts of the cathode will be traversed by different currents and will thus be raised to different temperatures, unless the heating current is so large that the heating effect of the emission current may be neglected. In order to avoid high heating currents and associated difficulties with vacuum seals, etc., the indirectly heated cathode is widely used in practice. The use of an independent heater allows of a small heater current at higher voltage to obtain the required heat dissipation, so that smaller lead-in wires may be used.

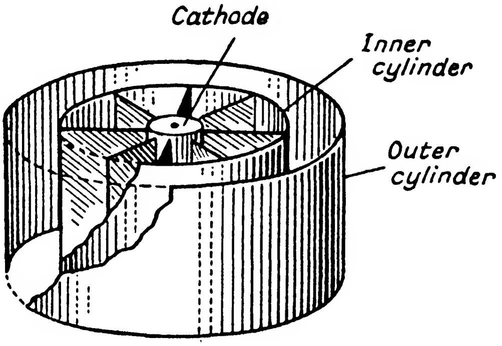


FIG. 12.—Part section of heat-screened indirectly-heated cathode.

In cases where high emission current and high efficiency are required, use is made of heat-screened indirectly heated cathodes, in which the heater and emitting surfaces are enclosed in a cylindrical metal sheath with smooth interior reflecting walls, designed to reduce heat loss to a minimum. A cathode of this type is shown in Fig. 12. The choice of a cathode in a given case naturally depends upon the conditions of operation, and several factors usually have to be taken into account.

SECONDARY EMISSION

If a particle such as an ion or an electron strikes a solid metal with sufficient energy it may cause the ejection of one or more electrons from the boundary region where, as we have already seen, they sometimes occupy rather precarious positions. In some way the energy of the bombarding particles is communicated to these electrons so as to enable them to overcome the restraining forces tending to keep them within the metal. As in other kinds of electronic emission the behaviour of the ejected electrons will depend upon the electric or magnetic fields to which they may be subjected after leaving the metal. A positive electrode in their vicinity will tend to attract them and thus prevent their return. The ejected electrons are called secondary electrons to distinguish them from the primary electrons, or ions as the case may be, which cause the emission.

Ions are relatively inefficient as compared with electrons in causing secondary emission, and in most practical cases the effect is due to electron bombardment. In many kinds of electronic apparatus the effect is unwanted, and steps have to be taken to prevent it interfering with the normal operation. In the case of valves, for example, it represents a disturbing factor which may require special means to prevent or minimise it, as by coating with graphite or some other poorly emitting material any electrodes subject to the effect. Secondary emission can also occur from the inside walls of a glass bulb, thus introducing charges which may interfere with the required distribution of charges in the tube. In such cases the effect may be overcome by an interior metal coating connected to the cathode, or a film of carbon which gives the blackened appearance noticeable in some kinds of valves. Unless such precautions are taken it is possible in high voltage tubes for hot spots to form on the inside of the glass wall owing to the formation of a positively charged region caused by the loss of secondary electrons by primary electron bombardment. This positive region attracts any free electrons and thus forms a focus for bombardment which continually increases owing to the increasing positive charge on the glass, and may lead to the eventual softening of the glass wall and consequent destruction of the vacuum in the tube.

Experiments have shown that under suitable conditions and with specially prepared surfaces a very high rate of secondary emission may be obtained ; as many as eight or ten electrons being ejected by each primary electron. These results have been put to practical use in a device known as the electron multiplier, which is described in Chapter V.

FIELD EMISSION

Electrons can be extracted, even from cold surfaces, by the application of a sufficiently strong electric field, ~~which~~ appears to pull the electrons out rather than boiling them out, as in the case of thermionic emission, or knocking them out as in the case of secondary emission. As the effect is obtainable at temperatures far below that necessary for thermionic emission it is also known as the cold cathode effect, and the terms cold emission or auto-electronic emission are sometimes used to differentiate from other kinds of emission.

If the field strength is expressed as a voltage gradient, i.e., as a certain voltage per unit distance from the surface, a value of the order of a million volts per centimetre is required to produce the effect, depending upon the material. This does not necessarily imply the use of very high voltages, as the requisite gradient can evidently be obtained by a sufficiently short gap. The electric field is strongest at sharp points or corners, and by using a very sharply pointed electrode it is possible to produce a sufficiently strong field with a voltage of the order of 200.

The laws governing field emission are strikingly similar to those for thermionic emission, the electric field strength playing a similar part to that of temperature in the respective cases. The importance of field emission in practice arises from the disturbing effects it is likely to introduce in high voltage valves or other devices unless precautions are taken to avoid sharp edges or corners on electrodes where the effect might occur.

FOR FURTHER READING

Thermionic and other forms of emission are generally dealt with in books on thermionic valves and in general works on electronics. The following deal specifically with the subject :—

1. T. J. Jones. *Thermionic Emission.* (Methuen), 1936.
2. A. L. Reimann. *Thermionic Emission.* (Chapman and Hall), 1934.

CHAPTER IV

ELECTRONS *IN VACUO* AND IN GASES

The surroundings in which electrons may find themselves after being ejected from a substance by any of the different forms of emission will vary according to the local conditions, and these will have a profound effect on the behaviour of the electrons. The space into which they are projected may be a vacuum or may contain a gas and in either case may be acted upon by electric or magnetic fields, or combinations of both. When it is considered that the degree of vacuum or the amount of pressure may vary between wide limits and that different kinds of gases may be used it is evident that the conditions are very variable and that only general results can be discussed unless great detail is introduced.

ELECTRON CONDUCTION *IN VACUO*

It is only in a vacuum that electrons play the sole part in electrical conduction because the presence of gas molecules of any kind results in the formation of ions which also take part in the conduction, as will be explained later.

The simplest arrangement for demonstrating conduction *in vacuo* is shown in Fig. 13, where two plane electrodes sealed into the ends of an evacuated glass tube form the anode and cathode respectively. The anode is the electrode connected to the positive terminal of the supply source, and the cathode is assumed to emit electrons by any of the processes already considered. The degree of vacuum is taken to be sufficient to prevent any gas molecules taking part in the proceedings.

The electrons emitted from the cathode, being of negative charge, are attracted by the anode and this action sweeps the electrons continuously from cathode to anode as long as voltage is maintained between the electrodes. The electrons entering the anode pass through the external circuit and restore the balance of electrons leaving the cathode. It will be noticed that the corresponding direction of electron flow is opposite to that of the current assumed by convention to flow from the positive terminal of the source. There is no mystery about this. The current flow conventions were fixed long before the nature of electricity was established, and it so happened that in this matter of even chance the wrong choice was made.

The continuous passage of electrons constitutes the current in the external circuit, and this current has the same value at any point in the circuit although the current density may vary widely in the different parts of the circuit. Since each electron carries a definite charge of quantity of electricity, the current (i.e., the quantity of electricity per unit of time), is determined by the number of electrons passing any given point in the circuit in unit time. It follows that any means of increasing the rate of flow of electrons through the tube will also be a means of increasing the current.

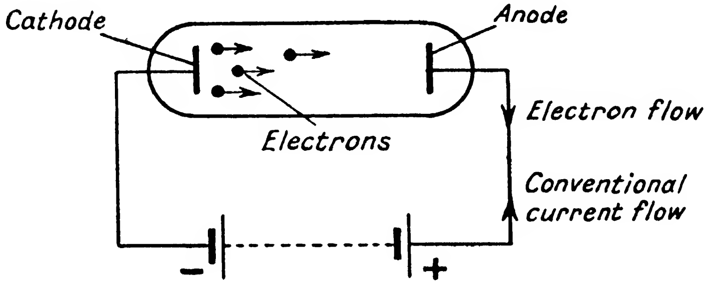


FIG. 13.—Circuit for demonstrating conduction in a vacuum.

One way of doing this which is effective up to a point is to increase the rate of emission from the cathode, e.g., by increasing its temperature in the case of a thermionic cathode. As the emission increases, so also does the current through the tube and at the same time there forms in front of the cathode an increasing swarm of electrons. They are being produced too fast for all of them to get clear of the cathode and those left behind hang there like a cloud which is continuously increased as the emission increases. The negative charge represented by the combined effect of all the electrons in the cloud is called the *negative space charge.* With increasing emission this accumulating charge exerts more and more repelling effect on electrons leaving, or trying to leave, the cathode, with the result that eventually very few succeed in getting to the anode and the current through the tube reaches a saturation point beyond which increase of emission causes no appreciable increase in current. This effect is known as space charge limitation of current. The space-charge plays an important part in the operation of vacuum valves.

CONDUCTION IN GASES

The presence of even a slight trace of gas changes the conditions of conduction appreciably, owing to the fact that gas molecules and atoms are present in the tube along with the electrons. There is also the risk of contamination of the electrodes by the chemical action of some gases. Both these effects introduced difficulties in the early days of electronic experiments, when the degrees of vacuum obtainable were seriously limited by the comparative crudeness of the pumping equipment then available. With modern methods it is possible to obtain pressures of less than 10^{-8} atmosphere (a hundred-millionth of normal atmospheric pressure) in which the effects of residual gas are negligible for all normal purposes. For some purposes, i.e., the numerous types of gasfilled tubes, a gas or mixture of gases is introduced into the tube, usually at a pressure of about a hundred-thousandth of an atmosphere. The presence of this gas completely transforms the action of the tube, and the considerations pertaining to conduction *in vacuo* no longer apply. In order to avoid chemical actions only inert gases such as argon, neon, helium, krypton or xenon are used in practice. Argon (from Greek *argos*, inert) was isolated by Rayleigh and Ramsay in 1894 and is present to the extent of about one per cent in the atmosphere. It is widely used in some forms of gasfilled tubes. Neon (from Greek *neon*, new) also has a wide sphere of use, notably in luminous discharge tubes ; it forms about two parts in a hundred thousand of the atmosphere. Helium (from Greek *helios*, sun) was first detected in the sun's atmosphere by the spectroscopic observations of Lockyer, the astronomer, in 1868 and isolated by Ramsay in 1895. It is not widely used in electronic apparatus. Krypton (from Greek *krypto*, I hide) was identified from its spectrum by Ramsay and Travers in 1898, and is very rare, being used in practice only to a limited extent. Xenon (from Greek *xenos*, stranger) also has an appropriate name, as it is present in the atmosphere to the extent of about five parts in a hundred million. It is the heaviest of all the inert gases and is used to some extent in special types of tubes.

The Conditions in a Gas

The interior of a gasfilled tube, even when no current flows through it, is the scene of intense and apparently meaningless activity. The gas molecules speed about in every direction, colliding with each other and with the walls of the tube and rebounding from these continuous collisions like tennis balls kept in random flight by some unseen agency. The chaos never ceases.

When electrons are projected into the space the conditions become even more wild, as in addition to the turmoils among the gas molecules themselves we now have electrons moving at very much greater speeds and buffeting their way through the mass of flying molecules in their efforts to obey driving impulses to which the molecules are insensitive. If we think of the molecules as tennis balls we must regard the electrons as like the smallest of small shot, travelling at immense speeds. On their journey they collide continuously with molecules and undergo corresponding changes of direction, so that their path through the gas, if it could be traced, would be found to consist of a series of short paths, not all the same length, pointing in random directions and representing the short journeys between successive collisions. The average length of path which the electron traverses between collisions is called the *mean free path*. It is dependent upon the number of molecules per unit volume, or the concentration of the molecules, which in turn depends upon the temperature and pressure, so that the mean free path is proportional to the temperature (increases with increase of temperature) and inversely proportional to the pressure (decreases with increase of pressure). It has been estimated that under the conditions normally applying to gaseous conduction the mean free path of an electron is only a few hundredths of an inch, so that its journey through the average length of path is decidedly eventful.

The Process of Ionisation

In its normal state, and without interference from external sources, a gas of any kind is a very poor conductor of electricity. The process by which it reaches the conducting state is that of ionisation, which may be brought about in either of two ways :—

(1) By subjecting the gas to radiation of sufficient energy, e.g., X-rays or light of short wavelength. (2) By arranging the electrical conditions so that the ionisation occurs in consequence of the conduction through the gas. The effect is the same in both cases. Some of the gas atoms lose one or more of their associated electrons and are thus transformed into positive ions. If more than one electron is lost, the ion is said to be multiple-charged. The electrons involved in these changes are assumed to be those circulating in the outer orbits around the atoms and requiring less energy for their detachment than those near the nucleus.

Returning to our tennis ball analogy, the onset of ionisation occurs when the particles shot into the gas from outside or the electrons in motion inside have sufficient energy to start knocking bits off the tennis balls. The energy which an electron must have to cause ionisation is called the *ionisation energy* and is measured in volts. Two values are usually recognised : the minimum, or first, ionisation energy required to remove only the outer electrons, and a higher value corresponding to the removal of other electrons.

The positive ions resulting from the ionisation process are very massive in relation to the electron as they have practically the same mass as the original atoms, the mass represented by lost electrons being negligible. Owing to the positive charge of the ions they become subject to the action of electric fields but are urged in the opposite direction to electrons. The advent of ionisation thus results in increased activity in the gas space, which is now the scene of the random movement of gas molecules, the one-way struggle of electrons and the opposite drift of positive ions. Owing to the relatively great mass of the latter their movement is altogether more stately and unhurried than that of the electrons, and they play a relatively small part in the creation of new ions by collision with gas molecules. When they arrive at the cathode, however, their continual impact raises its temperature, sometimes sufficiently to cause thermionic emission as well as the ejection of secondary electrons.

One highly important effect of the positive ions is to neutralise the negative space charge. Their action in this respect is particularly effective because owing to their slow movement they remain much longer in a given region than do the electrons. For this reason a relatively sparse flow of ions is able to neutralise the space charge of a considerable stream of electrons.

Owing to the electrostatic attraction between unlike charges the ions always tend to take on free electrons to replace those originally lost and thus to revert to the neutral atomic state. This process is known as *recombination,* and is continually going on in the gas, at the electrodes and at the walls of the tube. For a gas to remain ionised it is evidently necessary for the rate of production of ions to exceed the rate of recombination.

It will be seen that the net result of ionisation, in whichever way it is caused, is to multiply the number of carriers of electric charges in the tube and thus to increase the flow of current for a given voltage applied to the electrodes, i.e., to increase the conductivity of the gas space.

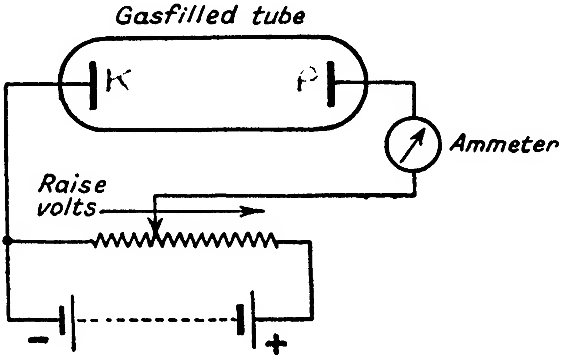


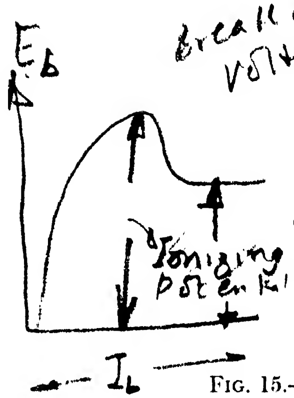
FIG. 14.—Circuit for varying the voltage across gasfilled tube.

It now remains to consider the different kinds of conduction which are possible, and the processes taking place in each. Experiment has shown that there are three clearly defined forms of conduction, or discharge as it is generally called. They are as follows :—

1. The non-self-sustaining discharge, relying upon supplementary effects, such as photoelectric or thermionic emission or X-radiation of the gas space, to enable it to continue.
2. The self-sustaining discharge, which requires no such means of excitation to maintain it, the required conditions for continuance being produced within the discharge itself.
3. The arc discharge, a special form of self-sustaining discharge, capable of occurring in various forms but always luminous and characterised by a high current and low voltage drop, i.e., a particularly low resistance path.

The Non-self-sustaining Discharge

If two metal electrodes enclosed in a gas are connected externally so that a variable voltage may be applied between them as shown in Fig. 14 it is found that the current bears a rather curious relationship to the applied voltage. Commencing at zero voltage, the current rises with increase of voltage at first, but the increase then becomes more gradual until no further variation takes place and the current remains at a constant value which is independent of voltage. This saturation value of current is very small: usually much less than a microampere (millionth of an ampere). If the voltage is increased sufficiently a stage is reached where further increase results in a sharp rise in current.



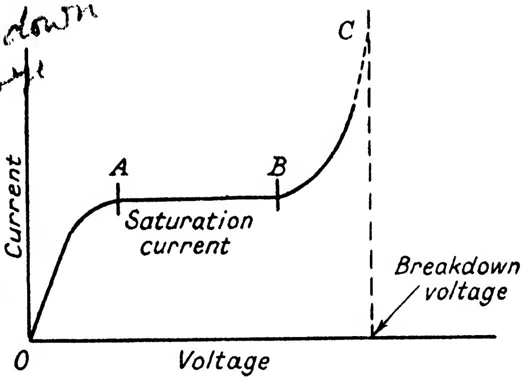


FIG. 15.—Variation of current with voltage across gasfilled tube.

The conditions are shown by the curve in Fig. 15, where *A* represents the start of the saturation effect. At this stage the absorption of ions and electrons by their arrival at the respective electrodes and by recombination balances the rate at which they are produced. This condition is maintained until a higher voltage is reached, corresponding to point *B* where the electric field attains a value sufficiently high to accelerate the electrons enough for them to cause ionisation by collision with the gas molecules. Increase of voltage beyond this point is accompanied by a sharp increase of current, and a point *C* is eventually reached where any further increase of voltage results in an indefinitely large increase of current. This is the commencement of the self-sustaining discharge, and the region between *B* and *C* corresponds to what is called the *Townsend discharge*. The voltage at the transition point *C* is variously known as the *breakdown, starting, initial, ignition or beginning* voltage. After breakdown has occurred the discharge is self-sustaining but has a variety of possible characteristics, depending upon the kind of gas, its pressure and the electrode arrangements.

The Self-sustaining Discharge

The various forms of this kind of discharge or conduction are best studied and distinguished by arranging for variation of the current over a wide range by means of an adjustable external

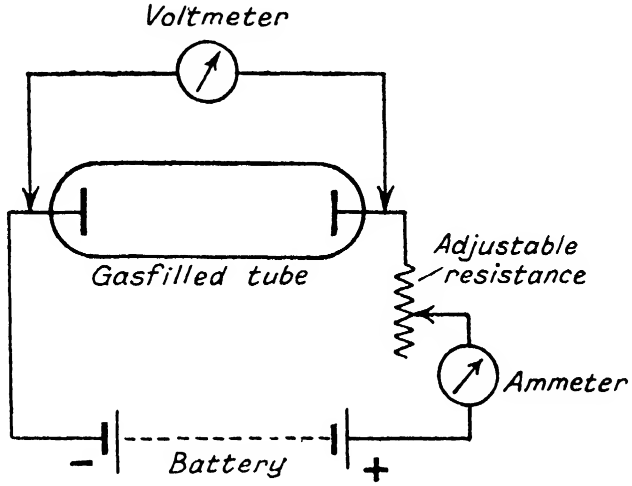


FIG. 16.—Circuit for varying the current through gasfilled tube.

resistance. As the current is gradually increased, a stage is reached where the discharge path begins to appear luminous and the voltage between the electrodes, i.e., the voltage drop across the discharge path measured as shown in Fig. 16 becomes lower than that corresponding to the dark conditions. The region of transition cannot be sharply defined. Sometimes it is fairly gradual, as represented by the curve in Fig. 17, but the drop in voltage indicating the change may also occur suddenly, as shown by the dotted lines. The voltage and current values shown on the curve ordinates are merely representative of average conditions.

As the current is increased the luminosity of the discharge increases, and also changes its position. If the current is sufficiently increased the glow completely covers the cathode and any increase of current beyond this point is accompanied by an

2. J. A. Crowther. *Ions, Electrons and Ionizing Radiations* (Arnold), 1938. Intended as a class-book for degree students, dealing with theoretical and experimental principles ; not applications. There is a minimum of mathematics, simplified where possible.

More Advanced :—

3. Sir J. J. Thomson and G. P. Thomson. *Conduction of Electricity through Gases.* (Cambridge), 2 vols., 1928–33.
4. E. L. E. Wheatcroft. *Gaseous Electrical Conductors.* (Oxford), 1938. Intended for those with previous training in physics. Deals with general principles and their application in certain branches of practice, e.g., valves, mercury-arc rectifiers, circuit breakers and luminous discharge tubes. Treatment largely mathematical.
5. J. D. Cobine. *Gaseous Conductors.* (New York, McGraw-Hill), 1941. A class-book dealing almost entirely with theory and devoted mainly to gas-conduction. Some practical applications are dealt with briefly. Mathematical.
6. F. A. Maxfield and R. R. Benedict. *Theory of Gaseous Conduction and Electronics.* (New York, McGraw-Hill), 1941.

CHAPTER V

PHOTOELECTRIC EFFECTS

Commencing about the year 1873 a number of basic discoveries were made concerning electrical effects caused by the action of light. These discoveries opened up an entirely new field of inquiry, and were at first studied purely from the academic standpoint. Many years elapsed before the effects came to be used in the form of so-called photoelectric cells or photocells, which were first put to such scientific uses as the measurement of illumination from stars, the changes of light during solar and lunar eclipses and other photometric applications where the electrical response obtained from the cell was taken as a measure of illumination and thus dispensed with the uncertainties of the human eye. This is probably also the reason why the photocell has sometimes been popularly referred to as the " electric eye."

From these early uses much knowledge was gained in the technique of applying photoelectric effects to practical purposes and the photocell is now not only a valuable and reliable electronic device with numerous industrial applications but has also made possible the advent of the sound film and television.

There are four different photoelectric effects, each corresponding to a different form of photocell ; the classification being as follows :—

1. Photoelectric emission, in which electrons are ejected from special metal surfaces by the action of light. The photoelectric cell or emission type photocell utilising the effect contains the sensitive surface and an electrode for collecting the electrons within an evacuated or gasfilled glass envelope. The current produced is measured in microamperes, and an amplifier is required in the majority of applications.
2. Photoconductivity, in which a change of resistance takes place in a thin film of selenium when the illumination varies. The effect is utilised in the photoconductive cell, also called the selenium cell or selenium bridge, to operate auxiliary devices such as relays either directly or through the medium of an amplifier.

3. The Photovoltaic effect, in which illumination of the region of contact between two specially prepared conducting surfaces causes a flow of current between the respective materials. Owing to the rectifying properties of the contact the current flow is in a particular direction. The photovoltaic cell or rectifier type photocell which utilises the effect generally produces a current of a few milliamperes and may be used with or without an amplifier.
4. The Becquerel effect, in which the illumination of a cell containing two special electrodes immersed in a liquid electrolyte produces at the terminals a small voltage. The arrangement utilising the effect is called a Becquerel cell, after its discoverer, or a photoelectrolytic cell.

There has been a great deal of confusion in the terms used to distinguish the various photoelectric effects and types of cell, but the foregoing terms are in general use and seem satisfactory for the purpose of general description.

PHOTOELECTRICITY

It is among the most remarkable discoveries of science that under suitable conditions an electric current in the form of a stream of electrons may be drawn from a metal merely by allowing light to fall upon it. Like so many other important discoveries it was largely accidental, and arose out of experiments carried out by Hertz in 1887 when it was found that a negatively charged body lost its charge rapidly when subjected to the action of light, particularly the light of short wavelength from the ultra-violet region of the spectrum. Positively charged bodies did not show the effect.

We now know that a negative charge is simply an accumulation of electrons, so that the loss of negative charge in the experiment means that in some way the electrons representing the charge are driven from the body by the action of light falling on its surface. At the time of the discovery, however, the idea of electrons had not yet appeared, and no reasonable explanation of the effect could be given.

Theory of Photoelectric Emission

Although the greater part of this work was completed before the end of the nineteenth century, it was not until 1905 that

Einstein was able to put forward a satisfactory mathematical theory of the photoelectric effect. This theory was based on the famous quantum theory, published by Max Planck in 1901.

According to Planck, the energy of radiation of any kind is dependent only upon its frequency, to which the energy is directly proportional. It follows from this that light of short wavelength should be the most active photoelectrically, as was found by the previous experimenters.

The theory also fitted in very well with the observed fact that for a given light-sensitive substance there was a definite *threshold frequency* of the incident light, below which there was no observable photoelectric effect. This fact was interpreted by Einstein to show that below this critical frequency the incident light energy was insufficient to liberate electrons from the substance, as required in order to produce the photoelectric effect.

According to quantum physics, light energy consists of an assemblage of photons, or light-quanta, each of which has associated with it a certain definite amount of energy, determined by the frequency of the light. The theory put forward by Einstein suggests that if the energy is sufficiently great, i.e., if the frequency is sufficiently high, the photons will possess enough energy to dislodge electrons from the photoelectric substance when the light falls on it.

It does not follow that every photon will be successful in dislodging an electron, or that an electron once dislodged from its normal association with the atoms of the material must inevitably be forced outside the boundaries of the material. In fact, the proportion of the incident light energy actually liberated in the form of photoelectrons is very small ; of the order of two or three per cent.

Emission Type Photocells

It was soon found that a particular group of metals, the alkali metals, gave a strong photoelectric effect with light corresponding to the visible part of the spectrum. The metals in question are very active chemically, and the technique of manufacture of photocells employing them is a highly specialised one calling for the greatest chemical purity of component materials, very high degrees of vacuum and absence of any form of contamination. Particulars of the principal alkali metals are given in the following table :—

Metal	*Chemical Symbol*	*Melting Point (deg. C)*	*Density grams per cubic cm.*	*Work Function, volts*	*Threshold Wavelength, Angstrom units* (10^{-8} *cm.*)
Barium	Ba	850	3.78	1.7	7,300
Caesium	Cs	26	1.87	1.36	9,110
Lithium	Li	186	0.534	2.36	5,250
Potassium	K	62.3	0.870	1.55	8,000
Rubidium	Rb	38.5	1.53	1.45	8,550
Sodium	Na	97.5	0.970	1.82	6,810
Strontium	Sr	800	2.60	2.0	6,200

As in the case of thermionic emission, the work function is an index of the degree of difficulty with which electrons are ejected : the metals with the lowest work functions being the most efficient photoelectric emitters. It is seen that caesium has the best performance in this respect, and as its response to light of different wavelengths is similar to that of the human eye it is widely used in emission photocells.

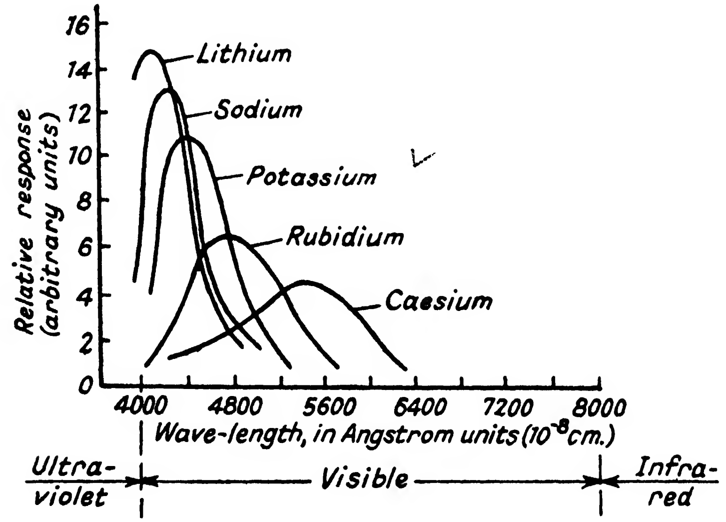


FIG. 19.—Variation of photoelectric emission with colour of incident light.

The manner in which the photoelectric emission varies in relation to the colour (i.e., the frequency) of the incident light in the case of a few typical materials is shown in Fig. 19. Such curves are generally called spectral response curves, and it will be noticed that each metal has its maximum response at a different part of the spectrum: also that some metals give response over a greater range of colours than others. Caesium has the greatest range within the visible part of the spectrum.

In general, the emission photocell consists of a plane electrode covered with a thin film of one or other of the alkali metals and sensitised by a special process to obtain the maximum photoelectric response in terms of the current produced with a given intensity of illumination. This electrode is either flat, or shaped to catch the maximum amount of light, and forms the cathode. It is faced by the anode, which usually takes the form of a wire loop so as to obstruct as little as possible the light directed on to the cathode. Both electrodes are arranged for external connection, and are enclosed in a glass bulb as shown in Fig. 20, representing a typical construction. For special purposes there are sometimes radical departures from this simple type: the anode may take the form of a wire gauze cylinder surrounding the cathode and in some cases the sensitive cathode surface is deposited as a film on the inside walls of the bulb, in which case a window is provided for admitting light. In all cases the anode is made positive with respect to the sensitive cathode so as to attract and collect the emitted electrons.

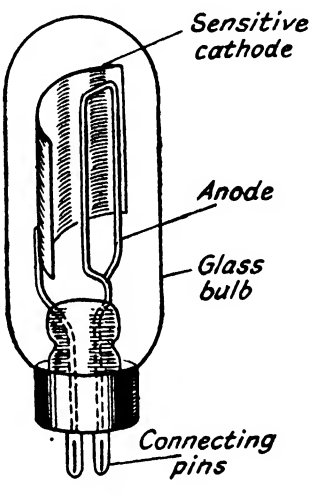


Fig. 20.—Arrangement of typical emission type photocell.

Vacuum and Gasfilled Cells

There are two essentially different types of cell, according to whether or not a gas filling is used in the bulb. In either case the bulb is highly exhausted, but in the gasfilled cell a small amount of inert gas, such as argon, is introduced after the evacuation process. The presence of this gas alters the relationship between the output current and the amount of light falling on the cathode and gives a higher sensitivity, i.e., a higher current for a given illumination.

The difference between the two types is shown by the curves in Fig. 21, where current output is plotted against applied voltage for a fixed value of illumination. Curve 1 applies to the vacuum cell, in which the current reaches a constant or saturation value in the region of about twenty volts. In the gasfilled cell, represented by curve 2, the same general features are retained at low voltages but at about the region of fifty volts an increase in voltage is accompanied by an increasingly sharp rise in current. This is the region of the *Townsend* discharge which we have already considered in Chapter IV (see p. 52), the current increase being due to the ionisation of the discharge path caused by the collisions between electrons and gas molecules. In practice the current must be limited, as by means of external resistance, as the cell may be seriously damaged if an excessive current flows. Although the increased sensitivity of the gas-filled cell enables it to detect lower values of illumination it is not so stable as the vacuum cell, which is preferable where measurement of illumination as distinct from detection is required.

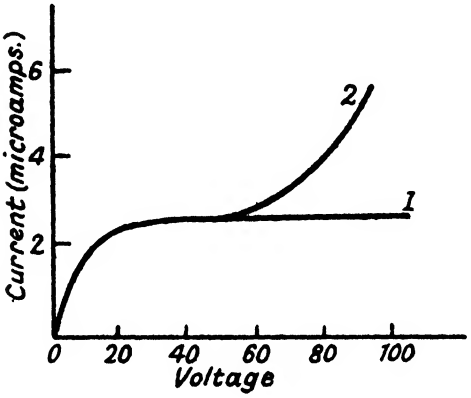


FIG. 21. — Comparison between vacuum (curve 1) and gasfilled (curve 2) photocells.

Valve Amplification

As the current output of an emission type photocell is only a few microamperes it is essential for most purposes to amplify the current very considerably before any practical purpose can be attained. It is here that the thermionic valve enables the emission photocell to be used in a wide variety of applications by multiplying its feeble current changes and following these changes without time lag. The speed of response of an emission cell is very high, especially for the vacuum type, in which the current follows the illumination within much less than a millionth of a second. With a gasfilled cell a certain amount of time lag is introduced by the fact that time is required for the ionisation process but it is only at high frequencies of operation that the time lag becomes noticeable. In either case the added lag due to the amplifying valve is altogether negligible for any ordinary purpose.

The circuit in Fig. 22 shows the basic principles. In this case the cell, either vacuum or gasfilled, is used to operate a relay through the medium of a triode valve, and batteries are used to provide the necessary voltages. The polarity of the batteries is important, since the respective voltages must act in the proper directions. Assuming that the cell is dark, so that no emission takes place, no current will flow between points *A* and *B*. The voltage of battery B_2 is selected so that the grid of the valve is sufficiently negative to prevent anode current flow or limit it to such a small value that the relay is not influenced by it. Under these conditions the cell could be entirely disconnected without affecting the situation. If now the cell is illuminated, electrons flow from cathode to anode, representing a current which to comply with convention must be represented as flowing in the opposite direction, as shown by the arrow. The result of this current, which is circulated by battery B_3, is to produce a voltage drop across resistance *R* and to make the grid of the valve (i.e., point *B*) less negative than it was before current passed between *A* and *B*. This may be seen from the fact that

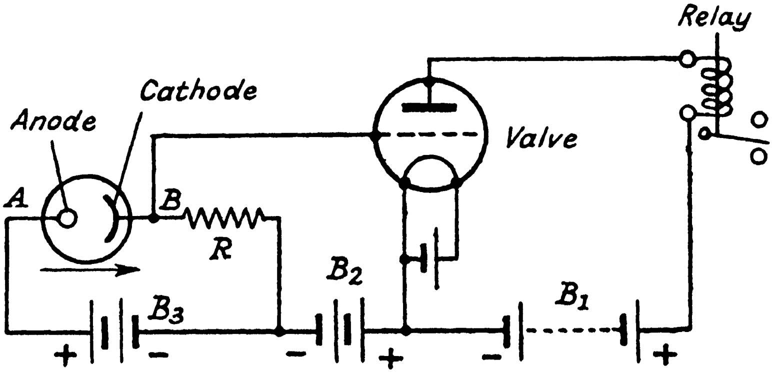


FIG. 22.—Simple form of photocell amplifier circuit.

the voltage across R and the voltage due to battery B_2 are in opposition, so that when current flows through the cell, point B (and, therefore, the valve grid) becomes less negative than it was before by an amount equal to the voltage across resistance R. The amount by which the grid voltage of the valve is reduced is, therefore, proportional to the amount of current passed by the photocell. The reduction of grid voltage causes the valve to pass anode current, which is provided by battery B_1, and operates the relay or whatever other device is connected in the anode circuit. When the cell is made dark again its emission ceases, resistance R no longer produces the corresponding voltage drop and the valve grid returns to its full negative potential as determined by battery B_2, thus cutting off the anode current.

It is left as an exercise for the reader to satisfy himself that if the photocell connections are reversed, i.e., the cathode connected to point A and the anode to point B, and battery B_3 also reversed, then the operation of the circuit will be inverted, so that the valve anode current will be cut off when the photocell is illuminated.

Practical Applications

The industrial and other uses of emission photocells, generally in association with some form of amplifier, have increased greatly during recent years and there is now an imposing list of applications in almost every thinkable branch of applied science. Among many other uses, the emission cell has been applied to the measurement of absorption of sunlight in the atmosphere, the penetration of daylight into seas and lakes, the photometry of all kinds of light sources, the opacity of liquids and solids, colour matching and the sorting of colours, temperature determination and control by colour response, the development of the sound film and television. Contrasted to these uses involving the measurement of light and sometimes also of colour are the applications which depend only on response to the interruption and establishment of light. The latter include the detection and timing of all kinds of movements, the counting of moving objects, the automatic opening of doors, and the protection of enclosures by light rays.

ELECTRON MULTIPLIERS

In cases where the output from a photocell must be greatly increased before it can be put to its intended use there may be considerable difficulty in applying a valve amplifier for this purpose. When several stages of amplification are necessary troubles arise from several sources, including microphonic noise in valves, variable leakage effects and the pick-up of spurious signals. The resulting distortions are amplified along with the proper signal, which tends to become lost against the background of random noise. In technical terms, the signal-to-noise ratio is said to be low, and this is a serious matter in such applications as television where the purest possible amplification of very weak light signals is required.

Difficulties of this kind may be largely overcome by using the electron multiplier, in which photoelectric and secondary emission effects are utilised within one envelope to produce a current several thousands of times greater than that corresponding to the photoelectric effect of the light signal received. There are several forms of multiplier, but all employ substantially the same principles. A photosensitive surface is arranged so that the electrons ejected under the influence of light are directed towards a positively charged electrode, sometimes called a target, having a specially sensitised surface with a low work function and high rate of secondary emission. The impact of the electrons liberates more electrons, as already explained in connection with secondary emission (see p. 43) and this increased flow is directed on to another positively charged target, where still more electrons are liberated on impact. This process can be repeated many times by using a number of targets, and results in a multiplication of the number of electrons, and hence the current, at each stage.

The principle is shown in Fig. 23, where the direction of the incoming light on to the photoelectric surface S is indicated by the dotted line. The electrons ejected from this surface are directed on to target T_1, which has a suitable positive potential with respect to S. Each electron striking the sensitised surface of T_1 produces several new electrons by secondary emission and this increased flow of electrons is directed towards target T_2, which has a suitable positive potential with respect to T_1. The secondary emission process is repeated at T_2, and a doubly magnified stream of electrons is aimed at the next target T_3. Theoretically, the multiplication can go on indefinitely, but in practice the repulsion between individual electrons causes a progressive spreading of the beams and places a limit on the number of stages used. A limit is also set by the heating of the final target anode.

There are several ways of directing the electrons from one target to another. Electric and magnetic fields, or a combination of both have been used. It will be understood that as each target must be positive with respect to the one before it, the respective targets must be at increasingly higher positive potentials with respect to the photoelectric cathode. As the voltage for each stage must be of the order of 100 in order to secure the high electron velocities necessary for secondary emission it follows that considerable voltages are required for operating the arrangement.

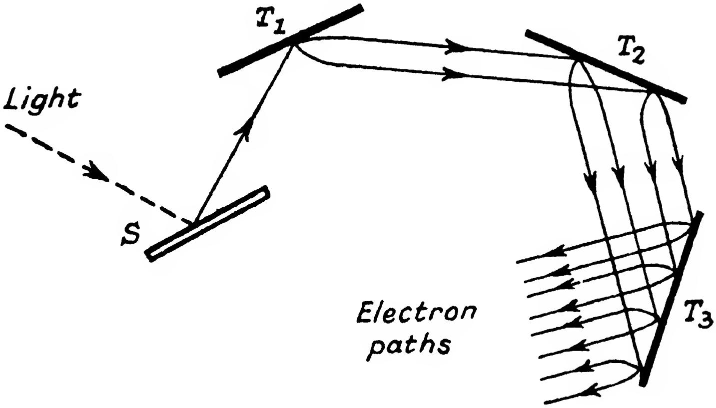


FIG. 23.—Principle of electron multiplier.

The conditions as drawn in Fig. 23 represent the case where the number of electrons falling on each target is doubled, and it is seen that for each electron leaving the cathode *S* no fewer than eight leave the third target, i.e., the original photoelectric current is multiplied by a factor of eight when three stages are used, each with a secondary emission factor of two. The total factor of multiplication is in fact simply equal to the secondary emission factor multiplied by itself the same number of times as there are stages. For instance, if there are eight stages (targets) and the secondary emission factor of each is three (i.e., three electrons emitted for each electron received), the total multiplication factor will be 6,561. Figures of this order are quite general for standardised types of multipliers, but for special purposes it has been possible to construct units with multiplication factors of the order of 100,000.

The arrangement of a modern form of multiplier is shown in Fig. 24. In this type the multiplying is done by electrons striking the respective grids G_1, G_2, etc., on their passage through the tube, which is generally about seven inches long and two inches in diameter. The photoelectric cathode C receives light and emits electrons in the usual way. These electrons pass into the open end of a cylindrical metal shield as shown, and encounter the first grid, which is usually about 100 or 150 volts positive with respect to the cathode and shield. Some of the electrons, perhaps half of them, pass through the grid openings and cause no secondary emission, but the remainder strike the sensitised grid wires and eject secondary electrons, the number of which depends upon the secondary emission factor of the wires. A similar process occurs at each successive grid, with a repeated multiplication of electrons, until those leaving the last grid G_6 shoot through the wide-mesh collector grid and strike the secondary emission anode A, liberating there the final stream of electrons. This stream is collected by the collector grid and corresponds to the current in the load resistance R. It will be

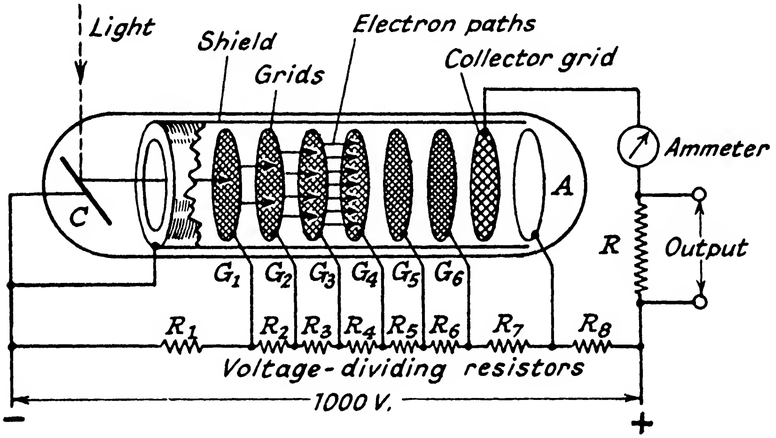


FIG. 24.—Arrangement of grid type electron multiplier showing first three stages of multiplication.

noticed that the collector grid has a higher potential than the anode. The potentials of the various electrodes are determined by the values of the respective resistances R_1, R_2, etc., and the difference of potential between adjoining grids is generally of the order of 100 volts. The high voltage required for operating the unit is readily obtained from A.C. supply mains by means of a transformer and rectifier with an output of about five milliamperes at 1,000 volts.

Electron multipliers have been developed extensively for television work, where they are valuable for amplifying very small photoelectric currents due to minute changes of illumination without introducing the complications associated with high gain valve amplifiers. A tube with an overall multiplication of 5,000 will produce an output current of one milliampere from an initial photoelectric current of a fifth of a microampere, i.e., one five-millionth of an ampere. If the photoelectric cathode in the tube had a sensitivity of forty microamperes per lumen this would mean that the output current of one milliampere would be obtained with an illumination of five-thousandths of a lumen.

PHOTOCONDUCTIVITY

On 4th February, 1873, the prominent telegraph engineer Willoughby Smith addressed a letter to Latimer Clark, Vice-President of the Society of Telegraph Engineers (now the Institution of Electrical Engineers) describing experiments carried out by himself and his assistant May in the course of which it was found that the electrical resistance of selenium underwent changes in accordance with variations of illumination. At about this time telegraph electricians were looking for a substance with a high electrical resistance for use in tests on submarine cables, and Smith was using some sticks of selenium for the purpose. His instruments showed very erratic results, which were found to be due to changes of resistance of the selenium with changes of daylight intensity. When exposed to bright sunlight the resistance of the sticks was much lower than when the sunlight was obscured.

Here was a most surprising and unaccountable discovery, and in view of the state of electrical knowledge at the time, long before any idea of electron theory, it is not surprising that it caused a stir. At a meeting of the Society of Telegraph Engineers some few years later, the newly invented microphone of David Hughes was being discussed and someone remarked that with its aid the walking of a fly could be heard like that of a horse crossing a wooden bridge. Smith, who was present at the meeting, mentioned some of his own work on microphonic effects and then, thinking of his experiments with selenium, went on " But I can tell you something which to my mind is still more wonderful ; by the aid of a telephone I have heard a ray of light fall on a bar of metal."

It is now known that many other materials show a change of resistance with illumination, but selenium has the most marked properties in this respect and is used exclusively in photoconductive devices. The element selenium was discovered by Berzelius in 1817 as an impurity in sulphuric acid, but although it was extensively studied because of its rather peculiar chemical properties there was very little electrical investigation until Willoughby Smith's discovery. Following this there was a great revival of interest in the substance itself and an immense amount of experimental work directed towards making practical use of the photoconductive effect.

Properties of Selenium

As selenium has several chemical forms it is always necessary to know which kind is meant when discussing its properties. It is only the grey crystalline form which is a conductor (though a very poor one) and sensitive to light. Some of its physical properties are given in the following table :—

Density (grams per cm. cube)	4.8
Melting Point (^{0}C)	220
Boiling Point (^{0}C)	690
Resistivity (ohms per cm. cube)	70,000
Linear expansion (per ^{0}C)	4.9×10^{-5}

It may be pointed out that the resistivity is very high in relation to ordinary metallic conductors. Copper, for example, has a resistivity of 1.7 *microhms* (millionths of an ohm) per cm. cube, so that the resistance of a piece of crystalline selenium is something like forty thousand million times that of a piece of copper of the same dimensions. Even this metallic form of selenium should properly be regarded as a semiconductor, and it will be understood that the resistances associated with selenium cells are high ; generally several million ohms.

Theory of Photoconductive Effect

The observed effect of light on selenium may be explained readily on the basis of modern physics, which regards a beam of light as a shower of corpuscles or *quanta* ; the energy of each of these entities, known as *photons*, being determined by the wavelength and frequency of the light, in accordance with Planck's hypothesis. According to this view we may suppose that the energy represented by the incidence of photons comprising a beam of light directed on to selenium may so disturb the surface atomic structure that some of the bound electrons are removed from their normal atomic orbits and added to the normal number of free electrons which exist under conditions of darkness and which determine the dark conductivity of the material. This is a reasonable explanation of the increase in conductivity caused by illumination.

It is reasonable to suppose that the number of free electrons produced in this way with light of a given wavelength will be proportional to the light intensity. The production of each free electron will also naturally result in the creation of a residual positive ion. It must be supposed that the recombination of electrons and ions is continually taking place in the material while it is illuminated, and it is evident that under these conditions a steady value of conductivity is reached only when the rate of production of ions and electrons is equal to their rate of recombination.

It may be doubted whether the simple theory we have outlined is truly representative of all the phenomena taking place in selenium. Recent work on the subject has shown that it is necessary to recognise the existence of a primary effect and resulting secondary effects of considerable complexity which are not yet fully understood. The secondary effects are comparatively slow, and severely restrict the use of photoconductive devices where rapidity of response is important.

Types of Cells

There is some confusion of nomenclature referring to photoconductive devices, which are referred to both as cells and bridges. The latter term describes more accurately the construction of most devices of the kind, in which the photoconductive material (usually selenium) bridges a gap between two electrode structures. The change of resistance between the electrodes under the effects of changing illumination may be utilised in a variety of ways to produce an electrical response governed by these changes.

Early types of bridge were inconsistent in their behaviour, and subject to sudden failure. A wide variety of forms appeared at different times, embodying various principles of manufacture, and it became very evident that one of the most important considerations in the manufacture of efficient types was the development, with long experience, of a satisfactory technique.

A typical modern construction is shown in Fig. 25. In this type the electrodes are formed by two thin metal grids with projecting teeth fused on to a thin glass plate. The teeth of the respective grids are interspaced without causing electrical contact or producing a high capacity. Molten selenium is applied to the surface and then thermally treated to obtain the required crystalline structure. The thickness of the selenium film is normally only a few hundredths of a millimeter, so as to eliminate as far as possible the shunting effect of any inactive material. The glass bulb is evacuated and then filled with an inert gas. Bridges of this type are made in different grades, according to the intended application, and have dark resistances ranging between 0.5 and 20 megohms.

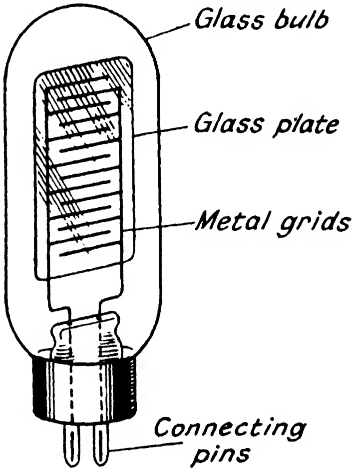


FIG. 25.—Arrangement of selenium bridge.

Use of Amplifiers

As it is possible with a selenium cell to obtain a current change of the order of a tenth of an ampere between dark and light conditions there are many applications, such as the operation of sensitive relays, where no amplification is necessary, and the control system may, therefore, be made very simple. In cases where a greater output is required a valve amplifier may be used in much the same way as with an emission photocell.

A simple form of circuit is shown in Fig. 26, where the bridge or cell is shown as a resistance *Se*. With the connections arranged in this way the voltage at the grid of the valve for a given position of the tapping on the battery B_1 will be determined by the relationship between the resistances of *Se* and R_1. If it be assumed that the bridge resistance is high in relation to R_1 then the greater proportion of the voltage will be dropped across the bridge. This will normally correspond to the dark condition. As the bridge becomes illuminated its resistance decreases and the voltage across it falls, while that across R_1 increases. It can be seen that this action changes the voltage between the cathode and grid of the valve, thus controlling the anode current in accordance with changes of illumination on the selenium bridge. By suitable arranging the polarity of battery B_1 and the value of R_1 it is possible to make the anode current of the valve either increase or decrease with increasing illumination on the bridge.

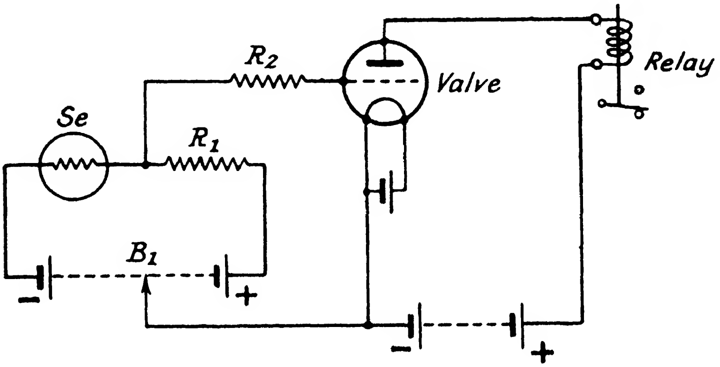


FIG. 26.—Amplifier circuit for selenium bridge.

Similar principles are equally applicable to amplifiers working from D.C. or A.C. mains, and more than one valve can be used where greater amplification is needed. If considerable power output is required, the grid-controlled rectifier or gasfilled triode can be used in place of, or in addition to, the vacuum valve. (See p. 128.)

The range of applications of the selenium bridge is similar to that of emission cells, except for cases requiring a high speed of response, as in sound film and television work, where the time-lag of selenium is a severe handicap.

PHOTOVOLTAIC EFFECTS

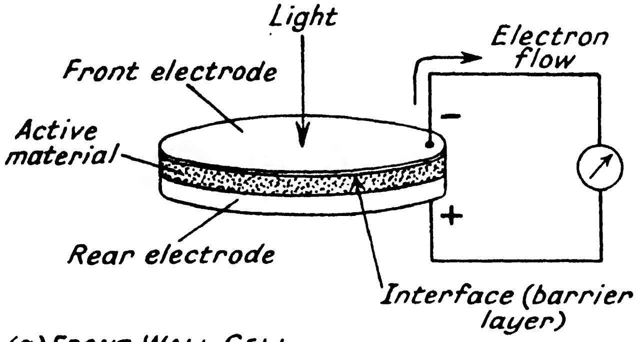
In 1877 it was found by W. G. Adams, Professor of Natural Philosophy and Astronomy at King's College, London, that a small electromotive force appeared in a piece of crystalline selenium when it was illuminated. This was yet another electrical effect of light to be studied by scientists, but some years elapsed before any practical results were obtained. A lot of work on the subject was done by C. E. Fritts in America, and he made some cells which operated on the newly discovered principle. He did a great deal towards developing the proper technique of manufacture, and seems to have been about the first experimenter to use a rotating perforated disc for admitting light to a cell and to observe the changes of tone due to the resulting intermittent current in a telephone receiver when the speed of the disc was varied. A considerable number of other experimenters appeared on the scene, and there was also a good deal of theoretical work, in the course of which a similar effect was found to occur with cuprous oxide—a substance allied to the film which forms on the surface of copper when heated in air.

It was not until about the year 1930 that the copper-oxide rectifier type photocell as we now know it first made its appearance, shortly followed by a similar type employing selenium. These cells differed from the emission and photoconductive types in developing a voltage when illuminated without requiring any external sources of supply. This obvious advantage over the other types is however offset to a large extent by other features, and the use of these latest types of photocell is rather severely restricted in practice to cases where simplicity of circuit arrangements is a primary consideration.

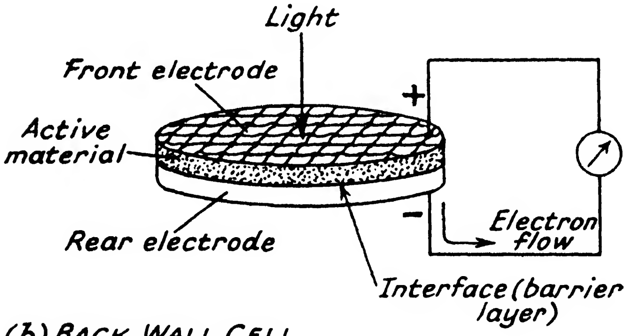
The effect with which we are now dealing may conveniently be described as photovoltaic, and the cells utilising it are often called rectifier photocells owing to the fact that on illumination the current corresponding to the generated voltage flows in a particular direction because of the rectifying properties of the arrangement.

General Theory

As in other photocells, a great deal depends upon manufacturing technique and the preparation of materials. The properties of selenium and copper-oxide as used in rectifier photocells are by no means fixed, and vary considerably according to treatment. It is possible, for example, by thermal treatment to change the electrical resistance of copper-oxide by a factor of more than a thousand without causing any physical change in the material. As the causes of this behaviour are not properly understood it is evidently necessary to exercise considerable caution in putting forward theories to account for the



(a) FRONT WALL CELL



(b) BACK WALL CELL

FIG. 27.—Alternative types of rectifier photocell.

effects observed. All that is known with certainty is that when the junction or interface between the semiconducting material (selenium or copper-oxide) and the conductor with which it is in contact is illuminated, a current flow is set up in the combination, and this current flows in a definite direction.

There are two ways in which the combination may be arranged so as to produce the effect, as shown in Fig. 27. In the arrangement shown at (*a*) the active material is deposited on a metal plate (usually copper in a copper-oxide cell and iron in a selenium cell) forming one electrode. The other electrode is a very thin film of metal so thin as to be transparent, which is sprayed on to the surface of the active material. When light falls on this transparent front electrode its energy is able to loosen electrons in the active material in the vicinity of the interface and these electrons flow from the active material to the front electrode and through the external circuit as shown. Just *why* the electrons should behave in this manner is not easy to explain, but it is known that the resistance to their flow in the other direction is generally some thousands of times greater. This arrangement is known as a front-wall cell, and the interface where the action takes place is called the *barrier layer.* A cell arranged on the principle of Fig. 27 (*a*) is consequently known as a *front-wall* barrier layer photocell.

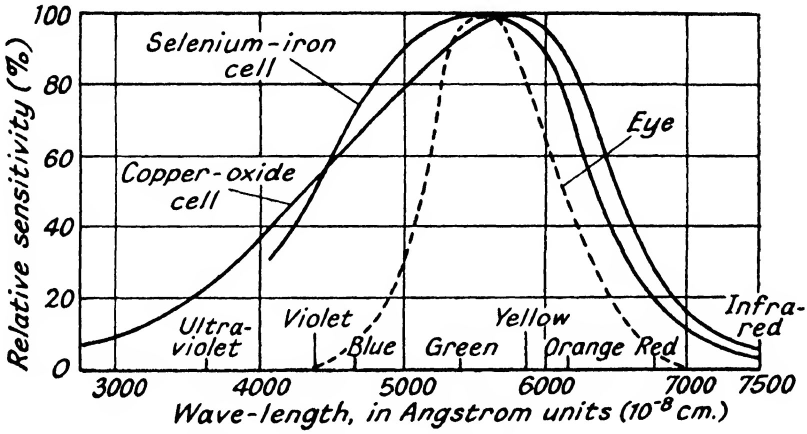


FIG. 28.—Relative sensitivity of selenium and copper-oxide rectifier cells.

An alternative arrangement is shown at (*b*) in Fig. 27, and represents what is called the *back-wall* construction. In this case the front electrode takes the form of a wire grid or a metal plate with openings for admitting light to the active material with which it is in contact. The seat of the action causing the electron flow in this type of cell is at the interface between the active material and the rear electrode so that the light has to penetrate the active material instead of only a transparent front electrode as in the case of the front wall cell. The direction of electron flow is from the active material to the rear electrode, which is accordingly negative according to convention whereas in the front-wall cell it is positive because the electrons flow from the active material to the front electrode.

Commercial types of rectifier cells are usually mounted inside a moulded case with a suitable window for admitting light, but any kind of mounting may be used to suit the particular application. The elements may be either circular or rectangular, and as they do not need to operate in a vacuum they are relatively robust. With modern methods of manufacture and ageing before use the cells are very stable in operation and have a practically unlimited life.

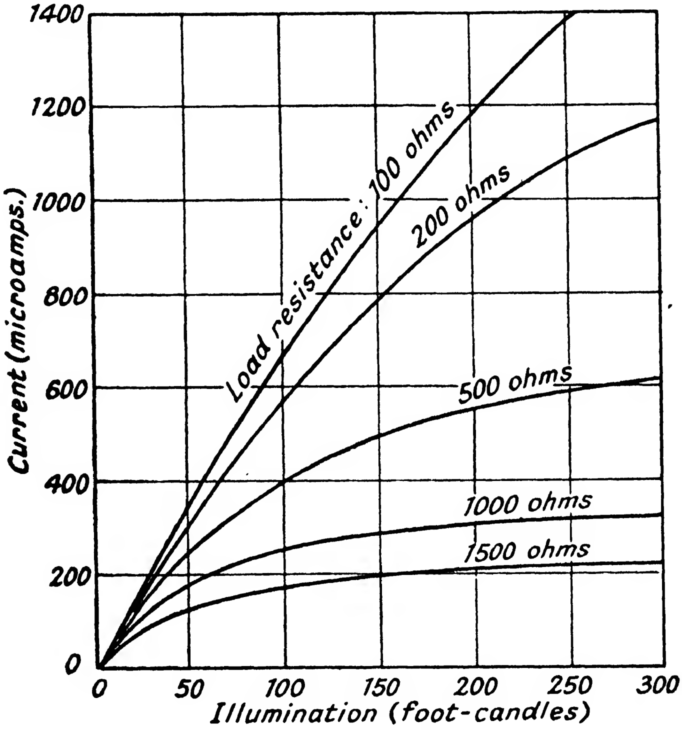


FIG. 29.—Output curves of selenium-iron cell for different load resistances.

Applications

One of the most successful uses for rectifier cells has been in portable illumination meters and exposure meters for photographic purposes. The spectral sensitivity, i.e., the relationship between the output and the colour of the incident light, is shown in Fig. 28 for both the selenium and copper oxide types of cell. The sensitivity curve for normal human vision is also given for comparison, and it is seen that the maximum response is similar in all three cases. The agreement can be improved by the use of suitable filters, and is a valuable feature of rectifier cells in colour matching and similar work.

Typical output curves for a modern front-wall selenium-iron cell, showing the relationships between illumination and output current for various values of load resistance, are given in Fig. 29. With low resistances the response is almost linear, so that equal increases of illumination give almost equal increases of current over a considerable portion of the curves. This feature is of value in illumination meters and simplifies the corrections necessary when a simple current-indicating instrument is used to show the amount of illumination. The power output with a given illumination depends on the load resistance, and is a maximum when this is equal to the cell resistance. Although the current is much greater than that obtained from an emission cell it is not sufficient to operate a robust relay directly. The current can be increased by connecting cells in parallel, but this is not often practicable.

Use of Amplifiers

Owing to the technical difficulties in producing a reliable amplifier for use with a D.C. signal voltage this arrangement is not used if it can be avoided. One way of avoiding it is to interpose between the light source and the photocell a perforated disc running at constant speed so that the illumination of the cell is intermittent and an intermittent voltage accordingly appears at its terminals ; the frequency being determined by the speed of the disc and the number of perforations or slots, which are of course equally spaced. A conventional type of A.C. amplifier may then be used.

Such an arrangement is not always practicable, and it may be necessary to use a special form of amplifier in which the best use is made of the very small input voltage. No particular difficulties are encountered if the output is only for operating a measuring instrument.

An important point arises in the use of rectifier cells owing to their construction. It will be seen from Fig. 27 that the respective surfaces between which the voltage is generated are of large area and very close together. The cell thus has considerable capacitance, the effect of which is to delay the voltage changes across the cell caused by changes of illumination. For steady or slowly varying illumination the effect is not apparent, but when the illumination is rapidly changing, as in sound film and television work, or when a rotating disc is used, the capacitance of the cell causes a loss of response which gets worse as the frequency of variation increases. The trouble may be overcome to some extent by making the cells as small as possible, so as to reduce the area and hence the capacitance for a given separation of the surfaces. There is evidently a limit to size reduction, and although rectifier cells for acoustical work have been made as small as two millimetres square it is usual to incorporate in the circuit special arrangements for increasing the output at the higher frequencies. Otherwise the high notes corresponding to these frequencies would be lost and the sound reproduction would be very imperfect.

THE BECQUEREL EFFECT

It was found many years ago by Henri Becquerel that a particular type of electrolytic cell had the curious property of producing a voltage between its terminals when illuminated. Its action is thus similar to that of rectifier type photocells, and it is indeed possible that the fundamental processes are the same.

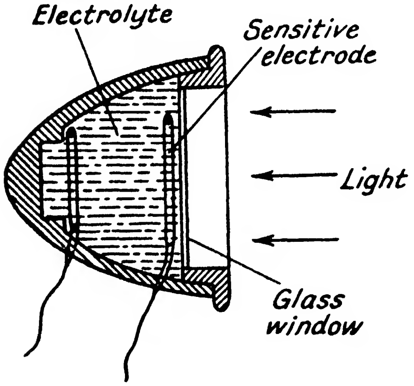


FIG. 30.—Sectional view of Becquerel photocell.

The electrolytic or Becquerel photocell is not widely used because of its inherent disadvantages in comparison with other types and its lack of any compensating advantages. The general arrangement of the cell is shown in Fig. 30 and is seen to resemble that of a small battery. The liquid electrolyte is generally cuprous oxide in a solution of potassium sulphate. Light entering the glass window causes a one-way flow of electrons between the sensitive electrode and the electrolyte and a corresponding flow of current in the external circuit connected to the electrodes. The cell is quite sensitive in comparison with other types, but is not stable and has the rather serious defect of deteriorating with time, even when not in use.

FOR FURTHER READING

The books listed below are representative of the various branches of photoelectricity :—

1. T. J. Fielding. *Photo-Electric and Selenium Cells.* (Chapman and Hall), 1940. A good introductory book.
2. R. C. Walker and T. M. C. Lance. *Photoelectric Cell Applications.* (Pitman), 1938. Gives general principles and describes wide range of applications of emission type cells.
3. N. R. Campbell and D. Ritchie. *Photo-Electric Cells.* (Pitman), 1934. Deals with physical aspect and theory rather than applications. Concerned only with emission cells.
4. B. Lange. (Trans. by A. St. John.) *Photoelements and their Application.* (New York, Reinhold ; London, Chapman and Hall), 1938. Deals exclusively with cells of the barrier-layer type.
5. A. L. Hughes and L. A. Du Bridge. *Photo-Electric Phenomena.* (New York, McGraw-Hill), 1932. A more advanced work, giving comprehensive treatment.
6. V. K. Zworykin and E. D. Wilson. *Photocells and their Application.* (New York, Wiley ; London, Chapman and Hall), 1932. Similar in scope to No. 5 above
7. W. Summer. *Photocells in Industry.* (" Electronic Engineering " Monograph), London, 1947. Gives numerous practical applications.

CHAPTER VI

THERMIONIC VALVES

We have already mentioned in Chapter III that early experimental work on thermionic emission led to the development of the first thermionic valve by Sir J. A. Fleming. In 1904 he took out his fundamental patent covering the two-electrode valve or *diode* and its applications, among which the most important was its use in rectifying the alternating currents induced in wireless aerial systems so that the resulting unidirectional currents might be used for producing intelligible signals in microphone receivers. This was one of the most important developments in applied science for many years, and the recognition of this fact led to increased interest in the possibilities of such arrangements.

The next important step came in 1907, when L. de Forest introduced a third electrode into the Fleming type of valve, thus transforming the diode into a *triode*. This electrode took the form of a metallic mesh, or *grid* as it was called, located between the cathode and the anode. It enabled the electron flow between these electrodes to be controlled, and thus enormously increased the practical possibilities of the device. It is probably not unfair to say that this development was in the nature of a step in the dark, as it transpired eventually that at the time of his invention de Forest had only the haziest of notions of what really took place inside his valve.

The thermionic valve, as represented by these early developments and as known in its numerous forms today, employs two principles : the production of electrons by thermionic emission and the control of these electrons by electric fields. From its beginnings in the hands of Fleming and de Forest during the early days of radio it has been the subject of continuous improvement, and several entirely new types have been produced, with the result that valves having two, three, four or even five grids are now a commonplace. The function of different types has also been included within a single envelope for special purposes, so that it is now possible to have, for example, a triode and a diode within one envelope performing different functions and yet utilising a common cathode.

Modern Construction Methods

There are various ways in which the electrodes may be arranged in a valve and sometimes special constructions are employed, according to the intended application. The construction shown in Fig. 31 is very much simplified, and is intended to show the principal parts of a triode rather than to illustrate any particular design. The anode and grid (or grids) are usually of nickel or molybdenum, and there is a wide variety of possible shapes for both these members. Instead of being a disc as shown, the anode may completely surround the grid and cathode, and in many modern valves the grid takes the form of a spirally wound cage enclosing the cathode. The various electrodes are carried on supports embedded in a glass pinch formed by closing together the ends of a tubular glass projection in the base of the bulb while in the plastic state during manufacture. The connections from the electrodes to the contact pins in the base of the valve also pass through this pinch. The cathode may be either directly or indirectly heated, as already explained in connection with thermionic cathodes (see p. 42).

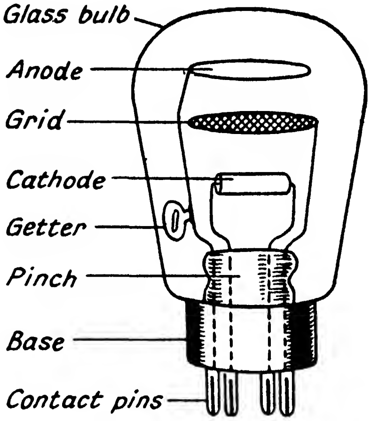


FIG. 31.—Simplified arrangement of triode valve.

One of the most important points in the manufacture of high vacuum valves is the extraction of every trace of gas before the bulb is sealed. The greater part of the gas is removed during the pumping operation, but some is left behind, locked up in the metal parts themselves and in the glass. This occluded gas, as it is called, has to be eliminated by special methods, including the application of heat and the process known as *gettering*. The actual technique naturally varies among different manufacturers, but it is usual to heat the electrodes to about 900° C in a vacuum furnace, after which the pumping process is commenced to evacuate the bulb. The final degree of vacuum depends upon whether gettering is afterwards employed. Baking takes place during pumping at a temperature slightly less than the melting point of the glass, and before gettering is done to clear up occluded gas the metal parts may be raised to a high temperature either by high-frequency eddy current heating or by heating the cathode and applying suitable potentials to the anode and grid to cause heating by electron bombardment.

The gettering process is carried out by fusing a short length of magnesium or barium metal inside the tube either after it is sealed or during pumping. The getter is attached to a metal support in the tube structure and fused from outside by the application of a high frequency alternating magnetic field which sets up sufficient heating by induced eddy currents in the strip to vaporise it. This action effectively absorbs any residual gases, and incidentally is the cause of the silvery appearance of the inside of some valves.

THE VACUUM DIODE

An elementary circuit for demonstrating the properties of the diode is given in Fig. 32. Attention may be drawn to the similarity of this circuit to the one in Fig. 13 (page 47) relating to the passage of electrons from one electrode to another through a vacuum. In both cases the anode is positive with respect to the cathode and collects electrons ejected from the cathode, but not all of them, because of the effect of space charge which we have already considered (page 48). We have also seen that if the emission is increased a stage is eventually reached where no appreciable increase of current takes place because of the space charge effect.

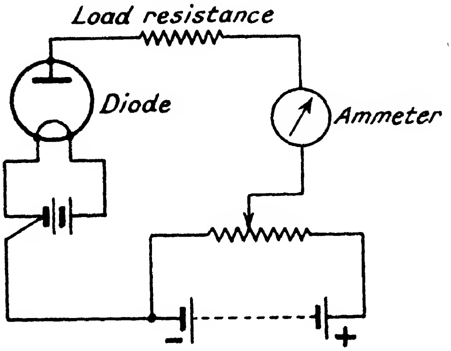


FIG. 32.—Circuit for showing properties of diode.

A similar limitation in anode current occurs if the emission is constant and the anode voltage is gradually raised. In the case of the diode, for example, the effect can be seen by gradually raising the anode voltage and observing the resulting increase in anode current as shown by the meter. With low anode voltages only a small proportion of the electrons emitted by the cathode manage to reach the anode because of the retarding effect of the negative space charge and the anode current is correspondingly small. As the anode voltage is increased the anode is able to exert an increasing attraction over the electrons until eventually even the electrons forming the space charge are swept to the anode. At this stage the anode current is a maximum, and corresponds to the total cathode emission. The conditions are shown by the curve in Fig. 33. The saturation current is sometimes called the *emission* or *emission current* and can be used as a criterion of performance of a valve although it is more usual to measure a lower value of emission so as to avoid damaging the valve. Most modern valves have a maximum cathode emission considerably in excess of the value required to give maximum rated anode current.

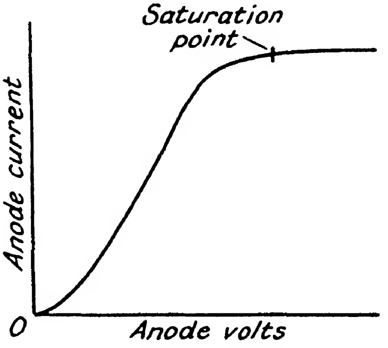


FIG. 33.—Characteristic of vacuum diode.

It is evidently desirable to have only a small space charge in a valve, because then the smallest value of anode voltage is required to produce a given anode current. This feature can be obtained by making the distance between anode and cathode as small as practicable. In some valves this distance is only about two hundredths of an inch.

The Diode as a Rectifier

It has been assumed up to now that the anode is always positive with respect to the cathode. If the anode is made negative, as by changing over the anode battery connections in Fig. 32, electrons will be repelled by the anode instead of attracted to it, and there will be no anode current. It follows that the diode has only one-way conductivity. If the anode battery in Fig. 32 is replaced by a source of alternating voltage it can be seen that anode current will flow in the valve only during the half of each alternating cycle when the anode is positive. As the alternating voltage is continuously changing in value with respect to time, rising from zero at the beginning of the cycle to a maximum and then falling to zero again, the resulting anode current in each cycle will also have a varying value. The resulting current in the anode circuit will thus be a series of pulses, as shown in Fig. 34, which also includes a simplified connection diagram. As the anode voltage rises from zero at the beginning of each positive half-cycle the anode current also rises, gradually at first, as will be seen from Fig. 33, and then more or less following the shape of the voltage wave because of the uniform relationship between the voltage and current over the straight portion of the curve, assuming that the region of saturation is not reached at any point. During the time when the anode is negative, i.e., during the negative half-cycles of anode voltage, no current can flow, and the successive pulses of current therefore occur regularly at half-cycle intervals. This is the most elementary form of *rectification*, the conversion from alternating to direct current. The kind of direct current it produces, represented by recurring pulses, is known as half-wave rectification, and although satisfactory for applications such as battery charging and electroplating is not sufficiently similar to the conventional D.C. produced by a battery or generator to be suitable in other cases.

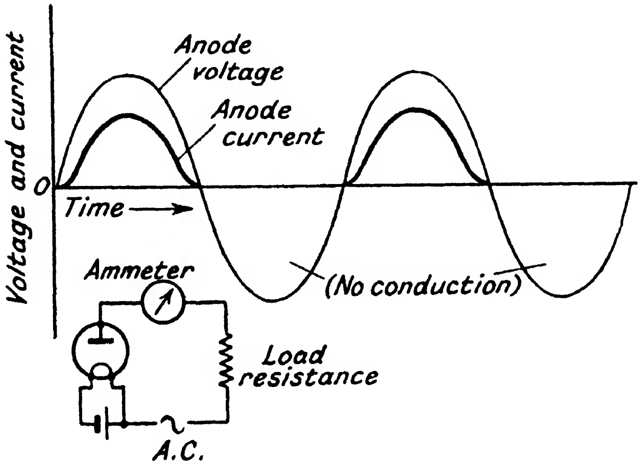


FIG 34.—Use of vacuum diode as half-wave rectifier.

Full-wave rectification, in which use is made of both half-cycles of each wave of alternating voltage, is a preferable and widely used arrangement. It requires two diodes, which may be contained in one tube (the so-called double-diode valve) and a transformer with a centre-tapped secondary winding. The circuit arrangement is shown in Fig. 35. In this case two diodes V_1 and V_2 are shown. The transformer has a primary winding P connected to the A.C. supply and two secondary windings S_1 and S_2 connected to the anodes of the respective valves. The load circuit it represented by the resistance R connected between

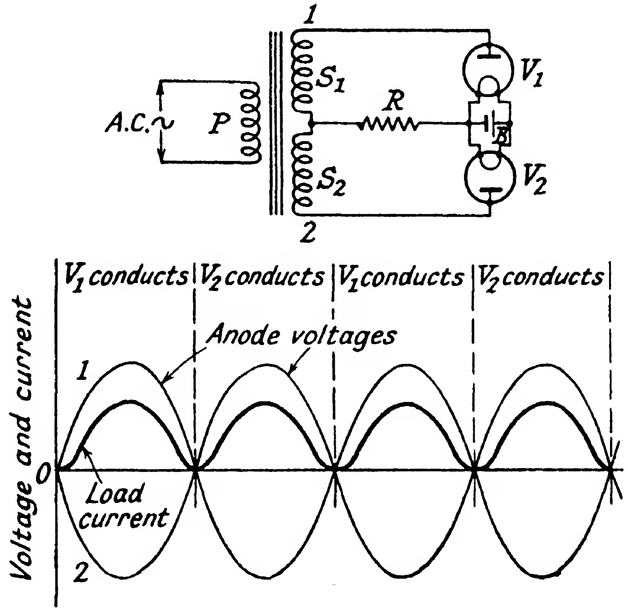


FIG. 35.—Use of two vacuum diodes in full-wave rectifier.

the centre tapping of the transformer secondary and the valve cathodes, which are joined together and supplied from a battery as shown. The cathodes may of course be directly or indirectly heated, and may be supplied from a suitable secondary winding on the transformer; this being the more usual arrangement. Windings S_1 and S_2 are arranged so that their outer ends (connected to the valve anodes) are always of opposite polarity. If at the commencement of a cycle, for example, point 1 is positive, then point 2 will be negative by the same amount at any given instant. This opposition will hold good throughout each cycle, a changeover taking place in the respective polarities at each zero of voltage.

The conditions are represented by curves 1 and 2 in Fig. 35. During the half-cycles when point 1 is positive, valve V_1 conducts, but not V_2, because its anode is negative during these periods. At the end of the half-cycle, however, point 2 becomes positive, and point 1 goes negative, so that valve V_2 now conducts while V_1 remains out of action. It is seen that with this arrangement current flows in the load circuit during each half-cycle instead of during alternate half-cycles as in the case of Fig. 34, and the process is accordingly known as *full-wave rectification.*

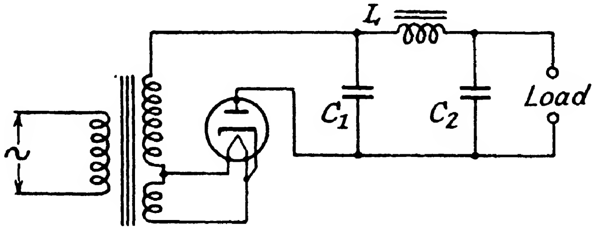
In all rectifier circuits the wave shape of the rectified current and voltage is determined by the circuit arrangements. In the cases we have so far considered, the load circuit consists only of resistance, so that the relationship between current and voltage is always constant and the curve of rectified voltage has the same shape as that of the current. When the circuit contains inductance or capacitance, however, this relation no longer holds owing to the property of the circuit of storing energy, so that the current at any instant is no longer directly proportional to the voltage at that instant. These properties of inductance and capacitance are in fact utilised to smooth out the current and voltage peaks by suitably arranging choke coils (inductances) and condensers (capacitances) in the form of *filter* or *smoothing circuits.*

Filter Circuits

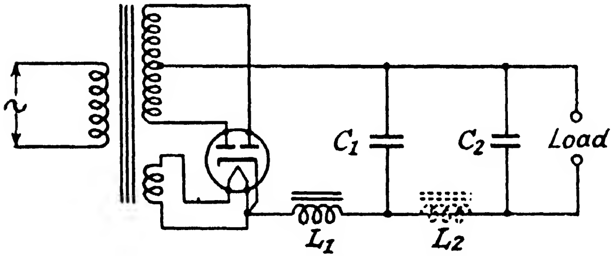
A condenser has the property of offering its minimum opposition to the flow of current when the voltage applied to its terminals changes at a rapid rate. Conversely, when the voltage changes slowly the opposition to the current flow is very high, so that in the extreme case where the voltage does not vary at all, no current can flow. It follows that a condenser connected across two points between which the voltage varies will tend to stabilise the voltage by smoothing out the changes. If a rise of voltage occurs, the condenser absorbs a charge of electricity, and this is released again when the voltage falls. In absorbing the charge, the condenser provides a path for current flow between the two points, the resistance of the path being low for rapid changes and high for slow changes.

A choke coil has similar properties in relation to current changes, and when connected in series with a circuit will tend to stabilise the current by smoothing out any changes. This action arises from the fact that in a choke coil, which is in effect an electromagnet, the change in the magnetic field produced by a change of current generates a voltage in the winding which always opposes the change of current.

In a filter circuit these properties are employed to smooth out current and voltage changes in the output of a half-wave or full-wave rectifier by connecting one or more condensers across the output and one or more choke coils in series with the output.



(a) Condenser input filter



(b) Choke input filter

FIG. 36.—Alternative forms of filter or smoothing circuit.

Two typical filter circuits applied respectively to a half-wave and full-wave rectifier are shown in Fig. 36. The circuit at (*a*) has a condenser C_1 connected across the rectifier, followed by a choke coil L and a second condenser C_1. This arrangement is called a *condenser input filter*, to distinguish it from the *choke input filter* in circuit (*b*), where the choke L_1 comes before the first condenser. The choice between the two arrangements is usually determined by design considerations involving the amount of power rectified and the voltage drop caused by the filter circuit. In Fig. 36 both circuits employ indirectly heated valves, the heaters being operated from a secondary winding on the transformer in each case. The valve in circuit (*b*) is a double-diode.

THE VACUUM TRIODE

By introducing a grid between the cathode and anode of a vacuum valve and applying to this grid a suitable variable voltage with respect to the cathode it is possible to control or vary the rate of flow of electrons from cathode to anode and thus control the anode current in the external circuit merely by variation of the grid voltage. When the grid is negative with respect to the cathode, this being the usual condition, it repels any electrons in its vicinity. This repelling action, which tends to prevent electrons reaching the anode, is proportional to the negative grid voltage. If this voltage is sufficiently high, no electrons will succeed in penetrating the region of the grid and the anode current will be zero. As the electrons are also acted upon by the attraction of the positive anode, the grid voltage required to cut off the anode current, or in fact to reduce it to any particular value, will depend upon the anode voltage. If the anode voltage is high, so that electrons tend to be strongly attracted to the anode, a correspondingly higher value of negative grid voltage will be necessary to reduce the current to a given value. From this it will be seen that in every valve definite relationships exist between the anode voltage, the grid voltage and the resulting anode current.

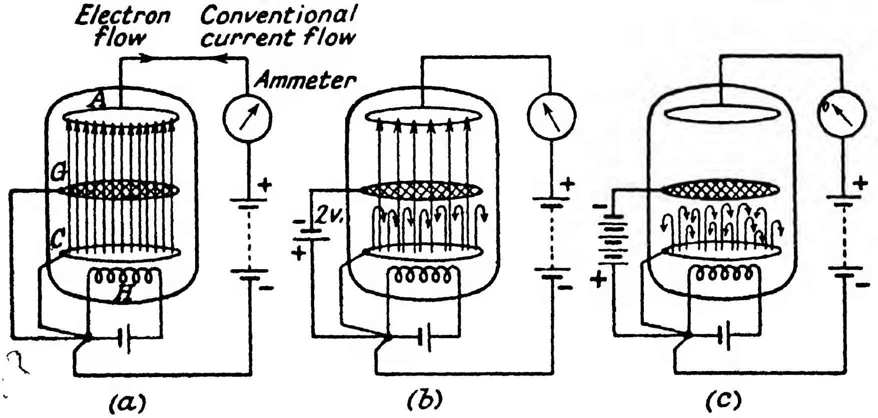


FIG. 37.—Effect of grid voltage on electron flow and anode current in vacuum triode.

The drawings in Fig. 37 will help to explain the interactions resulting from the different voltages and should make clear the basic principles of the triode valve. The diagrams represent successive internal conditions caused by changes in the grid voltage with the anode voltage kept constant. In diagram (*a*) the grid *G* is connected directly to the cathode *C*, which is of the indirectly heated type, the heater being shown at *H*. The direct connection of grid to cathode means that there is no grid voltage with respect to the cathode, from which reference point all the voltages are measured, and that consequently the grid has no effect on the electrons leaving the cathode. The result is an unimpeded stream of electrons from cathode to anode *A*, as shown by the arrows and through the external circuit, causing the current meter (usually a milliammeter) to give an appreciable reading. The anode current thus indicated is a maximum for the particular anode voltage and, assuming a fixed emission rate from the cathode, can only be increased by increasing the anode voltage; a limit being reached by the saturation point as Already described in connection with the diode. A triode connected in this way is in fact virtually a diode for the purpose of the present explanation.

If now a 2-volt battery is inserted in the grid lead so as to make the grid negative by two volts with respect to the cathode as shown in diagram (*b*) the grid is able to exert a repelling effect on the electron stream and some of the electrons are turned back from the region of the grid as shown by the curved arrows. Only those electrons which approach the grid with a sufficiently high velocity are able to shoot through it and complete their journey to the anode under the effect of its attraction. They constitute a smaller current in the anode circuit, and the meter reading is accordingly lower than in diagram (*a*). If further batteries are added in the grid circuit as shown in diagram (*c*) a stage is reached when *all* the electrons approaching the grid are turned back by its repelling action. None of them reach the anode, there is consequently no anode current, and the meter reading is zero.

In the foregoing it is assumed that the anode voltage remains constant. If it is raised, the attraction of the anode for electrons is increased and an increased anode current will result. The original conditions can then be restored only by a suitable increase in the grid voltage. Conversely, a reduced anode voltage gives less anode current, but the original value can be regained by a suitable reduction of grid voltage.

It may be mentioned here that the response of the valve to changes of voltage is for all normal purposes instantaneous, being represented by the extremely minute interval of time required for an electron to traverse the space between cathode and anode. In ordinary valves the speed of response is such that the anode current will faithfully follow changes of grid voltage occurring several million times per second.

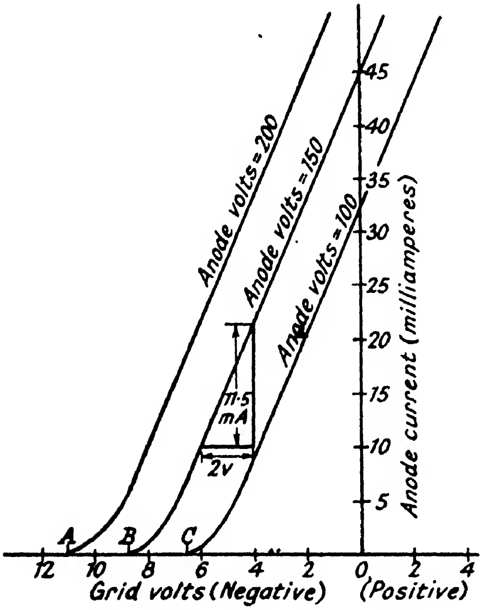


FIG. 38.—Mutual characteristics of vacuum triode.

General Characteristics

The effect of grid voltage on anode current is conveniently shown by a set of curves of the type shown in Fig. 38, generally called the *mutual characteristics.* They are determined by a series of measurements with a circuit similar to that in Fig. 37 except that means are provided for smooth variation of grid voltage so that it may be adjusted to any desired value. It is also usual to continue the curves into the region corresponding to positive grid voltages, although it is the negative region that is almost exclusively used. If the grid of a valve is allowed to become positive it acts as an anode, attracting electrons and passing a considerable current through the grid circuit. This action would be highly undesirable in many cases as it may cause serious disturbance in the circuit. Under normal conditions the grid is always negative, and passes only a minute current, generally much less than a millionth of an ampere (microampere). It is usual to give the mutual characteristic for several different values of anode voltage, and it is seen from Fig. 38 that the respective curves are substantially parallel with each other.

It will be seen from the curves that a change of grid voltage causes a change of anode current by an amount which depends upon the steepness of the curves, so that this factor, representing the rate of change of anode current with grid voltage, is a very important one. It may be specified exactly for the straight part of the characteristic, which is the part generally used in practice, by measuring horizontally a distance corresponding to a definite grid voltage, say 2 volts as in Fig. 38, and observing the resultant change in anode current as represented by the vertical line. In the case shown this amounts to 11.5 milliamperes, and the rate of change of anode current with grid voltage is thus 11.5 ÷ 2 = 5.75 *milliamps per volt.* This is the *mutual conductance* of the valve, sometimes called the *slope*, and is denoted by the letter g_m. A valve with high mutual conductance is highly sensitive to grid voltage changes and is employed in cases where maximum anode current change for a given input voltage is required.

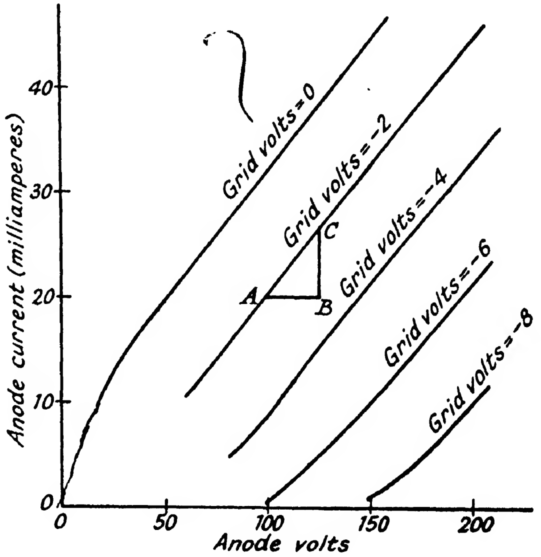


FIG. 39.—Anode characteristics of vacuum triode.

The points *A*, *B* and *C* in Fig. 38 represent the grid voltage for the respective curves which will just reduce the anode current to zero. Such a point is called the *cut-off* point, and a valve which is required to pass no anode current when no input voltage is present will require to have its grid maintained at the cut-off voltage. The valve is then said to be biased to cut-off. The cut-off bias increases with increasing anode voltage.

There is another and sometimes more convenient way of showing valve characteristics. This is by means of curves of the kind shown in Fig. 39, showing the relationship between anode current and anode voltage for various values of grid voltage.

Such curves are called the *anode characteristics*. They may quite readily be drawn by converting the values from a set of mutual characteristics, and simply represent a different way of presenting the same information. The curves in Fig. 39, for example, will be seen on inspection to include the same particulars as those in Fig. 38 as well as giving readings for lower anode voltages.

The rate at which the anode current varies with a change of anode voltage at a given point in the curve when the grid voltage constant is called the *anode conductance*. It is found by the construction shown on the minus 2 volt grid curve in Fig. 39, where *AB* represents a change of anode voltage and *BC* the corresponding change of anode current. The ratio *BC* ÷ *AB* is the anode conductance for that particular point on the curve. Unless straight parts of the curves are considered it is necessary to make *AB* and *BC* very small. A more useful factor is the inverse of the conductance, and is defined by the ratio *AB* ÷ *BC* in Fig. 39. It represents the resistance of the valve at a particular point of operation on the anode characteristic, and is expressed ohms. It is called the *dynamic* or *variational* anode resistance and is denoted by r_a.

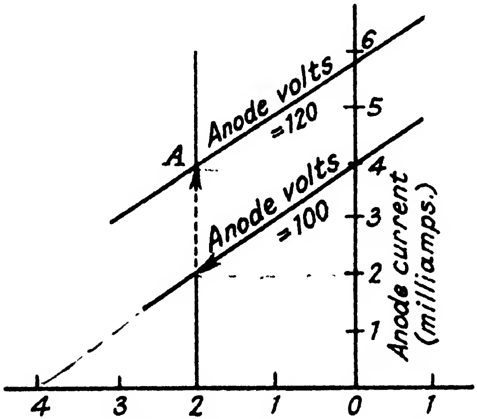


FIG. 40.—Determination of amplification factor.

We have seen that a change of anode current can be brought about either by (1) a change of grid voltage (with constant anode voltage) or (2) a change of anode voltage (with constant grid voltage). The first method gives a much greater change *per volt* than the second. If, for example, a valve has the mutual characteristics shown in Fig. 40 the anode current will be 4 milliamperes with 100 volts on the anode and zero grid volts. Changing the grid voltage to 2 volts negative reduces the anode current from 4 milliamperes to 2. In order to restore the current to its original value (keeping the grid at − 2 volts) the anode voltage must be raised to 120 volts, as shown at point *A* on the 120 volt characteristic, the current changes during the whole operation being as shown by the arrows. It is seen that a 20-volt change of anode voltage is required to compensate a 2-volt change of grid voltage, so that the grid is ten times more effective in producing a given change of anode current in this particular case. The ratio of the change of anode volts to the change of grid volts to produce a given change of anode current is a measure of the effectiveness of the grid control and is called the *amplification factor*. The usual symbol is μ (Greek *mu*) but sometimes m is used. As the factor is the ratio of two voltages it is a pure number and has no units.

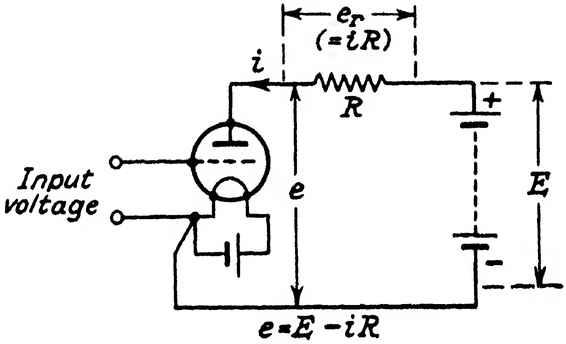


FIG. 41.—Effect of load resistance in triode valve circuit.

Dynamic Characteristics

The curves we have so far considered represent readings taken with steady voltages on grid and anode : a condition which rarely applies in practice as most valves operate with rapidly alternating voltages. It is obviously necessary to know how the operation of a valve is affected by these conditions, and for this purpose use is made of curves showing how the valve behaves under typical load conditions. Such curves are called the *dynamic characteristics*.

It may firstly be pointed out that mutual characteristics as given in Fig. 38 are generally measured with no resistance in the anode circuit, whereas in practice the valve always operates a load of some kind, so that its anode circuit must include resistance. The effect of this resistance can be considered with the aid of the simplified diagram in Fig. 41, where an input voltage applied to the grid causes a current i amperes to flow through the anode circuit. For any value of load current i the voltage drop across the load resistance will be equal to the product of current and resistance and will thus be expressed by $e_r = iR$. It follows that the actual voltage across the valve will always be less than E, the voltage of the anode supply, by the amount represented by e_r, so long as anode current flows. The only time that the voltage across the valve becomes equal to that of the source (so that $e = E$ in the diagram) is when the anode current is zero. There is then no voltage drop across the load resistance.

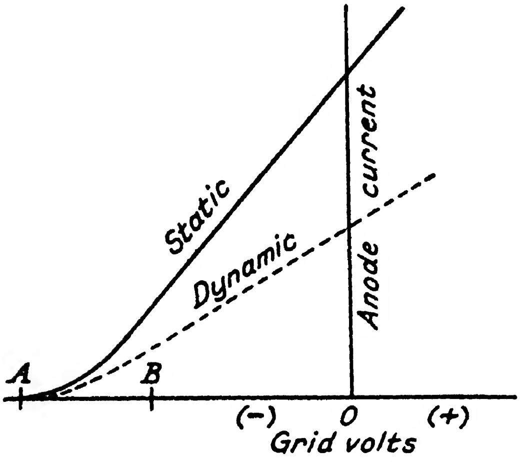


FIG. 42.—Change of triode characteristic due to load resistance.

It is evident that the conditions with a resistance in the anode circuit differ radically from those when no resistance is present, and to distinguish between the two cases it is usual to refer to curves of the kind in Fig. 38 taken without resistance as *static characteristics*, while curves corresponding to a loaded anode circuit are called *dynamic characteristics*. The difference between the two kinds of curves is shown in Fig. 42. With the grid voltage at the cut-off point *A* there is no anode current, and for either curve the voltage at the anode is equal to that of the source. If now the grid voltage is reduced, say to point *B*, the anode current which flows causes a reduction of voltage at the anode of the valve owing to the drop across the load resistance as already explained. There is consequently less anode current than would be the case without load resistance so that for given grid voltages the anode currents corresponding to the dynamic characteristic are always lower than those for the static characteristic; the deviation becoming wider with increasing anode current owing to the increased voltage drop across the load resistance. Both curves are fairly straight except at the beginning, and it is the straight portion which is used in practice, so that anode current changes are as nearly as possible in direct proportion to the changes of input voltage which cause them. It may be noted that the steepness of the dynamic characteristic increases when lower values of load resistance are used, so that eventually the two curves would become identical when the resistance is zero.

Similar considerations apply to the anode characteristics, in which the effect of load resistance on the relationship between anode current and anode voltage may be shown by means of a *load line.* This is a straight line drawn across the curves as in Fig. 43, the slope of the line depending upon the value of load [illegible]t-ance. The lower end of the line, where it meets the horizontal axis, corresponds to zero anode current, and at this point the actual voltage at the valve anode will be equal to that [illegible] supply source. This fixes one point *A* on the load line, which is at 200 volts in the present example. Assuming that the load resistance is 3,000 ohms, the voltage drop across it with say 10 milliamperes anode current will be 3,000 × 0.010 = 30 volts, and the anode voltage at the valve will accordingly be 200 − 30 = 170 volts. This gives another point *B* on the load line corresponding to 170 volts and 10 milliamperes. A straight line drawn through *A* and *B* is the load line for a resistance of 3,000 ohms. For comparison, a load line for 5,000 ohms with an anode voltage of 230 has been drawn on the same curves. It may be noted that the load lines increase in steepness for lower values of resistance and that the position of a load line on the curves is determined by the anode voltage.

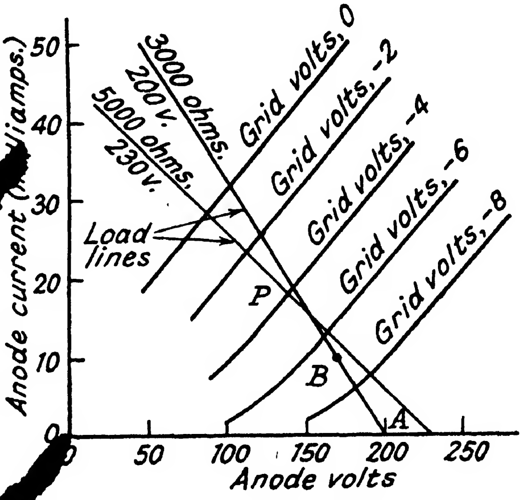


FIG. 43.—Construction of load lines on anode characteristics.

One of the chief uses of a load line is to determine how the anode current varies with respect to time when an alternating voltage is applied to the grid. In the case represented by Fig. 43 the grid would be biased, by a battery or other means, so as to be 4 volts negative with respect to the cathode with no input voltage. The point *P* where the load line intersects the bias voltage is called the *operating point* and is chosen so that as the input grid voltage alternates equally on either side of it the resulting changes of anode current are uniform and free from distortion. In general, this condition is met by not allowing the input voltage to swing beyond the zero grid voltage curve at the upper end (i.e., not allowing the grid to become positive and thus draw a large current) and avoiding the non-linear parts of the curves at the lower end. It can be seen from the curves in Fig. 43 that with the operating point at *P* as shown the corresponding current is about 19 milliamperes. With an input grid voltage having a peak value of 4 volts in either direction the anode current with a 5,000 ohm load will alternate between the limits of 8 and 29 milliamperes. With a 3,000 ohm load the limits are 6 and 32 milliamperes. In both cases the anode current swing is uniform on either side of the operating point.

TETRODES AND PENTODES

The possibility of improving the operation of the triode valve by introducing a second grid was first examined in 1916, but it was not until about ten years later that the advantages of the arrangement became fully realised. The original experiments aimed at improving the valve for radio applications, but the new types of valve which resulted from the work now have uses extending beyond the radio field and are of interest as an alternative to the triode for several purposes. The valves with which we are now concerned are the tetrode, the beam tetrode or beam power tube, and the pentode.

The Tetrode

This valve derives its name from the fact that it has two grids, making four electrodes in all. An early type used one grid for control in the ordinary way, the second being located close to the cathode and maintained at a suitable positive voltage with the object of neutralising the negative space-charge. The use of this *space-charge grid* increased the anode current for a given anode voltage and also increased the mutual conductance. This form of tetrode is not widely used now.

The more usual arrangement is shown in Fig. 44, where G_1 is the control grid and G_2, called the *screen grid*, acts as an electrostatic shield between the control grid and the anode. This screen or shield is maintained at a voltage positive with respect to the cathode, but not so high as the anode voltage. Its function is to reduce the capacitance between the anode and the control grid. This capacitance is very undesirable in high frequency applications as it constitutes an A.C. path between the circuits connected to anode and grid (the conducting power of a capacitance increasing as the frequency increases) and thus provides an undesirable *coupling* between the anode and grid circuits which is very liable to cause instability in operation.

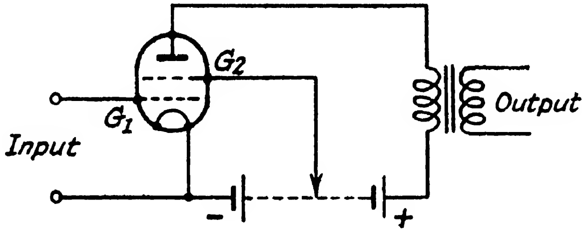


FIG. 44.—Connections of screen grid tetrode.

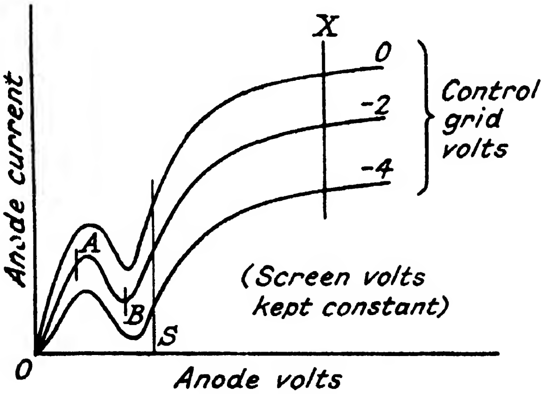


FIG. 45.—Anode characteristics of screen grid tetrode.

As the anode current in the *screen grid tetrode* is determined by three voltages (anode, control grid and screen) a single diagram cannot show all the effects. It is usual to plot curves showing the relationship between anode current and anode voltage for several values of control grid voltage and a particular value of screen voltage. A typical set of curves is given in Fig. 45 for three different values of control grid voltage, the screen voltage being maintained constant at a value corresponding to point *S*. The curious kink at the beginning of the curves is due to secondary emission effects when the anode voltage is less than the screen voltage, that is, over the region *OS*. The commencing portion *OA* of the curve is of normal shape but beyond the anode voltage corresponding to point *A* the electrons arrive with sufficient velocity to cause secondary emission at the anode.

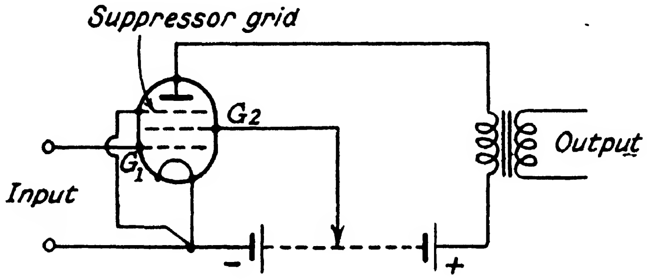


FIG. 46.—Connections of pentode.

As the screen is more positive than the anode it attracts to itself these secondary electrons, which are consequently lost to the anode. This loss is greater than the rate of arrival of electrons from the cathode, and the result is a progressive reduction of anode current with increasing anode voltage over the region *AB*. At point *B*, however, the anode voltage is approaching that of the screen, which therefore fails to collect so many secondary electrons. The draining away of electrons to the screen thus falls off and the anode current rises again. It may be noted that over the region *AB* the valve has a negative resistance. When operated in this region the valve is capable of delivering power to a load and becomes an active instead of a passive circui element. It is then capable of maintaining oscillations in specia forms of circuit for the generation of high frequency waves and is called a *dynatron*.

In normal applications the region *AB* must be avoided and in practice it is usual to operate the valve in the region to the right of *X*, where the anode current is practically independent of anode voltage. In this region the anode resistance is very high, because of the high ratio of a change of anode voltage to the corresponding change of current, and the valve has the property of preventing wide changes of anode current with fluctuations of voltage in the anode circuit.

The Pentode

The name of this valve indicates that it has five electrodes, there being three grids. It may be regarded as a screen grid tetrode with an additional grid, called the *suppressor*, located between the anode and the screen grid and having the same potential as the cathode as shown in Fig. 46. The suppressor grid shields the anode from the screen grid and the negative potential of the suppressor prevents the screen grid attracting secondary electrons away from the region of the anode. These secondary electrons are instead pushed back to the anode, thus eliminating the kink in the anode current curve which occurs in the tetrode.

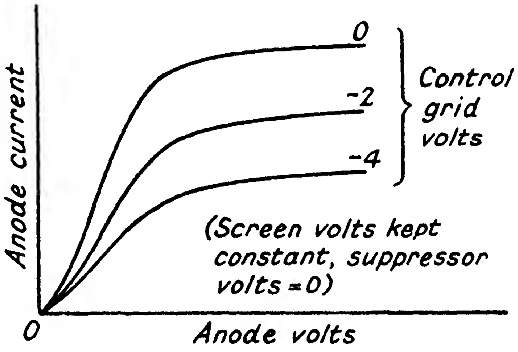


FIG. 47.—Anode characteristics of pentode.

Typical anode characteristics of a pentode are given in Fig. 47, from which it will be seen that the useful range of anode voltage is considerably extended by getting rid of the negative resistance region. The curves correspond to fixed voltages on the suppressor and screen grid, this being the usual method of operation.

The Beam Tetrode

One of the latest developments in valve design arises from recognition of the fact that it is not essential to provide an electrode for the purpose of suppressing secondary emission. The effect may be obtained quite readily by utilising the space-charge of the cathode-to-anode stream of electrons for repelling electrons produced at the anode by secondary emission. To obtain this effect the electron flow is condensed into beams by a special electrode construction such as that shown in Fig. 48. Surrounding the central cathode is the wire cage forming the control grid and outside this is a similar cage forming the screen grid. These cages are aligned so that the paths between them from cathode to anode are not impeded by the wires. Focusing electrodes maintained at cathode voltage serve to confine the electron stream to the form of flat concentrated beams between the cathode and the two anodes ; the latter being interconnected. This concentration of electrons in the space between the screen grid and the anode represents a negative space-charge forming an effective barrier to prevent secondary electrons released at the anode being able to reach the screen grid.

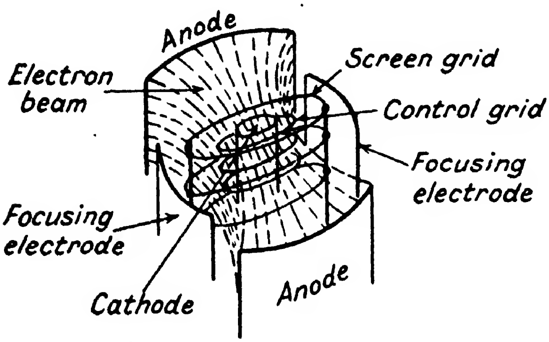


FIG. 48.—Electrode construction of beam tetrode.

The anode characteristics of beam tetrode are similar to those of a pentode except that the anode current rises more rapidly at the beginning of the curves and the knee of the curves is much sharper. The range of anode voltage over which the curves are almost flat is appreciably greater than with a pentode. The mutual conductance can be made very high and the valve has a relatively high power output.

FOR FURTHER READING

Among the books dealing specifically with thermionic valves the following give a well graduated treatment of the subject :—

1. F. E. HENDERSON. *Introduction to Valves.* (Wireless World). 1943. Intended for readers with little or no previous experience in electronics. Deals with all the usual kinds of valves and explains their uses.
2. E. V. APPLETON. *Thermionic Vacuum Tubes.* (Methuen), 1941. Describes action of usual types of valves and their use in various circuits, giving basic mathematical theory and references to supplementary literature.
3. E. L. CHAFFEE. *Theory of Thermionic Vacuum Tubes.* (New York, McGraw-Hill), 1933. An advanced and comprehensive treatment of valve theory and characteristics.

CHAPTER VII

AMPLIFIERS AND OSCILLATORS

THERMIONIC AMPLIFIERS

So far we have considered the properties of some of the principal types of thermionic valve and have seen that this device represents a highly efficient means of controlling power. The grid serves as the control element by determining how much power passes through the valve and its associated circuit but the power comes from the independent source represented by the anode supply, and the grid itself uses negligible energy in performing its function. When used in this way the valve acts as an *amplifier*; the controlled power being very much larger than that absorbed by the grid circuit.

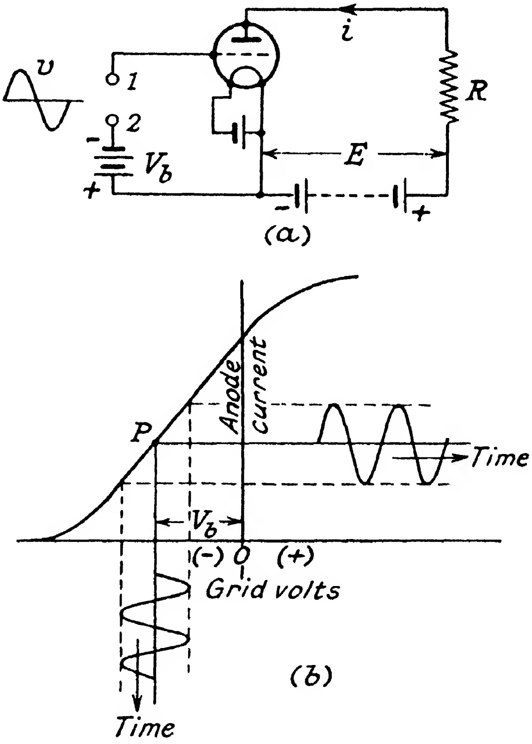


FIG. 49.—Elementary form of amplifier.

It may also be noted that in the general case where an alternating input voltage is applied to the grid the resultant anode current is also alternating, and an alternating voltage accordingly appears across the load although a direct current source is used for the anode supply. The conditions are represented in Fig. 49, where (*a*) is a highly simplified diagram of connections. The valve is a directly heated triode, and the supplies for anode, cathode and grid bias are taken from batteries. This arrangement would not normally be used in practice owing to the inefficiency of battery supplies, but is shown here for simplicity. The anode circuit is completed through the load, represented by resistance R, and the anode battery, which produces a voltage E. The anode current is represented by i. In the grid circuit the bias battery produces a voltage V_b which determines the region of the characteristic over which the valve will operate. The alternating input or signal voltage v is applied across terminals 1 and 2 in series with the bias voltage.

The conditions in the circuit are shown at (*b*) in Fig. 49. The sloping line is the dynamic characteristic for the particular valves of E and R employed. The bias voltage V_b is drawn in the negative direction from O, representing zero (i.e., cathode), potential, and the point thus located represents the origin or axis of the alternating input voltage v. If there were no bias voltage the input voltage would alternate about the point O, but the grid would become positive during every other half-cycle, resulting in a high grid current which, as we have already noted, is generally undesirable. The bias voltage is accordingly used to transfer the alternations of the input grid voltage to a region where this does not occur, and where the characteristic is straight, so that equal changes of grid voltage result in equal changes of anode current. This ensures that the output current is a true reproduction of the input wave form, and the arrangement forms a *linear amplifier*. It can be seen from the diagram that as the grid voltage alternates about the operating point P the anode current undergoes similar changes and at the same frequency.

This is the simplest form of amplifier, and there are many variations of this basic arrangement according to the intended application. They range from the simple circuit using one valve to operate a small relay or instrument on a D.C. or 50 cycle A.C. supply to complex amplifiers using numerous valves and special circuits for purposes such as television, where frequencies of several million cycles per second are employed. It is obviously not possible to consider all those special applications here, but attention will be given to the general principles underlying amplifier design.

Multi-stage Amplifiers

There are several ways in which valves may be connected in *cascade* for the purpose of increasing the output of either power or voltage as compared with that of a single valve. Any such arrangement constitutes a *multi-stage amplifier ;* each stage being represented by one valve and its associated circuits. The various forms of amplifier can be classified according to the method used for connecting one valve to the next, the principal forms of *coupling* being as follows :

1. Direct. For use with very slowly alternating or slowly changing input voltages.
2. Resistance-capacitance. Having wide general application on all but the highest frequencies.
3. Transformer. Generally used for *audio* frequencies (up to 20,000 cycles per second) and having the advantage that the grid voltage of successive valves may be stepped up by transformer action, giving a high *voltage gain* in relation to other types.

The voltage gain, which is an important relationship in amplifiers, may be defined as the ratio of the alternating voltage appearing across the load resistance to the alternating input voltage, both voltages being measured by suitable A.C. instruments indicating effective values.

In general, a high voltage gain is necessary in the first stages of an amplifier so as to obtain an adequate swing of grid voltage to operate the final or output valve. For a given input voltage the number of stages required is largely determined by this consideration.

Direct Coupling

A basic circuit for a two-stage direct coupled amplifier is shown in Fig. 50. The first valve V_1 is biased in the normal way by the battery B_1 and on the application of voltage to the input terminals a current which changes with changes of input voltage flows through resistance R_1. The resulting voltage drop is passed on directly to the grid of the second valve V_2 through the second biasing battery B_3 which keeps the grid voltage changes of V_2 in the required region to give linear operation and prevent the flow of grid current. The output voltage appears across resistance R_2 in the anode circuit of V_2.

Amplifiers of this kind have several disadvantages. In particular, they are unstable because any changes in the first stage, resulting for example from changes of supply voltages or in the valve characteristics, are passed on to the second stage where they cause changes which are independent of input voltage. Special forms of balanced circuits have been developed to overcome this trouble, but direct-coupled amplifiers are used only in cases where the technical difficulties are outweighed by the advantages obtained. These include the amplification of varying D.C. or very low frequency A.C. input and the preservation of slowly changing wave-shape. Direct coupled amplifiers have been used for magnifying and recording the small voltages of varying frequency associated with the action of the heart in one form of electrocardiograph. They are also used for special measurements and control systems, for relays and sounders in telegraphy and telephony and in conjunction with photoelectric cells.

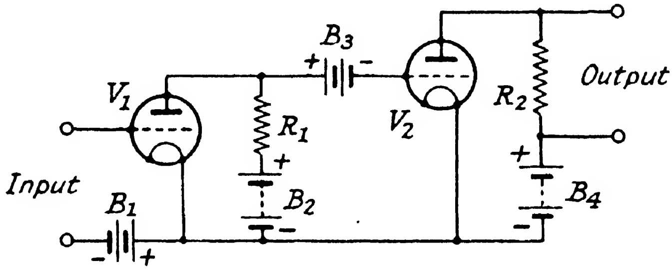


FIG. 50.—One form of direct coupled-amplifier.

Resistance-Capacitance Coupling

The principle of this arrangement is shown in Fig. 51. The alternations of input voltage cause a corresponding alternating voltage drop across resistance R_1 which is passed on through condenser C_1 to the grid of the second valve. As the condenser conducts A.C. but prevents the passage of D.C., the grid of V_2 is isolated from the anode supply (H.T.) although it receives an amplified reproduction of the input voltage. It will also be understood that for similar reasons the amplifier connected in this way is responsive only to an alternating input voltage. The load resistance is represented by R_2 in the anode circuit of V_2. For a further stage of amplification another condenser C_2

would be connected as shown dotted to the grid of a third valve. The resistances r_1 and r_2 are included to prevent the accumulation of negative charge on the respective grids which would otherwise change the bias. The relative cheapness of this type of amplifier owing to the use of resistors and condensers for coupling is often an important factor, but apart from this the amplifier is well suited to a wide range of applications.

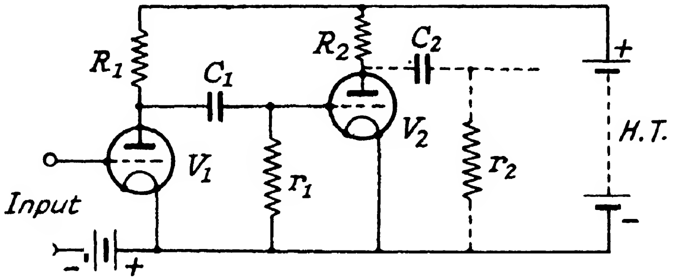


FIG. 51.—Resistance-capacitance coupled amplifier.

Transformer Coupling

As an alternating current appears in the anode circuit of a valve when suitably connected and supplied with an alternating grid voltage, it is possible by connecting the primary of a transformer in the anode circuit to obtain from the secondary an alternating voltage which may be used either as output or for supplying the grid of another valve for amplification purposes. The arrangement is shown in elementary form in Fig. 52. The alternating voltage induced in the secondary of transformer T_1 is applied directly to the grid of the second valve V_2. The alternating anode current of this valve excites the primary of the second transformer T_2 and the secondary voltage may be used either to supply the load as shown or as grid voltage for a further stage of amplification if required. As there is no direct connection between the anode circuit of V_1 and the grid of V_2 there is no need of coupling condensers and grid resistors.

One advantage of the arrangement is that the transformers can be designed to give the best operating conditions for a given application, but they are expensive in comparison with resistance-capacitance coupling. There is also the consideration that stray magnetic fields are produced by the transformers, and magnetic screening is sometimes necessary to prevent the interactions of these fields causing interference with operation.

Transformers are classified as *input*, *interstage* (coupling) or *output*, according to the purpose for which they are used in an amplifier. In some cases a transformer output stage is used with a resistance-capacitance coupled amplifier to obtain the advantage of arranging the transformer windings to match the particular load so as to give maximum power output.

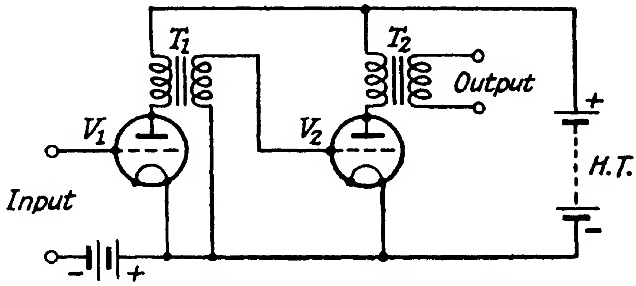


FIG. 52.—Transformer coupled amplifier.

Push-Pull Circuit

When more power is required from a single stage of amplification than is obtainable with a single valve, use is often made of a so-called push-pull circuit employing two valves connected as shown in Fig. 53. This arrangement is more satisfactory than the alternative of connecting two valves in parallel with the consequent difficulty in equal sharing of the load between them. With the push-pull circuit each valve contributes half of the total output power, which is consequently twice that obtainable from one valve. Each valve operates independently of the other, and the circuit is symmetrical, being equivalent to two single-stage transformer coupled circuits connected back-to-back. Double-triode valves containing two independent triodes with a common cathode are often used instead of two separate valves.

Referring to the diagram (Fig. 53) it is seen that the input transformer T_1 has a centre-tapped secondary and the output transformer T_2 a centre-tapped primary. Grid bias is supplied by battery B_1, and battery B_2 supplies the anode current. The opposite ends A and B of T_1 secondary are always at opposite potential when an alternating input voltage is applied, so that while the anode current in valve V_1 is passing through a positive half-cycle the anode current in V_2 is passing through a negative half-cycle ; the respective currents being equal and opposite at any instant. The primary windings of transformer T_2 are arranged so that these opposing currents produce fluxes in the *same* direction in the magnetic core, and a normal alternating voltage output is obtained in the secondary winding.

A circuit of this kind can be operated in several ways according to the degree of biasing. In many cases each valve operates as in Fig, 49 (*b*). This is called Class A operation. In other cases the valves are biased almost to the cut-off point, and each acts as a rectifier, passing half-cycles of current. This gives the advantage of reduced drain on the battery when there is no input voltage. Each valve operates for only half the input cycle and as the corresponding half-cycles of anode current are spaced half a cycle apart the result in so far as the output is concerned is the same as a continuous alternation. This is Class B operation.

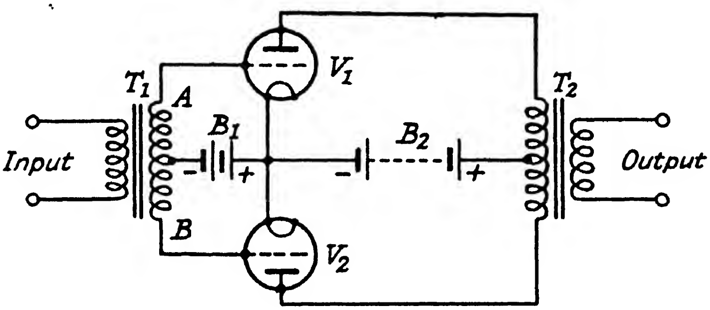


Fig. 53.—Basic push-pull circuit.

OSCILLATORS

Apart from its use in various forms of amplifier the thermionic valve has highly important applications as a means of generating oscillations for a variety of purposes. When used as an *oscillator* in this way the valve uses a source of D.C. power to produce A.C. power at a frequency ranging from a few cycles per second (or even per minute) to several hundred million cycles per second, according to the type of circuit used.

The phenomenon of oscillation is well known as an unwanted form of disturbance in radio receivers. It arises from the valve picking up a spurious signal which when superimposed on the normal circuit conditions makes the circuit currents oscillate at a particular frequency called the *resonant frequency*, which is usually very high. In the various forms of oscillator these conditions are purposely introduced so that oscillations at the required frequency are efficiently produced. Means of adjustment are generally provided for selection of frequency over a wide range.

The use of oscillators was formerly confined solely to communications, but they are now also used extensively in industry, notably in measurements and as a source of heating for various purposes by high frequency currents.

There are three principal types of oscillator, representing the three chief methods by which oscillations may be produced :—

1. Negative resistance, using this special property of valves such as the magnetron.
2. Feedback, using the property of regeneration, retroaction or reaction obtained by supplying part of the output of an amplifier to the input circuit in such a way that the amplifier provides its own excitation for producing sustained oscillations.
3. Piezoelectric or crystal control, using the properties possessed by certain crystals of producing a voltage when mechanically stressed and of changing their shape when suitably subjected to an applied voltage. These properties are used as a means of sustaining oscillations in a valve circuit as an alternative to the feedback method.

Attention will here be given to a representative oscillator of each type so as to illustrate the principles of operation in the respective cases.

The Magnetron Oscillator

The basic arrangement and circuit of a magnetron oscillator are shown in Fig. 54. The device contains two anodes in the form of a cylinder split lengthwise and having a thermionic cathode coincident with its axis. A solenoid coil or permanent magnet is arranged outside the tube so as to give a magnetic field parallel with the cathode as shown by the dotted arrows. An oscillatory circuit comprising a centre-tapped inductance L and a capacitance C is connected in push-pull fashion to anodes 1 and 2 and the H.T. source. A circuit of this kind produces oscillations of current and voltage at a particular frequency determined by the relative values of inductance and capacity, the oscillations being due to the interchange of energy stored in the two circuit elements. The frequency of oscillation can be selected by using particular values of L and C or adjusted by making L and/or C variable, the latter being more usual.

Some rather curious things happen to the motion of an electron when it is projected across lines of magnetic force, as will obviously happen to electrons leaving the cathode of the dynatron and attracted by either anode. To simplify the matter we will imagine that the anode is a proper cylinder instead of being split in two, and that it is maintained at a constant positive potential with respect to the cathode. The sequence of diagrams in Fig. 55 represents end-on views looking along the cathode, and, therefore, along the lines of magnetic force. Diagram (*a*) corresponds to zero magnetic field, and with this condition the path of the electrons leaving the cathode is radially outwards, as shown by the arrow. As the field strength is increased the electron experiences a force which curves its path, as at (*b*), during its passage from cathode to anode, and with further increase of field strength a critical value is eventually reached where the electron path is curved right round on itself as at (*c*) and the electron just glances the anode without being captured by it. If the field strength is reduced even very slightly from this critical value the electron is captured by the anode, and an increase of anode voltage has the same effect. Increase of field strength beyond the critical value reduces the radius of the circular orbit in which the electron travels, as shown at (*d*).

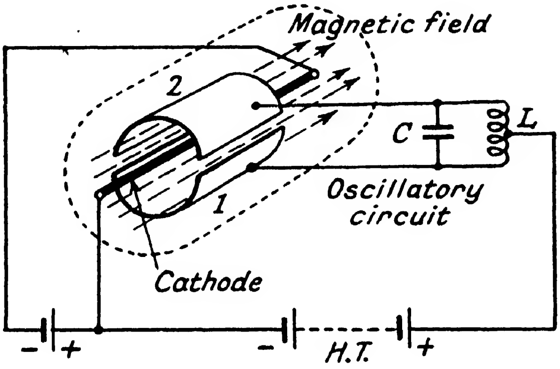


FIG. 54.—Magnetron oscillator circuit.

This sequence of events shows that with such an arrangement the anode current can be switched off suddenly, at the critical stage (*c*) by sufficiently increasing the magnetic field strength, the anode voltage being constant. A device of this kind was first described by A. W. Hull in America in 1921, and received the name *magnetron.* It has since been largely superseded by the *split-anode magnetron* which we set out to describe and to which we now return.

Referring again to Fig. 54 it will simplify the understanding of the complete circuit if we imagine first of all that one of the anodes (say No. 2) is maintained at a constant positive voltage with respect to the cathode. It is further assumed that the magnetic field strength is above the critical value as defined in connection with Fig. 55 (*c*). If now an increasing voltage is applied to the remaining anode (No. 1) and the resulting current from this anode is measured, the characteristic will have the peculiar shape shown in Fig. 56. With no voltage on anode No. 1 all the electrons flow to anode No. 2, as shown at (*a*) in the sequence of figures under the curve. As the voltage is increased, more and more electrons reach anode No. 1 and its current consequently increases, reaching a maximum at (*b*). At this stage the voltages on the two anodes are approaching equality. When this condition is actually reached, at (*c*), the action of the critical magnetic field is fully effective and the circular paths followed by the electrons miss both anodes. The part of the curve between (*b*) and (*c*) is a *negative resistance region* in which the current falls with an increase of voltage. As already mentioned in connection with the tetrode (page 98), this property allows the device to supply energy to an external circuit under suitable conditions and it may consequently be used to support oscillations in such a circuit as that shown connected to the valve in Fig. 54.

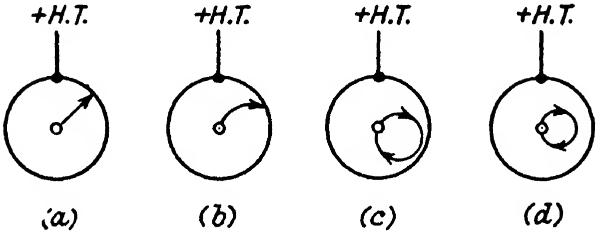


FIG. 55.—Effect of increasing magnetic field strength on path of electron in the magnetron.

When both anodes are at the same potential the electric field is radial from the centre, like spokes in a wheel. When an oscillation occurs in the external circuit, this state of affairs is upset because one segment receives a positive and the other a negative increment of potential. There is now an electric field between the respective segments, owing to their difference of potential, and this field is superimposed upon the radial field, causing a distortion which is most marked in the region of the gaps between the anodes. This causes a disturbance of the circular orbits which would otherwise be followed by individual electrons and results in electrons being caught by one or other anode, according to their relative potential, instead of missing both anodes. This action is purely automatic, and the so-called *dynatron oscillations* which result have a frequency largely determined by the natural frequency of the oscillatory circuit.

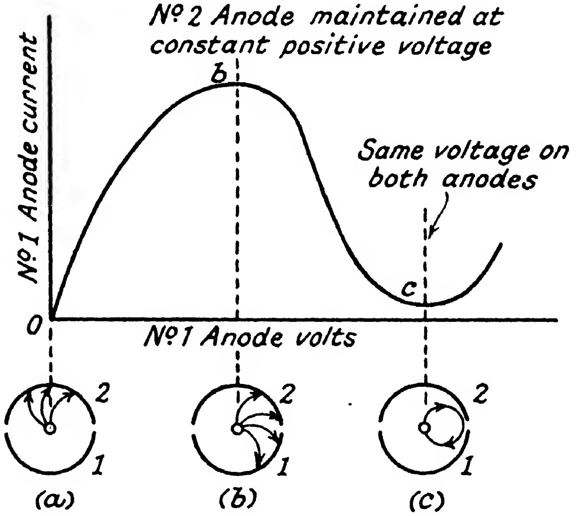


FIG. 56.—Anode characteristic of magnetron with one anode at constant positive voltage.

Circuits of the kind just described are capable of producing oscillations ranging from about 100,000 cycles per second (3,000 metre waves) to 500 million cycles per second (60 centimetre waves). Even higher frequencies up to 30,000 million, corresponding to waves only 1 centimetre long, are also obtainable with magnetron oscillators, but these *resonance oscillations* require a different explanation involving ideas for which as yet there is no substantial experimental evidence.

Feedback Oscillators

The basic principle in producing oscillations by means of feedback circuits is that part of the output of an amplifier is fed back to the grid circuit; the alternations of this input having the proper timing (phase) and magnitude to sustain the alternations of the output. There are many analogies of this kind of action. Many users of the older kinds of telephone sets with separate transmitter and receiver will be familiar with the howling noise which can be made by holding the receiver near the transmitter. In this case the transmitter acts as an amplifier and part of the alternating power is fed back through the air between the two diaphragms by way of the receiver into the transmitter circuit, thus causing self-supporting oscillations.

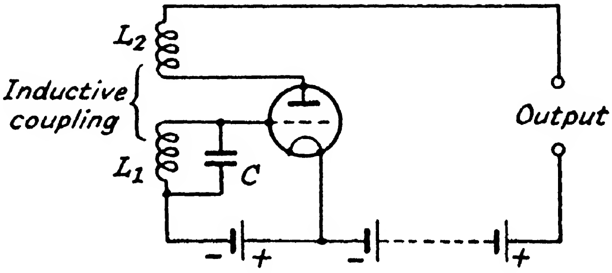


Fig. 57.—Feedback oscillator with tuned grid circuit.

The feedback principle may be more clearly grasped by imagining an amplifier with an A.C. input and producing an A.C. output at the same frequency in the usual way, the input being supplied from some external source. If now a part of the output is taken so that it is in every way similar (with regard to voltage, frequency and phase) to the original signal, then this part of the output may be substituted for the original signal without causing any difference in the conditions. By this change the amplifier is transformed into a feedback oscillator, part of its output being used to feed the input.

The numerous practical forms of feedback oscillator employ as the means of producing the self-exciting input an oscillatory circuit comprising inductance and capacitance, tuned to a particular frequency. Some form of coupling is also provided between the anode and grid circuits so that the feedback input may be suitably introduced. The tuned circuit may be in the grid or anode or both, and the coupling may be either inductive, as between two coils, or capacitive, as provided by condensers.

An inductively coupled circuit of the type generally called a *tuned grid circuit* is shown in Fig. 57. The coupling is provided by the coils L_1 and L_2, between which there is mutual inductance, i.e., their respective magnetic fields are interlinked. The coil L_1 and condenser C comprise the tuned oscillatory circuit, and the oscillations set up in it produce oscillations of grid voltage which in turn give rise to oscillations of anode current. Owing to the inductive coupling between L_1 and L_2 a voltage proportional to the anode current variations is fed into the grid circuit, and since the timing or frequency of these variations was originally determined by the relative values of inductance L_1 and capacitance C, the oscillations continue indefinitely. Similar explanations apply to the numerous other feedback oscillator circuits.

Piezoelectric or Crystal Control

Difficulties arise with feedback oscillators in cases where the frequency must be maintained with great accuracy. Owing to gradual changes in the properties of the circuit, caused mainly by temperature variations, the frequency is also liable to change.

Much greater stability may be obtained by using the piezoelectric properties of certain crystals, notably quartz, Tourmaline and Rochelle salt. As Tourmaline is a semi-precious stone its use is limited on account of expense and Rochelle salt has the disadvantage of being very fragile, although it is more sensitive than quartz and can be manufactured. The piezoelectric properties of crystals were discovered and investigated by M. and Mme Curie in 1880, the name of the effect being taken from the Greek word *Piezo*, to press. It was found that the application of mechanical pressure resulted in the production of a very small voltage across the crystal and that this voltage reversed if the pressure was replaced by tension. Conversely, the application of a voltage caused expansion or contraction according to its direction. Maximum effects are obtained when the application of the pressure and measurement of the voltage are made in particular directions, corresponding to what are called the mechanical and electrical axes respectively. There are numerous ways in which crystals may be cut relative to the axes in order to secure greater sensitivity.

A crystal has a natural frequency of oscillation, so that if an oscillating voltage with a frequency at or near the natural frequency of the crystal is applied to it a state of resonance will be set up in which the mechanical oscillations have a maximum value and the resulting oscillatory voltage developed across the crystal is correspondingly high. If a crystal operating in this way is suitably incorporated in a valve circuit it will provide the means of sustaining oscillations with much greater accuracy than is otherwise possible.

A representative circuit is shown in Fig. 58. It has a tuned anode circuit, and the crystal is in series with the grid. In this arrangement the capacitance between grid and anode within the valve itself is of importance and its effect on the outside circuit is represented by the connection of an imaginary condenser C_1 as shown dotted. This capacitance provides the link by which oscillations in the tuned circuit LC are conveyed to the crystal, so that part of the output of the valve is used to maintain the resonant vibrations of the crystal. The resulting voltage impulses act on the grid and are magnified in the anode circuit.

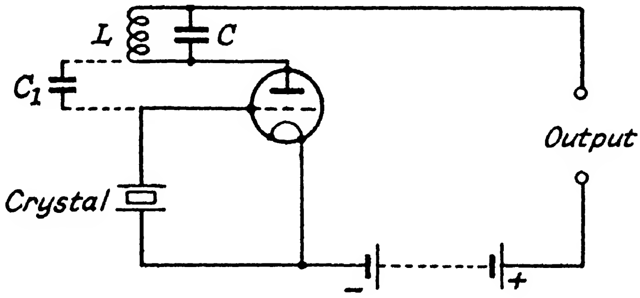


FIG. 58.—Piezoelectric or crystal controlled oscillator circuit.

As the resonant frequency of a crystal changes with temperature it is customary, in cases where the complication is warranted, to enclose it in a small oven with thermostatic temperature control. With such precautions the frequency may be held constant within a few parts in a million over long periods, as is necessary with many classes of oscillators used for radio communication. Frequencies up to about six megacycles may be obtained by direct crystal control but the output power is only a few watts. A limit to the frequency is set by the fact that it is inversely proportional to the thickness of the crystal, which thus becomes excessively fragile for very high frequencies. Both the frequency and power may be increased by special auxiliary circuits.

Uses of Oscillators

The choice of an oscillator generally depends upon the amount of power required and the frequency ; these being, as we have seen, opposing factors because of the difficulties which tend to limit power as the frequency is raised. Applications which call for only small power at high frequency, as in electromedical uses, can usually be met by ordinary valves of suitably modified design as developed for such purposes, but special types incorporating artificial cooling arrangements are needed when high power outputs are required. In some applications, especially communications, frequency stability may be the chief consideration, in which case a temperature-controlled crystal oscillator offers a solution, the power output being increased by auxiliary circuits if necessary. When very high frequencies are required the choice of valves is limited to the magnetron and velocity modulated types appropriate to the microwave region.

Apart from radio applications, the chief uses of oscillators are in connection with industrial heating, electromedical work and electrical measurements. Heating by high frequency currents can be done in two ways ; either by induced eddy currents in the case of metallic objects or capacitance currents in the case of non-conducting materials. Between them these methods have a wide and rapidly extending sphere of use in industry. The frequencies used range according to the application between about twenty kilocycles and upwards of a megacycle in eddy current heating and up to as much as 200 megacycles in some cases of capacitance current or dielectric heating. Eddy current heating results from the circulating currents induced in a conducting material subjected to an alternating magnetic field as shown in Fig. 59. The effect is familiar in practice as the cause of heating in the magnetic circuits of transformers and other apparatus operating from an A.C. supply, and is reduced in such cases by making the magnetic circuits from thin laminations which restrict and offer a high resistance to the induced eddy currents. As the heating effect increases with the frequency and the electrical resistance of the material it can be increased at will in a given case by using a sufficiently high frequency. There is, however, an effect known as the *skin effect* which tends to confine the current, and hence the heating, to the surface region of a conductor carrying alternating current, the restriction effect increasing with the frequency. It follows that if the frequency is high enough the heating will be practically confined to the surface of the metal ; a fact which has been utilised to good effect for surface hardening and tempering processes in industry. There are many interesting examples of eddy current heating in industry, especially in cases which would be very awkward if the heat was applied from outside. Metal canisters can be automatically soldered as they pass on a conveyor belt through the high frequency magnetic field produced by a suitably shaped stationary coil. Tin-plate can be produced in a fraction of the previous time by roughly depositing a layer of tin electrolytically on the iron sheet and then exposing it on a conveyor to the high frequency field. The rapid and uniform heating melts the tin deposit, which flows in a thin film over the iron surface. The use of eddy current heating in valve manufacture as a means of driving unwanted gases out of metal parts and in the gettering of valves has already been mentioned (page 82).

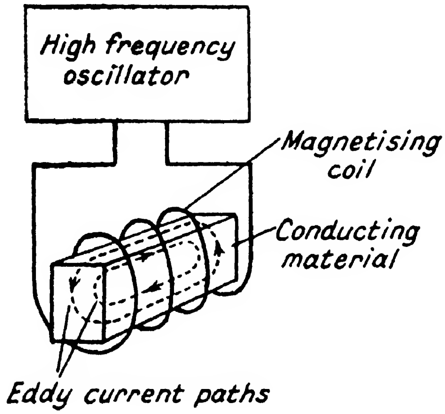


FIG. 59.—Principle of eddy current heating.

In the case of dielectric heating the substance to be treated is enclosed between the plates of a condenser connected to a high frequency oscillator as shown in Fig. 60. Under the rapidly changing electric field between the plates the outer electrons of the atoms and molecules of the substance are violently disturbed, and their movements against the frictional forces tending to restrain them are regarded as the source of heat within the material. Perhaps the widest industrial use of dielectric heating is in the preparation and treatment of plastics and the bonding of laminated materials such as ply woods which can be

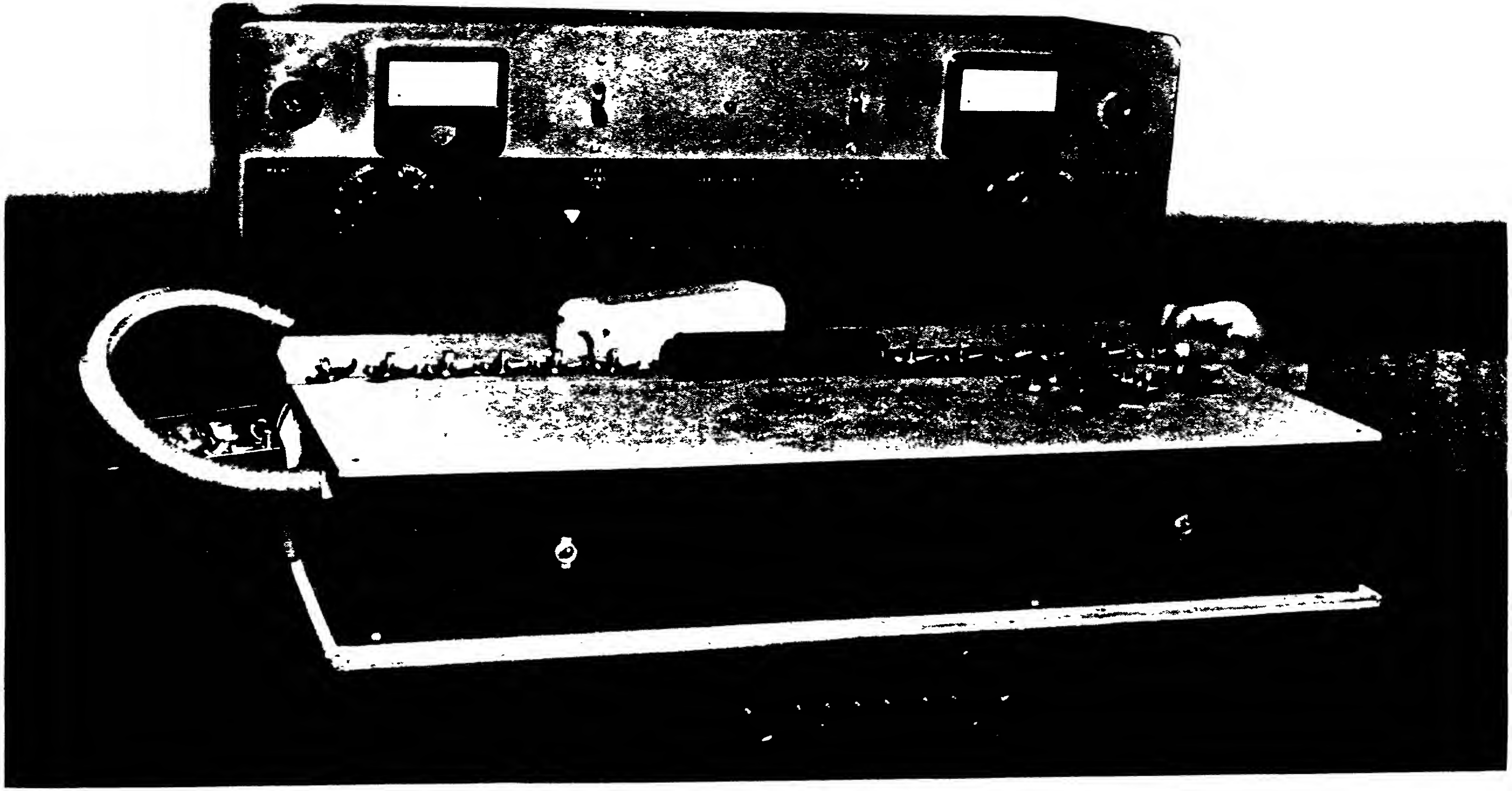


PLATE 2

Mass production of soldered joints by high frequency heating. Switch knobs are shown passing through the coil which induces high frequency currents in the metal. The power available is 2.5 KW. and the operating frequency is 500 Kcs.

(*Philips Electrical Ltd.*)

set into desired shapes in suitable moulds while the treatment is applied. The dehydration of food can also be efficiently and rapidly effected without deterioration or loss of food value.

In medical practice, currents at a frequency ranging between about 500 kilocycles and one megacycle are applied to the body in diathermy (heat treatment) and are known to destroy some types of infection. High frequency currents also have a coagulating effect on the blood, and have rendered possible electric cautery, otherwise known as bloodless surgery. In electrotherapy (curing by electricity) the frequencies used range from 10 to 100 megacycles and the electric field is applied between condenser plates enclosing portions of the patient's anatomy so as virtually to give dielectric heating of these parts. Recent research shows that it is not only the heating effects of high frequency currents which are lethal to germs. Experiments made with an oscillator in which the energy was delivered in pulses separated by rest periods so as to prevent any observable rise in temperature showed that various kinds of bacteria and viruses were destroyed by some action other than the heating effect.

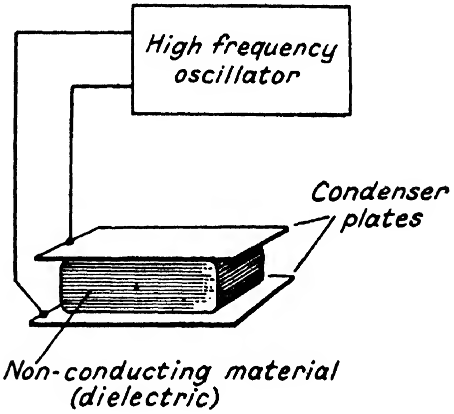


Fig. 60.—Principle of dielectric or capacitance current heating.

Oscillators with adjustable frequency ranges are now a standardised product and find routine use in many laboratories as a means of frequency measurement and comparison and also for accurate measurement of inductance and capacitance. In the latter case they are generally embodied in a completely calibrated measuring instrument from which direct readings are obtained.

THE HIGH FREQUENCY DOMAIN

During recent years there has been a marked increase in the uses of high frequency currents for various purposes, especially in communications, and a great deal of work has been done in connection with the design and manufacture of valves for use at these high frequencies. With ordinary valves a variety of troubles manifest themselves when high frequencies are used, and some very radical changes become necessary in order to obtain satisfactory operation. The higher limits of frequency beyond which troubles appear are different for different types and sizes of valves, but in general the special forms of construction are necessary when the frequency extends well into the megacycle range, i.e., when the alternations which the valve has to reproduce occur several million times per second.

With frequencies of this order the time of a single oscillation is so short that it may be comparable with or even less than the *transit time,* i.e., the time required for an electron to pass from cathode to anode, and the grid will be prevented from exerting its normal control. The transit time may be shortened by increasing the anode voltage, so as to speed up the electrons, and putting the electrodes close together, but there are obvious limits to both these measures, especially as the electrodes would need to be greatly reduced in size to offset the increase of capacitance caused by putting them close together. Capacitance is especially undesirable in high frequency tubes because of the unwanted coupling it gives between electrodes and the resulting increase of grid circuit power to operate the tube. Reduction in anode size reduces the amount of heat it can dissipate and therefore lowers the power output, so that several difficulties arise with this procedure.

A solution to these difficulties has been found in an entirely new principle of operation employing a form of construction differing completely from that of the conventional triode. Although the use of these tubes is at present limited to very high frequencies in communications and allied spheres of application, and thus lie outside the scope of this book, a description is given here as an example of advanced modern technique introducing some new conceptions of applied electronics.

Velocity Modulated Tubes

There are several high frequency tubes employing the principle of *velocity modulation* of electron beams in order to overcome the inherent difficulties with triodes, but there are different ways in which the principle can be applied. These differences relate to the method of extracting from the velocity modulated beam the power representing the amplified output.

The principle of velocity modulation in its simplest form may be explained by the diagram in Fig. 61, where G_1 and G_2 are two grids located between a thermionic cathode C and an anode or collector electrode A, all being enclosed within an evacuated envelope. A constant voltage V_2 is applied between the cathode and the first grid G_1 and an alternating voltage V_2 of lower value between the two grids. The voltage V_2 is of high frequency and is obtained in practice from some form of resonant circuit. The constant voltage V_3 between cathode and anode is lower than V_1.

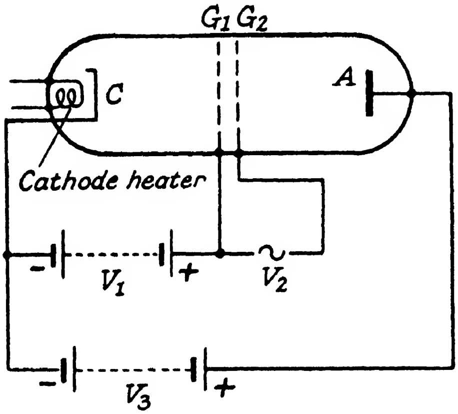


FIG. 61.—Special tube for showing principle of velocity modulation.

Owing to the high value of V_1 the electrons from the cathode reach grid G_1 and pass through it at such a high speed that the transit time of the electrons between the two grids is short in relation to the time of one alternation of the voltage V_2. The alternations of V_2 cause changes in the velocity of electrons passing between the grids because at those instants in the cycle when G_2 is positive with respect to G_1 these electrons are speeded up, whereas in the other half of each cycle, when G_2 is negative with respect to G_1 they are slowed down. Any electrons which pass through the grids while voltage V_2 is passing through zero are neither speeded up nor slowed down, but pass on at the same speed. We must thus try to visualise the velocity modulation process as producing in the space beyond G_2 a flight of electrons in which some are travelling at constant speed, others are tending to overtake them, and others are in turn falling farther behind, so as to be overtaken by the other two groups. A velocity modulated electron beam of this kind is of no use without some means of extracting energy from it, but before considering the ways of doing this an important fact may be noted. When G_2 is positive, energy is taken from the source of V_2 in imparting acceleration to the electrons, whereas when G_2 is negative, and the electrons are decelerated, energy is returned to the source of V_2. This cancellation of energy considered over a complete cycle explains the important fact that the energy delivered by the source of V_2 is in practice extremely small, which means that almost negligible power is required for velocity modulation of the electron beam by this means. All the electrons are collected by the anode, and as its voltage V_3 is low the energy dissipated in this electrode is relatively small.

Density Modulation

Before the energy contained in the velocity modulated electron beam can be utilised, some means must be found for conversion to *density modulation* in which the electrons are grouped in such a way as to deliver uniformly spaced pulses of current in the collector system. There are three principal ways of doing this. One method is to deflect the beam through an arc by an electric or magnetic field in the manner of a cathode ray tube so that the high speed electrons take a longer path than the slow speed ones, which are more sharply deflected. By placing electrodes so as to intercept the respective beams each electrode receives a series of electron bursts or pulses which alternate with those received by the other electrode because of the differences of timing imparted to them on passing through the modulating grids. The power is utilised by connecting a push-pull form of circuit between the two anodes. A second method of collecting power from the beam may be explained with the aid of Fig. 61 by assuming that the anode is connected directly to the cathode instead of having the positive voltage V_3. In this case the high velocity electrons reach the anode but the slow ones fail to do so and are eventually picked up by any electrodes having a positive voltage. Under these conditions a usable intensity modulated current flows in the anode lead. The third method of power extraction is more widely used than the others and will therefore be explained in more detail.

Reverting again to the velocity modulated beam, an analogy may help to show the basic principle of this particular method, sometimes called the *drift method* of density modulation. We may first imagine a long straight stretch of clear road. Along this road comes a group of cars, all travelling at the same speed and keeping close together. These represent the electrons which were slowed down in passing through the modulation grids. On the road behind this first group of cars comes a second group, close together like the first but travelling faster and thus gaining on the first group. These represent the electrons which shot through the modulating grids while the voltage was changing. Next on the road comes an even faster group of cars, representing the electrons which were accelerated on passing through the modulating grids. It is clear that at some point down the road one of the faster groups of cars catches up with those in front and that if there is no overtaking the other fast group will catch up with this combination. The point at which all the cars come close together we may call the concentration point, and in our analogy this corresponds to the concentration of energy represented by a group of electrons. If we place a suitable collecting device at the concentration point the pulses of energy corresponding to the continuously repeated process we have just described can be extracted and utilised. Owing to the entirely different conditions prevailing at very high frequencies as compared with general radio frequency practice the conventional methods of dealing with the output power are no longer applicable, and use is made of the technique of cavity resonators and wave guides of tubular construction in which the electric and magnetic fields associated with the waves are transmitted after the manner of sound waves in organ pipes. The full description of this technique belongs to the subject of high frequency engineering, but reference must be made to it here because the cavity resonator is the device normally used to extract the high frequency power from the density modulated electron beam in tubes using the drift principle of modulation.

Cavity Resonators or Rhumbatrons

The cavity resonator may be regarded as the high frequency equivalent of the conventional resonant circuit consisting of separate inductive and capacitive branches. In the resonator the magnetic and electric fields exist together inside the enclosure, which may have a variety of shapes. One form is shown in Fig. 62.

Cavity resonators may be excited or put into resonance by passing electron beams through them or applying suitable high frequency currents to small electrodes inserted in the cavity, and an excited resonator will deliver energy to an external circuit through similar electrodes. The name *rhumbatron* has been given to resonators of this kind, from the Greek for rhythmic oscillation, and they may properly be called rhumbic enclosures because of the oscillations taking place inside. Each form of enclosure has its own particular features as a resonator, and the frequency of oscillation depends upon the mechanical dimensions. It may be noted that the fields are entirely confined to the inside of the resonator, so that there is no external radiation loss.

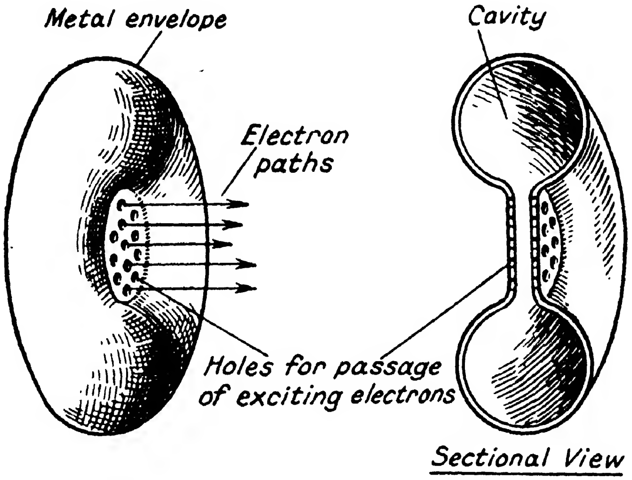


FIG. 62.—One form of cavity resonator (rhumbatron), excited by electron beam.

The Klystron Tube

The cavity resonator forms a convenient means of velocity modulation and may also be used as a means of extracting the high frequency energy from a density modulated beam. The use of cavity resonators in this way is shown in Fig. 63, representing a particular form of tube known as the Klystron, the name being derived from the Greek for waves breaking on a beach. This alludes to the wave action of the groups of electrons.

The tube consists essentially of two cavity resonators, or rhumbatrons connected together mechanically and each provided with grid-like apertures for the passage of the electron beam from a cathode at one end to a collector or anode at the other. In some tubes the end cap is of metal and acts as the collector without the provision of a separate electrode.

If we suppose that the first resonator (called the buncher) is excited from an external source through the input terminals an alternating voltage will exist between its grids and electrons passing through these grids into the drift space will be velocity modulated in the manner already explained the resonator grids behaving in the same manner as G_1, G_2 in Fig. 61. The length of the drift space is such that groups or bunches of electrons form on arrival at the grids of the second resonator (called the catcher) which is excited by the successive pulses due to the sequence of electron groups. Energy is taken from these groups only when the oscillations of the resonator are so timed that the field between its grids causes retardation of the groups as they pass through. This effect is automatically obtained by tuning the catcher to the frequency of the input signal. The tuning is done by slight distortion of the shape and volume of the catcher by means of screw adjustments. When used in this way the drift tube or Klystron acts as an amplifier of very high frequency oscillations up to the order of 3,000 megacycles per second, corresponding to radio waves only 10 centimetres long. Very short waves of this kind have a wide and increasing sphere of use in radio communications, and among their properties are the valuable ones of penetrating power and directivity which

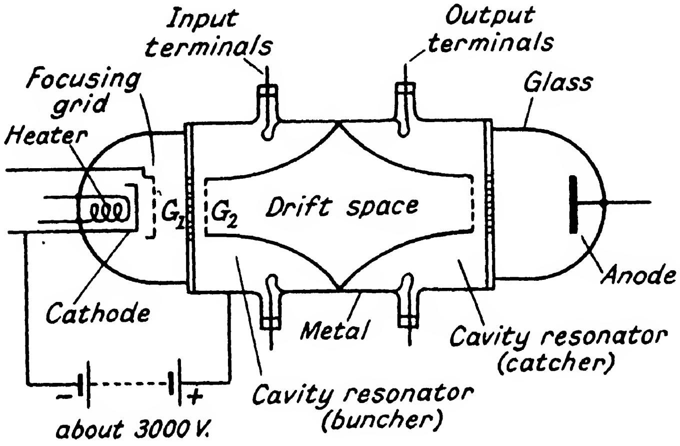


FIG. 63.—Sectional view of Klystron tube using drift principle and cavity resonators.

enable them to be used as guiding beams for the blind landing of aircraft in fog and other navigational purposes.

By using the feedback principle the Klystron may be used as an oscillator. In this case part of the output energy is returned to the input by interconnecting the input and output resonator terminals. The arrangement then operates as a self-maintained generator of very high frequency oscillations. It might in fact be said that the versatility of these new types of tubes makes them as important for the future development of high frequency engineering as was the simple triode valve in relation to the development of radio more than a quarter of a century ago.

FOR FURTHER READING

A treatment of amplifiers and oscillators is generally included in books dealing with valves, and reference may be made to the books by Henderson, Appleton and Chaffee given on p. 100. There are several good reference books on radio engineering in which the subject is dealt with more fully :—

1. F. E. TERMAN. *Radio Engineering.* (New York, McGraw-Hill), 1937.
2. R. HENNEY. *Radio Engineering Handbook.* (New York, McGraw-Hill), 1941.

Reference may also be made to :—

3. H. A. THOMAS. *Theory and Design of Valve Oscillators.* (Chapman and Hall), 1939.

On high frequency valves and technique generally, the following may be recommended :—

4. A. F. HARVEY. *High Frequency Thermionic Tubes.* (Chapman and Hall), 1944. Deals with properties of various tubes and circuits, including magnetron and klystron types, at high frequencies. Surveys published work and gives extensive bibliographies. Advanced.
5. M. S. KIVER. *U.H.F. Radio Simplified.* (New York, Van Nostrand ; London, Macmillan), 1945. Develops from low-frequency technique the principles of ultra-high-frequency engineering and gives elementary exposition of modern practice in this sphere.
6. J. G. BRAINERD (Ed.) and others. *Ultra-High Frequency Techniques.* (New York, Van Nostrand ; London, Chapman and Hall), 1942. An advanced and comprehensive treatment of modern u.h.f. equipment, including generation and transmission. Extensive bibliography.

GASFILLED VALVES

We have seen that in the production of vacuum valves elaborate precautions are taken to ensure that every trace of gas is removed during manufacture so that the vacuum is as perfect as possible. It now remains to deal with a series of valves in which gases are purposely introduced with the object of obtaining properties of use for particular applications.

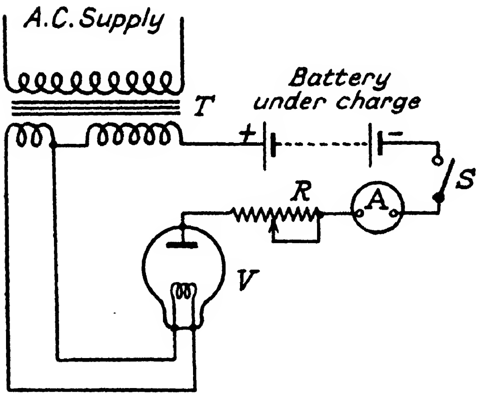
The pronounced differences between electrical conduction in a vacuum and in a gas were considered in Chapter 4, and it was seen that in general the vacuum discharge is characterised by a small current at an appreciable voltage, whereas in gas conduction the current is much higher and the voltage relatively low. The presence of positive ions in the gas discharge also introduces the possibility of ionic bombardment of the cathode ; an effect which may become very destructive if the voltage drop between anode and cathode exceeds a particular value.

THE GASFILLED DIODE

This tube is widely used as a rectifier, and in one or other of its commercial forms is generally familiar in small accumulator-charging equipments designed to operate from A.C. supply mains. An example familiar to many users of battery operated radio sets is the trickle-charger incorporating a gasfilled rectifier valve.

Owing to the presence of the gas and the resulting ionisation of the conducting space between anode and cathode the gasfilled diode has a very low resistance as compared with a vacuum tube. The ions in the conducting path also neutralise the negative space-charge, so that the anode current is not limited by space-charge effect as it is in the case of vacuum valves (see page 48). The voltage drop from anode to cathode is generally between 10 and 20 volts, and is almost independent of current. Its actual value depends upon the particular gas used and its pressure. Typical gases are the inert ones, including argon, helium and neon, but some tubes employ mercury vapour.

Two important considerations arise in connection with the presence of positive ions owing to their damaging effect on the cathode if they reach it at too high a speed. Firstly, it is necessary for the cathode to be at working temperature before it is called upon to supply current, i.e., before the anode voltage is applied ; otherwise, owing to the excessive voltage drop under these conditions, the ions arrive at the cathode at a speed which may be sufficient to cause disintegration of the coating. This condition makes some form of delay necessary between the cathode being switched on and the anode circuit being completed. The *cathode heating time* which determines the delay period depends upon the size of the tube and whether the cathode is directly or indirectly heated. It is rarely less than 30 seconds, and may be as much as 15 minutes for very large tubes. The cathode is also liable to damage if an excessive current is passed through the tube, even if the duration of this excess is very short.



A = Ammeter. Rf = Variable resistor.
S = Switch for anode circuit.
T = Transformer. V = Gasfilled diode.

FIG. 64.—Simple battery-charging circuit using gasfilled diode.

The gas pressure has a rather critical effect on the properties of the gasfilled tube, and varies according to the intended class of service. A relatively high pressure, such as about one-twentieth of atmospheric, reduces the normal rate of evaporation of the cathode, but too high a pressure reduces the *inverse voltage* which the tube can withstand. This is the voltage which can be applied with the cathode positive without anode current flowing and which must obviously be known fairly accurately if the risk of conduction during negative half-cycles of anode voltage is to be avoided. Too low a gas pressure may result in an insufficient supply of ions to neutralise the negative space charge and will cause an increased voltage drop with consequent risk of excessive evaporation of the cathode. This effect can be offset by reducing the anode current but such a tube would be inefficient. It follows that for every tube in a given application there is a fairly definite maximum and minimum of gas pressure. During the life of a tube there is a constant, though normally very gradual, fall of pressure owing to absorption of the gas into the metal parts and the glass walls ; a process which is technically known as *clean-up*.

It is worth noting that in the case of mercury vapour the pressure varies considerably with the temperature. For this reason the characteristics of a mercury vapour tube are less stable than those using inert gases.

The manner of using a gasfilled diode as a rectifier in a simple battery charging circuit is shown in Fig. 64. By providing tappings on the transformer and a selector switch it is possible for one equipment to deal with alternative batteries requiring different charging voltages. The charging current is indicated by the ammeter and can be adjusted to the required value by the variable resistor.

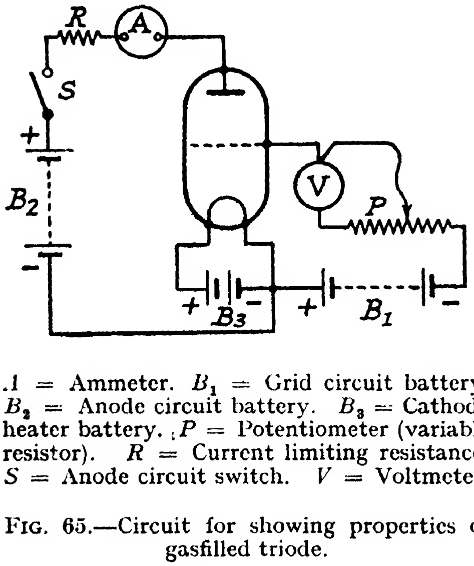
THE GASFILLED TRIODE

About the year 1913 some research was carried out in America on triode valves with a gas filling instead of a vacuum. Such an arrangement might be regarded alternatively as a gasfilled diode rectifier with the addition of a control grid. The early experiments showed that the device had some very interesting possibilities and it was not long before numerous suggestions were made for its application in various branches of electrical practice. Foremost in this work were G. W. Pierce and I. Langmuir, who used a tube having a mercury vapour filling provided by a globule of mercury enclosed in the evacuated bulb. Apart from this the tube was similar to an ordinary triode valve, but it was found that the presence of the mercury vapour had a profound effect on its behaviour. The grid could give only partial control of the anode current, being able to prevent it altogether when made sufficiently negative with respect to the cathode, but when once anode current started to flow, quite unable to exert any influence on this flow, regardless of the potential applied to the grid.

Because of this action, the name *thyratron* was given to the device, coined from the Greek word *thyra*, meaning a door ; the idea being that it would either allow or prevent the flow of current in the anode circuit. There was no true valve action of the kind associated with the vacuum triode. The name thyratron strictly applies only to the product of the American company in whose laboratories Langmuir's experiments were made. Other firms making their own version of the device have used other names, such as *gasfilled relay* and *grid-controlled rectifier*, but for general purposes the term gasfilled triode is more descriptive and avoids commercial implications.

Principles of Operation

A circuit for demonstrating the action of a gasfilled triode is shown in Fig. 65. The anode voltage is fixed, but the grid voltage may be varied by moving the slider of the potentiometer resistance. In the extreme right-hand position of the slider the grid receives the maximum negative voltage and as the slider is moved to the left the grid voltage, although still negative, is progressively reduced until in the extreme left-hand position of the slider the grid voltage becomes zero with respect to the cathode. Several interesting facts about the gasfilled triode may be found with this simple circuit.



A = Ammeter. B_1 = Grid circuit battery. B_2 = Anode circuit battery. B_3 = Cathode heater battery. P = Potentiometer (variable resistor). R = Current limiting resistance. S = Anode circuit switch. V = Voltmeter.

FIG. 65.—Circuit for showing properties of gasfilled triode.

With a sufficiently high voltage grid battery B_1 and the slider of potentiometer P in the maximum voltage position (extreme right) the switch S maybe closed, thus applying the full voltage from the anode battery B_2, without any flow of anode current taking place. Under these conditions the electrons emitted from the cathode are so strongly repelled by the negative field of the grid that none of them can penetrate the grid structure. A cloud of electrons, forming a negative space-charge, thus congregates in front of the cathode.

If now the potentiometer slider is moved slowly to the left so as gradually to decrease the grid voltage, a point is reached where some of the electrons emitted from the cathode pass through the grid region. On their way to the anode these electrons cause ionisation of the gas space by collision with the gas molecules and thus precipitate the breakdown process leading from the no-current condition to that in which a steady anode current flows. This process is extremely rapid, occupying only a few millionths of a second. The value of grid voltage at which it occurs depends upon the anode voltage, being higher for higher anode voltages, and is called the *critical grid voltage* for the particular anode voltage in question.

A very important condition arises owing to the appearance of positive ions when conduction commences. As the grid has a negative voltage it represents a source of attraction to these ions, which therefore settle all over its surface, thus forming a positive layer or ionic sheath which prevents the grid having any controlling effect over the discharge, irrespective of how high the grid voltage is made. In fact it will be seen that the higher the negative grid voltage, the greater will be the attraction for ions in the conduction path. When once it has been started, the anode current flow can be stopped only by reducing the anode voltage to a value below that necessary for sustaining the discharge or opening the anode circuit altogether. The discharge and its attendant ionisation collapse rapidly under these conditions, the period required for complete deionisation, or the *deionisation time*, being measured in millionths of a second. The period is largely determined by the internal structure of the tube and the polarities of the members during the deionisation process. Thus a large negative grid surface close to the discharge path very much facilitates the clearance of ions, and their rate of recombination is more rapid with low gas pressures.

After the lapse of the deionisation time the grid is able to regain its control of the anode conduction, which cannot start again if the grid is sufficiently negative when the anode voltage is restored. It may be noticed that if an *alternating* voltage is applied to the anode the gasfilled triode not only acts as a rectifier, passing anode current only when the anode is positive with respect to the cathode, but that the grid has the opportunity during each negative half-cycle of anode voltage of regaining control of the striking or breakdown process leading to the flow of anode current in the subsequent positive half-cycle. Before the full implications of this statement can be grasped it is necessary to understand quite clearly the relationship between the anode voltage and the critical grid voltage.

We have seen that for a given anode voltage the critical grid voltage can be determined by gradually reducing the grid voltage from some high value until anode current passes, as indicated by a sudden reading of the ammeter and the appearance of a glow inside the tube. The grid voltage at which this anode conduction occurs is the critical value. By repeating the process for various values of anode voltage and plotting the results as shown in Fig. 66 a curve may be drawn showing the negative grid voltage which will just allow conduction with any given value of anode voltage. Such a curve is called the *grid control characteristic*, and a considerable portion of it is straight, indicating a linear relationship between anode volts and critical grid volts. From this part of the curve may be calculated the *grid control ratio*, which is simply the slope of the straight part of the curve. In Fig. 66, for example, the two points *A* and *B* correspond to a difference of 100 anode volts and 6 grid volts. The slope, and therefore the grid control ratio of the tube in this particular case, will be given by 100/6 = 17 approximately. It should be understood

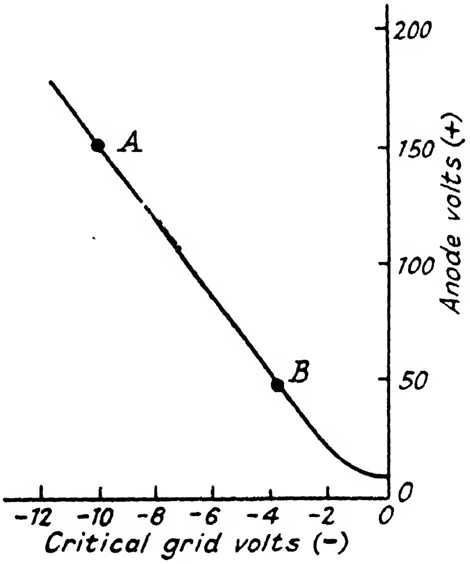


FIG. 66.—Grid control characteristic of gasfilled triode.

that as the determination of the grid control characteristic is affected by various factors the conditions of measurement have to be carefully controlled if consistent results are to be obtained. The grid control ratios of different types of tubes vary considerably, according to the intended class of service, and even between tubes of the same type there are usually small differences caused by unavoidable variations in manufacture.

Operation as a Rectifier

The full line in Fig. 67 represents a cycle of alternating voltage which we are to imagine is applied between anode and cathode of a gasfilled triode. Underneath the positive half-cycle is drawn dotted the curve of critical grid voltage, obtained from a curve such as that in Fig. 66 showing for every instantaneous value of anode voltage the grid voltage which will allow the striking of anode current as the grid voltage is reduced from some more negative value. The crest value *a* of anode voltage corresponds to the maximum critical grid volts a_1. Lower anode voltages such as *b* require a correspondingly lower grid voltage b_1 to prevent conduction. All voltages are measured from the central zero line, positive values being above and negative values below this line. These relationships allow of variation of the instant of conduction in each positive half-cycle of anode voltage by suitable variations of grid voltage and thus allow a corresponding variation in the anode current.

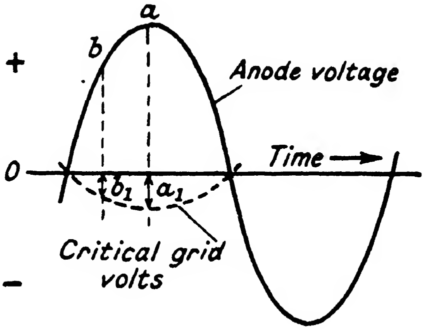


FIG. 67.—Relation between critical grid voltage and alternating anode voltage.

One way of achieving this result is shown in Fig. 68, in which the upper diagram shows the circuit arrangement while the lower diagram represents the voltage and current relationships with respect to time. In this case a steady alternating voltage is applied to the anode while the grid receives a direct voltage varying in magnitude according to the position of the potentiometer, but always negative with respect to the cathode. With a sufficiently high grid voltage, as represented by line 1 there is at no time any crossing of the critical grid voltage curves, and therefore no anode current flows. If now the grid voltage is lowered gradually it will eventually reach the critical value at a point *p* corresponding to the peak value of anode voltage (line 2). Conduction will accordingly commence at this point and flow for the remaining quarter of a cycle, being stopped only when the anode voltage passes through zero because, as we have seen, the grid has no further control while anode current flows. With still further reduction of grid voltage the conduction point or striking point becomes earlier and earlier in the cycle (points 3 and 4) until in the limit, when the grid voltage is zero, current flows during the entire half-cycle. As the effective value of the rectified anode current, as indicated by the ammeter, is proportional to the duration of current flow in each successive half-cycle it can obviously be varied at will with this circuit simply by varying the grid voltage. As the minimum and maximum conduction periods are a quarter-cycle and half-cycle respectively, it follows that a two-to-one variation of anode current is obtainable with this arrangement, but the current

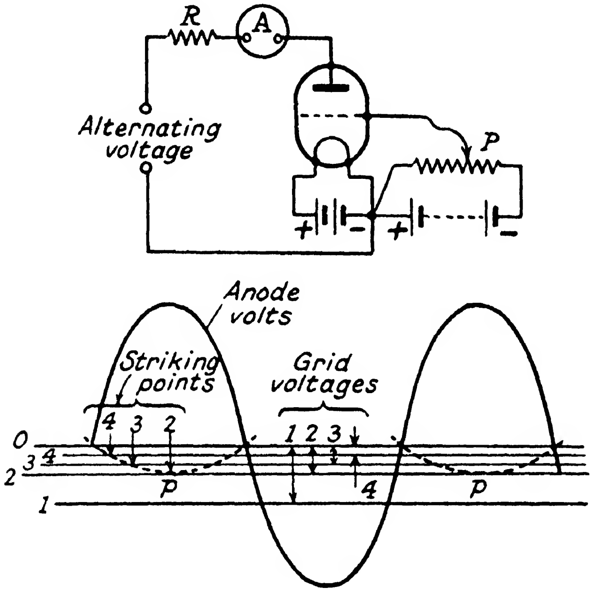


FIG. 68.—Control of striking point in anode cycle with D.C. on grid.

cannot be brought below the value corresponding to quarter-cycle conduction.

With the circuit shown in Fig. 69 it is possible to vary the striking point over the entire half-cycle and thus to control the anode current between zero and maximum. Alternating voltages are used on both anode and grid, and arrangements are made for the *phase* of the grid voltage to be varied in relation to that of the anode voltage, i.e., so that the instant at which the grid voltage reaches its peak value, for example, may be changed with respect to the instant at which the anode voltage reaches its peak value. The frequency of both voltages is the same, so that there is a constant relationship between them throughout successive cycles. There are several ways of varying the phase of the grid voltage, and in the arrangement shown it can be done

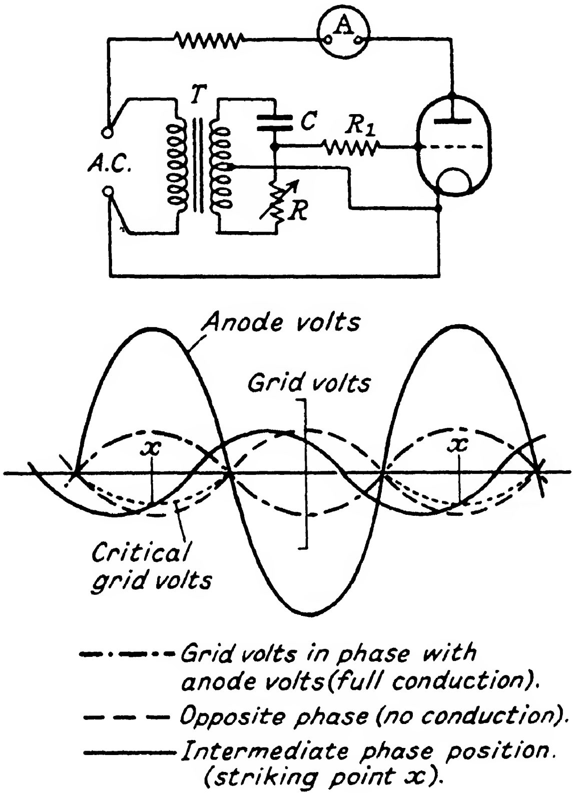


FIG. 69.—Control of striking point by phase-shift control.

by varying either the resistance *R* or the capacitance *C*. In either case the effect is to change the phase of the grid voltage and to shift the striking point, thus changing the anode current. When the voltages are of opposite phase the grid voltage is at all times greater (more negative) than the critical grid voltage, and the anode current is zero. When the voltages are of the same phase the grid voltage becomes positive at the beginning of each positive half-cycle, and maximum anode current flows. The current can be varied between zero and maximum, provided that the required range of adjustment of resistance and/or capacitance is available. A grid circuit resistor is always used when there is any possibility of the grid becoming positive with respect to the cathode, so as to limit the grid current.

Circuits of this kind have a very wide range of application in all kinds of automatic control systems where the control element can be used to produce the required grid circuit changes. Sometimes the control is indirect, as when a selenium bridge is used in place of the variable resistor in Fig. 69, making the anode current passed by the gasfilled triode dependent upon changes of illumination affecting the resistance of the bridge.

MERCURY ARC RECTIFIERS

The type of cathode we have hitherto considered, with a coated metal surface and means of direct or indirect heating from an independent power source, is not well suited to providing very high emission currents. It is liable to damage in the event of excessive currents, even of short duration, and is thus unsuitable for heavy industrial service where high current ratings are required. In such cases use is made of the *mercury pool cathode.* This form of cathode necessitates special features of design and has produced a particular form of construction typified by the *mercury arc rectifier.*

The general principle is shown in elementary form in Fig. 70, where *C* is the pool of mercury forming the cathode, *A* is the anode and *E* the excitation or ignition anode which is necessary to bring the cathode into operation. If a suitable voltage is applied to electrode *E*, positive with respect to the cathode, and the electrode is then brought into contact with the mercury, as for example by tipping the tube to the right, an arc will be drawn between the electrode and the mercury when they are separated again. With a positive voltage on the main anode the conditions are thus favourable to conduction and rectifier action, so long as the cathode is maintained operative, i.e., supporting an arc during the periods when the main anode is negative. Under the action of the arc, the *cathode spot* or the point of the mercury surface to which the arc attaches itself is raised to a high temperature, of the order of 600° C., and forms a copious source of electrons and mercury vapour for supporting the arc between cathode and anode. At the end of each positive half-cycle the arc fails as the anode voltage approaches zero and cannot re-strike in the opposite direction (as the anode becomes negative) because the conditions at the anode are not conducive to emission. In practice the anode temperature is kept down by making it massive and cooling it if necessary by artificial means, and a poorly emitting material such as graphite is used as a further deterrent to possible conduction in the wrong direction which would of course upset the rectifier action.

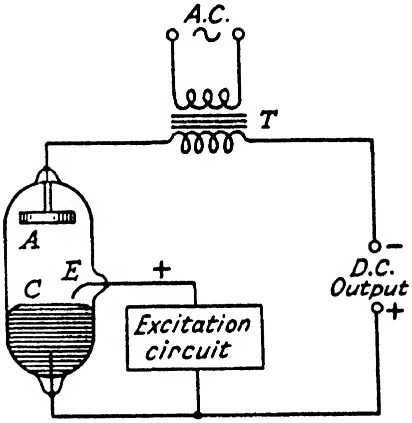


FIG. 70.—Elementary form of mercury arc rectifier.

It will be seen that the mercury pool cathode is extremely durable, the active region or cathode spot being continuously and automatically replaced. The size of the spot varies with the arc current, and generally corresponds to a current density of about 25,000 amperes per square inch. The mercury vapour liberated by the action of the arc condenses in the cooler parts of the tube and runs back into the cathode pool. In one form of glass-bulb rectifier the cooling action is increased by making the upper part of the bulb in the form of a large dome.

In an alternative form of construction the electrodes are enclosed in a dismountable steel tank, so that during operation nothing can be seen from the outside except the auxiliary apparatus used for water cooling and the special pump for producing and maintaining the necessary vacuum. This construction is well suited to intensive cooling of the electrodes the space surrounding the arc, so that relatively high currents and can safely be dealt with. This form is known as the *ironclad* mercury arc rectifier, and there are several variants, including the *steel bulb* type in which the vacuum is permanent, as in the glass bulb type, although arrangements are incorporated for renewing the vacuum with a portable pumping set if necessary

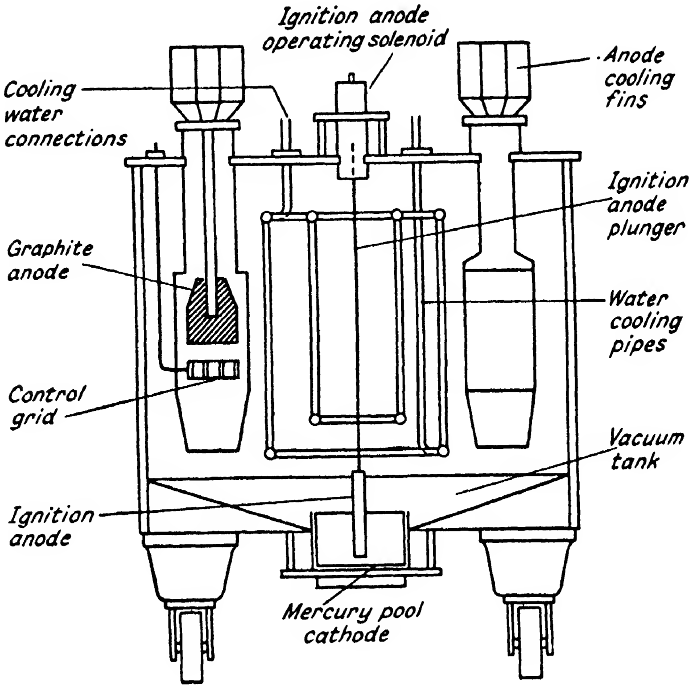


FIG. 71.—Sectional view of ironclad mercury arc rectifier (simplified).

at any time during service. Fig. 71 gives a sectional view of an ironclad rectifier and shows the location of the principal parts.

In practice, mercury arc rectifiers are usually operated from three-phase A.C. supplies and may have three, six or twelve anodes according to the arrangement of the secondary windings on the main transformer. In multiple anode rectifiers of this kind each anode makes its contribution to the total rectified current in sequence as determined by the order in which the anode voltages, i.e., the voltages of the respective secondary windings, become positive with respect to the cathode during the alternations of voltage.

The principles of grid control which we have considered in relation to the gasfilled triode are applicable to mercury arc rectifiers and are used in cases where it is desired to control the rectified anode current. The great advantage of the method is that the control is virtually without losses, where as if control of the current were obtained, for example, by means of resistors, the losses in the form of waste heat would be prohibitive in the majority of cases.

The Ignitron

The necessity of having auxiliary anodes for maintaining the cathode spot during negative half-cycles has been overcome in the *igniter-type rectifier* or *Ignitron* by arranging for the arc to be struck afresh at the desired point of each successive positive half-cycle without the use of grids. As shown in Fig. 72, the ignitron consists of an evacuated bulb containing an anode *A*, mercury pool cathode *C* and an igniter electrode *K*. This electrode is made of semiconducting material and its tip is permanently immersed in the mercury pool. If a moderate current is passed between the igniter electrode and the mercury the heat generated at the region of contact is sufficient to form a cathode spot and thus prepare for the striking of the anode arc without the igniter electrode being withdrawn. In the diagram the igniter current is shown applied by means of a simple switch *S* from the battery *B*, but in practice a timing circuit is used to deliver pulses of current to the igniter at the same frequency as the main A.C. supply. By varying the timing of these current pulses in relation to the positive half-cycles of anode voltage the striking point of the anode arc is varied, as in the case of grid control, and the effective value of the rectified output current thus con-

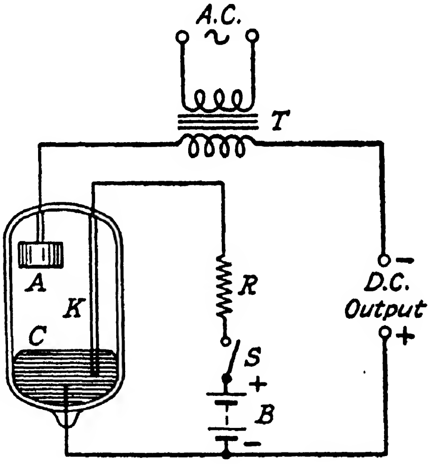


FIG. 72.—Arrangement of igniter type rectifier (ignitron).

trolled. Rectifiers of this type are available with ratings up to several hundred amperes, the large units being generally water-cooled by means of an outer jacket through which the cooling water is circulated. Forced draught or natural cooling is used for small units.

The principal sphere of use of the ignitron is in connection with electric welding, owing to the ease with which the short pulses of high anode current required for this work may be controlled. The timing circuit usually employs a gasfilled triode for energising the igniter, so that grid control of the triode at a very low power level gives control of the rectified output current.

FOR FURTHER READING

The majority of books on rectifiers deal mainly or exclusively with the mercury arc type as applied in heavy electrical engineering practice. The following are representative :—

1. H. Rissik. *The Rectification of Alternating Current.* (English University Press), 1938.
2. H. Rissik. *The Fundamental Theory of Arc Convertors.* (Chapman and Hall), 1939.
3. J. Rosslyn. *Power Rectifiers* (Newnes), 1941.
4. A. Guntherschulze. *Electric Rectifiers and Valves.* (Chapman and Hall), 1947.

For valves of the thyratron type reference may be made to :—

5. G. Windred. *The Gasfilled Triode (Thyratron) and its Applications.* (" Electronic Engineering " Monograph), 1947.

CHAPTER IX

X-RAYS

The discovery of X-rays was made in 1895 by W. C. Roentgen in the course of his studies of the ultra-violet light produced under suitable conditions by an electric discharge through an evacuated tube. For detecting this light Roentgen was using crystals of platinum-barium-cyanide on a paper screen which became fluorescent under the action of the ultraviolet rays, and he noticed that the effect continued when the tube was covered with opaque paper so that the discharge could not be seen. The effect could be stopped only by interposing heavy solid objects between the tube and the fluorescent screen. In view of the early period of these discoveries, it is not surprising that Roentgen gave the name X-rays to the mysterious emanations from the tube which could penetrate appreciable thicknesses of different materials without their source being visible.

As might be expected, this interesting discovery was taken up by numerous other workers, and their efforts soon revealed new facts. It was found that in addition to producing fluorescence in certain salts the rays affected photographic plates and were also capable of causing the ionisation of gases. As we have already seen, the latter property was made use of in early experiments on the conduction of gases (see page 49). It was thus possible to study the rays by visual, photographic and electrical means, and a great deal of information was assembled as the result of experimental work by numerous investigators. The chief discoveries were that the rays were not subject to reflection and refraction like light and were not affected by a magnetic field, that they could penetrate solid substances in varying degree and that they were scattered and partially absorbed in doing so.

The X-ray Tube

The basic arrangement of an X-ray tube and its operating circuit is shown in Fig. 73. In this case the bulb *B* is a highly evacuated glass tube, containing the anode *A*, also called the ***anti-cathode*** or ***target***, usually made of tungsten and representing the source of the X-rays. The cathode *C* is of the thermionic type, introduced into X-ray tubes by W. D. Coolidge in 1913. The emission of the cathode is controllable by means of the regulating resistance *R*. The D.C. supply to the anode is of high voltage in order to obtain sufficiently active radiation. Voltages ranging from about 10,000 to several hundred thousand are usual and with the most modern tubes, differing radically in detail from the simple glass bulb construction, voltages up to 1,000,000 are being used.

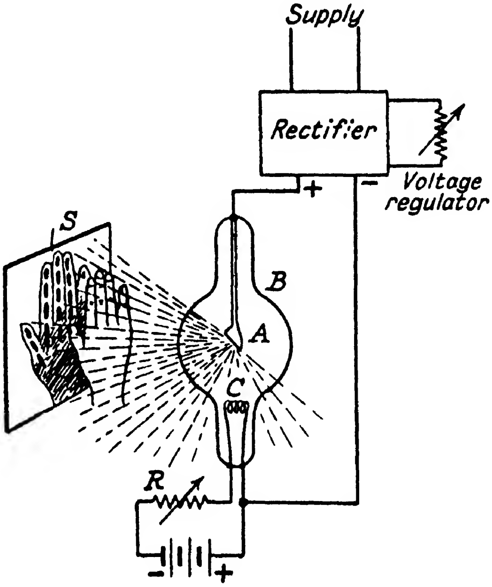


FIG. 73.—Basic arrangement of X-ray tube.

The electrons emitted by the cathode are accelerated to the anode owing to its high potential and strike it at very high speeds. It is in this collision process that the X-rays are produced, although the exact mechanism of the action is not yet fully understood. Part of the energy of the bombarding electrons appears as heat in the anode material but the remainder is transformed into the radiation of short wavelength which we call X-rays. It is interesting to note that the process is virtually a reversal of the photoelectric effect, in which the incidence of light (visible or otherwise) liberates electrons from a photosensitive material. In the case of X-rays the incidence of electrons of sufficient velocity on a given material produces radiation having the nature of light but of such short wavelength as to be invisible.

As the anode voltage is increased, the penetrating power of the rays increases—the rays are said to become harder. The majority of equipments have means for varying the anode voltage so that the hardness of the rays may be adjusted to suit individual conditions.

Various modifications of design have been introduced from time to time owing to the tendency for more and more powerful tubes and their application to an increasing variety of uses. In some cases the anode is continuously rotated while the tube is in use so as to minimise the destructive effect of the local and intense heating caused by the electron bombardment. Some of the large tubes are arranged for continuous evacuation by pumps which form part of the auxiliary equipment. Water-cooled and oil-immersed tubes are also widely used.

Medical Uses of X-Rays

There is little doubt that the most generally familiar applications of X-rays are the medical ones. For many years past the medical profession has made extensive use of X-ray methods in the diagnosis of ailments and in the treatment of certain bodily disorders. In the case of diagnosis the penetrating action of the rays is used to obtain what might be called a shadow-picture in which internal structures may be distinguished by contrast due to the different degrees of penetration of the various parts according to the nature of their substance. The principle as relating to visual observation is shown in Fig. 73, where a platinum-cyanide screen *S* produces the shadow-picture of the part to be examined, which is placed between the tube and the screen. In this case the picture is formed by the varying degrees of fluorescence of the screen due to the varying degrees of penetration of the X-rays through the different parts of the object. When a photograph is required, a plate or film is substituted for the fluorescent screen and exposed to the action of the rays for sufficient time to obtain the required degree of contrast. It will be noticed that as a photograph shows the accumulated effect of the rays during the time of exposure it gives greater contrast than a screen picture for a given hardness of the X-rays, so that when using the screen it is generally necessary to use higher anode voltages than in photography. An advantage of visual observation in many cases is that the object may be moved while under observation. It may be mentioned here than an X-ray photograph is technically called a *radiograph*, and that the technique of producing these pictures is called *radiography*.

In X-ray therapy use is made of the damaging effects of the rays on living tissue in order to destroy harmful growths and eliminate diseased conditions of the skin. A wide variety of skin diseases can be successfully treated by X-rays and their use in the treatment of cancer is well known. In early cases the growths may be completely destroyed so as to render surgery unnecessary and in other cases the size may be sufficiently reduced to make surgical removal a simple matter. The name *deep therapy* has been given to the use of highly penetrating rays for reaching deeply situated parts of the body which may be difficult of access to the surgeon. In all cases it is necessary to protect healthy tissue against excessive doses of the rays; otherwise the characteristic X-ray burns may appear. These burns do not respond well to treatment and may develop into bad sores which may even necessitate amputation, when this is possible. Excessive exposure to X-rays also causes a reduction in the number of red and white corpuscles in the blood (especially the latter) and may cause serious forms of anaemia.

X-ray Examination of Materials

During recent years there have been striking advances in the application of X-ray technique to the examination of various kinds of structural materials, especially metal castings. This method of examination has the great advantage that it is non-destructive. When once the necessary technique has been established for interpreting X-ray observations or radiographs the method is also quick and reliable.

It should not be thought that the X-ray examination of materials is an entirely new subject. Roentgen himself, shortly after his initial discovery of the rays, suggested their use for examining the properties of alloys, for discriminating between real and artificial jewels, for finding flaws in metallic welds and for examining the insulation of cables. In 1896, within a few months of Roentgen's discovery, Dr. Obach was using X-rays to control some of the processes in cable manufacture by detecting foreign bodies in the gutta percha used in submarine cables, by checking the centralisation of cable core and by locating faulty joints and internal air bubbles in finished cables.

At the present time special X-ray equipments are available for routine inspection of materials in industry and are widely used for the examination of parts and castings for aircraft construction and for the examination of welds in ship-building and other structural work. The advantage of being able to determine the interior condition of an article without cutting it open is obviously very great, and there is little doubt that X-ray methods in industry will rapidly extend in the future. This applies particularly to foundries, where X-ray examination can not only reveal internal faults in castings but is a valuable guide to the elimination of such faults by changes in the casting process. Equipments working at 250,000 volts can be used for the examination of steel up to about three inches thick, and with the latest million-volt equipments the thickness of steel that can be examined is increased to about eight inches. It thus appears that X-ray inspection can be effectively applied in the great majority of industrial cases.

Industrial X-ray Equipment

The chief components of an X-ray equipment are the tube, the high-voltage generator (generally a special transformer and rectifier arrangement) and the control panel. This panel includes as essentials a milliameter for indicating the anode current, a regulator for selecting the required anode voltage and a regulator for varying the cathode heater current. Other devices, such as a mains voltmeter and safety arrangements, indicator lamps, etc., for the protection of water-cooled tubes are generally also included, according to the type of installation. Specially insulated and metal-sheathed cables are necessary for safety in view of the high voltages employed.

A typical equipment intended for general use in industry is shown in Fig. 74. The tube, which is mounted on a carrier adjustable both as to angle and height, is of ironclad construction and designed so as to confine the rays to the funnel-shaped opening. It is connected to the high-voltage unit by special flexible cables. The control desk can be used on the trolley as shown in the figure, or removed to a distance if required by the operator. Such an equipment has numberless applications, ranging in this case from the radiographic examination of steel welds up to five-eighths inch thick and light alloy castings up to

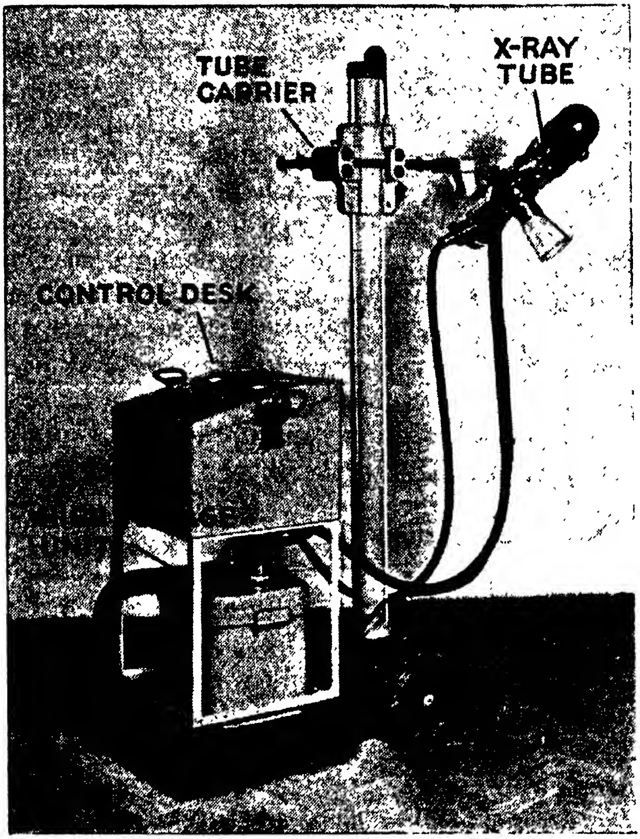


FIG. 74.—Industrial X-ray equipment (by courtesy of Philips Electrical Ltd.).

four inches thick to the detection of impurities and deterioration of tinned foods.

When a visual method of inspection is used it is important that the operator shall not be exposed, at any rate for long periods, to the rays passing through the viewing screen. The requisite protection is given by the simple arrangement in Fig. 75 using an inclined mirror and horizontal screen. For routine inspection

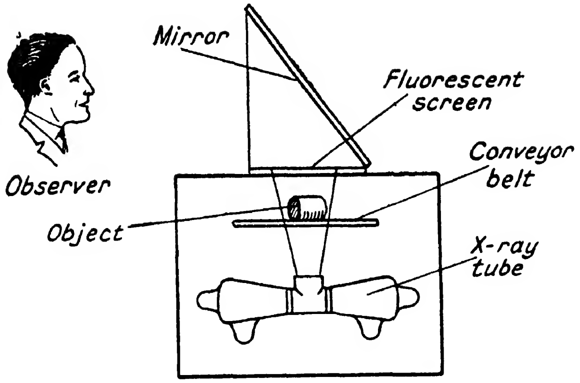


FIG. 75.—Arrangement for X-ray inspection of articles on conveyor belt.

of large numbers of articles they may be placed on a conveyor belt located as shown, the speed of the belt being adjusted so that the shadow-picture seen in the mirror as each article passes by can be thoroughly examined.

FOR FURTHER READING

The following books deal with general principles of X-ray theory and technique :—

1. H. M. TERRILL and C. T. ULREY. *X-Ray Technology.* (New York, Van Nostrand ; London, Chapman and Hall), 1930. Gives practical aspects of X-ray technique, mainly for users of apparatus, and describes usual equipment. Very little mathematics.
2. A. H. COMPTON. *X-Rays and Electrons.* (New York, Van Nostrand ; London, Macmillan), 1927. An advanced and mainly theoretical treatment of experiments on the structure of matter as revealed by X-rays. Mathematical.

CHAPTER X

THE CATHODE RAY TUBE

The modern cathode ray tube may be regarded as a highly elaborate development of the many forms of tube used by early experimenters for studying the so-called cathode rays. This work commenced with Hittorf's discovery of the rays as long ago as 1868, and it soon became known that they could be deflected by a magnetic field. Suggestions were then put forward concerning the use of this effect as a means of measurement, and the first successful oscillograph, incorporating for the first time a fluorescent screen, was produced by Braun in 1897. The Braun tube usually contained nitrogen at a pressure of a few millionths of an atmosphere, as it was found that under these conditions a very fine pencil of rays was produced. New features were added by several other workers until the modern type of cathode ray oscillograph was evolved, differing in almost every point of detail from the original Braun tube.

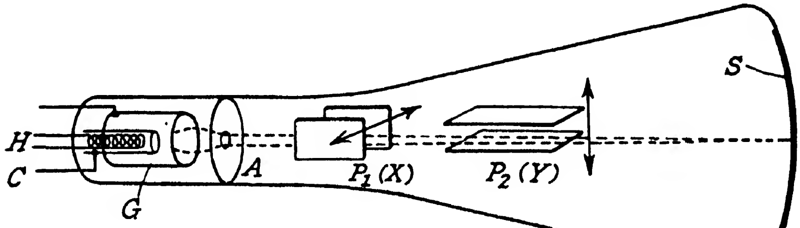


FIG. 76.—Diagrammatic arrangement of cathode ray tube.

The essential parts of a modern form of cathode ray tube are shown in Fig. 76. The enclosing bulb is of glass and contains at the narrow end a thermionic cathode *C*. Around this cathode is a metal cylinder *G* which is made negative and has the effect of concentrating the electron beam from the cathode, thus preventing divergence. This cylinder is known variously as the grid, shield, negative cylinder, Wehnelt cylinder or, in television

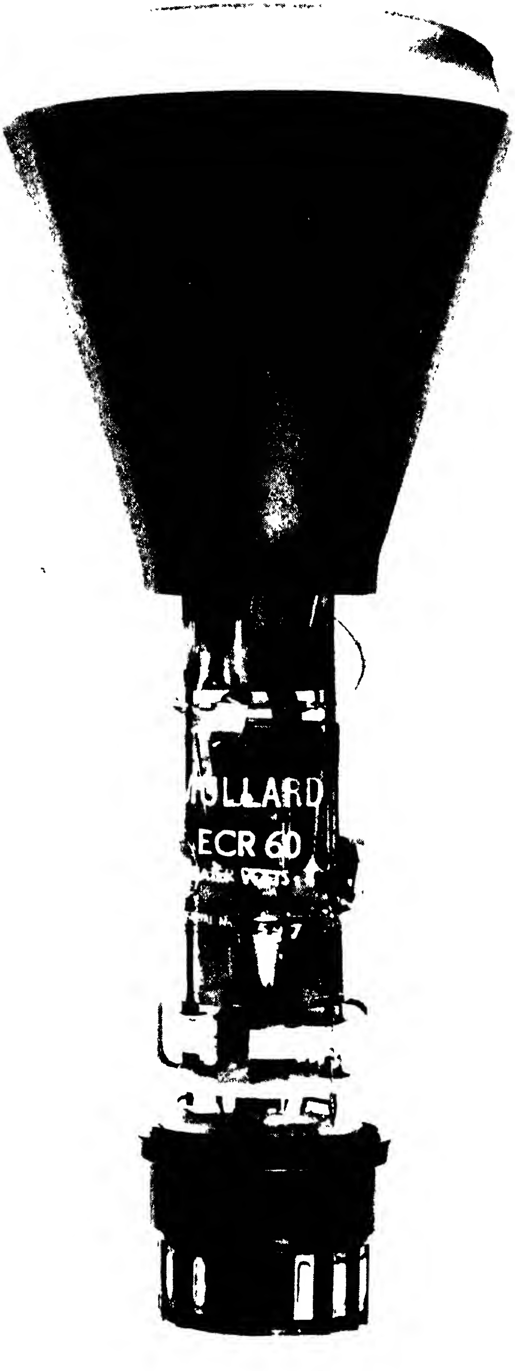


PLATE 3

A modern electrostatically focused cathode ray tube of a type widely used during the last war. The beam is electrostatically deflected by the plates shown under the edge of the black coating in the bulb.

(*Mullard Radio Valve Co., Ltd.*)

tubes, the modulator. The anode *A* for accelerating the electrons is in this case a single metal disc with a central hole about one mm. diameter for the beam to pass through. The voltage applied to the anode depends upon the purpose for which the tube is intended, but values of 1,000 volts or more are quite usual. After leaving the anode the electron beam passes through two pairs of deflecting plates P_1 and P_2 arranged at right angles. The voltages applied to these plates determine the position of the beam at any instant. The inside of the enlarged end of the bulb is coated with a special compound such as Willemite or zinc silicate forming the screen *S* which becomes luminescent where the electron beam strikes it.

It is obviously important for clearness of definition that the luminous spot where the beam strikes shall be as sharply defined as possible. As a beam of electrons always tends to diverge because of the mutual repulsion of individual electrons it is necessary to incorporate in the tube some means of preventing this and keeping the beam in the form of a thin pencil. There are three ways of obtaining the desired effect : (1) By means of electrostatic fields produced by a system of auxiliary anodes ; (2) by a magnetic field produced by a cylindrical coil surrounding the neck of the tube (mainly for television) ; (3) by introducing a small amount of gas, usually argon or helium, into the evacuated tube. The presence of the gas causes ionisation by collision and the positive charge represented by the ions neutralises the repelling charges due to the electrons. There are limitations to this method at high frequencies owing to distortion and loss of focus caused by the presence of the ions.

Principles of Operation

From the foregoing brief description it will be understood that the cathode ray tube is essentially an indicating instrument. Instead of a pointer it employs an electron beam, the position of which depends upon the voltages applied to the deflecting plates, and is indicated by the luminescence of the screen where the beam strikes it. If the applied voltages are steady the beam gives a luminous spot in a fixed position on the screen, but if, as is usually the case, the applied voltages are continuously varying, then a luminous line or trace is seen instead of a spot. The shape of the line or the figure which it traces is determined by the variation with respect to time and with respect to each other of the applied voltages. Special circuits are employed in conjunction with the cathode ray tube so that the figure traced out by the moving spot gives clearly intelligible and precise information concerning the applied voltages.

Returning to Fig. 76, it will be seen that the respective pairs of deflecting plates P_1 and P_2 have their planes at right-angles. Their action on the electron beam when subjected to suitable voltages may be understood by remembering that the beam is composed of negative charges (electrons), all of which will be attracted by the plate which is positive and repelled by the plate which is negative. This means, for example, that if when looking at the front of the tube a voltage is applied between plates P_1 so that the left-hand plate is positive (and the right accordingly negative) the luminous spot on the screen will move to the left. If the polarity of the plates is reversed, then the spot will move to the right. With no applied voltage the spot is in the centre of the screen. Similar reasoning can be applied to the plates P_2. If the upper plate is positive the spot moves upward and *vice versa.* It is seen that the plates P_1 give horizontal deflections of the spot as shown by the corresponding arrows. For this reason they are often referred to as the X-plates ; the other pair, giving vertical deflections, being called the Y-plates. When voltages are applied to both sets of plates the position of the spot depends upon the relative magnitude and polarity of the voltages. If when looking at the front of the tube the left-hand X-plate and the upper Y-plate are both positive at a given instant, then at that instant the spot will occupy some position in the upper left-hand part of the screen. Its exact position will be determined by the relative values of the voltages at the particular instant.

Time Bases

In the most general manner of use the X-plates receive by means of a special circuit a continuously recurring pulse of voltage which sweeps the spot across the screen (say from left to right) at a uniform rate, returning it very rapidly to the left at the end of each sweep. This rapid *flyback* as it is called is essential for clear results. If there is no voltage on the Y-plates the observer sees a luminous line passing horizontally across the centre of the screen. If now the Y-plates receive a voltage varying with respect to time, the trace is no longer horizontal but assumes a shape corresponding to the variations of the Y-voltage with respect to time. For this to be true it is necessary that the periodic voltage applied to the X-plates shall sweep the spot across the screen at a perfectly uniform rate, i.e., through equal distances in equal times. This arrangement, which is secured by ingenious circuit devices, forms what is called a *linear time base.*

In most cases the uniform increase of voltage with respect to time across the X-plates is obtained by connecting them across a condenser which charges through a series resistance. The conditions can be arranged so that the charging voltage approaches very closely the ideal *saw-tooth* shape shown in Fig. 77. The rising portion of each pulse is the voltage across the condenser (and the X-plates) while the condenser charges, and corresponds to the forward linear sweep of the spot. The falling portion of each pulse is the condenser voltage during its discharge across a suitable short-circuit and corresponds to the flyback, which is made as rapid as possible. As an oscillograph has to deal with measurements of different frequencies it is necessary to include in the circuit a means of changing the *sweep frequency,* which is generally done by combined selection of appropriate values of capacitance from a group of condensers and adjustment of the corresponding charging resistance.

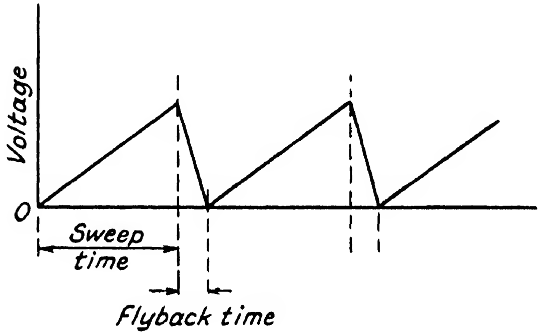


FIG. 77.—Ideal form of voltage variation for sweep circuit (linear time base).

A typical linear time-base circuit employing vacuum valves is shown in Fig. 78. The valves V and V_2 are pentodes and V_1 is a triode. It may be noticed that the condenser C, which provides the pulses of voltage, is connected directly across the X-plates and also across the circuit formed by the triode V_1 and its anode resistance R_1. This is the discharge circuit for condenser C, and controls the flyback process. When the circuit is connected to the D.C. supply, condenser C commences to charge through the pentode V which owing to its constant impedance gives the required linear rise of voltage across the condenser as it charges. During this action, while the spot is swept across the screen of the tube, valve V_1 is biased to the cut-off point by the anode current of valve V_2 passing through resistance R_2. The progressive rise of voltage across condenser C as it charges eventually imposes sufficient voltage on the anode of V_1 to overcome the bias on V_1, which then begins to pass anode current. The resulting voltage drop across resistance R_1 is conveyed through condenser C_1 to the grid of V_2, making it more negative and hence reducing its anode current. This in turn causes a drop in the voltage across resistance R_2, thus reducing the bias on V_1 which accordingly passes progressively more current until the condenser C is discharged. When this happens, the voltage across V_1 is zero and the process starts over again. The discharging action corresponds to the flyback period, the rate of which can be controlled by varying resistance R_1. The adjustment of R_2 determines the voltage at which the condenser begins to discharge, and therefore also the amplitude of sweep of the spot across the screen.

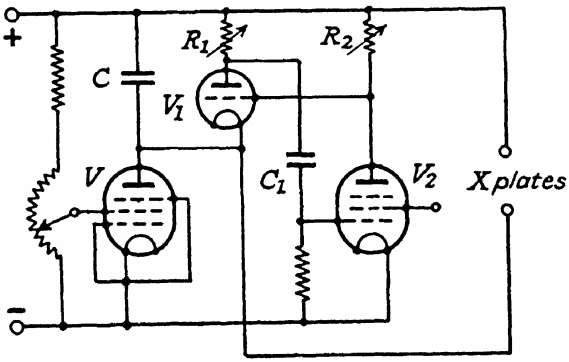


FIG. 78.—Linear time-base circuit.

In view of the somewhat complex actions taking place in a circuit of this kind it may be difficult to realise that the entire process just described may take place many thousands of times per second. The cathode ray oscillograph is widely used for measurements at high frequencies and many ingenious refinements have been incorporated in modern designs. One of these makes use of a double beam, obtained by arranging an electrode so as to split the beam into two sections, each of which is influenced independently by *one* of the Y-plates. This arrangement allows of the simultaneous showing of two independent traces on the screen and has greatly extended the usefulness of the cathode ray oscillograph.

It is usual to provide an amplifier for increasing the voltage to the Y-plates so that very small voltages can be observed and measured. The amplifier is linear, so that equal vertical distances on the screen represent equal voltages. A ruled screen is often provided so that the relative or absolute values of deflections may be estimated. Traces on the screen may also be photographed with suitable apparatus, including the appropriate kinds of photographic plates and paper which are available for oscillographic work. Several types of oscillograph are available for a wide range of purpose. The frequency range of these instruments is normally from about 10 cycles per second to 250,000 or more.

Special Applications

Apart from the very numerous applications as a routine measuring instrument in electrical laboratories the cathode ray tube has important uses for special purposes. The sensitiveness of the electron beam to the action of a magnetic field forms the basis of a proposal to use the tube as a compass. In this case the deviations relative to the earth's field are indicated by the movements of the spot on the screen. Another form of the idea uses small metallic bucket electrodes placed close together in the centre of the screen. Deviation in either direction then causes one or other electrode to receive the electron beam and the resulting charge is utilised to operate auxiliary apparatus, such as relays.

Another use is in conjunction with one of the many devices which convert mechanical pressure or force changes into electrical variations. These include the piezo-electric crystal (see page 118) which produces a voltage when mechanically stressed, the carbon pile which changes its resistance when subjected to a force, special compound materials which change their resistance when bent or flexed, and electromagnetic pick-up devices of various kinds for converting mechanical into electrical variations. Any of these devices may be used, according to the nature of the problem, to provide an electrical signal, usually in conjunction with an amplifier, for measurement and study in the cathode ray tube. In this way much valuable information may be gained regarding the stresses involved in high-speed mechanical processes which cannot readily be studied by other means.

The cathode ray tube also has electromedical uses as an indicator of the potentials which accompany the action of muscles and nerves and from the records of which valuable information is provided for the physician. The branch of medical science known as electroencephalography, for example, uses the cathode ray tube as an indicator of varying potentials arising from the action of the brain. The state of this organ has a marked effect on the nature of the variations. Specially designed amplifiers are necessary in this work, as frequencies as low as one or two cycles per second may be recorded and an amplification of several millions is necessary to make the minute electrical signals usable.

Perhaps the most familiar use of the cathode ray tube is in television receivers. This application is considered separately in Chapter XI.

FOR FURTHER READING

Many general works on electronics include a treatment of the cathode ray tube. The following books may be recommended from the large number dealing specifically with the subject :—

1. G. Parr. *The Cathode Ray Tube and its Applications.* (Chapman and Hall), 1943. An introductory treatment with an extensive bibliography. Only essential mathematics used.
2. M. A. Bly. *Guide to Cathode Ray Patterns.* (New York, Wiley ; London, Chapman and Hall), 1943.
3. W. E. Miller. *Cathode Ray Oscilloscope.* (Iliffe), 1942.
4. S. K. Lewer. *The Cathode-Ray Tube Handbook.* (Pitman), 1945.
5. O. S. Puckle. *Time Bases.* (Chapman and Hall), 1947.
6. J. H. Reyner. *Cathode Ray Oscillographs.* (Pitman), 1947.

CHAPTER XI

TELEVISION

The transmission of intelligible information across distances by electrical means has occupied attention from the earliest beginnings of electrical science. The telegraph, the telephone and radio were the outcome of widespread and continuous research by many investigators dating back to the seventies and eighties of last century, but it has taken even longer for the modern technique of television to evolve, even though experimenters in the later stages were assisted by means not available to earlier workers. In just the same way that the development of radio had to await the advent of new kinds of valves—in particular the triode—so also television had to await the evolution of radio technique, the perfection of photoelectric apparatus and especially the production of electron cameras and special forms of cathode ray tubes.

As long ago as 1847, Bakewell produced an electro-chemical telegraph with which it was possible to transmit a message in the sender's handwriting. With this arrangement the sender used the extremity of a conducting wire in the manner of a pen, and at the distant point a reproduction of his writing appeared on a sheet of prepared paper. Any kind of outline drawing could be transmitted in the same way. A further step came in 1909, when H. Knudson used wireless means to transmit a portrait of King Edward VII across a room in the course of a demonstration at the Hotel Cecil. The telegraphic transmission of pictures also received considerable attention at this time, and by 1911, these methods had been put into service for transmitting pictures to the Press. By 1925 photoelectric methods were in use in America for transmitting portraits and other material over telephone lines. In 1926 J. L. Baird gave the first demonstration of practical television in England.

These were the beginnings of television, and by May, 1934, there had been sufficient progress to justify the appointment of a Television Committee by the House of Commons " to consider the development of television and to advise the Postmaster-

General on the relative merits of the several systems under which any public service of television should be provided."

Television Systems

In its simplest elements a television system comprises, at the transmitting end, some form of scanning device by which the differences of light and shade forming the scene or object to be televised are converted, with the aid of auxiliary apparatus, into electrical signals which are amplified and broadcast from an aerial by the normal radio technique. At the receiving end is another aerial which conveys the radio signals to an amplifying and reproducing system which converts them into light pulses having an intensity corresponding to the variations of light and shade of the original scene. These light pulses are projected successively on to a viewing screen in the positions corresponding to the original scene so as to reproduce at the receiving end the visual impression which an observer would have in viewing the actual scene at the transmitting end.

There are many ways in which this rather elaborate process may be carried out, but all of them have so far used the principle of *scanning* in which the scene is virtually divided up into a number of small parts, called *elements*. The radio signals between transmitter and receiver carry these elements in the sequence in which they are taken from the scene, and if the arrangements are such that each successive signal element causes at the receiving end an illumination of a corresponding element of the viewing screen at the correct intensity and in the proper position relative to the rest of the picture, then an impression of the original scene is conveyed *provided that the process is sufficiently rapid to give the eye the impression that all the elements of light appear together.* The quality of the picture produced in this way will obviously depend upon the grain or texture and will be better the greater the number of elements used, so that for satisfactory results two things are essential ; a sufficiently large number of elements and a sufficiently rapid rate of scanning.

Mechanical Scanning

The principles involved are perhaps best understood by considering the simplified diagram in Fig. 79, representing an early form of equipment using mechanical scanning. The subject to be televised is powerfully illuminated by the light sources S_1

and S_2, and by means of the lens system L_1 an image of the subject is projected on to the scanning disc. This disc has a series of holes at varying radius arranged so that as the disc is rotated by the motor M_1 at constant speed the whole of the picture is scanned in successive elements corresponding to the size of the holes. As it is known that the human eye cannot distinguish between a continuous picture and one which is repeated or interrupted at a sufficiently rapid rate, it follows that with sufficiently rapid scanning the impression of a continuous picture can be obtained. The effect of continuity is lost and flicker becomes noticeable when the complete scene is viewed less than twelve times per second.

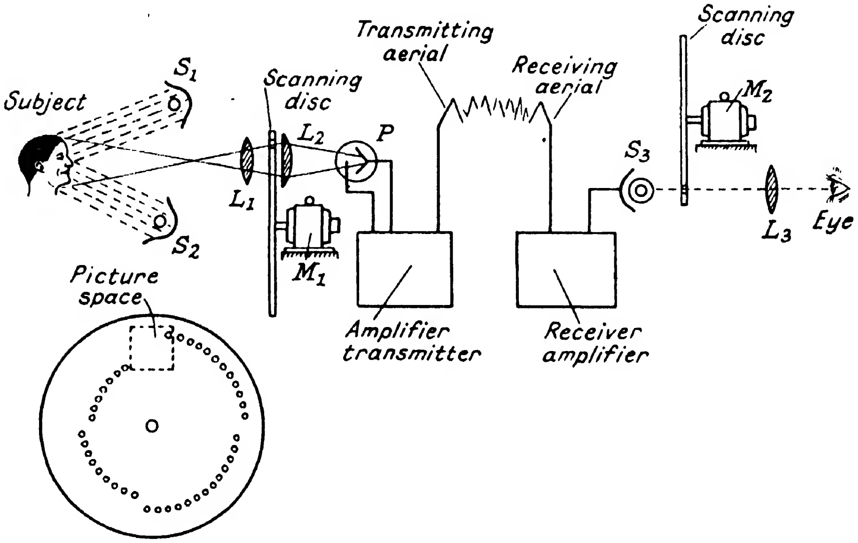


FIG. 79.—Early form of television system (mechanical scanning).

The light variations caused by the rotation of the scanning disc are passed through the lens system L_2 on to the photocell P, the current from which at any instant is a measure of the brightness of the particular element of the picture uncovered by a disc aperture at that instant. The varying photocell output is amplified and radiated by the transmitter. The receiving aerial conveys these signals to a receiver-amplifier which modulates the intensity of the light source S_3 (usually a form of neon lamp) in conformity with the instantaneous variations of the signals. The changes in light intensity are viewed through the lens system L_3. If the scanning disc at the receiving end runs at precisely the same speed as the one at the transmitter and the corresponding holes in each disc cover the same relative element of the picture at each instant then the variations of light intensity presented in succession to the observer's eye convey the impression of the original picture. It may be noted that in a system of this kind the transmission of the picture takes place through two processes; the light variations and the radio signals. Both these have the same speed of propagation—that of the speed of light, or 186,000 miles per second, so that the entire process of signal production, transmission and reception is virtually instantaneous for all ordinary purposes. This is an important factor which makes television possible.

There are many modifications of the simple scheme we have just described. The light source and subject can be interchanged, so that the subject is continuously scanned by a spot of light as the disc revolves. The photocell is placed so as to register the changes in illumination reflected from the successive parts of the subject. A screen may be used at the receiving end instead of a direct viewing system.

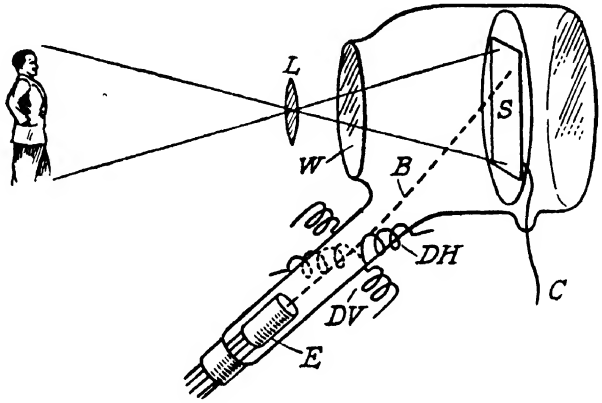
Modern Methods

There are many objections to mechanical systems of television, and numerous technical difficulties had to be overcome before such systems became reasonably practical. The modern tendency is to abandon mechanical methods in favour of what might be called all-electronic methods in which scanning discs and synchronously running motors are eliminated.

These new systems employ, at the transmitting end, an *electron camera,* of which several forms have recently been developed and in which the picture to be transmitted is scanned by an electron beam. At the receiving end is a *picture tube* resembling a cathode ray oscillograph and employing an electron beam to reproduce on the fluorescent screen a duplicate of the original picture.

One form of electron camera, known as an *iconoscope* is shown in Fig. 80. The bulbous end of the glass tube contains a special photoelectric screen *S* consisting of a thin mica plate having on the back a thin metallic layer and on the front, to receive the picture, a finely grained mosaic formed by minute particles of silver oxide, each coated with an extremely thin layer of caesium or rubidium, which as we have seen are photosensitive materials (see page 59). The mosaic thus forms a host of tiny independent photocells, numbering some hundreds of thousands per square inch.

The picture or scene to be televised is focused by an optical system *L* through the window *W* on to the mosaic screen. The variations of light intensity comprising the image thus formed on the screen cause a varying degree of emission among the elements of the mosaic; those which receive most light giving off most electrons. By this action the picture on the screen is converted into a distribution of positive charges (due to the loss of electrons), the charge on each element being proportional to the brightness of the picture at that point.



B = Electron beam. *C* = Connection to amplifier. *E* = Electron gun. *DH* = Horizontal deflection coils. *DV* = Vertical deflection coils. *L* = Lens. *S* = Mosaic screen. *W* = Window.

FIG. 80.—Electron camera for television transmission (Iconoscope).

The tubular extension of the bulb at *E* is virtually a cathode ray tube producing a fine electron beam *B* which, by means of the horizontal and vertical deflection coils *DH* and *DV*, is made to scan the mosaic screen. The scanning is done horizontally, as in the reading of a printed page, the process being controlled by a special circuit giving the effect of a cathode ray time base, but lowering the electron beam by a suitable amount after each sweep so that the entire picture is covered in a series of horizontal strips. The flyback of the beam is suppressed so as not to interfere with the reproduction of the picture. The present British system of television uses 405 lines per picture and this picture is repeated 25 times per second. A system of *interlaced scanning* is employed to reduce flicker. Every alternate line is scanned in the first half of the period and the remaining lines in the second half. Each half-period occupies one-fiftieth of a second, so that a complete scan takes a twenty-fifth of a second and there are thus 25 complete pictures per second.

Some idea of the rapidity of action involved in the scanning process may be gained from consideration of the fact that with a picture image four inches wide scanned in 405 lines 50 times per second, the scanning speed is more than 4,500 miles an hour! Even a few years ago such a process would have been quite beyond the bounds of possibility, and its perfection in such a relatively short time is an outstanding example of development assisted by progress in several branches. In the present case we can realise the dependence upon chemical deposition technique in making the delicate mosaic, the unusual application of photoelectric principles and the employment of cathode ray oscillograph practice for manipulating the scanning beam.

As the scanning beam passes over the successive elements of the mosaic the positive charge of each element is neutralised by the electrons in the beam. This loss of positive charge of each element induces a corresponding change of potential on the metallic film behind the mosaic, and this change, which is proportional to the intensity of light produced by the picture on that element, is passed on to the input of the transmitter amplifier through the connection *C* in Fig. 80. It is perhaps difficult to grasp the fact that these input signals follow each other at the rate of many millions per second. When the picture is completely scanned, all the original charges have been neutralised, but a new set, representing any minute changes of lighting or movement in the picture in the meantime, is immediately available for the repetition of the process during the next scanning sweep.

The radio signals representing the changes of light intensity detected during the scanning process are passed from the receiving aerial to an amplifier and thence to the picture tube, where they must be set up on the viewing screen in the form of light variations having the proper relative position to give a clear reproduction of the original scene or picture. The picture tube may be regarded as a cathode ray tube with a suitably large screen for viewing purposes. The movements of the electron beam are controlled by a synchronising and scanning circuit so that they are in exact agreement with those of the scanning beam in the electron camera at the transmitting end. The intensity of the beam in the picture tube is varied in accordance with the strength of the successive picture signals so as to give the appropriate variations of intensity for building up the picture. This variation of beam intensity is obtained by variation of the potential on the grid or modulator of the tube (see Fig. 76, page 146) by the picture signals. Proposals have also been made for *velocity modulation*, in which the intensity of the spot on the screen would be varied by changing the velocity of scanning movement of the beam. With this method the beam would move slowly for bright elements and rapidly (giving a fainter trace) for darker elements.

Sound Channels

So far nothing has been said about the picture transmission being accompanied by sound effects. As these are a necessity for most subjects they are normally regarded as part of the television system. The provision of sound obviously requires the addition of a separate broadcasting system on the usual lines, with a microphone to pick up the sounds accompanying the

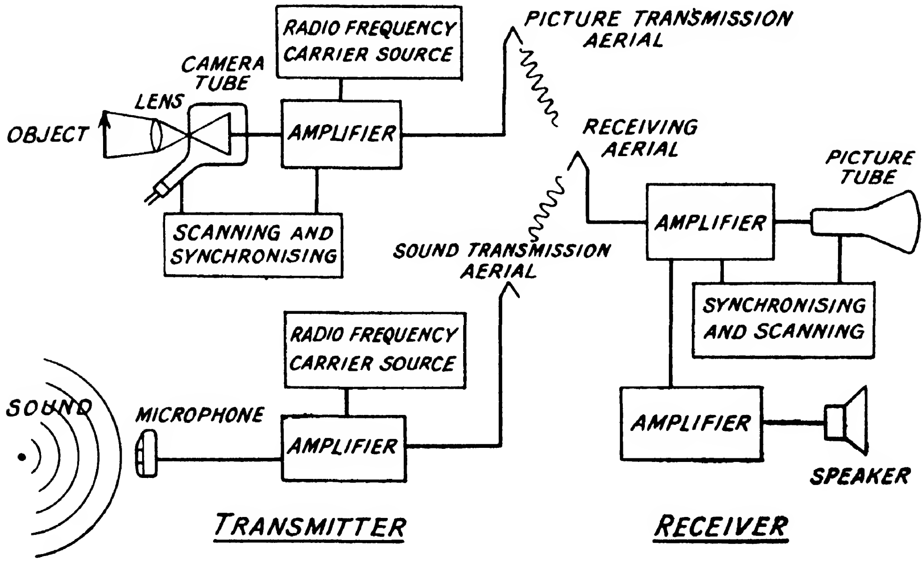


FIG. 81.—Complete television system.

movements at the transmitting end and a loudspeaker for reproducing these sounds at the receiving end.

The layout of a complete television broadcasting system is shown in Fig. 81. The radio transmission of both picture and sound requires a carrier source in accordance with normal broadcasting practice and these two channels use independent aerials. The picture or *video* transmission uses high frequencies, of 100 million cycles per second or more, with the object of avoiding interference with other broadcasting and from static sources.

The receiving amplifier system separates the audio and video signals by the superheterodyne conversion process, familiar to those acquainted with radio technique, and amplifies these signals independently for use in the picture tube and loudspeaker.

At the present time (October, 1947) developments are contemplated in several directions. The foremost is an increase in the size and brightness of the received picture and the possibility of projecting it on to a viewing screen in the manner of a cinematograph. There are also proposals for colour-television and stereoscopic projection to give the impression of depth to the picture. Although television practice has progressed enormously during recent years there remains much to be done before these anticipated improvements are realised.

FOR FURTHER READING

Only the relatively recent books on television deal with non-mechanical scanning systems such as electron cameras and cathode ray picture tubes. The following give up-to-date treatments :—

1. L. DE FOREST. *Television Now and Onwards.* (Hutchinson), 1946. A very easily readable account for readers having a minimum of technical knowledge. Explains circuits and apparatus used in modern systems and deals with every aspect of technique and future possibilities.
2. D. G. FINK. *Principles of Television Engineering.* (New York, McGraw-Hill), 1940. A more advanced book designed to give radio engineers a familiarity with television engineering. The treatment is comprehensive, and numerous references are given.
3. V. K. ZWORYKIN and G. A. MORTON. *Television.* (New York, Wiley), 1940.

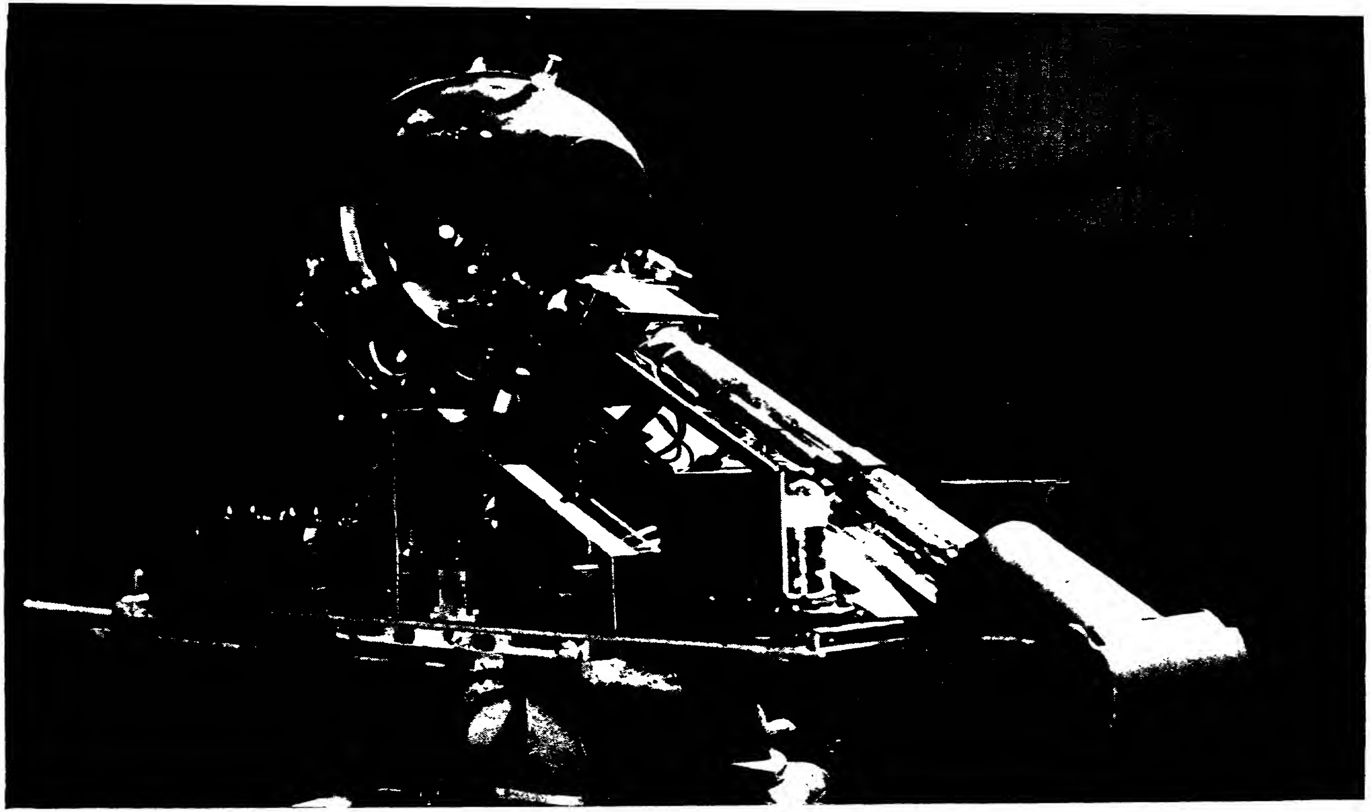


PLATE 4

View of the interior of an Emitron television camera showing the photo-sensitive plate inside the bulb and the pre-amplifier for the picture signal.

(*B.B.C. photo*)

facing p. 160

CHAPTER XII

THE SOUND FILM

Attempts by experimenters to produce moving pictures accompanied by appropriate sound date back to at least 1888, when Edison worked on a " talkie " system comprising a picture cylinder working in conjunction with a wax cylinder of his famous phonograph. According to a contemporary account there was " one cylinder of fairly good noises and one full of frightful pictures," so it appears that his efforts were not very satisfying, however ingenious they may have been.

The distinction of making the first celluloid motion picture film appears to belong to Friese-Green and dates from 1896. It was Ruhmer who succeeded about this time in actually making a sound record on a strip of film, so that between them these two experimenters made vital contributions to the art as it is known today.

The work of Eugene Lauste at the beginning of the present century was remarkable for its insight into the problems of the sound film. In 1906 he produced a talking film machine based to some extent on the work of Ruhmer and Friese-Green and using a selenium cell as the light-sensitive element, but without the means of amplification represented by the thermionic valve there could be little progress with such an arrangement. It was not until 1907 that the triode valve was invented, and several years more elapsed before it became generally used, so that during this time the progress of sound film experiments was held back by the absence of an essential element.

The moving picture films, which science was as yet unable to endow with a voice, achieved in spite of this shortcoming a tremendous popularity and caused, as is well known, the appearance of a new industry giving employment to many thousands in a wide variety of professions. The arrival of the sound film was destined to give an immense stimulus to this already great industry.

Prior to 1914, C. M. Hepworth produced the Vivaphone system using gramophone records synchronised with the film. Although the system was very ingenious it was not very successful because of the difficulty of maintaining synchronism and the lack of means for amplifying the sound. The system was sold to America where it appeared under the name Vitaphone and was used to produce the first films good enough for showing to a public audience. These films used a 16-inch record with each reel of film, and it could not be said that the arrangement was very convenient. However, a start had been made, and it was evident that the " talkies " had come to stay. Their popularity gave the necessary impetus to the improvements which were obviously required, and improved systems soon began to make their appearance.

Sound Track Methods

It soon occurred to the film technicians that the logical arrangement was to have the sound-producing track on the film

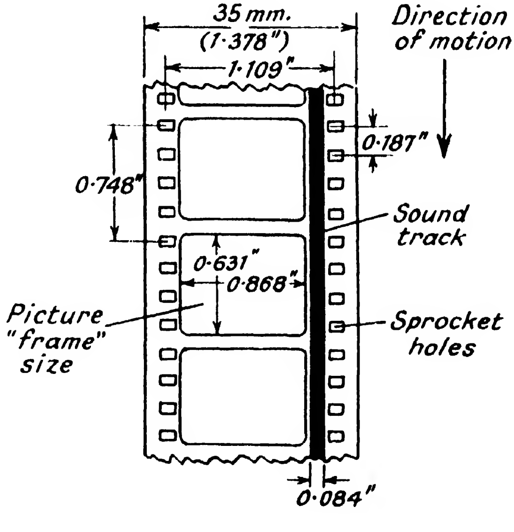


FIG. 82.—Arrangement of 35 mm. sound film.

itself, so as to eliminate any possibility of the sound getting out of step with the picture, as so often happened with records. It is now universal practice to have the sound track on the film, located on one side of the pictures as shown in Fig. 82, which also gives the principal dimensions of the standard celluloid film used in cinemas. Photography of the sequence of pictures is at the rate of 24 separate pictures or *frames* per second, corresponding to about a thousand feet of film every ten minutes.

When the film is made, the sound which accompanies the action is picked up by microphones and the resulting electrical signal, after amplification operates a special recording instrument which modulates a light beam focused on the sound track of the film in such a way that the light reaching the film undergoes changes precisely in accordance with the variations of sound entering the microphone. After being developed and fixed, the sound track contains a permanent record of the variations. By virtually reversing the recording process it is possible to reproduce the original sounds from the film track.

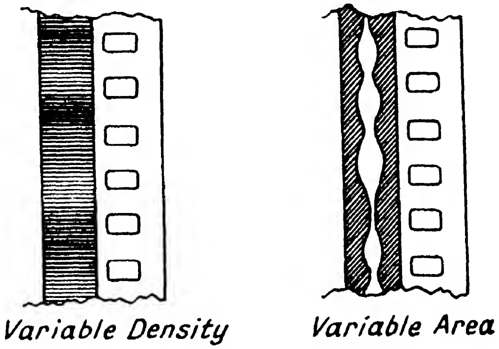


FIG 83.—Alternative forms of sound track.

There are two different methods of making the sound track, as shown in Fig. 83. In the *variable-density* method the intensity of the light is modulated in accordance with the sound, whereas in the *variable-area* method it is the area of the recording light beam which varies. It is claimed for the latter method that printing and developing are easier than with the variable-density method owing to the absence of fine graduations of light and shade.

Sound Recording

For various technical reasons the negatives of the pictures and the sound track are usually taken separately and afterwards combined to make the complete film. Among other things, this allows of the sound track being advanced in relation to the pictures so as to allow for the fact that the projector lens and sound unit are in different positions on the projector, and thus to preserve synchronism between sound and action.

The arrangement of a typical recording system is shown in Fig. 84. The microphone and picture camera are on the stage, taking in sounds and scenes respectively, the camera being driven at a constant and closely regulated speed by a special electric motor operating from the master generator of a synchronous driving system. Sound currents from the microphone are amplified so as to yield sufficient power to opeate the light-modulating instrument in the recorder. The exact arrangement of the recorder depends upon the system, but the general principle is that the sound film negative is driven from one drum on to another and passes on its way the exposure point where the light beam, either brightness or area-modulated, shines upon it and records its pattern corresponding to the sound vibrations. The driving motor of the recorder operates from the same master generator as the camera motor, so that perfect synchronism is preserved between the picture and sound recordings. We do not need to consider here the highly specialised technique of developing, fixing, combining and printing the two records in making the complete film.

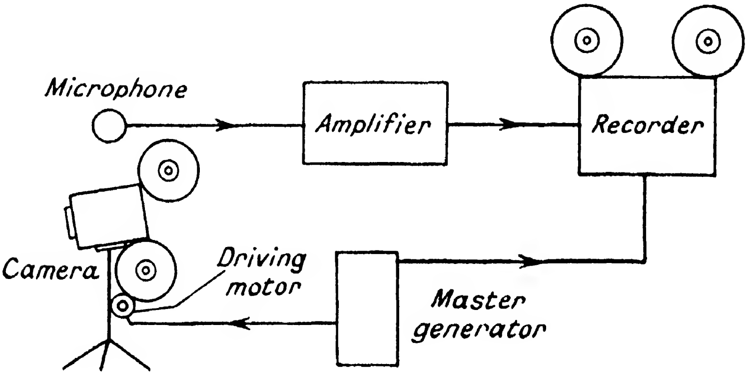


Fig. 84.—Elements of sound recording system.

The amplifiers used for sound recording are generally of the resistance-capacitance coupled type with a push-pull output working into a transformer which can be matched to the load circuit (the light-modulating instrument) so as to give maximum efficiency.

Sound Film Projection

The modern sound film projector is an elaborate mechanism embodying, in addition to the basic parts, a large number of refinements for improving the quality of the reproduction. It consists essentially of an ordinary moving film projector with the addition of a *soundhead* through which the film passes in the process of having its sound track reproduced in terms of electrical oscillations for producing sound in the loudspeakers in the auditorium. The essential parts of the soundhead are a lamp, generally called the *exciter lamp*, an optical system for focusing the lamp beam on to the sound track of the film, and a photoelectric cell for receiving the modulated light passed through the sound track. As the current obtained from the photoelectric cell is far too small for direct use an amplifier is used to operate the loudspeakers in accordance with these minute current changes.

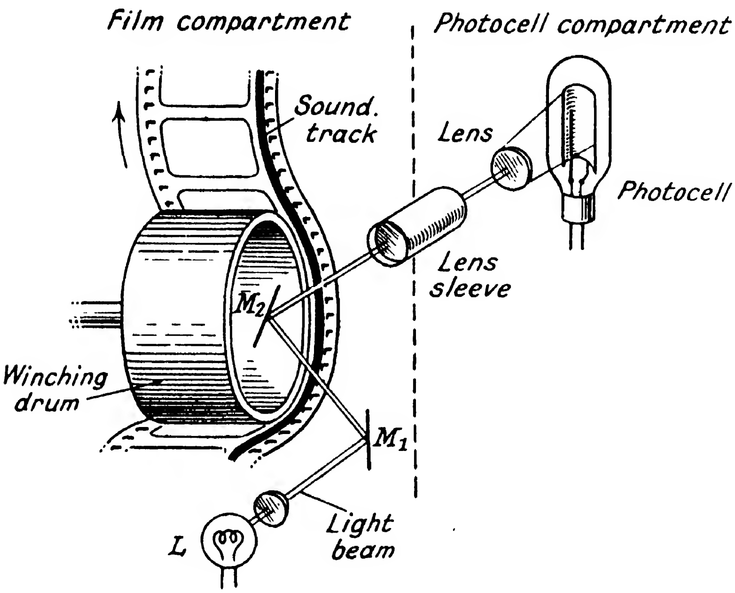


FIG. 85.—Simplified arrangement of projector soundhead.

The general principle of the soundhead is illustrated by Fig. 85, where L is the exciter lamp and M_1, M_2 a mirror or prism system by which the light beam is directed through the sound track on the edge of the film overhanging the winding drum. After passing through a lens sleeve in the wall between the film compartment and the photocell compartment the light beam is spread out by a further lens so as to cover the cathode of the photocell. The arrangement of the winding drum is such that the film itself drives the drum, which maintains a perfectly uniform speed by damping out any effects tending to change the rate of motion.

The complete projector and loudspeaker installation is shown in simplified form in Fig. 86. The driving motor on the projector winds the film from the upper to the lower drum or magazine so that the film passes downwards through the picture projection mechanism and then through the soundhead. The photocell in the soundhead usually operates at about 90 volts D.C., and is of the emission type (see page 59), giving the very rapid response to light changes which is essential in sound reproduction. The photocell output is taken to a resistance-capacitance coupled amplifier with a push-pull output stage matched to the loudspeaker load. It will be understood that the loudspeakers are usually concealed about the auditorium in such a way as to give the best possible acoustic effects.

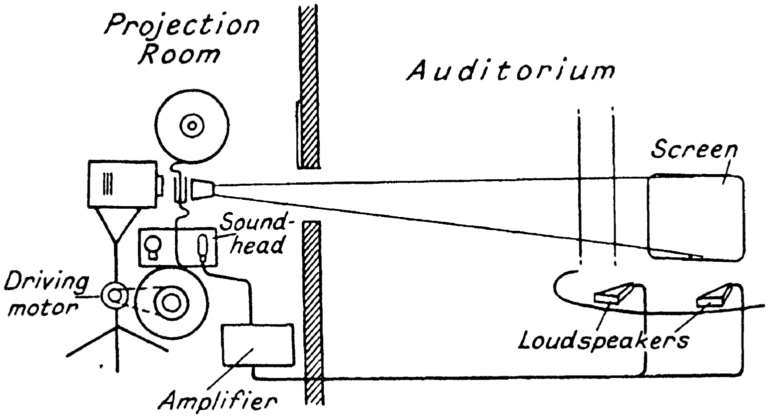


FIG. 86.—General arrangement of projector and loudspeakers in cinema..

The projection room in a modern cinema always has two projectors so that continuity can be maintained in the presentation of the films. Somewhat elaborate arrangements of change-over switches are included in the installation, so that if necessary the projector in operation can be connected to another amplifier. Provision is also made for volume control of the loudspeakers and for the equalisation of volume when a changeover is made from one set to another. In most equipments all the components are easily accessible, so that in the event of breakdown of any part the finding of the trouble is facilitated. In view of the complex nature of the apparatus, the number of breakdowns experienced in practice is very small, and complete loss of sound is now a very rare occurrence. The remarkable efficiency and reliability of the modern sound film installation give clear testimony to the effectiveness of applied electronics in this particular sphere.

FOR FURTHER READING

Sound film technique is touched upon in some of the books on photoelectricity (see p. 79). For specific works, reference may be made to :—

1. F. W. Campbell and others. *Sound-Film Projection.* (Newnes), 1945. Describes principal types of sound film projection equipment, with particular reference to installation and maintenance.
2. *Motion Picture Sound Engineering.* (Research Council of Acad. of Motion Picture Arts and Sciences, Hollywood, Cal.), 1938.
3. *The Technique of Motion Picture Production.* (Society of Motion Picture Engineers, Hollywood, Cal.), 1945. A collection of papers presented at the 51st Convention of the Society.

CHAPTER XIII

ELECTRON OPTICS

There is a striking similarity between the changes that can be made in the directions of a ray of light by means of lenses and the changes that can be made in the directions of a beam of electrons by means of electrostatic and magnetic fields. The similarity is in fact so close and so important that a special branch of electronics known as *electron optics* has been brought into being. The subject has been developed mainly through studies connected with the cathode ray tube, particularly in its application to television, and in the design of electron cameras for this purpose.

Basic Principles

Considering first the action of electrostatic fields on moving electrons, we know that an electron will always be attracted by the positive electrode or anode of the system. The path it takes towards this electrode will depend upon its direction and velocity and upon the strength and distribution or shape of the field. As it is not possible to visualise the shape of the field except in simple cases it is usual to employ the principle of *equipotential surfaces* as a method of field plotting, in much the same way that contour lines can be drawn on a map to show variations of altitude. The simplest case, represented by two plane metal surfaces connected to a source of constant voltage, is shown in Fig. 87. The field between the plates is uniform, and its intensity varies uniformly from one plate to the other. The equipotential surfaces are imaginary planes represented, when seen edgewise, by the dotted lines. Before going further it will be necessary to give more definite meaning to the terms we employ and to visualise the relationships of the different factors.

If we imagine a very small charge of electricity (an electron will do) placed in an electrostatic field, it will experience a force in a certain direction. The force is a direct measure of the *intensity* of the field and the direction is that of the field at the point where the charge is located. In general, for charges (or electrons) initially at rest, the path along which the charge is urged by the force due to the field is a line the tangent of which at every point is the direction of the field intensity at that point. This is the definition of a *line of force.* The case represented by Fig. 87 is special, in so far as the lines of force are straight lines representing the shortest distance between the plates. In most cases lines of force are curves because the field is not uniform throughout. The reason for specifying a *very small* charge in the foregoing is that otherwise it would affect and distort the original field. It should also be noted that a charge *initially at rest* is specified. If the charge (or electron) enters the field with an appreciable velocity its direction is not that of the intensity at every point (i.e., along a line of force), but is intermediate between this and the direction in which it is travelling. The direction of motion tends to become that of a line of force as the motion continues.

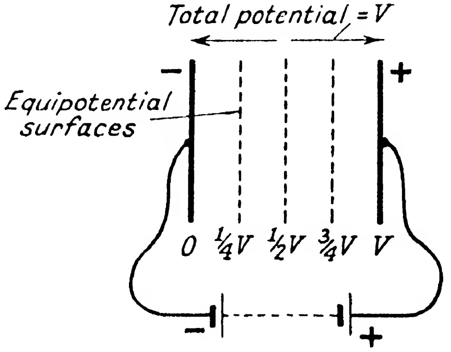


FIG. 87.—Distribution of potential between plane electrodes.

The *potential* at any point in the field is defined as the work which is done on a particle with a very small charge (such as an electron) in bringing it to that point from somewhere outside the field. This work is proportional to the intensity of the field and is positive or negative, i.e., work done or work reclaimed, according to whether the motion is opposed or assisted by the action of the field. An *equipotential surface* is defined as a surface in the electrostatic field having the same potential at every point, i.e., on which there is everywhere the same action on a given elemental charge of electricity. Finally, we have the law that lines of force and equipotential surfaces are always at right angles to each other.

Equipped with these basic laws and definitions we may now consider some of the arrangements used in electron optics for focusing electron beams by means of electrostatic fields.

Electrostatic Lenses

The system shown in Fig. 88 consists of a thermionic cathode *C*, a plane anode *A* and an intermediate electrode *D* with a central hole. The dotted lines represent the equipotential surfaces seen in section. As the anode itself is an equipotential surface, the field distribution close to it is represented by parallel flat surfaces, but as the distance from the anode increases the surfaces become distorted owing to the aperture in the electrode *D*. As the electrons emitted from *C* tend to follow lines of force and thus cross equipotential surfaces at right angles it can be seen that their path is such as to bring them to a focus at point *F* on the anode. Even those electrons which are directed outwards from the cathode as shown by the arrow are acted upon so as to bring them to the focus through a longer path.

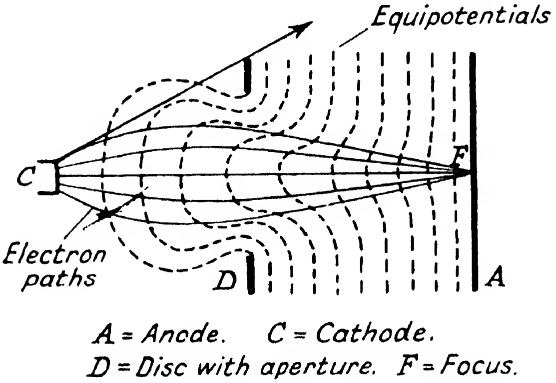


FIG. 88.—Simple form of electrostatic lens.

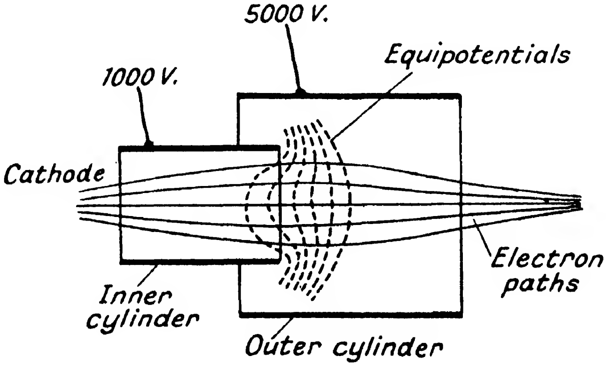


FIG. 89.—Cylindrical electrostatic lens.

Another type of electrostatic lens is shown in Fig. 89. It consists of two concentric cylinders displaced axially and having different voltages. The dotted lines show the sections of the equipotential surfaces. This form of lens, which is widely used in cathode ray tubes, acts upon a divergent beam of electrons so as to reduce it to a fine focus. There are many other electrostatic lens systems, but the two we have described are representative and illustrate the general principles.

Magnetic Lenses

Electron focusing effects may also be obtained with suitable magnetic fields, and this method is widely used in the electron microscope.

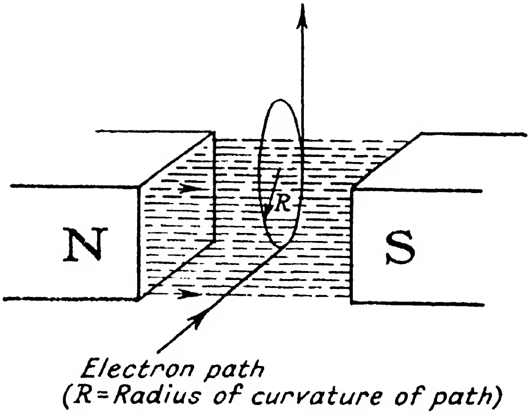


FIG. 90.—Curvature of electron path caused by magnetic field.

It is only electrons in motion which are affected by magnetic fields. When such an electron passes across the lines of force, as shown in Fig. 90, its path becomes deflected into a curve, the radius of which is proportional to the velocity and inversely proportional to the magnetic field strength. This means that the higher the velocity the more gradual is the curve and the greater the field strength the more sharp is the deviation from the original course. There is a definite relationship between the direction of the field, the direct on of travel of the electron relative to it and the resulting deflection. If the electron entered from the other side of the picture in Fig. 90 its deflection would be downward instead of upward.

When the electron is projected slantwise across the lines of force its motion is no longer in one plane, as in Fig. 90, but becomes spiral, representing the combination of the original motion with the twist imparted by the magnetic field. If the field is strong enough, and the electron velocity not too high, the electron does not escape from the field but describes a spiral path through it.

Special considerations arise when the source of electrons is located in the axis of the field so that the electrons are projected lengthwise through the field. This happens in the arrangement shown in Fig. 91, which is in fact the basic form of electromagnetic lens. The cathode *C* is the electron source, giving a divergent beam of electrons, and the cylindrical focusing coil *F* surrounds the beam as shown. A steady current passing through the coil winding produces a field which is substantially uniform inside the coil and directed along the axis. Any electrons which happen to leave the cathode in a perfectly axial direction, so as to pass through the centre of the coil, will experience no force from the field because in this direction they do not pass across lines of force. The majority of electrons, however, will be directed outwards, away from the axis, and they will follow spiral paths as already explained.

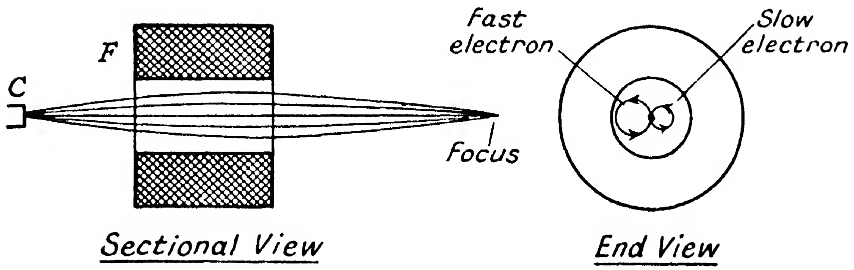


FIG. 91.—Focusing action of magnetic field (electromagnetic lens).

At this stage we may note a very important fact in connection with these spiral paths. Since, as we have already seen, the radius of the circle through which the electron is turned is proportional to the velocity of the electron, the faster electrons have proportionately farther to travel and the time for making a complete turn is *the same for all electrons, regardless of their velocity*. This means that all electrons directed outwards at the cathode undergo a rotation which brings them in a given time back to the axis. As this time is the same for all the electrons they arrive together at a definite point along the axis, after which they again diverge. This point is the *focus* of the electromagnetic lens, and a screen placed there will show an image of the cathode. This arrangement has been used for studying the emission of cathode surfaces and has yielded much information on the factors affecting thermionic emission.

The Electron Microscope

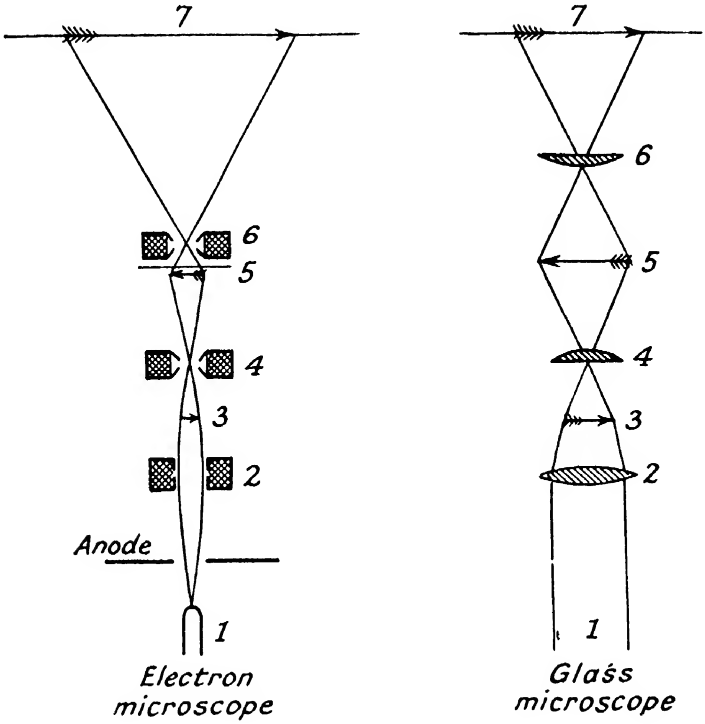
From our brief consideration of electron optics it is apparent that with suitable electric and magnetic fields it is possible to deflect and focus electron beams in the same way that light beams may be deflected and focused with glass lenses. These facts in themselves suggest the use of electron optical systems for doing what is normally done by a microscope, but as yet we have put forward no good reason for replacing the relatively simple conventional microscope with a highly complicated arrangement of electron lenses using a thermionic cathode, operating in a vacuum, and requiring an external source of power for its operation. The justification for such measures arises out of a fundamental limitation of the glass microscope—the limit set to the useful magnification of a lens by its *resolving power*, i.e., its ability to distinguish as separate two points very close together on the object being viewed. The resolving power thus determines the amount of detail that can be seen. In a glass microscope the ultimate limit of resolving power is set by the wavelength of the light used, and with ordinary light it is impossible to resolve two points which are less than about one fifty-thousandth of an inch apart. After using every artifice and refinement of design, the maximum magnification of the ordinary glass microscope is limited to the order of 2,000. While this figure allows a very useful range of microscopic investigation it is still a long way short of the value necessary for reaching into the molecular domain represented especially by the numerous bacteria and viruses, about which the ordinary microscope can discern little or nothing of value.

According to the wave theory of the electron, which is used for explaining results such as the diffraction patterns obtained when electrons pass through a very thin metal foil (and which cannot be explained if we assume that the electron is merely a lump of something, a corpuscle) an electron in motion has a train of waves associated with it. The wavelength is inversely proportional to the velocity of the electron and can, therefore, be made small by using high accelerating voltages on the anode. This is the basis of the high resolving power of the electron microscope : by using sufficiently high voltages the wavelength can be reduced to a fraction of that corresponding to the shortest visible light waves. To use a rather crude analogy, the electron wave train is able to wriggle through between points which would be too close together to admit the passage of light rays. The overall effect is an increase of resolving power allowing useful magnifications up to ten times that obtainable with the best glass microscopes, or a figure in the region of 20,000. Exact comparisons are impossible because of the special features that can be incorporated in either instrument.

Electron microscopes are made with either electrostatic or electromagnetic lenses. The instruments are quite massive, standing perhaps taller than a man, and require considerable auxiliary apparatus, including means for maintaining a high vacuum and closely regulated supplies of electrical power. Perhaps the chief disadvantage of the instrument arises from the fact that the electron beam must penetrate the object in order to form the picture, and the object must, therefore, be very thin. The preparation of specimens for examination by an electron microscope is in fact a highly specialised procedure, as the thin specimen has to be supported on a transparent film and then positioned accurately in the instrument through a dismountable vacuum-tight joint.

The general principle of a magnetic electron microscope is shown in Fig. 92 in comparison with a glass microscope, the relative parts being numbered. It will be noticed that in place of a source of light the electronic instrument uses a source of electrons consisting of a thermionic cathode and that the refractive medium is a magnetic (or electrostatic) field instead of a glass lens. In the glass instrument the image is viewed either directly or on a glass screen, whereas in the electronic instrument it is received on a fluorescent screen or a photographic plate. The lens magnets are flat cylindrical coils enclosed in an iron shell with very short air gaps so as to give a powerful and concentrated field. High voltages up to the order of 50,000 volts are used on the anode so as to improve penetration of the electron beam through the object.

The electron microscope has already been used for important investigations in many branches of science and it seems likely that the most outstanding results may be obtained with the instrument in due course. The examination of dusts and smoke particles has revealed hitherto unknown facts about small-scale structures of this kind. Some smoke particles are seen under the electron microscope to have complex forms, sometimes with formations of needle-like projections. In the domain of bacteria the results so far obtained are most striking. Photographs have been taken showing not only individual bacteria, but also



1 Source of electrons
2 Condenser magnet
3 Object
4 Objective magnet
5 Image of object
6 Projector magnet
7 Final image

1 Source of light
2 Condenser lens
3 Object
4 Objective lens
5 Image of object
6 Projector lens
7 Final image

FIG. 92.—Comparison between electron and glass microscopes.

bacteriophages (the anti-bacterial organisms which give resistance to infection) in the actual process of attacking and destroying the bacteria. Viruses about which medical science formerly knew relatively little are well within the range of the electron microscope; even the excessively small but highly destructive influenza virus can now be studied by this means. It is unnecessary to emphasise the humanitarian importance of this work, the furtherance of which is worthy of every care and effort so as to bring to medical science perhaps the greatest help it has yet received in the battle against disease.

To glimpse the future, it may well prove possible eventually to bring individual molecules of matter within the range of the electron microscope and thus to see and study the very bricks of which the universe is built. These speculations, which are well within the bounds of possibility, show very clearly the kind of developments which may reasonably be expected to result from future advances in the technique of electron optics.

FOR FURTHER READING

Perhaps the best introductory books are :—

1. E. F. Burton and W. H. Kohl. *The Electron Microscope.* (New York, Reinhold), 1942. This is a very readable and accurate non-mathematical treatment dealing fully in elementary style with basic principles and including several reproductions from photographs taken with the electron microscope. Gives a bibliography.
2. O. Klemperer. *Electron Optics.* (Cambridge), 1939. A theoretical treatment.
3. L. M. Myers. *Electron Optics.* (Chapman and Hall), 1939. A very detailed and comprehensive work, with copious references.
4. D. Gabor. *The Electron Microscope.* (" Electronic Engineering," Monograph [2nd Edition.]), 1948. An introductory treatment, with advanced chapters on theoretical principles.

CHAPTER XIV

THE CYCLOTRON AND BETATRON

Among the numerous special devices which have been made possible by the use of oscillators, particular mention may be made of the cyclotron. This is an instrument of research with which highly important discoveries have been made in the realm of atomic physics.

It was found by Rutherford in 1919 that high velocity protons were emitted from nitrogen subjected to the alpha-rays from radium. From what has already been said about atoms and their constituents on page 33 we know this must mean that the nitrogen atoms were undergoing the drastic change of having the constituent parts of their nuclei knocked out, thus transforming them into atoms of another substance. In technical language, Rutherford had achieved the artificial nuclear disintegration of a stable element. Since his time the process has been applied, mainly by means of the cyclotron, to nearly all the known elements, changing them into other elements, some stable and some unstable, i.e., radioactive.

The Acceleration of Ions

The extension of this work led eventually to the idea of producing the high speed ions by artificial means, as by accelerating them in suitable electric fields. The original application of this idea was in the *linear resonance* method of acceleration. The ions to be used for bombardment purposes were accelerated through a succession of separate tubular electrodes fed in alternate pairs from a high-voltage high-frequency oscillator in such a way that the ions are acted upon by the same voltage during their passage between the successive tubes, which have to be progressively longer because of the increasing speed of the ions and the fixed frequency of the oscillator. The obvious advantage of the arrangement is that the ion reaches a speed corresponding to the applied voltage multiplied by the number of gaps (regions of acceleration) and by using the same voltage over and over again in this way the technical difficulties associated with very high voltage oscillators are avoided.

In the search for more and more effective means of disrupting atomic nuclei it was found that protons and *deuterons* gave good results. The deuteron in the nucleus of the so-called heavy hydrogen atom which is twice the weight of the ordinary hydrogen atom although, like ordinary hydrogen, it only has one satellite electron. The heavy hydrogen atom thus has a nucleus with twice the weight but the same single positive charge as the ordinary hydrogen nucleus, i.e., the charge of one proton. This double-weight nucleus, the deuteron, made a valuable addition to the armoury of the modern alchemists.

Owing to limitations of the linear resonance method which became apparent as the work extended, and which included the unwieldy length of the tubular accelerator required for high velocities, fresh thought was given to the problem. It occurred to E. O. Lawrence, an American physicist closely engaged in this work, that the same effect could be obtained by using one pair of accelerating electrodes supplied by an oscillator in such a way as to drive the particles to and fro while imparting to them an ever increasing velocity. This is the basic principle of the cyclotron which has become famous in recent times as one of the most important tools by which man may, for good or ill, or both, change the fundamental arrangement of nature.

The Cyclotron Principle

The arrangement of the cyclotron is shown in Fig. 93. There are two hollow metal chambers, generally known as " dees " because of their shape, separated so as to give a short gap in which a powerful electric field is produced by the voltage derived from the oscillator. The space well inside the dees is virtually devoid of any electric field but the entire assembly is permeated by a powerful magnetic field acting as shown and provided by a large electromagnet system (not shown). The whole arrangement is in a vacuum chamber which is highly exhausted when in operation.

The ions, protons, deuterons or alpha-particles to be used in the bombardment are produced at the centre of the space between the dees by means of the filament, which is operated from an external supply. Any such positively charged particle released in the gap will immediately be attracted into the opening of the dee which happens to be negative at that instant. Once inside the cavity it will cease being acted upon by the electric field (the cavities being field-free) but the action of the perpendicular magnetic field will be to curve its path into a circle. The conditions are similar to those we considered on page 171 in connection with an electron passing through a magnetic field and, as in that case, the time of traversing the path is independent of the velocity. Owing to the circular path the particle emerges into the gap again after the lapse of its travelling time, and if during this time the dees have changed polarity under the oscillating voltage the particle will be attracted into the opposite dee, receiving at the same time an increment to its velocity corresponding to the applied voltage. It thus traverses the inside of the dee with a velocity corresponding to two applications of voltage and a correspondingly larger radius of travel. At the end of the fixed travelling time, however, it finds itself once again in the gap and owing to the change of polarity of the

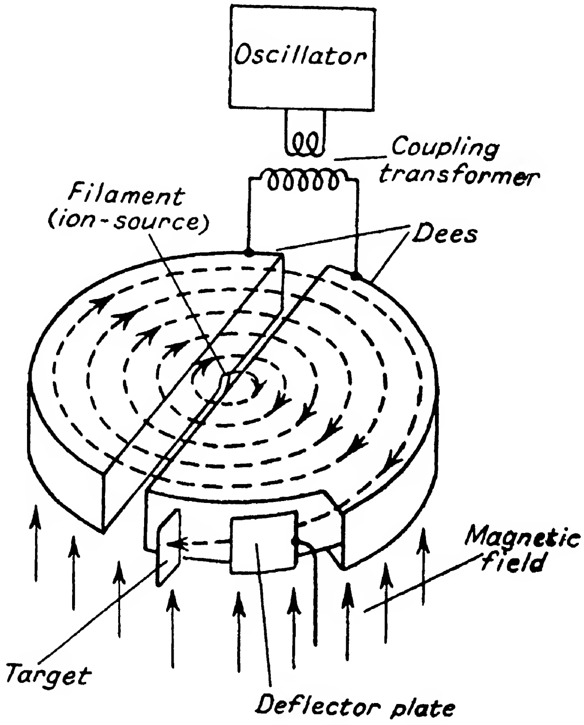


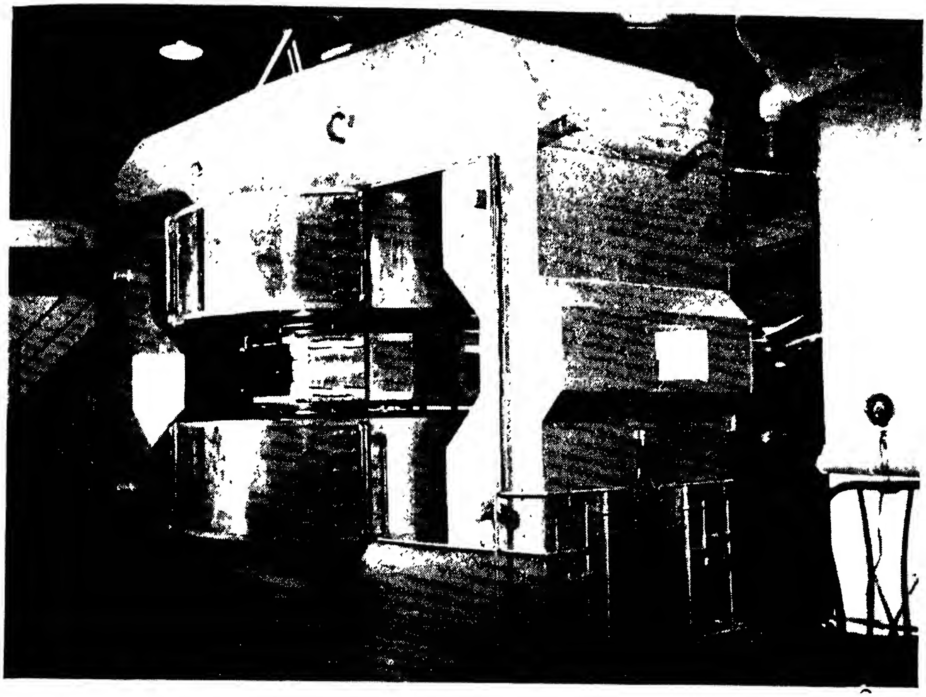
FIG. 93.—Arrangement of cyclotron. (Enclosing case and electromagnet not shown.)

dees during this time is urged in the same direction with a further increase of speed which causes a further increase in the radius of its path. This process goes on, with the beam of particles taking an ever widening path, until it shoots through an aperture on the outside edge of one of the dees. Thereafter it is guided by a highly charged deflector plate so as to impinge in the desired position on the target where the material to be bombarded is placed. The products of such bombardment, especially the radioactive materials produced in this way, are of the greatest interest and value in biology and medicine as they allow the extent and rate of assimilation of various substances in the body tissues to be observed. These substances, when subjected to bombardment in the cyclotron become radioactive in varying degrees. Ordinary salt (sodium chloride) for example, is converted into radiosodium and gives off the same rays as radium although the effect disappears entirely after about 200 hours. Such a substance is readily and harmlessly absorbed by the body but during its radioactive life may be located by various means. It is also possible to produce in the cyclotron a radioactive sodium which is even more effective than radium. These and numerous other developments give great hopes for the use of the cyclotron as an aid to medical research in addition to its applications in physics and chemistry.

The cyclotrons in use in various laboratories throughout the world follow the same general lines of construction as the original one we have described, the magnetic system of which alone weighed 75 tons. The construction is complicated by the necessity of providing entry for the filament and oscillator connections, for the insertion of probes for measuring the high frequency energy, for observation windows, for vacuum pumping and cooling system connections and the provision of a vacuum-tight trap for the target so that this can be changed without losing vacuum. The largest cyclotron yet constructed has dees five feet in diameter and the oscillator impresses on them a voltage of the order of 250,000 volts. Experience with this cyclotron shows that there should be no insuperable difficulties in building and using even bigger and higher voltage types.

The Betatron

A somewhat similar arrangement, used for the acceleration of electrons instead of ions, has been developed recently, and is



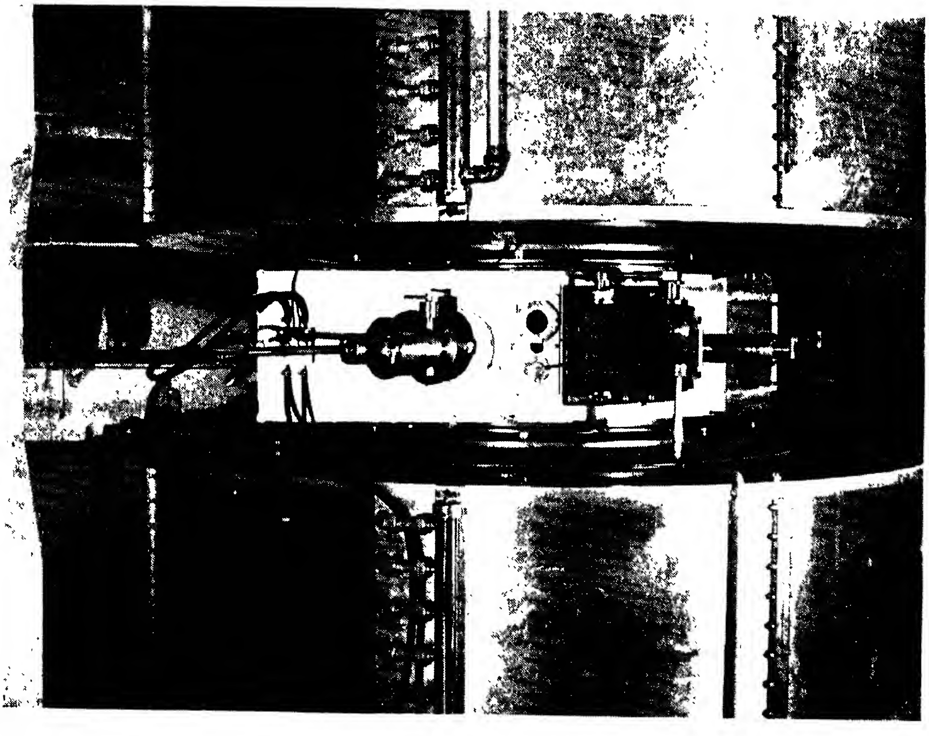


PLATE 5

The Carnegie Institute Cyclotron. General view of cyclotron showing the cast steel pole structure, field coils, and magnetic gap in which the particles are accelerated.
Below : Close-up of the target electrode and accelerating chamber.

known as a betatron, rheotron or electron induction accelerator. A massive electromagnet surrounds a chamber comprising an evacuated glass toroid (hollow ring) inside which the electrons rotate in circular orbits. The magnetic field is alternating, so that during each cycle it grows from zero to a maximum, reduces to zero again and repeats this process in the opposite direction.

Electrons are injected into the glass toroid at a tangent from outside by means of an electron " gun " comprising a thermionic filament and high-voltage accelerating anodes which impart to the electrons a velocity corresponding to several thousand volts. Inside the toroid the electrons are acted upon by the field, which confines them to a circular orbit and greatly augments their velocity. By means of external circuits the injection of electrons is timed to correspond to an early part of each cycle of the alternating magnetic field, which in rising to its peak value imparts an enormous acceleration to the electrons. At the peak of each wave, corresponding to maximum velocity of the electrons, the field is switched off, and by means of an auxiliary field the electron orbit is suddenly expanded so that the electrons strike a tungsten target electrode located in the toroid, liberating very intense radiation having the nature of X-rays.

One of the largest betatrons yet made has an electromagnet which alone weighs 130 tons, and the apparatus is capable of imparting to electrons a speed corresponding to 100 million volts. The toroid is 74 inches diameter and has a radial width of 8½ inches. The initial velocity of the electrons injected into the toroid corresponds to between 50 and 70 thousand volts, and inside the toroid they are accelerated at the rate of 400 volts for each circuit of the orbit. In the quarter of a cycle during which the magnetic field is applied the electrons make 250,000 circuits of the orbit, corresponding to a distance of some 800 miles and giving a final speed equivalent to acceleration through 100 million volts.

Apart from its use as a means of changing atomic structure the betatron should appear to have very promising applications in the medical sphere. It has been found that the rays not only deeply penetrate tissue, but also have their maximum effect well below the surface, so that surface burns, one of the chief drawbacks in ordinary X-ray technique, are avoided.

FOR FURTHER READING

The literature of the cyclotron and betatron is scattered in a large number of scientific and technical journals, mainly American. There is only one book dealing specifically with the subject :—

1. W. B. MANN. *The Cyclotron.* (Methuen), 1940. This is a small work, intended mainly for readers with some training in physics, dealing with the general principles and operation of cyclotrons. It refers mainly to the 37-inch and 60-inch cyclotrons built for the Radiation Laboratory of the University of California. Applications, largely medical, are also outlined, and there is a comprehensive bibliography.

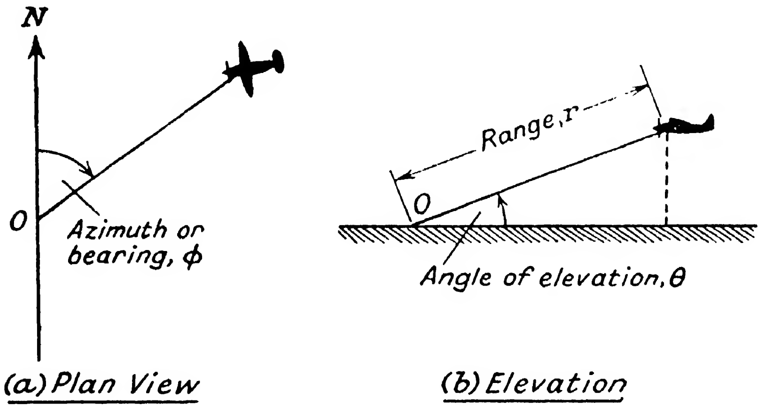
CHAPTER XV

RADIOLOCATION

At a fairly early stage in radio development it was found that wireless waves, like light, had the property of being deflected by bodies in their path. The effect was very noticeable at high frequencies and was well known to the early experimenters as a source of difficulty in the use of short-wave transmission. Under suitable conditions the waves were reflected back to the transmission point from a distant object and the existence of the object, which might otherwise be unknown, would in this way be revealed.

These effects were commented on as long ago as 1922 by Marconi in the course of a lecture to the Institute of Radio Engineers in America, and with the typical inventor's urge to find a use for any new discovery he added a suggestion which formed the basis of all subsequent work on radiolocation. Referring to his work with short radio waves he said, " In some of my tests, I have noticed the effects of reflection and deflection of these waves by metallic objects miles away. It seems to me that it should be possible to design apparatus by means of which a ship could radiate or project a divergent beam of these rays in any desired direction, which rays, if coming across a metallic object such as another steamer or ship, would be reflected back to a receiver screened from the local transmitter on the sending ship and thereby immediately reveal the presence and bearing of ships, even though these ships be unprovided with any kind of radio." It may be noticed that Marconi refers to the presence *and bearing* of the object which, as he clearly states, requires the use of a directed beam. Without this, only the *presence* of the object would be known, without any indication of its direction. Even with a knowledge of its direction or bearing an object is still not located in the strict sense ; we still want to know how far away it is. The distance of the object is technically known as the *range*, and complete radiolocation, which must include this factor, has been given the hybrid name *Radar*, derived from the words RAdio Detection And Ranging.

The essential factors involved in the radiolocation of an object in space, such as an aircraft, are shown in Fig. 94, where (*a*) and (*b*) are views in plan and elevation respectively. The angle between a straight line joining the reference point or origin *O* and the aircraft and a line pointing due North is called the azimuth or bearing, and fixes the direction in a horizontal plane. The range or distance in a straight line from the reference point *O* to the aircraft and the angle of elevation between this line and the horizontal complete the location in a vertical plane. The aircraft is completely located by the three factors : azimuth, elevation and range. In practice the azimuth and elevation are readily determined by the use of directive aerials by the range, involving the radio measurement of distance, calls for special consideration.



(a) Plan View (b) Elevation

FIG. 94.—Location of aircraft by azimuth, elevation and range.

Radio Measurement of Distance

The possibility of measuring distance by means of radio waves relies upon the fact that these waves travel through space at a constant speed, and no time is lost in the process of reflection from a distant object. It follows that if the time of the to-and-fro journey can be measured, then the total distance traversed by the waves (there and back) is equal to the time multiplied by the rate of travel of the waves. This rate is known to be 186,000 miles per second, so that the time periods to be measured in radiolocation are very short, usually being expressed in microseconds. Even in a microsecond (millionth of a second) the distance travelled is 0.186 miles or 327.36 yards, so that without great accuracy in time measurement the results would be very erratic.

As an example we may assume that a radio wave takes ten microseconds to reach an object and return again to the sending point. We infer that it takes half this time, or five microseconds, to reach the object and an equal time to get back. The distance of the object is then the distance travelled by radio waves in five microseconds, and from what has already been said the reader can satisfy himself that this distance is 1,637 yards, to the nearest yard. The simple relationship between time and distance enables us to draw a graph as in Fig. 95 showing the distance of the object for any given time interval. As it is the total interval between the sending and return of the signal which is measured in practice the distance values refer to half the corresponding time values ; the distance of the object being half the total distance travelled by the waves.

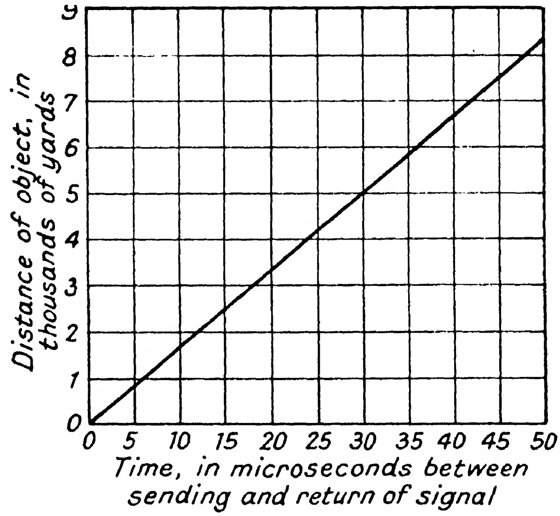


FIG. 95.—Relation between distance of object and time interval as measured in radiolocation.

So far we have taken it for granted that we can measure the time of transit of radio waves from one place to another, but as one wave in a train of waves is just like another, the question arises, how are we to identify particular points in the waves for the purpose of timing? The answer to this question is that we must *modulate* the transmission so as to produce in the waves a change which may be identified and used as a time-marking. There are two ways in which modulation can be effected; either by changing the frequency or the intensity of the transmitted waves. The frequency modulation method was actually used in this country in 1924 to determine the height from the earth's surface of the layers in the upper atmosphere which deflect radio waves and have a pronounced influence on radio transmission. The method is to use a transmitter, the frequency of which can be modulated or varied in a regular manner, and a local receiver. The broadcast waves contain high-frequency pulses which are equally spaced. If these waves encounter an object and are reflected by it the local receiver will detect *two* signals, representing the original transmitted wave and the reflected wave, the frequency modulated parts of the latter lagging in time behind those in the original wave by an amount depending on the distance of the object. The conditions are shown in Fig. 96, and the time-lag is used as a measure of the distance as already explained.

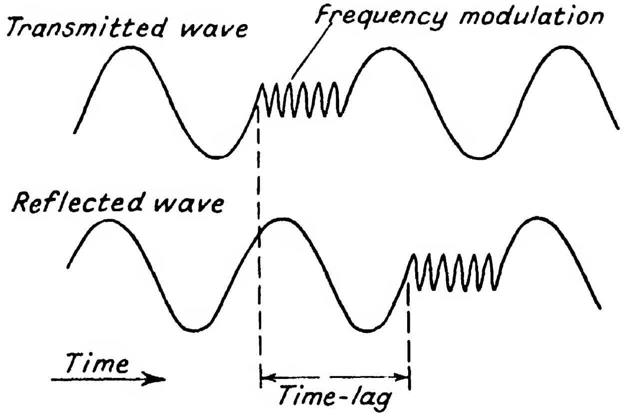


FIG. 96.—Principle of radiolocation using frequency modulation.

Special considerations arise if the object is moving relatively to the transmitter, in which case there will be a change (increase or decrease) of frequency of the reflected wave according to whether the object is moving towards or away from the transmitter. Most readers will have noticed the acoustic version of this effect, called the Doppler effect, when a locomotive sounds its whistle continuously while approaching and passing through a station. When the locomotive passes the observer there is a sudden lowering of the note, which is particularly noticeable when the locomotive is travelling fast. While the locomotive is approaching the observer the note is raised above the true note which the observer would hear if the locomotive were stationary or if he were travelling with it. A moving object reflecting radio waves behaves in a similar way to a moving source of sound. If it approaches the observer the waves reach him at a faster rate, giving an increased frequency (or raised pitch of sound) whereas if it moves away the time between the arrival of successive waves is lengthened, giving the effect of reduced frequency (or lowered pitch). With fast-moving objects the change of frequency is greater than for slow movement, so that objects moving at high speed may readily be detected and their speeds measured by the change of frequency of the reflected wave as compared with the original wave, which does not then require to be modulated. We have already seen, however, that an unmodulated wave is of no use for *distance* measurement.

To sum up, an unmodulated radio wave will not reveal the presence of a remote stationary (or very slowly moving) object, but an object travelling with sufficient speed can be detected by measuring the change of frequency of the reflected wave, which is also a measure of the speed of the object. By means of a radio wave modulated either with respect to frequency or strength it is possible to detect a remote stationary or moving object and by measuring the time-lag between the sending of a pulse (of increased frequency or increased strength) and the arrival of the corresponding part of the reflected wave the distance of the object may be determined.

Pulse or Amplitude Modulation

Owing to technical difficulties the frequency modulation method has been used only to a limited extent, and chiefly for the radiolocation of the layers in the upper atmosphere which reflect radio waves. For other forms of radiolocation the pulse modulation method is general.

The term pulse modulation as applied to radiolocation refers to the strength modulation of a wave at regular intervals, and can be regarded as the transmission of short but very intense bursts of energy spaced equally in time. The effect is achieved by special transmitter circuits, and the duration of each pulse is only a few microseconds, sometimes as short as one microsecond, owing to the necessity for keeping the transmitted pulse clear of the reflected one. The conditions in many ways resemble the behaviour of the sound waves when the hands are clapped sharply in front of a reflecting surface, and the pulse modulation method of radiolocation is in fact the radio counterpart of this method of echo production.

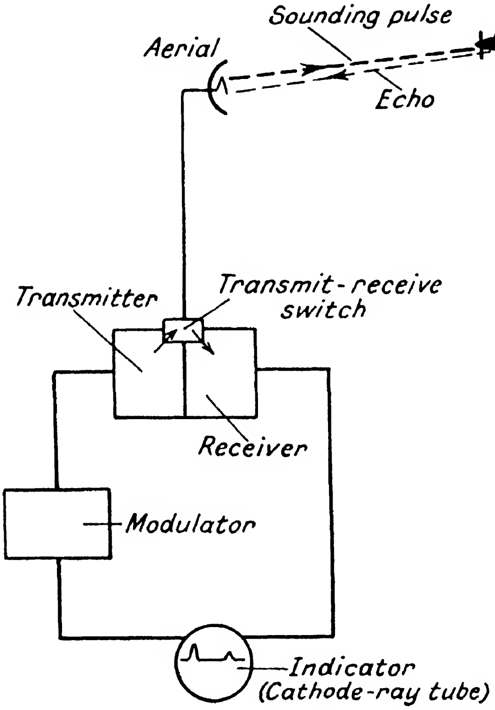


FIG. 97.—Basic elements of radiolocation system.

The development of apparatus for accurate radiolocation has occupied the period from December 1924 when the method was first used to find the height of the Heaviside layer in the course of experiments on radio propagation. A great impetus was given to the subject during the recent war in its application to the location of aircraft and ships, and the technique is now highly developed. Perhaps the greatest advance was the introduction, in 1931, of the cathode ray tube for the visual indication and time measurement of pulses and echoes. This instrument forms a vital part of all modern radiolocation systems.

The basic arrangement of a pulse modulated equipment is shown in Fig. 97. The single aerial is used for both transmitting and receiving; an automatic selector circuit effecting the necessary switching of energy from the transmitter and to the receiver. There are many forms of aerial, which vary according to the wavelength employed and the type of installation. Short wavelengths are invariably used owing to the greater efficiency and simplification of design, the frequency generally having a value between 100 and several thousand megacycles. This means that the special technique of wave guides and resonators replaces normal radio frequency practice. The transmitter valves, which have to produce the intense bursts of energy for the sounding pulses, are of the magnetron type. For location purposes the aerial must be able to move in both the horizontal and vertical planes when picking up aircraft, but for marine use a horizontal sweep alone may be sufficient. Many readers will have seen the elaborate aerial structures fitted on searchlights in the war. Owing to the evident difficulties in moving large aerials for sweep purposes, methods have been developed for exciting fixed aerials in such a way as to give a sweep or scanning effect.

Referring to Fig. 97, the modulator, which is generally a form of oscillator, sends simultaneously to the transmitter and indicator pulses of energy which are regularly timed. At the transmitter each pulse initiates the sharp burst of energy from the aerial representing the sounding pulse. The echo signal reflected from any located object is picked up by the aerial and amplified in the receiver before being passed on to the indicator. This is a cathode ray tube with a horizontal sweep operating from a linear time base in the usual way. Owing to their regular timing, the pulses from the modulator, which also represent the timing of the sounding pulses from the aerial, can be arranged to recur at the same point on the cathode ray sweep and thus form a stationary trace, generally of rectangular shape because of the sharpness of the pulses. The amplified echo pulses likewise appear as a humped trace and as they occur later in time this trace is displaced along the base line by an amount representing the distance traversed by the horizontal sweep in the time taken for a sounding pulse to reach the object and its echo to return to the aerial. That is, the horizontal spacing of the two traces is a direct measure of the *total* time interval.

A view of the indicator showing the appearance of the traces on the screen is given in Fig. 98. As the horizontal sweep speed is constant a scale can be provided, as shown, giving the time in microseconds corresponding to various positions from an arbitrary zero mark. If this mark is lined up with the pulse the total time interval can be read directly from the scale at the corresponding position on the echo trace. Owing to the direct relationship between time and distance, as already shown in Fig. 95, a second scale can be provided, from which the distance of the object can be read directly in the same way. In practice the scale may conveniently be made of transparent material so that it can be placed over the trace on the cathode ray screen. If two or more objects come within range so as to give traces on the screen, each may be located by aligning the aerial to give maximum echo signal (amplitude of trace), thus fixing the bearing or azimuth.

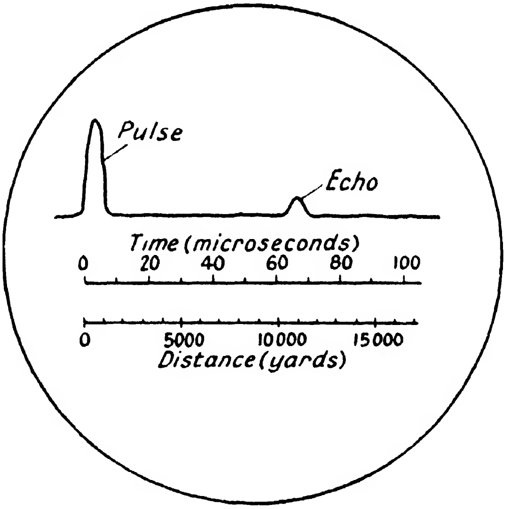


FIG. 98.—Typical radiolocation indicator screen, showing measurement scales.

It will be understood that the foregoing description applies only to essentials. There are of course numerous special features of detail in the various circuits employed, such as the ingenious device called the transmit-receive switch, also known as the

TR box or duplexing cavity, which enables the same aerial to act for both transmission and reception (see Fig. 97). The device consists essentially of two electrodes separated by a small gap and forming a resonant cavity which is tuned to the frequency employed. The transmitted pulses are applied to the cavity by means of a coupling loop and as each pulse appears the cavity is excited, producing a sufficiently high voltage across the electrodes to break down the gap between them. The resulting short circuit puts the receiver out of action and protects it from the powerful sounding pulses. At the conclusion of each pulse the gap rapidly resumes its original state so that, until the next pulse comes, the receiver is able to function normally. Owing to the relative weakness of the echo signal pulses they do not excite the cavity sufficiently to cause breakdown of the gap.

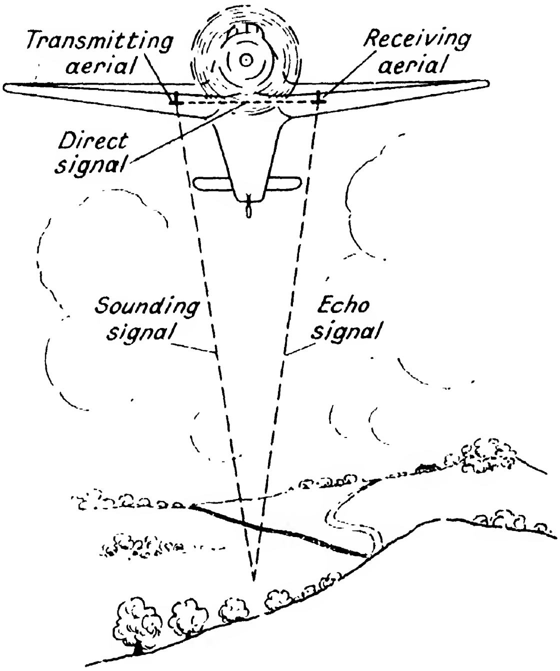


FIG. 99.—Principle of radio altimeter.

The Radio Altimeter

The radiolocation principle has also been applied to the measurement of the distance to the ground from aircraft so that under any weather conditions the pilot may have an accurate indication of his altitude. The method is shown in Fig. 99 and consists in sending out a sounding signal from an aerial under one side of the aircraft and receiving the reflected echo from the ground on a second aerial under the other side of the aircraft.

In one form of this device the frequency modulation method is employed. The echo signal reaching the receiving aerial combines with the direct signal from the transmitting aerial and a beat frequency is produced which can be indicated on a meter. This instrument is calibrated to show the altitude in feet, and the position of the needle at any moment indicates directly the height of the aircraft above the ground. The sensitivity is such that in flying over a town the meter reading fluctuates rapidly in accordance with the height of individual buildings.

The Plan Position Indicator

Another form of radiolocation, which has great possibilities in navigation at sea, makes use of a steadily rotating aerial and uses the cathode ray tube indicator in a rather unconventional manner. The aerial, or in some case only the reflector around it, is driven by a motor at constant speed so that the sounding beam sweeps round horizontally at a uniform rate, rather like the rotating beam from a lighthouse. The cathode ray indicator tube uses an electromagnet for producing the sweep corresponding to the time base and this electromagnet is rotated around the neck of the tube at the same rate of rotation as the aerial by means of a servo motor system. The sweep starts from the *centre* of the screen simultaneously with each sounding pulse and travels straight towards the edge of the screen ; each successive sweep being moved around the centre owing to the continuous rotation of the sweep magnet coil. In this way the entire surface of the screen is covered over and over again by the rapid radial sweeps. Any object which returns an echo signal before the sweep reaches the edge of the screen will be located on the screen by a bright spot, and the distance or range of the object can be determined, as already explained, by a time scale, except that the measurements are made radially outwards from the centre of the screen. If the apparatus is adjusted so that with the aerial directed straight ahead, in the case of a ship, or along the meridian in the case of a land installation, the sweep is vertically upwards, then an azimuth scale can be placed around

PLATE 6

Radar map of an area covered by a thunderstorm. The bright patches are produced by the traverse of the electron beam across a cathode ray tube screen. This type of display is based on "P.P.I." (Plan Position Indicator) developed during the war.

the screen as shown in Fig. 100 and used to read the bearing of any object directly. In the case of a ship such bearings would be relative to the ship, but absolute bearings (relative to the meridian) can readily be obtained by suitable incorporation with the ship's compass readings. As it is possible, with a fixed rate of sweep, to use concentric circles as range marks, as shown in Fig. 100, readings of range and bearing are brought conveniently together. Not only ships, but also islands and other objects are delineated on the screen, which gives in effect a plan representation of the surrounding objects. There is a lower limit to the resolution obtainable, i.e., so that separate objects are reproduced as such on the screen, but for navigation in restricted or unknown waters in the dark or under adverse weather conditions, such as thick fog, the immense advantages of the plan position indicator, or PPI system are evident. Close navigation can be effected from the readings of the equipment located below decks without lookouts, and evasive measures can be taken if necessary to prevent the possibility of a collision without actually seeing the other craft or obstruction.

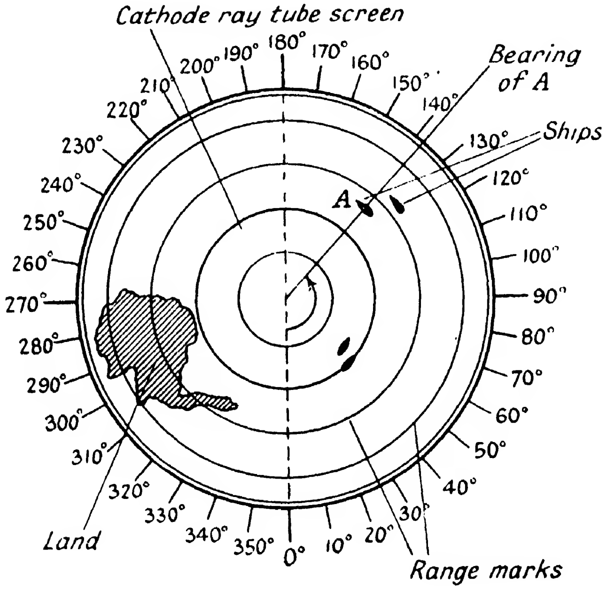


FIG. 100.—Viewing screen of plan position indicator.

FOR FURTHER READING

Numerous articles on radiolocation have been published recently in England and America, and reference may be made to the respective periodicals as follows :—

1. *Electronic Engineering.* (Morgan Bros., London)—monthly.
2. *Electronics.* (McGraw-Hill, New York)—monthly.
3. *Radio News.* (New York)—monthly.

Many recent issues of these journals contain articles, mostly of a general descriptive nature, on the various types of radiolocation equipment and installations.

The following publication makes a good general introduction :—

4. *Radar. A Report on Science at War.* (H.M. Stationery Office), 1945. Pamphlet reprint of a report published originally in U.S.A. reviewing war-time developments of radiolocation. No diagrams or technical details.
5. R. W. Hallows. *Radar : Radiolocation Simply Explained.* (Chapman and Hall), 1946. A very readable and complete elementary account of radar, mainly as applied to military purposes.

INDEX

Index