PRINCIPLES AND PRACTICE OF RADAR

BY

H. E. PENROSE

late Admiralty Signal Establishment

Member of the South African Institute of Electrical Engineers

WITH 515 *ILLUSTRATIONS*

LONDON

GEORGE NEWNES LIMITED

TOWER HOUSE, SOUTHAMPTON STREET

STRAND, W.C.2

FIRST PUBLISHED 1949

PRINTED IN GREAT BRITAIN BY THE WHITEFRIARS PRESS LTD.
LONDON AND TONBRIDGE

PREFACE

THE application of radar as an aid to navigation, and to the safety of all voyagers, whether seaborne or airborne, must necessarily interest all those responsible for the management of marine and air services.

The application of radar to commercial requirements presents no great difficulty as regards the rapid and efficient training of skilled operators. The provision of adequately trained maintenance staff is a much more serious consideration.

Complex machinery, electrical or mechanical, necessitates a high standard of technical knowledge for its efficient maintenance, not because of any abnormal unreliability factor, but because a sound technical knowledge is essential to efficient diagnosis when failures occur.

Thus, the need arises for assembly and co-ordination of the principles upon which radar systems have been developed, so that they may be presented to the reader in a progressive sequence, leading to a rapid understanding of the subject.

In writing this book, therefore, the object has been to extract from a number of available sources the principles upon which radar may be said to depend, and to assemble and present them in three stages. These three stages, as seen by the author, are :

(1) Principles directly connected with established radio practice.

(2) Principles connected with the modification to, or extension of, established radio practice.

(3) Principles associated with entirely new technique.

The starting point is fixed by this classification. Clearly, as a minimum, it must be at or about the standard of technical knowledge required by the Postmaster-General for a first-class certificate of proficiency in wireless telegraphy.

The information contained in this volume is the outcome entirely of the work of that large body of scientists, British and American, who laboured to perfect the radar systems which played such a conspicuous part in the late war.

H. E. P.

CONTENTS

INTRODUCTION

RADAR—meaning ranging and detection by radio waves—is not just an invention by a single person. Its development is due to the work of many thousands of people, directed in this country and the United States by eminent scientists, who have drawn upon established telecommunication practice, modified it, and added much new technique in course of development.

It has been known for many years that energy in the form of æther waves at radio frequency is reflected by objects in the wave path, but it required the stimulus of dire necessity to bring about the concentration of thought and effort which has resulted in the evolution of modern radar systems from knowledge of this phenomenon.

Radar Applications

The stimulus which brought about the evolution of radar technique has also influenced the trend of development, so that there are two clearly defined applications :—

(*a*) The detection of moving objects at extreme range with measurement of range accurate to within 2 or 3 per cent. and bearing approximate.

(*b*) Precise measurement of range, bearing and, in the case of aircraft targets, elevation.

The requirements for application (*a*) are unfortunately in conflict with the requirements for application (*b*). Thus radar systems for application (*a*) are known as "warning" sets, to distinguish them from range, bearing and elevation (accurate) measuring systems or precision systems.

This means that a radar set suitable for detection of an object (in radar terminology a *target*) at, say, 200 miles, with an accuracy of range measurement of, perhaps, plus or minus 2 to 3 per cent. of the total range, is quite unsuitable for measuring the range of a target at, say, 10,000 yards with any reliable degree of accuracy—say plus or minus 50 yards. Also, the ***minimum*** distance at which such a set would be capable of detecting a target at all might be in the region of 1,000 yards.

The Display

The data obtained by radar technique is presented on the screen of a cathode ray tube. This form of presentation is continuous and follows the movements of all targets within range, under observation, from instant to instant. The presentation is known as a " display," which may be arranged in a variety of forms to suit individual applications.

By continuous observation of the display and comparison of the data presented from moment to moment, it is possible to plot the course and speed of an unseen target far beyond the range of normal vision. It is also possible, in some cases, to deduce certain characteristics of the target itself.

Radar coverage is affected by meteorological conditions, in many respects similar to those which influence W/T and R/T communication, so that under abnormal weather conditions ranges for operation may exceed or fall below the normal. Since, however, meteorological conditions affecting radio communication generally have been fully dealt with in standard works on radio communication, the reader is referred to suitable books for information on this subject.

Radar, as an aid to navigation, is an obvious application. The data which a suitable radar equipment can supply should greatly minimise the danger of collision and wreck. It can also facilitate the recognition of landmarks and the navigation of estuaries and channels under conditions of poor visibility. These are by no means all the possible applications, but they serve to show that radar can be employed to expedite travel, by sea or by air, with a greatly increased measure of safety and efficiency.

Essential Functions

The fundamental principles upon which are based the measurement of range, bearing and elevation are simple and easily understood. Approach to the study of radar is further simplified by the fact that certain essential functions must be performed in all radar systems. Here, however, common ground between (*a*) warning systems and (*b*) precision systems abruptly ceases.

These essential functions can be performed by a variety of circuit arrangements. A suitable selection depends upon the performance specification, the application and the conditions under which the equipment is to be operated.

Each alternative combination of the major units requires a more or less complete rearrangement of the subsidiary units which co-ordinate the essential functions. Thus the composition and layout of radar systems differ widely, according to the requirements and conditions that must be met.

Clearly, therefore, it is necessary to examine :—

(1) The factors governing (*a*) range, (*b*) bearing, (*c*) accuracy of measurement and discrimination.

(2) The basic radar system.

(3) Alternative combinations of major units with particular reference to sub-paragraph (1) and new technique.

(4) The circuit arrangements employed, circuits and new technique involved in co-ordinating the functions of the major units for alternative combinations of these units.

Assumed Standard of Knowledge

Since the carrier frequencies for (*a*) warning sets and (*b*) precision sets may vary from about 100 megacycles or less for application (*a*) and up to 10,000 megacycles or more for application (*b*), it can be seen that the R.F. technique of radar, alone, covers a wide field. It seems therefore essential that the reader should approach the study of the subject equipped with a knowledge of the characteristics of series and parallel resonant circuits and the elementary theory of magnetism and electricity to that standard. It is also important that he should have some knowledge of the thermionic valve and the principles involved in the superheterodyne receiver.

In general, and until the R.F. stage is reached, the reader is concerned with waves and pulses which are mainly non-sinusoidal in shape, wherefore, a study of the various ways in which the major functions can be performed, is concerned largely with the production, shaping and electrical dimensioning of non-sinusoidal waves and pulses, at repetition rates that are generally determined by a master synchronising source.

Because these waves and pulses are for the most part non-sinusoidal, it is important to note that A.C. technique for sinusoidal waves does not apply. It is, therefore, necessary to consider, for example, the output waveform for a square wave input to a capacity-resistance circuit. In this connection, it is the distribution of voltage in the C.R. circuit for the duration of the applied pulse with which the reader will be chiefly concerned.

It is of the utmost importance that the behaviour of inductance, capacity and resistance in shaping and dimensioning circuits should be understood; consequently, Chapter IX (Differentiation) is of great importance.

Layout and Treatment of Subject

Radar covers such a wide field of electronics and radio engineering that it has been necessary to give a great deal of care and attention to the layout of this book in order that established telecommunication practice, modifications thereto and new technique may each be clearly distinguished.

Accordingly, for the convenience of those requiring to refresh, early chapters, following upon the initial discussion of the principles of range, bearing and elevation measurement, are devoted to brief recapitulation of the elementary principles of D.C. and A.C., with which the reader is assumed to be already acquainted.

This is followed by Chapters recapitulating vital principles connected with the operation of the thermionic valve, amplifiers oscillators and the cathode ray tube, as applied to radar.

Before discussing special circuits and new technique as used in radar, a chapter is allocated to a discussion of the basic radar system, and this is followed by a chapter devoted to the functions of pulse-forming circuits. In chapters which immediately follow some attempt has been made to classify pulses and the circuits by which they may be produced.

Separate chapters each deal with time bases and time base circuits, calibration and calibrator circuits and ultra-high-frequency technique.

The function of generating high peak power R.F. pulses at micro-wave frequency, is performed almost exclusively by the Resonant Magnetron. This valve is so important that a long chapter is devoted to the subject. However, it must be borne in mind, when reading it, that the theories advanced are based on knowledge at present available. Consequently, whilst many of the phenomena connected with this valve can be satisfactorily accounted for, there is still much to be learned about the conditions from which they arise.

Following upon a detailed survey of transmitters, receivers and display units (indicators), a few selected combinations of units (shown in block diagram form) are discussed. Here, it should be noted, that the combinations reviewed are not actual combinations taken from any particular equipment. They do serve, however, to illustrate the principles on which radar systems are built up for different applications.

The chapters on aerials incorporate feeder systems, switching or duplexing systems, beam switching, together with electro-mechanical systems for rotating the aerial and for transmitting information relating to instantaneous bearing in azimuth and/or elevation to the operating position.

The theory of transmission lines, wave-guides and cavity resonators as applied to radar is treated in Appendices I, II and III respectively. These Appendices have been separated from the main book sequence because each chapter is long and would disturb the continuity of study for those already acquainted with these sections of the work. On the other hand, those readers wishing to refer or refresh can do so easily at any stage of progress.

A few notes on test equipment have been added as Appendix IV. It is only possible in this volume to give a very brief outline of some of the test equipment commonly used in radar, and it is not at all possible to enter into any detailed description of its use.

CHAPTER I

FUNDAMENTAL PRINCIPLES OF MEASUREMENT OF RANGE, BEARING AND ELEVATION

VELOCITY (directed speed) =

$$\frac{\text{length}}{\text{time}} = \frac{\text{distance travelled}}{\text{time taken to do the journey}}.$$

It therefore follows that when velocity is a known constant and time can be measured, distance can be calculated = velocity × time.

This is the basis of range measurement. The velocity of æther waves in free space is known to be constant at 300,000,000 metres per second, and since very small fractions of time can be measured with the aid of the cathode ray tube, the distance between a W/T transmitter and a remote receiver is the velocity constant × t (in seconds) where t is the interval between the instants of transmission and reception of a wave.

For the measurement of comparatively small distances up to 200/300 miles or thereabouts, the velocity unit of 300,000,000 metres per second is much too large. It is more conveniently expressed as 328 yds. per microsecond (1 millionth sec.) approx.

The Echo

It is a well-known fact that when an electromagnetic wave travelling in free space encounters a solid body, some of the energy of the wave is scattered or reflected. This process may be thought of as re-radiation and, consequently, a redistribution of part of the energy of the transmitted electromagnetic wave. Some of this redistributed energy travels back to the transmitter site and may therefore be looked upon as an incoming wave emanating directly from the reflecting object. It can be received as a normal signal and is generally spoken of as the "echo."

Measurement of Range by the Echo

In order to calculate the range of a reflecting object or target from the transmitting site, it is only necessary to measure the time taken by the reflected wave to travel from the target back to the transmitter.

Since the time taken by the echo to travel from the target to the receiver must be equal to the time taken for the transmitted wave to travel from the transmitter to the target, the time interval between the start of the echo from the target and its receipt at the transmitting-receiving site, must be half the time interval between the start of the transmitted wave and the receipt of the echo. Consequently the range of the target must be $\frac{328 \times t}{2}$, or 164t yds. approx., where t measured in microseconds is the time interval between the instant of transmission and the instant of receipt of the echo.

An essential requirement for the measurement of range by this method, is identification of the echo with the transmission by which it is produced. This is accomplished by pulsing the transmitter at some convenient but constant rate, the duration of each successive pulse being identical and extremely short.

The transmitter output is therefore a succession of short sharp pulses or bursts of radio frequency energy.

The echo signals returned by a target are identical in shape but greatly reduced in amplitude, and since they are received in the intervals between each successive pulse of transmitted energy, they can be automatically identified with the pulse of transmitted energy by which they are produced.

As a result of the above arrangement the instants at which the transmitter " fires " (*i.e.*, the instant at which each successive pulse commences) can be used as a reference from which to measure the time intervals between it and echoes received from targets at varying ranges. These ranges can then be calculated from the formula :—

$$\text{Range (yards)} = \frac{328 \times \text{interval time (microseconds)}}{2} \text{ (approx.)}.$$

Pulse Repetition Rate, Pulse Width, Carrier Frequency and Wavelength

The number of times per second that the transmitter is stopped and started is the pulse repetition rate. The time or width of the pulse is its duration. The frequency and wavelength of the R.F. energy transmitted during each pulse is determined by the electrical constants of the R.F. generator as for normal W/T or R/T signalling.

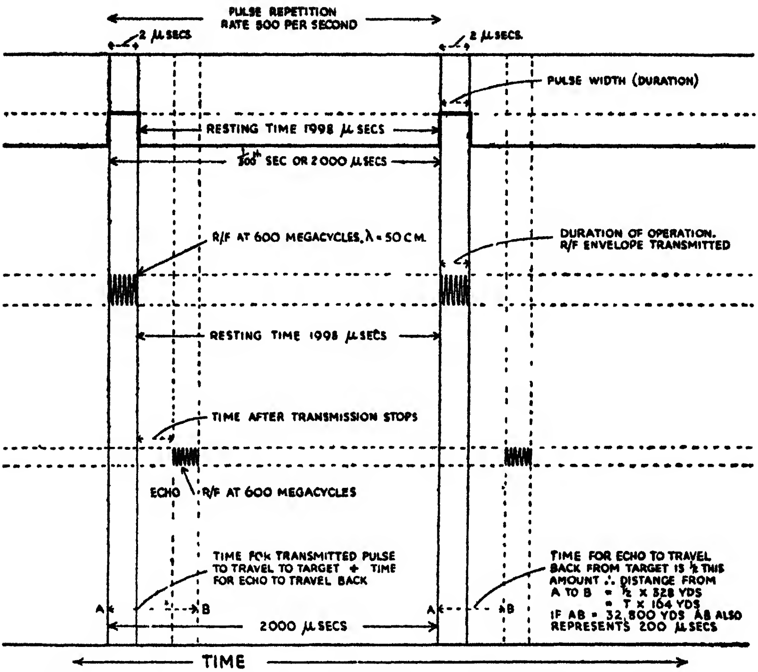


Fig. 1.—ILLUSTRATING RELATIONSHIP IN TIME OF PULSE WIDTH, PULSE REPETITION RATE, R.F. ENVELOPE AND ECHO.

Range is measured from the "leading" edge of the transmitted pulse to the "leading" edge of the echo.

Fig. 1 illustrates the relationship in time of pulse width, pulse repetition rate, carrier or radio frequency and received echo for the following conditions, all of which may vary within comparatively wide limits for different radar systems and conditions.

Pulse width (duration of pulse)	= 2 microseconds
Pulse repetition rate Transmitter started and stopped 500 times per second	= 500 per sec.
Radio frequency	= 600 megacycles per sec.
Wavelength	≑ 50 cm.
Range of target	= 32,800 yds.

(1) *Note* that the resting time of the transmitter (the time

elapsing between each stop and start) is long compared with the duration of each pulse, *i.e.* :—

Pulse width	= 2 microseconds
Pulse repetition rate	= 500 per sec.
Pulse time period	$\begin{cases} = \frac{1}{500} \text{ sec.} \\ = 2{,}000 \text{ microseconds} \end{cases}$
Resting time of transmitter	$\begin{cases} = 2{,}000 \text{ microseconds} - 2 \text{ microseconds} \\ = 1{,}998 \text{ microseconds} \end{cases}$

(2) There is ample time between each stop and start of the transmitter for all echoes from targets within the maximum range of the equipment to travel back and to be received without being masked by the next successive pulse from the transmitter.

The Basic Unit of Range Measurement and Calibration

Referring to Fig. 1, the time, AB, measured from leading edge to leading edge, is clearly the total time for the outgoing pulse to reach the target plus the time for the reflected echo to travel back to the transmitter-receiver position. Since both outgoing and incoming waves travel the same distance at the same velocity, the time for the return echo must be exactly half the total time, *i.e.* $\frac{AB}{2}$. Consequently the distance of the target from the transmitter in terms of time will be $\frac{AB}{2}$ microseconds. The wave velocity is 328 yds. per microsecond, $\frac{AB}{2}$ microseconds $\times$ 328 yds. equals the distance of the target from the transmitter in yards. This is more conveniently expressed as

$$AB \text{ microseconds} \times 164 = \text{range (in yards) approx.}$$

The above expression forms the general basis for calibrating the *screen* of a cathode ray tube.

This is accomplished by causing the electron beam of a cathode ray tube to trace a bright line horizontally across the diameter of the screen, *each time the transmitter fires*; the *commencement* of each trace, being *synchronised* with the *instant of firing* of the transmitter. Since each line occupies the entire working diameter of the screen, the *length* of each trace is *constant*, for a particular tube say—6 in.

This 6-in. line can be made proportional to any desired time by varying the rate at which it is traced by the electron beam, *i.e.*, by varying the speed at which the electron beam is moved or deflected across the screen. Thus if each trace is completed in say 200 microseconds, the 6-in. line is proportional to 200 microseconds. If the trace is made at half that speed, it will take 400 microseconds to complete the same length of trace and the 6-in. line will then be proportional to 400 microseconds.

This means that a line traced in the above manner across the screen of any C.R.T., no matter of what length, can be made proportional to any desired *time* by adjusting the speed at which it is traced.

This line is called the *time base*, and the circuits which control the speed of tracing are the *time* base circuits.

From the foregoing, and from Fig. 2, it is evident that *time* base can be calibrated in target range yards, *e.g.*, in Fig. 2 let the length of the time base be proportional to a time of say 366 microseconds. $366 \times 164 =$ approx. 60,000 (target) yards. Therefore, the time base can be proportional to either 366 microseconds or approx. 60,000 yds.;

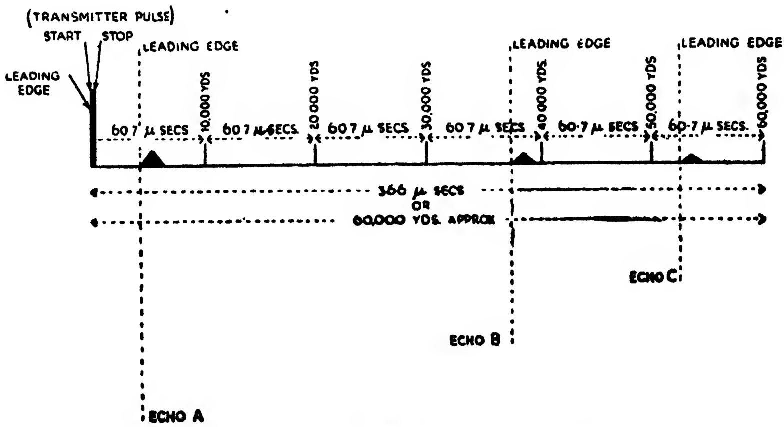


Fig. 2.—CALIBRATION OF A TIME SCALE IN TERMS OF YARDS.

Echo A appears at about 5,000 yds., or 30·55 μ secs. (approx.) after the transmitted pulse.

Echo B appears at about 38,000 yds., or 231·7 μ secs. (approx.) after the transmitted pulse.

Echo C appears at about 52,000 yds., or 317 μ secs. (approx.) after the transmitted pulse.

366 can be divided into equal lengths, marking the time base by some means at each point of division. The time base (Fig. 2) is calibrated in six 10,000-yd. lengths, each of which is proportional to 60·7 microseconds (approximately).

Example of Simple Calibration

In Fig. 2, echoes A, B and C are shown on a time base at ranges of 5,000, 38,000 and 52,000 yds. respectively. (Note that the height of the echo diminishes as the range increases. This is true when the targets at B and C present identical *reflecting surfaces*. If, for example, however, the A target was very small and the B target was very large, it might happen that the B echo was comparable with the A echo.)

The effect, produced on the screen by the arrival of an echo signal, is achieved by the application of the output of the receiver (at video frequency) to the cathode ray tube, in such manner that echo signals deflect, vertically, the time base trace. This means that a " pip " appears on the time base for each echo received, marking by its position on the time base, the instant at which it is received relative to the commencement of the time base and to the firing of the transmitter.

The distance measured along the time base from the leading edge of the echo signal to the leading edge of the transmitted pulse, see Fig. 2, is proportional to the time interval between the firing of the transmitter and the receipt of the echo.

Since this length can be physically measured as a fraction of the total length of the time base scale, it has a proportional value in time and range, consequently, assuming that the time base scale has been calibrated at 164 range yards per microsecond of length, the range of the target can be read from the position of the echo on the time base.

Measurement of Bearing

In order to measure the bearing of a target it is necessary to be able to identify the direction from which a given echo arrives. This requirement means that it is essential to be able to direct the transmitted energy in any given direction at will, and to discriminate between signals received from a given transmission in a given direction and signals arriving from all other directions.

These requirements are met by concentrating the transmitted energy into a narrow beam, which can be directed at will, by rotating the aerial in azimuth,* and since the receiving pattern of an aerial is similar, fortunately, to its transmitting pattern, an aerial which transmits a narrow beam discriminates in favour of all signals received from targets within the beam. The aerial may therefore be pointed in any direction at will, or caused to sweep continuously through all points of the compass.

If the beam is symmetrical about the centre of the radiating system the distribution of R.F. energy in the target area will be maximum when the centre of the radiating system is aimed directly at the target. Since this factor determines the pattern for distribution of the transmitted energy over the target it must also control the collected or received pattern. Consequently, maximum energy is received or collected from a target when the target lies along the axis of the radiating system. In other words, the conditions applying to maximum transmission of energy to the target are identical with those for maximum collection of energy by the aerial system from the target.

Beam Width

The beam width is very important. The narrower the beam the greater is the concentration of energy upon the target for any given transmitted signal and, all other things being equal, a larger echo will be returned for any given range and target. Also, the separation (or definition) in the case of bunched targets is better with a narrow beam, and greater accuracy of the bearing reading is also rendered possible. In general, very narrow beams necessitate, for practical reasons, the use of centimetre waves and range is therefore limited. Hence, there is a line of demarcation between long-range warning sets and sets for precise measurement. Long-range sets generally employ frequencies of the order of 100 to 400 megacycles. The beam, with practicable aerials at these frequencies, is generally wide, but since a high degree of accuracy in measurement is not essential, this does not matter very much. Indeed, it may be an advantage for search purposes. Precision sets require to use narrow beams, and in order to obtain them with practicable aerials, ultra-high frequencies must be employed. Range is, however, limited.

* Azimuth, in which all "true bearings must lie," is defined by the angle formed by the intersection of the line of sight and the true meridian passing through the observer's position. It is measured "clockwise" at the position of the observer, from the north point of the horizon and from 0° to 360°.

Transmission of Bearing Data to Operating Position

By the use of rotatable and highly directional aerial systems, the direction from which echoes are received can be identified by rotating the aerial until maximum echo signal is received. This means that the aerial is pointing directly at the target. To enable the operator to ascertain the direction in which the aerial is pointed and, therefore, the bearing of the target, a device variously known as a Magslip or Selsyn is geared to the aerial shaft.* As the aerial is rotated, the rotor of the Magslip or Selsyn rotates with it and transmits electrically to a receiver and indicator at the operating position, the bearing of the aerial relative to the ship's head. It is also possible to read bearings from the handwheel with a properly designed Servo System.

By means of a repeater mechanism, where a gyro compass is fitted, the compass reading is also transmitted to the operating position.

Comparative readings enable the operator to give the bearing of the target, either as a bearing relative to the ship's head (relative bearing), or relative to compass north (true bearing).

Measurement of Elevation or Height

The measurement of height at low altitudes, particularly with beams that are comparatively wide in the vertical direction, is complicated by the reflected ground wave, which breaks up the free space pattern.

In order to plot the position of aircraft in space, relative to the observer, it is necessary to obtain the slant range, the angle of elevation and the bearing of the aircraft.

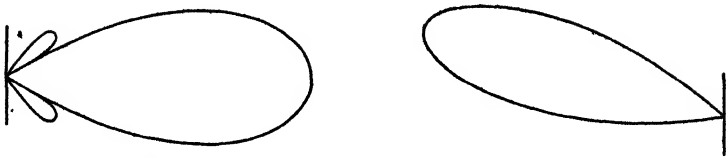
Polar Diagrams

The general patterns for the distribution of energy in free space from a highly directional aerial are shown in Fig. 3; these are developed from horizontal and vertical polar diagrams respectively. To construct the polar diagram, the aerial is taken as centre and observed field strength is plotted against radii.

The horizontal polar diagram (a) shows that the maximum radiation is along the aerial line of sight, *i.e.*, along a line projected perpendicularly from the electrical centre of the aerial system towards the observer.

The vertical polar diagram (b) is drawn as it would be plotted by an observer in an aeroplane flying round the aerial at

* Sometimes an "M"-type transmitter is substituted for the Magslip or Selsyn. The effect is the same.



(a) (b)

Fig. 3.—Horizontal and vertical patterns of a highly directional aerial suitable for direct measurement of altitude.

(a) Radiation pattern in horizontal plane (moderately wide beam).
(b) Vertical pattern in free space of another aerial (narrow beam).

different heights, with the aerial tilted, to direct the lobe into free space so that no part of the lobe touches the ground.

With the beam dimensioned as shown in Fig. 3 (b), the slant range of an aeroplane can be obtained by tilting the aerial in elevation and rotating it in azimuth until the target lies directly along the aerial line of sight. This condition is indicated to the operator by maximum echo. The tilt angle or angle of elevation, and the bearing on which the target is found, are both communicated to the operating position by repeater mechanisms under control of the aerial.

Modification to Free Space Pattern made by Ground Reflection

This method is only practicable, however, when the elevation is such that no part of the lobe touches the ground. When these conditions do not apply, reflections from the ground occur, and the lobe pattern is distorted. In consequence of this the polar diagram is considerably modified by the ground reflection factor.

Reference to Fig. 4 shows that the path ABC for reflected energy must necessarily be longer than the direct path ADC. Wherefore, at various points in the area covered by the pattern and according to the difference in length of the alternative paths to any given point, the reflected wave must arrive more or less in or out of phase with the direct wave. Thus, the vertical polar diagram for free space, without reflection, will be considerably modified when reflection occurs. The field strength at any point, covered by

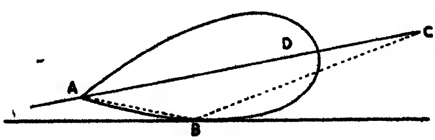


Fig. 4.—The direct and indirect reflected paths.

radiation from the aerial, will now be the algebraic sum of the energy transmitted to it by the direct path, and the energy transmitted to it by reflection from the ground. If the direct wave and the reflected wave arrive in phase they will add ; if they are in anti-phase they will cancel. The resultants at intermediate phase positions will be the vector sum.

Effect of Ground Reflection on Detection and Ranging of Aircraft

The effect of the reflected wave in breaking up the main lobe and the irregular pattern produced is shown in Fig. 5. When the free space lobe is broken up by ground reflection it can be seen from Fig. 5 that an aeroplane flying in at height A will be detected at A¹, whilst another plane flying in at height B

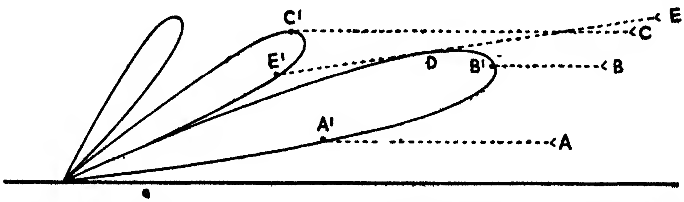


Fig. 5.—MODIFIED PATTERN DUE TO INTERFERENCE BY REFLECTED PORTION OF WAVE.

will be detected as it reaches B¹. Corresponding conditions apply to height C, and so forth.

Conclusions

Now, consider the case of an aircraft descending along the path of EDE¹. It can be observed at D but will be quickly lost until it is again picked up at E¹. The conclusions to be drawn are that when beams, wide in the vertical dimension, are aimed at low-flying aircraft so that the lower part of the free space pattern touches the ground, (*a*) the tilt method of measuring elevation cannot be employed. (*b*) The maximum range at which the plane will be picked up depends upon the height at which it is flying and the shape of the polar diagram as modified by the ground reflection factor. (*c*) Echoes will appear, fade and reappear if the line of flight is such that the plane passes out of one lobe and into that of the next. In such cases, height is computed from measured range and by reference to tables constructed from vertical polar diagrams. It is, therefore, advantageous to employ a narrow beam, with particular reference to the vertical pattern, when it is desired to obtain accurate data regarding the movements of aircraft in flight.

Note. When reading polar diagrams it must not be supposed that the lobe is drawn to define the boundaries of radiation. The observed field strength along the axis of the lobe is taken as 100 per cent. Radiation in other directions is proportional to radius. Consequently the lobe pattern of ratios thus obtained is constant for any distance measured along the axis. In other words, the shape of the lobe remains unchanged for any given range or distance measured along its axis. In imagination it merely expands proportionately in all directions, so that whilst actual values change, their ratios are constant.

Summary

The fundamental principles of the measurement of range involve the measurement of very short periods of *time*. Measurement of bearing and elevation is performed by " aiming " the aerial at the target and measuring the displacement angle from zero. The application of the cathode ray tube enables these functions to be carried out by means of a visual signal.

Coordination of the outputs of transmitter and receiver, together with control of the display circuits, is carried out by means of synchronising pulses, the frequency of which is the repetition rate of the system.

The shape of a pulse, *i.e.*, its width in *time*, controls the duration of operation of the circuit to which the pulse is applied.

Pulse-forming and amplifying circuits are widely used in radar. In some cases the utmost care must be taken to preserve, at the output, the shape of the input pulse, whilst in other cases controlled distortion is deliberately introduced.

Before proceeding to a detailed study of these circuits, therefore, it is essential to ensure that certain fundamental principles in D.C. and A.C. technique, together with the principles of the thermionic valve, valve amplifiers, oscillators and the cathode ray tube, are clearly understood.

For this reason, and for the convenience of those who wish to refresh, Chapters II, III, IV, V and VI have been written to recapitulate some of the more important conclusions. In all cases of difficulty it is recommended that a standard work should be consulted on the particular point.

CHAPTER II

SOME IMPORTANT ELEMENTARY PRINCIPLES OF D.C. AND A.C. CIRCUITS

THE following chapters have been written with the assumption that the reader is acquainted with the elementary principles of direct and alternating current, the thermionic valve and the general principles of radio transmission and reception. There follows, therefore, a brief recapitulation of some of the most important of these principles as employed in radar.

The Unit of Current

Any movement of an electron in a continuous uni-directional manner, or as an oscillation between two extreme points, or in course of displacement such as takes place when a condenser is charged or discharged, constitutes an electric current.

The magnitude of the current is measured by one or other of the effects exhibited when a unit quantity of electricity (1 coulomb) passes the measuring point in unit time (1 sec.).

The ampere is that rate which transports 1 coulomb past the measuring point in 1 second, and it can be measured by the electromagnetic, heat or chemical phenomena associated with the passage of electricity at unit rate.

Since the quantity of electricity, of which the electron appears to be entirely formed, is $1{\cdot}59 \times 10^{-19}$ coulombs, a unit current of 1 ampere is represented by electron movement past the measuring point at the rate of $6{\cdot}29 \times 10^{-18}$ per second.

Electronic and Conventional Signs

The minus sign is employed as a label to relate the effects produced by an electron in motion to the electrical phenomena associated with conventional current plus. It is, therefore, important to bear in mind, when thinking in terms of electron movement, that the conventional signs of + and − do not necessarily apply to magnitude but are labels indicating relative position, direction or state *e.g.*, a body which by convention is labelled "positively charged" possesses less than its normal complement of electrons but, since the electron itself exhibits all the characteristics of a body *conventionally* labelled "negatively charged," the body which is minus its normal complement of electrons must be considered to be positively charged in a conventional sense.

Throughout the following chapters the direction of current will be taken as the direction of electron movement. Since, however, the observed phenomena associated with the electron current are identical with effects associated with currents flowing under conventional signs, the latter can be employed as long as the signs are understood to be merely labels in the same sense as red and black might be used.

The Ratio $\frac{E}{I}$

When a source of e.m.f. is connected across its terminals, a disturbance is instantaneously propagated throughout a circuit ; this disturbance upsets the state of electrical equilibrium existing between the atoms and their electrons, consequently a current flows in order to re-establish it. Within the limits of temperature rise, and once established, the flow of current remains constant as long as there is no change in the voltage across the circuit. The value of the ratio $\frac{E}{I}$ for a D.C. circuit is stated in ohms and is termed "resistance." (E = volts. I = amperes).

Resistance

Resistance may be thought of as a measure of an effect of random collisions between electrons, resulting from the application of e.m.f. to the terminals of a circuit. These collisions generate heat and the electrons become more agitated, as a consequence more collisions occur and the temperature rises, until the rate of heat dissipation is equal to the rate at which heat is generated. Any increase in terminal voltage causes a rise in current, and consequently, more collisions ; this results in an increase of temperature. Since, however, resistance is a measure of an effect of collisions, it must clearly increase when the temperature rises, which is in fact the general case with exceptions..

Impedance in an A.C. Circuit

In an A.C. circuit the value of the ratio $\frac{E}{I}$ is known as impedance (Z). Impedance resolves into components of resistance and reactance.

H.F. Resistance and Skin effect

The resistance component in an A.C. circuit is due to similar causes and acts in the same manner as in a D.C. circuit.

When the frequency is high, the A.C. resistance of a conductor of given dimensions is greater than its D.C. resistance. This, whilst indirectly caused by inductive effects, is termed a skin effect which, by reducing the available cross-sectional area for conduction, causes the A.C. resistance of a conductor to appear greater than its D.C. resistance as determined by $\frac{\rho l}{a}$. The skin effect increases with frequency which is quite logical, since the inductive reactance which causes it also increases in like manner.

Reactance in A.C. Circuits

The reactance component (X) is the vectorial resultant of the composition of ωL and $\frac{1}{\omega C}$ where ω = angular velocity ($2\pi f$), and L = inductance (Henrys), and C = capacity (Farads).

Z is the vector quantity resulting from the composition of R and X or $Z = \sqrt{R^2 + X^2}$.

The significance of the ratio $\frac{E}{I}$ in the analysis of transmission line and matching problems is particularly marked.

Kirchoff's Laws

Voltage dividers of various types and differentiating circuits are widely used in radar ; Kirchoff's laws are of great importance when considering these circuits.

Kirchoff's First Law. At any junction of resistances the sum of the currents towards the junction is the sum of the currents flowing away from it.

Kirchoff's Second Law. In any closed circuit the sum of the e.m.f.s reckoned positive in the direction of the current and negative in the opposite direction, is equal to the sum of the products of current and resistance in every part of the circuit.

This can be reduced to a simpler form. It means that the total voltage drop round any complete path = zero.

Resistances in Parallel and Inductances in Parallel

Kirchoff's laws can also be applied to A.C. circuits, but in this case, since the quantities involved are vector quantities, for word "sum" read "vector sum."

When the equivalent resistance is required for two resistances in parallel, the reciprocal of the sum of the reciprocal formula may be simplified by the direct solution.

$$R = \frac{R_1 \times R_2}{R_1 + R_2}.$$

Inductances in parallel can be calculated by the same means, but inductances in parallel with resistances cannot be calculated in this way, since vector quantities are involved.

Energy, Power and Work

The work done in a D.C. circuit is 1 joule when 1 coulomb is moved through a potential difference of 1 volt.

The total work done therefore = QE joules. Power is the rate at which work is done, *i.e.*, the power concerned

$$= \frac{QE}{t} \text{ joules}$$

$$= \frac{Q}{t} \times E \text{ joules.}$$

Since $\frac{Q}{t}$ = the current I, then the power consumed is $E \times I$ watts or $\frac{E \times I}{10^3}$ kilowatts.

This conception has to be modified in an A.C. circuit unless the current and voltage are in phase.

When voltage and current are in phase the power is simply $E \times I$ watts, but where the current lags or leads the voltage, to obtain true watts the quantity EI must be multiplied by a power factor which is always less than unity.

Power Factor

The power factor is $\cos \phi$ where ϕ is the phase angle between current and voltage.

The ratio $\frac{X}{R}$ is the tangent of the angle ϕ. When the value of this ratio is known, the angle can be found from a trigonometrical table and the cosine can be ascertained in the usual manner. Since $\cos 90° = 0$, the phase angle can never be equal to or greater than 90°. Some component of I must always be in phase with E in order to supply the loss in resistance which must always be present in any complete circuit. This may be very small, but it cannot be zero. In an open circuit, however, it may approach infinity.

RESONANT CIRCUITS

Capacity and Inductance in A.C. Circuits

If the behaviour is compared of capacity and inductance in an A.C. circuit, the following results are obtained.

In the case of fixed values of capacity, the capacity reactance $X_C = \frac{1}{\omega C}$ varies inversely with the frequency, thus at high frequency the reactance X_C may become very small; in fact, if the frequency is made high enough, the value X_C may become negligible. Thus, if a resistance is shunted by a condenser and the frequency is made high enough, X_C will become negligibly small and will then, in effect, short circuit the resistance.

Conversely, as the frequency is decreased, X_C is increased and the resistance shunts the condenser more effectively. Thus, at zero frequency, *i.e.*, D.C., X_C becomes infinity and the resistance shorts the condenser.

The behaviour of inductance, under similar conditions, is exactly opposite to capacity.

X_L varies directly with frequency; consequently, if the frequency is made sufficiently high, X_L approaches infinity and behaves very nearly as an open circuit. Since the inductance must have some resistance, however, no matter how small, it cannot act as a perfect open circuit. There must be some drop across its resistance component and, therefore, a small in-phase current must flow.

As the frequency is decreased the reactance X_L becomes smaller until at zero frequency, *i.e.*, D.C., X_L becomes zero as long as there is no change in the voltage across it.

If a resistance is shunted by an inductance the shunting effect at very high frequencies is almost negligible. At zero frequency, however, the resistance is practically short circuited by the inductance (assuming the resistance component of the inductance to be very small compared with the resistance with which it is in parallel).

In a series circuit, at zero frequency, a condenser produces the effect of an open circuit. At extremely high frequencies the effect of the condenser is very small.

An inductance in a series circuit behaves as a low resistance at zero frequency and becomes a choke at high frequencies.

Resonant Frequency

From the foregoing, it follows that capacity and inductance in a series circuit will have reactances X_C and X_L, which are in opposition; consequently, if the values of C and L are constant and the frequency of the applied e.m.f. is varied, some value of frequency will be found which makes the reactance X_C = to the reactance X_L. When this condition is satisfied, the term X in the equation $Z = \sqrt{R^2 + X^2}$ becomes zero, and the current at this exact frequency is then determined by the value of R.

Conditions for Resonance

This is the condition for electrical resonance :—

$$(1)\ \omega L - \frac{1}{\omega C} = 0$$

$$(2)\ \omega^2 = \frac{1}{LC}$$

$$(3)\ \omega = \frac{1}{\sqrt{LC}}$$

$$(4)\ f = \frac{1}{2\pi\sqrt{LC}}.$$

Since reactances X_C and X_L cancel in a series resonance circuit, if R is kept small enough, the current can become very large.

Characteristics of Series Resonance and Parallel Resonance Circuits

Thus, the characteristics of a series resonance circuit are :—

Low impedance, large current in phase with the voltage, and a high voltage across the inductance proportional to L.

Inductance and capacity in parallel form a parallel resonance circuit at the resonant frequency. A circuit of this nature is sometimes referred to as a tank circuit.

In the case of a parallel resonance circuit, in accordance with Kirchoff's first law, the line current must be the vector sum of the currents in the two branches I_C and I_L.

At resonance these two currents are nearly equal, but are in phase opposition. They will consequently cancel relative to line junctions. Cancellation will not be complete, since a small in-phase current flows due to the resistance of the circuit. At resonance this small current only will appear as the line current.

At the resonant frequency a large current does, however, circulate internally in the tank or closed circuit.

Since currents I_c and I_L are in anti-phase and flow at maximum value simultaneously between the two branches of the parallel circuit but on opposite sides of it, there is, therefore, a large current circulating internally and a small current must flow in from the line to make up I^2R losses. At resonance frequency, Z approaches infinity and large voltages build up across L and C. At resonant frequency, series circuits admit maximum current, but parallel circuits present maximum impedance.

The characteristics of a parallel resonance circuit are therefore — high ratio of $\frac{E}{I}$ in the main line. This means that the resonant tank circuit offers a very high impedance at resonant frequency to the e.m.f. applied across it. The main line current is small and conditions for resonance can be adjusted by the dip reading of a meter in the main line. As the circuit goes into resonance the main line current will fall, rising again as the frequency passes through the resonant value.

Due to large internal circulating current, a large voltage appears across the condenser. This voltage may be thought of as arising out of the circulating current mentioned in the previous paragraph.

A.C. and H.F. Losses

The efficiency of a circuit $= \frac{\text{output}}{\text{input}}$, thus losses must be reduced to the minimum. Unless great care is taken in design, appreciable losses occur at radio frequencies other than those arising from the resistance of conductors alone.

In the case of a radio frequency inductance, the principal loss is generally due to the self capacity of the coil, which may be thought of as being shunted across the turns. At very high frequencies the shunting effect may be serious and greatly reduce the effectiveness of the coil and the output of the circuit. This also applies to stray capacity between leads, etc. Dielectric and insulation losses occur and are also important at high frequencies, consequently the total loss in a radio frequency circuit may be very much greater than that represented by I^2R losses.

For the above reasons, the apparent A.C. resistance of a circuit for high frequencies may, therefore, be very much greater than its ohmic resistance.

The Factor "Q"

The factor Q determines the goodness of a coil or circuit at a particular frequency and is equal to $\frac{\omega L}{R}$.

In certain cases it is desirable that a circuit should discriminate sharply at a given frequency. This can be done by using a parallel resonance circuit, having its constants adjusted for resonance at the selected frequency.

The impedance at resonant frequency is

$$Z = \frac{(\omega L)^2}{R}.$$

Since at resonance X_C is equal to X_L,

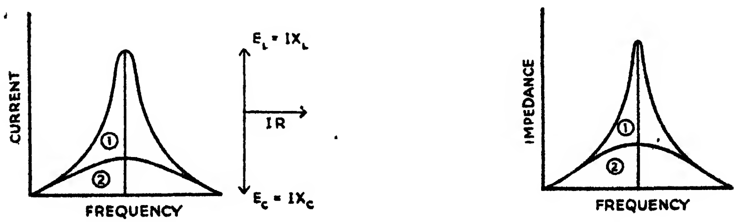
$$Z = \omega L Q,$$

I is proportional to EQ.

The value of Q may vary from 20 to 100 for coils with iron cores, and may go as high as 30,000 for resonant cavities.

Resonance Curves, without Resistance and with Resistance

From comparison of curves a and b (Fig. 6), it will be seen that the current curve for series and the impedance curve for parallel circuits resemble each other closely in shape.



(a) (b)

Fig. 6.—COMPARISON OF RESONANCE CURVES FOR SERIES AND PARALLEL CIRCUITS.

(a) Current curves in series resonant circuit.
  (1) Resistance small.
  (2) Resistance large.
(b) Impedance curve parallel resonant circuit.
  (1) Small resistance.
  (2) Large resistance.

Frequency-discriminating Properties of Series and Parallel Resonance Circuits

From curves 1 it will be seen that a very small departure on either side, from the resonant frequency, causes a large fall of

(a) current, (b) impedance. A circuit having either of these characteristics can discriminate very sharply against all frequencies off the resonant frequency. Curve 2 shows the effect of introducing resistance. The sharply peaked curve is now depressed and flattened. The resonant point is less clearly defined, and the circuit cannot discriminate so sharply against " off resonance " frequencies.

For maximum discrimination then, a circuit must have a high Q.

Fig. 7 shows how impedance varies relative to applied frequency for series resonance.

It will be noticed that as frequency is increased from the lowest value towards the resonant frequency, the value of X_C diminishes whilst that of X_L grows. At frequencies below the resonant value the circuit behaves capacitatively. At resonant frequency the circuit behaves almost as though it were a pure resistance. At frequencies above the resonant frequency the circuit behaves inductively.

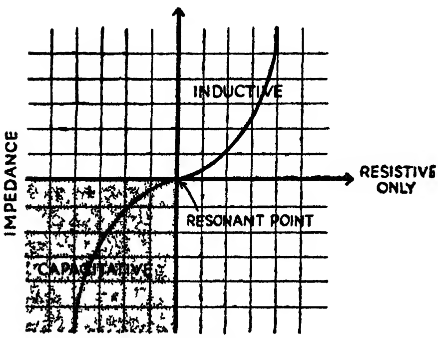


Fig. 7.—CHARACTERISTIC OF A SERIES RESONANT CIRCUIT

Below the resonant frequency, therefore, the current leads the voltage. Above resonance it lags the voltage.

Impedance Matching

Impedance matching is essential when it is desired to transfer the maximum amount of energy from one circuit to another.

When a transformer is used for this purpose the impedance of the primary $= \frac{E_p}{I_p}$. Similarly, the impedance of the secondary $= \frac{E_s}{I_s}$.

$$Z_s = \frac{E_s}{I_s}$$

Expanding
$$\frac{E_s}{I_s} = \frac{E_p \dfrac{N_s}{N_p}}{I_p \dfrac{N_p}{N_s}}$$
$$= \frac{E_p}{I_p}\left(\frac{N_s}{N_p}\right)^2$$
$$\therefore \frac{Z_s}{Z_p} = \frac{(N_s)^2}{(N_p)^2}$$
$$\left(\frac{N_s}{N_p}\right) = \sqrt{\frac{Z_s}{Z_p}}.$$

For unity coupling the impedance ratio is approximately equal to the square of the turns ratio.

When an air core transformer is used to couple energy out of one circuit into another, the impedance match can be controlled by varying the coefficient of coupling between primary and secondary.

The effect of this coupling on the primary is similar to that of adding resistance in series with the primary circuit. This apparent resistance can be considered as being reflected from the secondary circuit. Increasing the coupling increases the reflected resistance and so reduces Q.

When the coupling is loose the reflected resistance is small and the resonance curve is sharp. When the circuits are tightly coupled the resonance curve is very broad and the Q is low.

Matching is important for two reasons :—

(*a*) To obtain the maximum transfer of energy.

(*b*) To avoid the presence of reflected waves which tend to produce distortion of the input signal.

Oscillatory Circuits

If the condenser C in Fig. 8 is assumed to have been charged to V volts, the energy stored in static form is $\frac{1}{2}CV^2$ joules. If the switch S is opened, C will discharge through the circuit LC, and providing the resistance of the circuit is smaller than $2\sqrt{\frac{L}{C}}$, oscillations will occur at the natural frequency of the circuit.

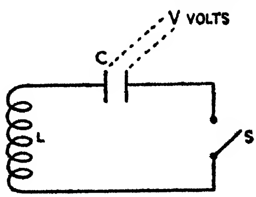


Fig. 8.—CONDENSER CHARGED PRIOR TO CLOSING S.

Actually the natural frequency of

the circuit may not be exactly the same as the resonant frequency, *i.e.*, the frequency at which maximum current is established by an e.m.f. applied at the frequency which establishes resonant conditions.

Natural frequency is $\frac{1}{2\pi}\sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}}$.

The resonant frequency $= \frac{1}{2\pi}\sqrt{\frac{1}{LC}}$.

When R is very small, however, as it naturally is in an oscillatory circuit, it is justifiable to accept $2\pi\sqrt{\frac{1}{LC}}$ as the oscillation frequency. The form of oscillation that occurs at the natural frequency is in reality a " forced oscillation."

To return to Fig. 8. At the instant after the condenser commences to discharge, the voltage across the condenser is engaged in establishing the instantaneous current and the magnetic field of the inductance L. V is, therefore, resolved into two components, *i.e.*, IX_L and IR, nearly 90° out of phase. IX_L is that component of E required to overcome the counter e.m.f. $-IX_L$, and IR is the in-phase component required to maintain it.

Energy Transfer, Voltage and Current Distribution

As the condenser discharges, the voltage V across it falls and V will be zero when the condenser is fully discharged. During the period of discharge the magnetic field associated with L grows and reaches maximum intensity at the instant the condenser is fully discharged. The energy originally stored in the condenser $\frac{1}{2}CV^2$ is now stored in the magnetic field of L and appears as $\frac{1}{2}LI^2$ less whatever loss has been incurred in making the transformation.

Losses

In a low frequency circuit this loss is mainly due to I^2R, I being the medium through which the transformation takes place.

Since the condenser is now fully discharged, V, across the condenser, falls to zero and the current tends to zero. The magnetic field of the inductance cannot now be sustained and, in collapsing provides the new e.m.f. which, by Lenz's law causes the current to persist and to recharge the condenser again in the opposite direction. The energy $\frac{1}{2}LI^2$ joules of the magnetic field is again transferred to the condenser as $\frac{1}{2}CV^2$. A further loss, however, must again be incurred, so that the new charge of magnitude $\frac{1}{2}$ CV^2 will be less than $\frac{1}{2}$ LI^2 by the loss of energy I^2R sustained in making the transfer. In high frequency circuits losses other than I^2R become important; also, of course, R is greater than its D.C. value.

Since the energy in the circuit becomes less at each transformation, the losses incurred at each change become less, since I^2 also decays. This decaying effect is mainly determined by the constants of the circuit at low frequencies, but at very high frequencies energy losses due to dielectrics, insulation, stray capacities and radiation become important. The effect of this is that R apparently increases in value, and some new and larger value must be given to the ohmic resistance R so that all losses may be made to appear as a *total* I^2R loss.

Continuous Oscillations

The circuit losses discussed above determine the number of oscillations which, for any given set of conditions, occur before they are damped out. Consequently, when undamped or continuous oscillations are required, energy must be fed into the circuit from some external source, of the correct value and in the right phase, to make up for circuit losses and to maintain constant and continuous oscillations.

The conditions for oscillations hold true as long as the total resistance (ohmic and other losses) does not exceed $2\sqrt{\frac{L}{C}}$.

Application

Damped oscillations excited in the manner previously discussed are frequently employed in radar circuits. This method of production referred to in Fig. 8 is frequently termed "shock excitation." Circuits generating continuous undamped oscillations are also largely employed.

CHAPTER III

ELEMENTARY PRINCIPLES OF THE THERMIONIC VALVE

It is assumed that the reader is familiar with the general construction and elementary principles of operation of the thermionic valve. In case this information has not been acquired, reference should be made to one of the standard text-books on radio communication.

If the anode volts are held at a fixed value, increase in temperature of the emitter does not increase the anode current. When the supply of electrons is in excess of the anode demands, a cloud of free electrons hovers in the vicinity of the cathode. This is the space charge, and for a fixed value of anode volts, its magnitude depends upon the rate of emission, modified by the rate at which electrons are returned to the cathode by mutual repulsion.

In this condition the anode current is drawn from this reserve or space charge, the boundary of which is established by the force of attraction of the anode, counter attraction by the cathode and mutual repulsion between the electrons themselves (diode).

Consequently the anode current that flows is just that

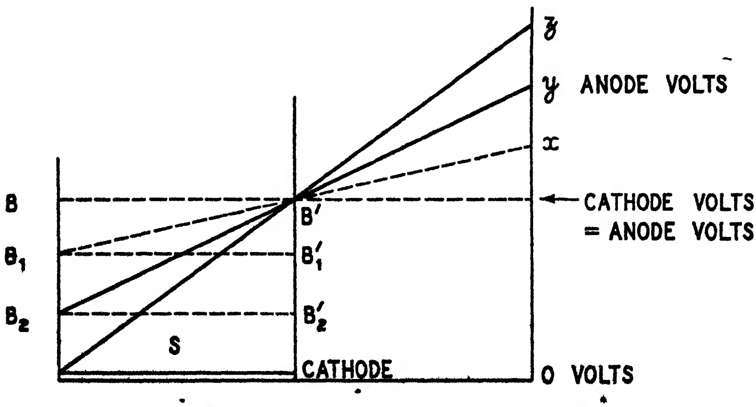


Fig. 9.—Diagram illustrating the effect on electrons of the space charge of increasing anode volts.

amount which establishes equilibrium between the various forces acting on the electrons in the space charge.

If anode volts are now increased, the state of equilibrium is upset and more current flows to the anode to re-establish equilibrium conditions. Consequently, the conditions obtaining when anode volts are increased continuously may be thought of as shown in Fig. 9.

S is the space charge bounded by B, B′. The anode volts are shown as though reaching over the boundary B, B′ to attract electrons from the space charge.

The first and smallest value of anode volts "x" will obviously only attract the outer layers of electrons down to the level $B_1B'_1$. y volts on the anode overcome the counter attraction still more and draw away electrons down to the level $B_2B'_2$. z volts on the anode are sufficiently large to overcome all counter attraction and draw away all the electrons as fast as they are emitted. If, at this point, the emitter is already operating at the optimum safe temperature, the valve is completely saturated. A further increase in emission is not possible with safety and therefore a further increase in anode volts can produce no further increase of anode current.

Space Current

Under normal conditions the total *space current, i.e.*, the total current flowing through the valve by way of one or more of its electrodes (according to whether it is a diode, triode, tetrode, etc.), *is drawn initially from the space charge.*

The Control Grid

The introduction of a grid between the space charge and anode has a profound effect upon the space current.

When the grid is positive it accelerates electrons coming from the cathode towards the space charge boundary so that they approach nearer than normal to the anode. Under the increased influence of the anode they are drawn to it and the space current is greatly increased.

When the grid is negative it decelerates electrons approaching the boundary, the result is that any space current flowing is decreased. If the deceleration is sufficiently great the space current will be suppressed entirely.

Small changes of potential on the grid ΔE may, therefore, cause large changes of anode current ΔI.

Also, if the space charge outer boundary is thought of as an extension to the radius of the cathode, it is clear from the above that the effective radius of the cathode grows when the grid is positive and decreases when it is negative, consequently the grid-cathode capacity must vary in a similar manner.

For low values of cathode or filament temperature, with normal volts on the anode, current can be increased by raising the temperature of the emitter until the rate of emission is equal to the anode demand. The anode current under these conditions is temperature limited.

Valve Constants

The most important constants of the valve are mutual conductance (g_m.), A.C. resistance (R_a) and amplification factor (μ).

These three constants are determined by the geometry of the valve but can be developed from its static characteristics.

Mutual conductance is the rate of change of anode current (I_a) with grid volts (E_g) and may be written $\frac{I_a}{E_g}$ symbol $= g_m$.

A.C. Resistance. This is sometimes spoken of as anode resistance, impedance or plate resistance. This constant is the reciprocal of the rate of change of anode current (I_a) with anode volts (E_a) when the grid volts E_g are held constant. Thus $\frac{1}{\frac{I_a}{E_a}} = \frac{E_a}{I_a}$. Symbol $= R_a$. This constant is usually understood to be the A.C. resistance of the valve measured at some constant value of grid volts.

Amplification Factor. The change in anode current brought about by a change in grid volts is much greater than for an identical change in anode volts. This constant is measured by the grid voltage required to change I_a from I_{a1} to I_{a2} as compared with the change which must be made in anode volts to restore I_{a2} to its original value (I_{a1}). Thus amplification factor $= \frac{E_a}{E_g}$ for some agreed change of I_a (I_{a1} to I_{a2}).

When E_a and E_g are held at constant values a steady current I_a must flow.

If the value of E_g is changed the value of E_a must also change because the current I_a now changes to a different value.

Thus, if the grid voltage is changed by an amount ΔE_g and the anode volts are held constant, the change in current would be

$$^{*}\Delta I_a = g_m \Delta E_g.$$

When E_g is changed, however, E_a must alter. Let this change be ΔE_a; therefore the change in I_a for this reason is

$$\Delta I_a = \frac{E}{\Delta E_a}.$$

The total change in I_a due to simultaneous changes in E_g and E_a is

$$\Delta I_a = g_m \Delta E_g + \frac{E}{\Delta E_a}.$$

Consequently, when the available anode voltage, the input conditions and the required output are known, the point on the $\frac{E_a}{I_a}$ characteristic at which the valve must be put to work can be determined for any desired performance.

Load Line

When a load line has been drawn, maximum and minimum values for output voltage and current for any given input can be obtained.

The working point of the valve for any chosen values of anode volts load resistance (R_L) and grid bias (G_B) can also be determined by construction of a load line. A family of curves representing the E_a/I_a characteristics is employed for this construction. The optimum value of load resistance for a particular value is usually specified by the manufacturers, but smaller values may be chosen.

The output voltage is the drop across R_L, and since the value of R_L is known, the value of I_a required to produce the drop (I_aR_L) can be determined.

When this current I_a flows, the corresponding anode volts will be $E - I_aR_L$. At the point corresponding to this value on the E_a axis, a line is drawn perpendicular to it, cutting the appropriate E_g curve. This point of intersection

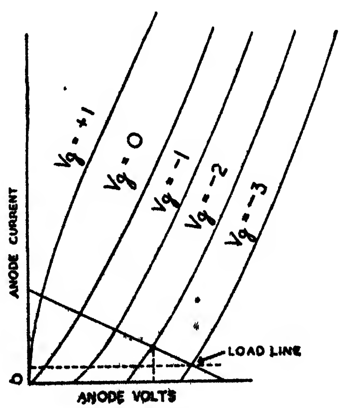


Fig. 10.—LOAD LINE CONSTRUCTION.

* For symbol Δ read, "some very small change of I_a".

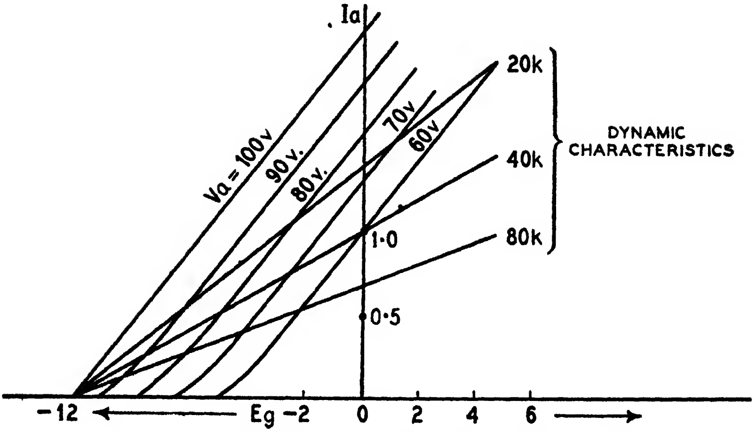


Fig. 11 —DYNAMIC CHARACTERISTICS.

is the working point for the circuit conditions specified above.

R_L is constant for all values of μE_g and μ is also constant as long as the valve is operated on the linear part of its $\frac{E_a}{I_a}$ characteristic; therefore another line drawn from the point corresponding to E on the E_a axis ($E_a = E$ $I_a = 0$), through the working point towards the I_a axis ($E_a = 0$, I_a = maximum) determines the values of E_a and I_a for the given circuit conditions at each point of intersection with the remaining curves of the family.

Dynamic Characteristics

Analysis of the load line figure provides the information shown in Fig. 11. The data thus obtained indicates the dynamic performance of this particular valve for the values of H.T., grid bias and anode load (R_L) employed to construct the load line diagram.

From consideration of Fig. 12, it will be seen that whilst E_g and I_a are in phase, E_a is in anti-phase relative to E_g and I_a, so that ($E_a + I_aR_a = E$) is a result strictly in accordance with Kirchoff's law.

Development of Grid Bias

When grid bias is required it may be applied in one of three different ways :—

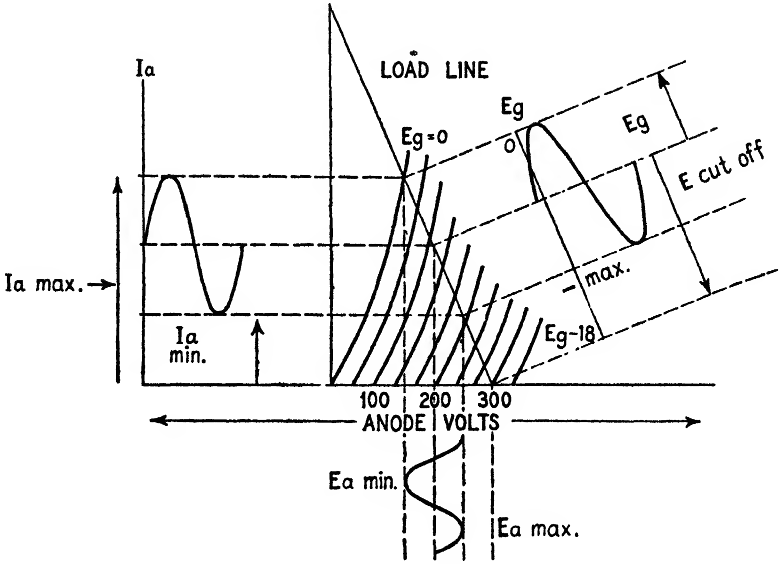


Fig. 12.—LOADING DIAGRAM.

(1) By a fixed voltage supply.
(2) By the IR drop across a resistance in the cathode.
(3) By coupling condenser and grid leak.

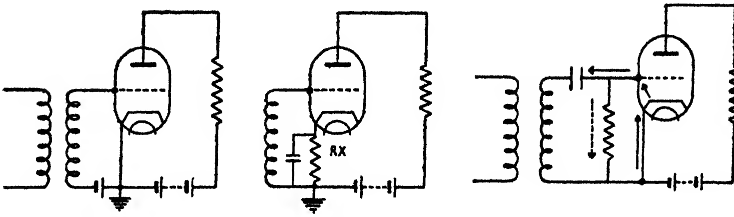
In (2), bias is produced only when current flows through the valve. The IR drop across the resistance in series with the cathode makes the cathode positive to earth. The grid being at earth potential must, therefore, be negative to the cathode when current flows by an amount = I × R.

Since with this arrangement bias is only applied when current flows through the valve, the valve cannot be held at cut off.

The condenser across the biasing resistance is made large enough to have negligible reactance at the input frequency, wherefore it shunts the cathode-dropping resistance for A.C. In consequence, the drop across R_x is due to the D.C. component of the current through the valve and the bias is steady.

There is an additional reason for shunting the cathode-dropping resistance with a large condenser. This is done in order to avoid negative feedback, when it is not required.

The combination of a grid condenser and grid leak can be used to produce a bias when the grid is being driven positive during each half cycle of the input.



(a) (b) (c)

Fig. 13 —ALTERNATIVE METHODS FOR OBTAINING NEGATIVE GRID BIAS
Charging path represented by full arrows, discharging path by dotted arrow
(a) Fixed bias (b) Cathode bias. (c) Grid leak bias.

When the grid is driven positive, grid current flows, therefore the ratio $\frac{E_g}{I_g}$, which was previously infinity, may now drop to the region of 1,000 ohms or so.

The R.C. time constant formed by the condenser and this low-resistance value $\frac{E_g}{I_g}$, as compared with the R.C. time constant of the condenser and grid leak, is very much less, consequently the condenser will charge very quickly when E_g is positive, but when negative it may discharge but very little through the much higher grid leak path. This is more particularly the case since the time constant of the grid leak path as will generally be the case, is comparable to the time period of the input signal. In this manner a bias is developed when the grid goes negative after a positive swing. Because of this residual bias the grid is not permitted to follow the positive swing of succeeding cycles.

Distortion of Input and Output

In the following circumstances the output wave will not reproduce the shape of the input wave :—

(*a*) If the valve is overdriven, *i.e.*, if a voltage wave is applied to the grid such that I_a approaches too closely to its maximum and minimum values, either or both. In either or both of these circumstances the valve will be forced to operate beyond the top and/or bottom bends of its $\frac{E_a}{I_a}$' characteristic curve. If E_g is sufficiently large, I_a may be driven into saturation by the positive grid swing, or to cut off by the negative swing.

(*b*) If the valve is biased either too high, too low or not at all, the effect on the output for moderate values of E_g may be similar to that produced by excessive values of E_g.

In effect, (*a*) and (*b*) may produce similar effects since E_g relative to earth $= G_B \pm$ input voltage.

Inter-electrode Capacity

Capacity between electrodes produces effects which in some circumstances are very useful and in others highly undesirable; for instance, in oscillation generators of high frequencies inter-electrode capacity may be used to couple the anode circuit to the grid circuit. On the other hand, this coupling effect is objectionable when sustained oscillations are not wanted.

At very high frequencies, about the order of 100 megacycles or so, *ordinary* triodes become very inefficient as R.F. amplifiers. At these frequencies, the inter-electrode capacities have small reactance and in consequence there is an important shunting effect.

Let C_{ag} = capacity between anode and grid
C_{ak} = ,, ,, ,, ,, cathode
C_{gk} = ,, ,, grid and cathode.

Of these, C_{ag} is probably the most troublesome in high-frequency amplification. As mentioned above, it can feed back some of the energy of the anode circuit in phase with the grid voltage and so cause sustained oscillations. This effect has been counteracted by using an external condenser to feed back a voltage in anti-phase, thereby neutralising the effect of C_{ag}.

It can be shown that the inter-electrode capacity of C_{ag} apparently decreases and that of C_{gk} apparently increases when the control grid potential increases in the positive direction. This has important effects on the input (grid-cathode circuit), particularly at high frequencies.

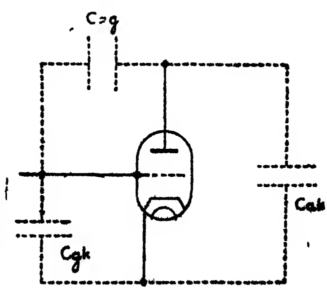


Fig. 14.—SCHEMATIC OF INTER-ELECTRODE CAPACITY.

When E_g is positive, *i.e.*, when electrons flow past the grid, to the anode, the effective capacity C_{ag} decreases. The current I_a at any point between grid and anode is proportional to the density and velocity of the moving electrons.

Due to the capacity C_{ag}, the total current across the grid-cathode space is the sum of the current represented by the electron stream and the current due to the action of C_{ag}.

Under the influence of the anode, electrons approaching it are accelerated, consequently the density of the electron stream is less near the anode than near the grid.

It appears, therefore, that the current near the anode is smaller than near the grid. By Kirchoff's first law, however, this is impossible. Therefore the capacity current must change so that it is less near the grid. But the anode volts are constant, therefore it can only mean that the capacity C_{ag} changes its effective value, decreasing with alternative half cycles of E_g, that is as each pulse of electron current passes the grid.

This variation of anode-grid capacity with change in polarity of E_g is such that energy from the anode circuit is fed back in phase to the grid circuit and thus produces instability. It is, in fact, the channel through which phase change, due to the nature of the anode load, reacts and modifies the grid input impedance.

When the grid is at a positive potential and C_{ag} decreases, the boundaries of the space charge approach nearer to the grid, consequently if the boundary of the space charge is considered to be an extension of the cathode surface in the direction of the grid, the grid-cathode capacity increases when the anode to grid capacity C_{ag} decreases, thus increasing the shunting effect on the input circuit (grid-cathode circuit) of the valve. This shunting effect increases with frequency.

Since C_{ag} and C_{gk} are effectively in series, any reactive effects due to the anode load are communicated *viâ* this channel to the grid-cathode circuit, thereby introducing a phase distortion with modification to the shape of the output wave.

When the anode load is of such a character that the above conditions would apply and when it is desired to preserve, at the output, the shape of the input wave as closely as possible electron coupling is used and a multi-grid valve is employed.

The inter-electrode capacity of the *ordinary* triode imposes limitations on its use for R.F. amplification at high frequencies. Another grid is therefore inserted between control grid and anode in order to shield the control grid from the electric field of the anode.

Screen Grid and Pentode Valves

This does not, however, take care of secondary emission from the anode which produces the negative resistance effect which occurs in the lower part of the screen grid valve characteristic ($I_a - E_a$) curve. In order to suppress this undesirable secondary emission and generally to improve the voltage amplifying characteristics of the valve, a fifth electrode called the " suppressor," is added.

Suppressor Grid

The suppressor grid performs its function either by repelling electrons of secondary emission back to the anode, or altogether prevents emission taking place. The result is that the anode current rises smoothly and the characteristic negative kink in the $\frac{E_a}{I_a}$ curve of the screen grid valve is avoided. The distribution of the space current in a pentode is dealt with more fully in a later paragraph.

The pentode may be used as an R.F. amplifier, an L.F. amplifier, or to give increased power output, but at ultra-high frequencies it again gives place to *special* triodes, since long internal leads tend to produce undesirable negative feedback.

Beam-Power Valves

A beam-power valve is a tetrode having its electrodes shaped in such a manner that electrons of secondary emission are concentrated in the space near the anode. As long as the anode potential does not fall below the low potential created in this space, electrons of secondary emission are drawn back to the anode. The advantage of this type of valve is increased power output efficiency.

Multi-grid Valves

The possibilities of suppressor grid modulation has led to the introduction of multi-grid valves. As the suppressor is made more negative the anode current is reduced, so also is the anode impedance. Multi-grid valves then are designed for control of the anode current by two separate grids. The control grid and screen grid function in the normal manner, but the space current on its way to the anode must pass the intervening electrodes. The intervening electrodes, which are grids, may form the input grid and anode of a separate oscillator, in which case the anode current is modulated by the action of these elements.

Alternatively, the additional grids may themselves be used to generate oscillations, in which case mixing or modulation is accomplished by electron coupling.

In order to reduce the number of separate valve units in a given assembly when space is limited, the electrodes of two valves are frequently enclosed in one glass envelope. They operate, however, as two separate valves.

Frequency Limitations of Standard Valves

The employment of standard valves at ultra-high frequencies is limited by inter-electrode capacity, inductance of internal leads and transit time.

At frequencies higher than about 100 megacycles, the inter-electrode capacity of standard valves is sufficient to by-pass the R.F., consequently *special* types must be used. The transit time for an electron in a standard valve is about one thousandth of a microsecond. This time is about one-tenth of the time for one cycle at 100 megacycles, and this can be tolerated, but at 400 megacycles it is one-quarter of the time period for one cycle and the reduction in efficiency is consequently very great. This decrease in efficiency is due to two causes :—

(*a*) Phase shift between grid voltage and anode current.
(*b*) Decrease of effective resistance between grid and cathode.

Transit Time and Phase Shift

When transit time is negligible, the total number of electrons approaching the grid is equal to the total number receding from it towards the anode. Therefore, their combined effect on the grid is zero.

When transit time is important, *i.e.*, at frequencies greater than 100 megacycles for standard valves, a current starting from the vicinity of the cathode flowing towards the anode is in phase with the grid volts and 180° out of phase with the anode volts at that instant of time.

(*a*) Since phase is a function of time, and because some time does elapse in transit, the electron current arriving at the anode must lag by some small amount due to change in relative I_aE_g phase positions, with time.

(*b*) Thus the current at the anode must lag the current leaving the cathode. This shift in phase, relative to the grid voltage, produces a loss of power and the grid dissipation increases.

(*c*) As the electrons forming the anode current move past the grid they induce grid currents. Consequently, when transit time is negligible, the number of electrons approaching the grid is considered to be equal to the number receding from it towards the anode.

(*d*) Relative to the grid, these are currents in opposite directions and therefore equal and opposite charges are induced in the grid and the resultant is zero grid current.

(*e*) When the transit time is an appreciable part of a cycle, the approaching and receding currents are not equal and (see pars. (*a*) and (*b*)), therefore, a grid current must flow.

The effect is, therefore, similar to a sudden decrease in resistance between cathode and grid. The effect increases with frequency, consequently when the frequency reaches a sufficiently high value, the effect will be similar to a short circuit from grid to cathode. Similarly, the inductance reactance of internal leads which may be small at low frequencies increases with frequency and may also become a limiting factor.

Development of Special Triodes for V.H.F. and U.H.F.

As a result of the limitations imposed by the above, special valves have been developed (triodes), such, for example, as the "Acorn," and other types.

The basis for this development is that if all the physical measurements of a standard valve are reduced by the same scale, the constants of the smaller special valve will be similar to those of the larger one, but the inter-electrode capacities will have been reduced as well as the transit time. A further improvement is effected by omitting bases and by bringing out the internal connections directly through the envelope by the shortest possible lead.

The upper frequency limit for special triodes of this type is about 600 M/cs. Valves for use in receivers as local oscillators for frequencies above 600 M/cs. include the lighthouse and klystron types. The first of these is a special kind of triode developed to function in grounded grid circuits and the second depends for its action on velocity modulation (see Chapter XVI).

Since the successful operation of these valves at ultra-high frequencies is closely associated with circuit conditions and ultra-high frequency technique, it is desirable to postpone to a later chapter any detailed consideration of the principles involved in their construction and operation. Similar remarks apply to the magnetron, the valve almost universally used in U.H.F. transmitters.

Gas-filled Valves

Valves into which small quantities of nitrogen, argon, neon, mercury vapours or the like have been introduced are very useful for certain purposes. The performance of gas-filled or "soft" valves differs considerably from the "hard" or high-vacuum varieties.

Ionisation

The mechanism of space current conduction is dependent upon ionisation of the gas molecules by free electrons. Free electrons emitted from a cathode and accelerated by an anode voltage, liberate by collisions with gas molecules, other electrons and these, in turn, perform a similar function. The action is therefore cumulative and results in a large number of free electrons and currents of considerable magnitude.

Firing Point

Before any conduction takes place the process of ionisation must commence. The voltage at which this occurs is called the "firing point."

When ionisation commences and conduction takes place, the current is maintained and the internal resistance falls to a very low value as long as the H.T. does not fall below the potential at which the gas de-ionises.

The de-ionising voltage is less than the original voltage required to start ionisation and conduction.

Neon Tubes

The neon tube can be ionised by an R.F. field and will glow when current passes across the gas-filled space as long as the R.F. field is maintained.

Conductance in a thyratron takes place in one direction only, viz., from the cathode to the anode, the anode being positive to the cathode. The electrodes in a neon tube may, however, interchange their functions according to which element is at the highest potential. This is possible, because no heater, such as is necessary to other types, is required to provide emission as a necessary prelude to ionisation. Ionisation in the case of the neon tube is performed by the R.F. electrostatic field established between cathode and anode. Neon tubes may therefore be used to indicate by glow the presence of an R.F. field, or when suitably connected, they may be employed as voltage regulators (volts remaining constant over a wide range of circuit values), oscillators or detectors.

Mercury Vapour Valves

Mercury vapour tubes are employed largely as rectifiers when comparatively large rectified currents are required. A small quantity of mercury at the cathode is vaporised by heat, ionised by electron emission, and heavy unilateral conduction takes place. When this type of rectifier is employed, a suitable time delay must take place before the H.T. is switched on, in order to allow time for the mercury fully to vaporise.

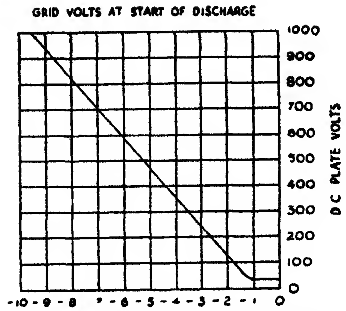


Fig. 15.—GRID CONTROL CHARACTERISTICS OF TYPICAL THYRATRON.

The Thyratron

Another useful form of gas-filled valve is the thyratron. In this case a control grid is introduced into the gas-filled envelope, between the cathode and anode. The function of this electrode is to control the firing potential.

The operation of firing may be said to have been completed when a small initial current starts to flow immediately after ionisation has occurred.

The potential of the grid relative to anode volts will therefore control the start of this initial current ; consequently, for any fixed anode voltage the firing point may be delayed by making the grid more negative or put forward by a positive grid voltage. Once the firing operation has been completed, however, the grid loses all control of the anode current. The anode current, after starting, can only be stopped by a reduction of the anode volts below the de-ionising value.

The thyratron has many applications, because it can be used to control large values of power. It has a very low resistance when conducting, and the firing point can be controlled by a suitable pulse applied to the grid. Thyratrons are often used in radar as a low-resistance pulse-controlled discharge path for a condenser in circuits where a condenser must be discharged, periodically, under control of a synchronising signal.

CHAPTER IV

IMPORTANT FACTS ABOUT AMPLIFIERS

AMPLIFIERS may be classified under three main headings :—

(*a*) Function.
(*b*) Bias.
(*c*) Response.

Under classification (*a*) amplifiers are either voltage amplifiers or power amplifiers.

When the main object is a high ratio of $\frac{\text{A.C. Output volts (anode)}}{\text{A.C. Input volts (grid)}}$ without regard to the power which can be delivered, the function of the amplifier is Voltage Amplification.

When the main requirement is maximum $\frac{\text{Output power}}{\text{Input power}}$ without regard to the voltage amplification obtained, the function of the amplifier is Power Amplification.

Voltage amplification for high gain requires a high resistance load. The load for power amplification, however, depends upon the degree of distortion in the output that can be tolerated. When an undistorted output is required the impedance of the load must be adjusted so that the permissible degree of distortion specified is not exceeded. When this is not a primary consideration, however, the load may be adjusted for maximum anode efficiency $\left(\frac{\text{output power}}{\text{D.C. input power to anode}}\right)$. In general anode efficiency is low when minimum distortion is specified.

Amplifiers are also classified according to their operating conditions. This grouping is determined by the portion of the cycle during which anode current flows under the control of grid bias.

Class A

Grid bias and A.C. voltages are adjusted so that anode current flows throughout the E_g cycle. Operating under these conditions, the output has minimum distortion, low power and a high amplification ratio. The anode efficiency is relatively low.

Class B

The valve is biased back to almost cut off, and I_a is approximately $= 0$ when $E_g = 0$. In consequence, I_a flows only during the positive half-cycle of E_g. Power output, anode efficiency and power amplification may all be classified as medium.

The alternating component of I_a is proportional to the amplitude of E_g. Therefore power output is proportional to $(E_g)^2$.

Class AB

Grid bias and E_g are such that I_a flows for more than half of each cycle but less than a full cycle. The output has low distortion at medium E_g input levels and medium efficiency at high levels.

Class C

Grid bias is larger than the cut off value and $I_a = 0$ when $E_g = 0$. Consequently I_a flows for less than a half of each cycle of E_g input. The alternating component of I_a is directly proportional to E_a and the output power is therefore proportional to $(E_a)^2$. Output is at high anode efficiency, with high power output, low power amplification ratio.

Classification under frequency response depends upon the frequency range over which the amplifier is designed to operate, *i.e.* :—

Radio frequency (R.F.).
Intermediate frequency (I.F.).
Video frequency pulse amplifiers.
Audio frequency (A.F.).

R.F. and I.F. amplifiers are generally designed for tuned circuit coupling.

Video frequency amplifiers are nearly always R.C. coupled. Their gain must be constant over a very wide range of frequencies.

A.F. amplifiers may be transformer coupled, impedance coupled, or resistance coupled.

Three different types of distortion may be distinguished; these are :—

Frequency distortion,
Phase distortion,
Amplitude distortion.

In radar it is sometimes essential that the shape of the amplified output wave shall resemble, as nearly as possible, the shape of the input wave. In other cases it is necessary that the shape of the output wave should be widely different from that of the input wave. In other words, a certain amount of predetermined distortion has to be deliberately introduced. It is, therefore, important that the factors causing each type of distortion should be clearly distinguished.

Frequency Distortion

Frequency distortion occurs when some frequency components of a signal are amplified more than others. Fig. 16 (a)

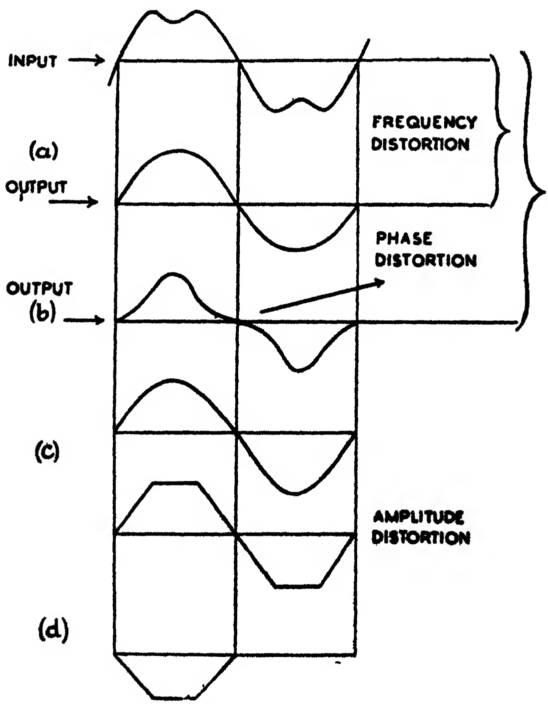


Fig. 16.—Typical distorted outputs caused by frequency, phase and amplitude distortion, respectively.

illustrates distortion of a signal consisting of a fundamental frequency and a third harmonic.

Comparison of input and output shows that only the fundamental frequency has been amplified. The third harmonic has been entirely lost. The conclusion is that the amplifier response was cut off at some frequency higher than the fundamental but lower than its third harmonic.

Frequency distortion may occur at low frequencies if the coupling condenser is too small because X_c is then large for the lower frequency components. At high frequencies distortion may be caused by distributed capacities.

Phase Distortion

Fig. 16 (b) shows an example of phase distortion. The input is similar to that in Fig. 16 (a). In this case the amplitudes of both component frequencies have been increased by identical ratios. There is, therefore, no frequency distortion present, yet the output wave differs considerably in shape from that of the input. This is due to the fact that the third harmonic has been shifted in phase relative to the fundamental.

Coupling circuits are usually responsible for this form of distortion.

Phase shift in the case of a sine wave input has little effect on the output but, with more complex inputs, the output may bear little resemblance to the input when the phases of component frequencies are shifted. To avoid phase shift, special coupling circuits must be employed.

Amplitude Distortion

Amplitude distortion occurs when E_g operates for some part of its cycle on the non-linear part of the E_a-I_a characteristic. In these circumstances equal changes of E_a do not take place for equal changes of E_g. Fig. 16 (c) shows the distortion that occurs when an amplifier is overdriven in both directions. Fig. 16 (d) shows the effect when overdrive takes place in both directions with " cut off " bias.

The definition of good response or " fidelity " is, therefore, that the output must contain all the frequencies that appear in the input but without any additional ones. Also all these frequencies must bear the same phase and amplitude ratio to the fundamental in the output, as they did in the input.

In order to satisfy these conditions, it is necessary to examine

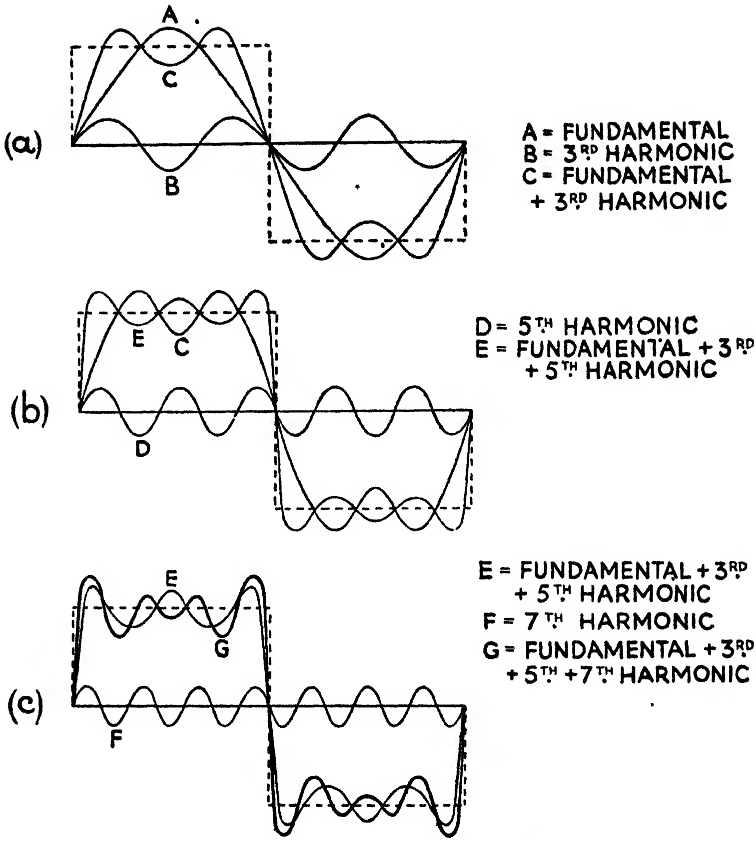


Fig. 17.—Composition of square wave.

the inputs to ascertain the maximum widths of the bands of frequencies which must be passed without distortion.

The square wave is most exacting in its demand in this respect, Fig. 17 (a) shows how a square wave can be built up by adding to the fundamental frequency an infinite number of odd harmonics. In Fig. 17, (b) and (c), the fundamental sine wave of Fig. 17 (a) is approaching nearer to a rectangular shape ; the addition of each harmonic causes the sides of the wave to become more vertical and the tops to become flatter.

It is clearly impossible to add odd harmonics to infinity, consequently the production of a perfectly square wave by a composition of sine waves is not possible. Similarly, it is impossible to design and construct a circuit that will pass a perfectly square wave without some distortion, but it is practical to design circuits to pass all frequencies from a few cycles to several megacycles with an amount of distortion so small that it can be tolerated.

The frequency limits between which a square wave or pulse amplifier must operate, in order to give an output which may be regarded as tolerably distortionless, are from a few cycles up to 3 or 4 megacycles. This is approximately the video band width, consequently the technique of such amplifiers must be closely allied to that employed to cover a similar band width in television.

When analysing an amplifier circuit it is convenient to think of it as a network corresponding to the load and coupling conditions when connected across the output terminals of an alternator.

The variations produced in the anode circuit of a valve by the application of a voltage to the grid are exactly the same as the variations that would be produced in the anode circuit

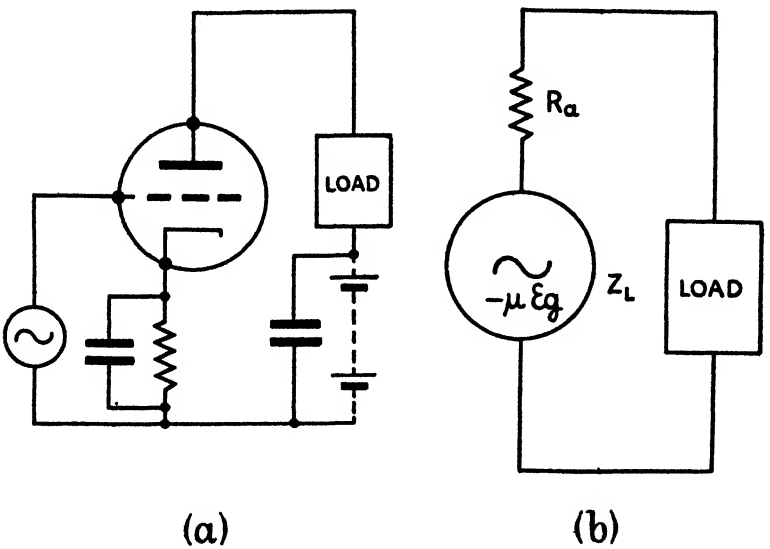


Fig. 18.—Simplification of valve circuits for analysis.

(a) Circuit to be analysed.
(b) Simplified equivalent.

by a generator developing a voltage of μE_g (grid voltage signal $\times$ amplification factor) acting inside the valve in a circuit consisting of the valve anode resistance in series with the load impedance. In other words, a generator developing a voltage $-\mu E_g$ (the minus sign denotes anti-phase),* in series with a resistance representing R_a, can be placed diagrammatically across a load of equivalent arrangement and corresponding impedance, in order to simplify analysis.

Consequently the circuit in Fig. 18 (a) can be reduced to that shown in Fig. 18 (b).

Applying Kirchoff's Law to this circuit,

$$(a)\ I_a = \frac{-\mu E_g}{R_a + Z_L}.$$

$$(b)\ \text{The output voltage} = I_a Z_L = \frac{-\mu E_g Z_L}{R_a + Z_L}.$$

It is sometimes convenient to think of R_a and Z_L as a voltage divider across which the voltage generated in the valve is distributed. This is reflected by equation (b).

Note. Total anode current is given by the sum of the A.C. currents determined by (*a*) and the D.C. standing current flowing through the valve when $E_g = 0$.

Resistance Capacity Coupling

Resistance capacity coupling is very widely used in radar circuits. The characteristics are :—

(*a*) Good response over comparatively wide range.
(*b*) Freedom from pick-up interference.

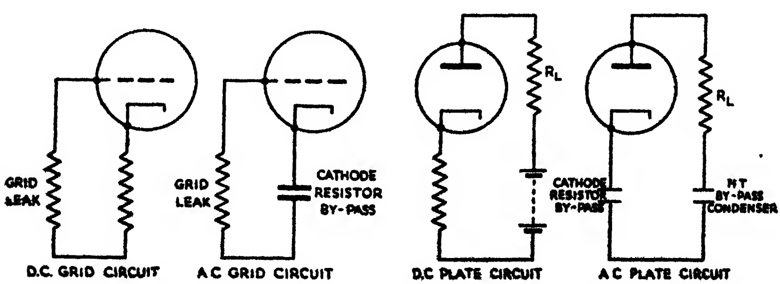


Fig. 19.—ANALYSIS OF D.C. AND A.C. CIRCUITS OF TRIODE.

* The imaginary generator $-\mu E_g$ must be shown in anti-phase to μE_g because E_a is in anti-phase to E_g.

(c) Suitability as a load for pentodes and high μ triodes.
(d) Low cost.

Analysis of D.C. and A.C. Circuits of a Triode

The D.C. and A.C. circuits of a triode are analysed in Fig. 19.

A limit is imposed upon the value of the load resistance for any given value of H.T. voltages. This limit is determined by the minimum permissible value for E_a when I_a flows. This value is $(E - I_aR_L)$, where $E =$ H.T. volts on open circuit.

Anode Circuit Voltage Distribution

Fig. 20 shows the manner in which the H.T. volts = E are distributed around a simple single stage circuit.

In Fig. 19 the condensers across the cathode resistor by-pass the A.C. component of the anode current and prevent negative feedback which would occur if this resistance was not by-passed. In order to perform their function, these condensers must present the lowest possible reactance to the lowest frequency to be by-passed. This means that the capacity of this condenser is generally large.

The H.T. is by-passed by a condenser in order to shunt the A.C. round the H.T. supply. It is usually referred to as a decoupling condenser.

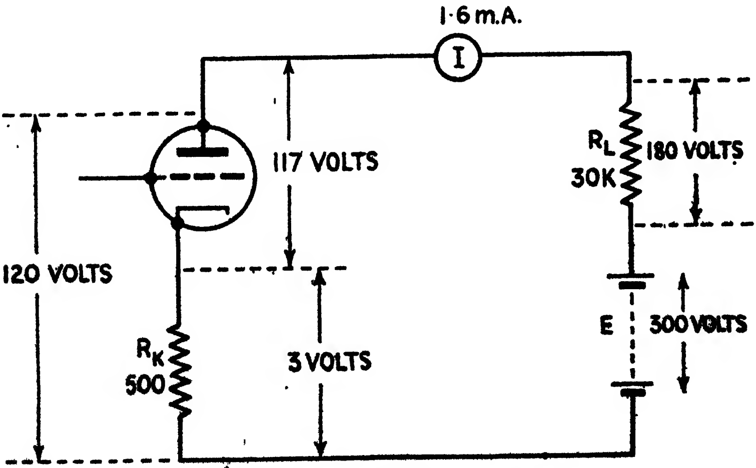


Fig. 20.—ANODE CIRCUIT VOLTAGE DISTRIBUTION.

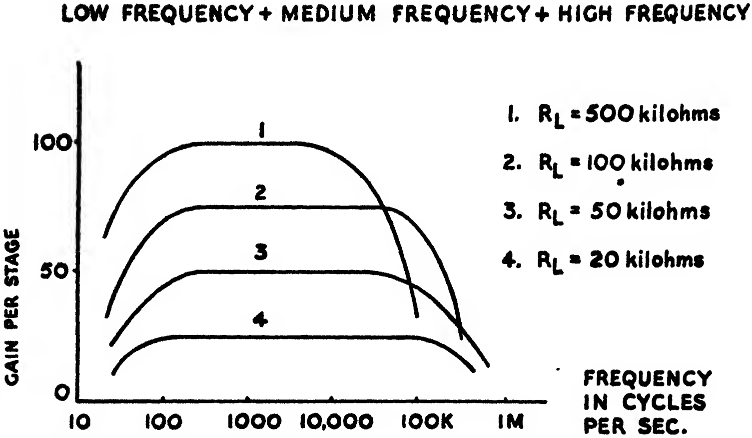


Fig. 21.—FREQUENCY RESPONSE CURVES FOR VARIOUS VALUES OF R IN AN R.C. COUPLED CIRCUIT.

Analysis of Response Curves

Analysis of typical response curves for R.C. coupling is shown in Fig. 21. For frequency response curves 3 and 4, the amplification is poor. Curve 1, the high frequency response, falls off because of capacity shunted across the anode load resistance. This capacity is made up from :—

(*a*) The output capacity C_c of the coupling condenser.
(*b*) The input capacity C_1 of the following stage.
(*c*) Distributed capacity of the coupling network.

The falling off in the response at low frequencies is due to loss of voltage across the coupling condenser, with consequent reduced gain.

Note. At high frequencies the coupling condenser reaction X_c is so small as to be almost negligible.

These factors are demonstrated by consideration of the actual and equivalent circuits of Fig. 22.

The equivalent circuits for low, medium and high response shown in Fig. 23, (a), (b) and (c) respectively, include only those components which have a direct bearing on the response for the particular range of frequencies undergoing analysis.

The values of C_o, C_d and C_i are very small, therefore, at low frequencies their reactances are such that the resulting shunt is a very high impedance and may be neglected.

The IR drop across R_L is applied across R_g *viâ* C_c to the input

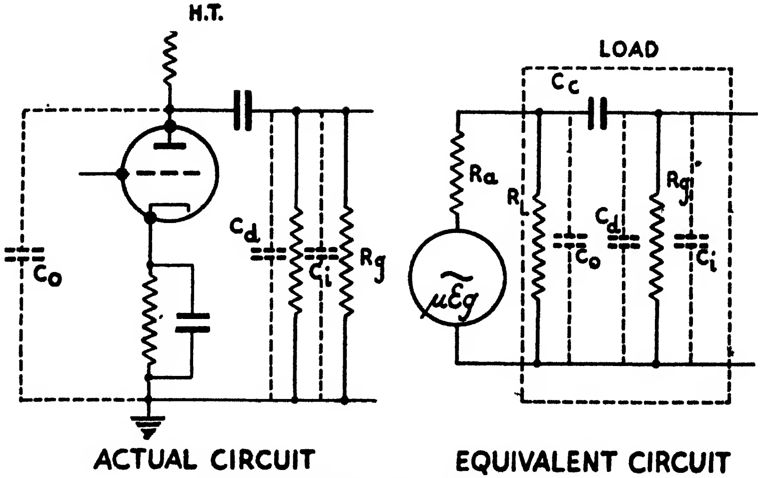


Fig. 22.—STAGE OF R.C. AMPLIFIER.

R_a = A.C. resistance.
R_L = load resistance.
C_o = inter-electrode capacity.
C_d = distributed capacity.
R_g = grid leak next stage.
C_i = capacity input to next stage, *i.e.*, C_{gk} V_2.
C_c = coupling condenser.

of the next stage. Since the reactance of X_c (coupling) increases as the frequency becomes lower, the voltage distribution across C_c and R_g becomes larger across C_c and less across R_g at the lower frequencies. The input to the following stage therefore decreases at the low frequencies unless C_c is made sufficiently large to offer negligible reactance.

At the medium frequencies the coupling condenser does offer negligible reaction and can be neglected as a response-determining factor. The equivalent circuit is therefore as shown in Fig. 23 (b). In the medium frequency range, gain is practically constant throughout.

At high frequencies the inter-electrode capacities and the distributed capacity of the network become important. At low and middle frequencies X_{cs} (reactance inter-electrode and distributed capacities) is very high compared with R_L, but as frequency increases X_{cs} falls in value and becomes comparable to R_L. This introduces a serious phase displacement angle, as a consequence of which the effective voltage across R_g becomes less.

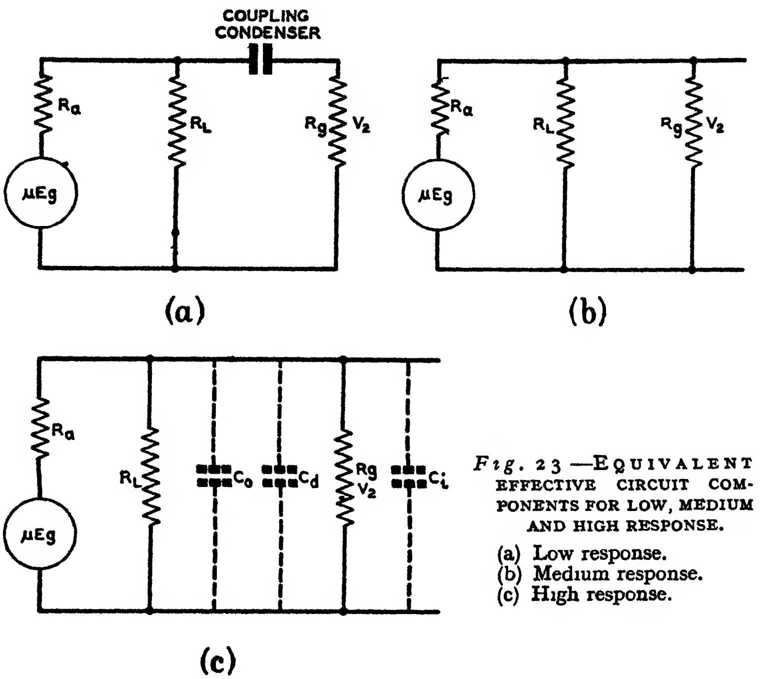


Fig. 23—EQUIVALENT EFFECTIVE CIRCUIT COMPONENTS FOR LOW, MEDIUM AND HIGH RESPONSE.

(a) Low response.
(b) Medium response.
(c) High response.

Modifications for H.F. and A.F. Amplification

Thus when it is desired to amplify high frequencies, valves with low inter-electrode capacity must be used in order to keep the shunt impedance high as compared with R_L, and R_L is also reduced so that the impedance at high frequencies offered by R_L is appreciably less than the impedance of the capacity shunt.

Reducing R_L has the effect of reducing the overall gain for all frequencies, but this is the price that must be paid in order to secure good response at high frequencies.

R.C. coupled amplifiers give a fairly uniform amplification for frequencies in the audio-frequency range of from 100 to 20,000 cycles. By making certain modifications and changes in the values of the coupling condenser and load resistance, the frequency range can be extended to cover the video-frequency band. This, however, can only be accomplished with a loss of gain.

Inductance-Capacity Coupling

Impedance or inductance-capacity coupling is obtained by substituting an inductance for R_L as the anode load.

For the high frequencies the inductance constitutes a high impedance load, but its D.C. resistance is low; the IR drop is less and the valve can operate at a higher anode potential than for resistance-capacity coupling. Consequently, at the higher frequencies, the amplification obtained with inductance-capacity coupling will be greater than with resistance-capacity coupling. On the other hand, amplification is no longer uniform, since X_L varies with frequency and amplification must also vary with frequency.

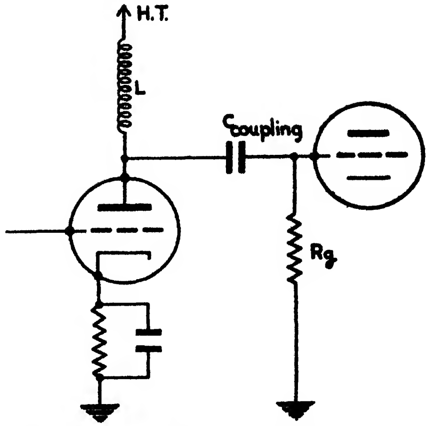


Fig. 24.—INDUCTANCE-CAPACITY COUPLING.

Amplification in the upper ranges of the frequency for inductance-capacity coupling is limited by shunt capacity. The distributed capacity is increased by the capacity between adjacent turns of the inductance. There are also resonances formed by the anode inductance and its distributed capacity to contend with.

Impedance-coupled amplifiers can be designed to give a reasonably satisfactory response over a short frequency range. Within this limit the degree of amplification obtained is higher than for R.C. coupling.

Transformer-Coupled Amplifier

Approximately equivalent circuits for medium-, low- or high-frequency response of the transformer-coupled circuit of Fig. 25 are shown in Figs. 26 (a), (b) and (c). Complete equivalent circuits are complex but the simplified versions given embody the major factors influencing response, and serve as a basis for a very general analysis.

The voltage output from the first stage across the transformer primary is approximately μE_g when the reactance of the transformer is high enough to load the anode effectively.

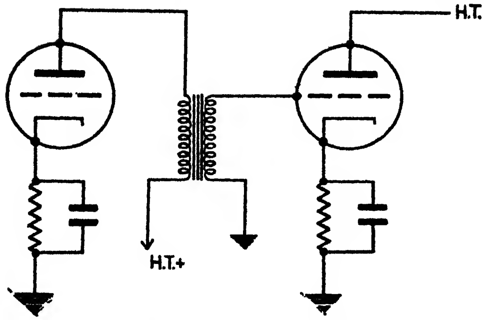


Fig. 25.—TRANSFORMER-COUPLED AMPLIFIER.

Neglecting magnetic losses and other leakages, the mid-frequency voltage applied to the grid of the second stage, is approximately $N\mu E_g$ where N is the ratio of turns, secondary/primary of the transformer.

At low frequencies, the inductive reaction of the primary falls off, the loading is reduced and the voltage across the primary is also reduced.

At high frequencies the shunt capacities, in particular that due to the capacity between the transformer winding turns, is large, and the effective voltage across the primary is accordingly reduced.

In Fig. 26 (a), R_a is the A.C. resistance of the valve and the resistance of the primary of the transformer. At medium frequencies the value of X_L for the primary is very high, so is the value X_0 for the shunt capacities. These values are so high that conditions approach those of an open circuit, the voltage across which is $N\mu E_g$.

At low frequencies the effect of shunt capacity can be entirely neglected, but the value of X_L for the primary inductance falls directly with the frequency. As the low-frequency region is reached this decrease in the value of X_L (primary), causes a reduction of voltage across the primary, and hence there must be a falling off in gain at the lower frequencies (see Fig. 26 (b)).

At high frequencies the value X_L becomes so great that its effect on response can be neglected. The value of X_c, however, decreases inversely with frequency, consequently X_c falls and therefore exercises an important shunting effect upon the input to the grid circuit of the next valve as the higher frequencies are approached.

The equivalent circuit is as shown in Fig. 26 (c).

R_L = resistance of the transformer secondary.

L = The equivalent value of inductance that would produce effect similar to those caused by magnetic (or iron) losses of the transformer.

C_s = The distributed capacities.

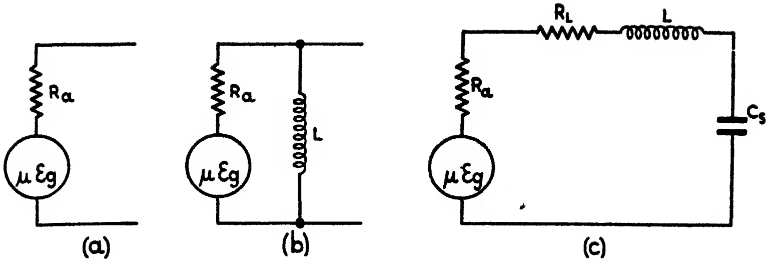


Fig. 26.—EQUIVALENT CIRCUIT OF TRANSFORMER-COUPLED AMPLIFIER FOR THREE FREQUENCY RANGES.

(a) Mid-frequency range.
(b) Low-frequency range.
(c) High-frequency range.

As might be expected, resonance effects occur despite the low Q of the circuit. Special measures must be taken to flatten out such resonance peaks and to broaden the response. Unless great care is taken in design, the gain falls off very rapidly on either side of the resonance condition. In general, it may be said that for optimum results the transformer should be designed to suit the particular valve with which it is to be used.

Even so, response curves for inter-valve transformer coupling compare unfavourably with R.C. coupling. Transformer coupling is used in push-pull circuits and is often employed for matching purposes.

The Direct-coupled Amplifier

In the direct-coupled amplifier shown in Fig. 27, no coupling device whatever is employed between the anode of one stage and the grid of the succeeding stage.

It is consequently necessary to provide a suitable voltage divider, so that the operating potentials of the anodes, grids and cathodes may be correctly adjusted.

Analysis. Commencing at the negative H.T. input terminal E_g, V_1 is connected at A to the extreme negative potential of the H.T. supply. The proper grid bias is established by the drop B—A.

The anode of V_1 is connected to a point D along the voltage divider *viâ* R_L, corresponding approximately to half the voltage drop of the H.T. supply between A and E, and the anode of V_2 is connected to the most positive terminal of the H.T. supply.

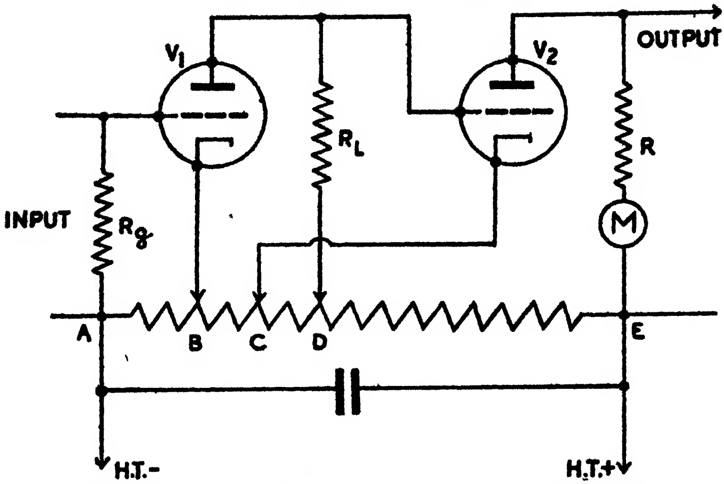


Fig. 27.—DIRECT-COUPLED AMPLIFIER OR LOFTEN-WHITE CIRCUIT.

The point along AE at which the cathode of V_2 is to be connected is determined by the bias required to produce the required operating condition of V_2.

Thus the total H.T. voltage must be approximately twice the value of that required for any single valve.

Since no coupling device is employed, the full signal variation *i.e.*, D.C. and A.C. components) of the anode of V_1 is applied to the grid of V_2.

When adjusted for class A amplification, the output is free from distortion and the response is uniform over a wide frequency range. There is no impedance variation with frequency, consequently this type of amplifier is very suitable for amplifying slow voltage variations. It is also employed whenever the *minimum* distortion introduced by coupling devices is objectionable.

Some of the uses of a D.C. amplifier are for the control of mechanical loads—such as relays, meters or counters, alternatively to control the gain of an amplifier.

The absence of a condenser or other coupling device between anode and grid alters the grid input conditions, because the D.C. component of the anode output is no longer isolated from the grid of the preliminary stage.

When V_2 is operated as a normal R.C. coupled amplifier with a negative grid bias, and by a capacity coupled input, the coupling condenser charges to an average value, and then passes to the grid of V_2, the *changes* of the average value established by the charge which it has acquired. These changes vary *about* the grid reference, which is adjusted by bias. If the grid of V_2 is driven positive it draws grid current and the coupling condenser accumulates a negative charge which reduces the load of the positive swing.

In direct coupling this is not the case, both D.C. and A.C. components of the anode output of V_1 are coupled to the grid of V_2. Consequently positive inputs tend partly to cancel a negative bias, thus permitting a larger anode current to flow. D.C. amplification has therefore been brought about. This is the principle on which a valve voltameter operates.

Feedback

In valve amplifier circuits it is possible to feed back part of the output to the input. The results obtained depend upon whether the energy from the output is fed back in phase or in anti-phase.

When energy is fed back in phase the circuit may be made to oscillate at its natural frequency. If the energy of feedback is equal to the circuit losses, oscillations can be maintained.

If the energy is fed back in anti-phase, a larger driving signal must be fed to the input of the amplifier in order to maintain the same output, but any distortion which is normally present in the output can be rectified by feed back in anti-phase (negative feedback) and a distortionless output obtained.

The process is perfectly simple and logical.

Any distortion between input and output must occur as a result of one or more distorting forces operating somewhere in or between the circuits of the amplifier. If part of the distorted output is deliberately fed back to the input in *anti-phase*, the distorting forces again operate upon it but in the *reverse direction*. The nett result at the output is therefore a corrected or distortionless output.

The price paid for this is a reduction in the overall gain, but at worst it usually means nothing more than the addition of another stage to preserve the overall gain of the amplifier.

When positive or negative feedback is introduced,

$$\text{Gain} = \frac{A}{1 - BA}$$

where A = gain of the amplifier without feedback. B = feedback.

When the value $1 - BA$ is less than 1, the gain of the amplifier is increased. This is positive feedback, *i.e.*, feedback in phase, but any distortion that may be present is also increased proportionately. This condition is sometimes distinguished by the term "regeneration." If $1 - BA$ is increased to 1, the energy fed back is sufficiently large to maintain the amplifier in a state of oscillation.

If $1 - BA$ is greater than 1, the feedback is negative, sometimes called inverse or degenerative feedback, and the gain is reduced.

Generally the amount of negative feedback is so large that the quantity $1 - BA$ may be considered equal to BA. The gain of the amplifier is then small and =

$$\text{Gain} = -\frac{1}{B}.$$

Note. Negative feedback can have no effect on amplitude distortion caused by grid current in the input stage. This occurs at the grid and is therefore amplified in the same ratio as the desired signal.

Noise generated within the amplifier may be reduced by feedback, but not if it is present at the grid input.

Again, feedback cannot affect noises due to thermal agitation, shot effect and microphonics unless they occur in the early stages.

Amplifiers can be designed within these limits to have any desired frequency response by varying the quantity B (the amount of feedback). If feedback is obtained through a resistance network, it is independent of frequency and therefore the response of such an amplifier can be improved.

The gain of an amplifier employing negative feedback can be made independent of the load. Normally an increased load is accompanied by a fall in output volts. When negative feedback direct from output to input is employed, any tendency for the output to fall is compensated for by the reduction in feedback. A reduction in feedback immediately puts up the gain. Thus reduced load is compensated for by reduced gain, and the output volts are maintained at a steady value despite any reasonable fluctuations of the load.

There are several ways of obtaining negative feedback. Some of these are illustrated in Figs. 28, 29 and 30. In Fig. 28, assuming a positive going feedback signal, I_a increases and E_a at the top of T_1 falls. This change is fed *viâ* C_1, R_1 to R_2.

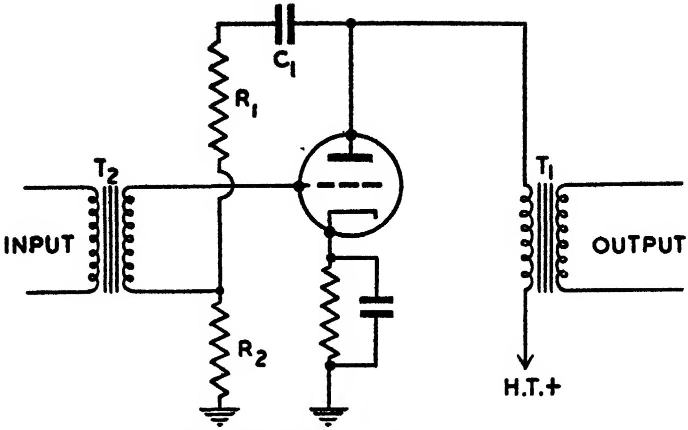


Fig. 28.—SIMPLE AMPLIFIER WITH NEGATIVE FEEDBACK.

This decreases the potential at the top of R_2, and E_g then becomes E_g signal — E_g feedback. The effective grid voltage is their difference.

In Fig. 29, negative feedback is developed across R_K by I_a; any change of voltage across R_K must appear in series with the input signal to the grid.

Negative feedback can be applied to more than one stage provided due attention is paid to phase (see Fig. 30). In this case R_1 and R_2 form a voltage divider to adjust the amount of feedback.

A positive pulse applied to the grid of V_1 causes the grid of V_2 to become more negative, therefore I_aV_2 decreases, making the top of R_K more positive, thereby increasing the bias on V_1, which is equivalent to a voltage applied in anti-phase to the incoming signal.

R.F. Amplifiers

Amplifiers in which part of the coupling is a parallel-tuned circuit are used for radio and intermediate frequency amplification.

Because of the discriminating properties of a resonant circuit, tuned amplifiers are particularly suitable for amplifying a narrow band of frequencies to the exclusion of all others.

Special modifications are necessary to R.F. and I.F. amplifiers when used for radar purposes (see Fig. 31).

The tuned circuit consists of a small coil which resonates with its distributed capacity, the capacity of the wiring and

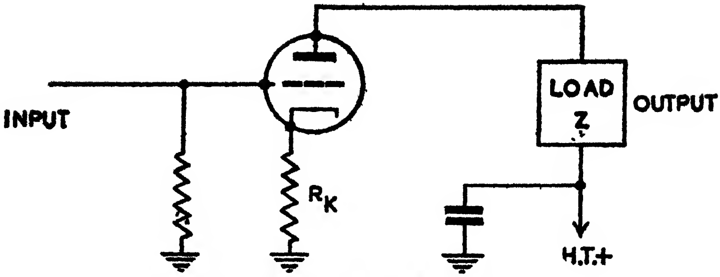


Fig. 29.—CURRENT FEEDBACK.

the inter-electrode capacity. Tuning may be performed by means of a movable slug or powdered iron core.

The valve must have low inter-electrode capacity to permit the use of an inductance of reasonable size.

A low value of anode resistance is used in order to obtain uniform response over a wider band, consequently the gain is lowered. This is offset to some extent by using a valve with high mutual conductance. Since the gain is low, a high value of H.T. is not required for optimum operation.

The screen is operated at approximately the same voltage as the anode.

Some degree of negative feedback is provided by leaving part of the cathode resistance unby-passed. This is done in order to correct distortion introduced by variations in the input capacity of the grid. The latter is brought about by the voltage excursions of E_g (see pp. 31 and 32).

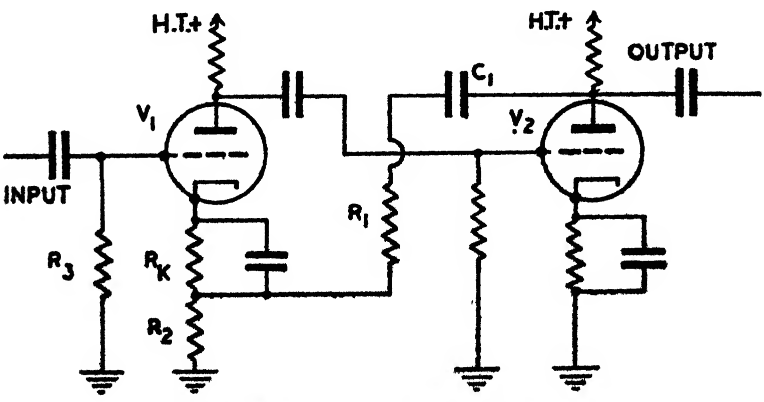


Fig. 30.—FEEDBACK (TWO STAGES).

A resistance-capacity-coupled amplifier for video frequency differs from a similarly coupled radio-frequency amplifier only in respect of component values.

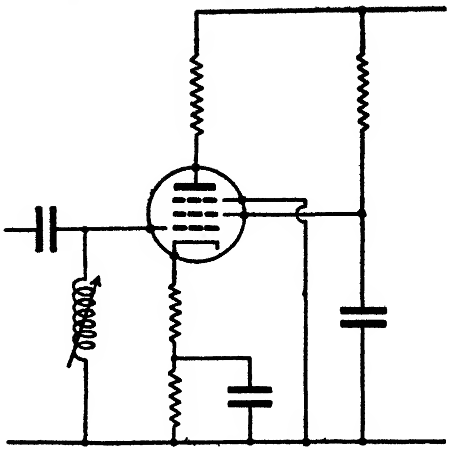


Fig. 31.—TUNED AMPLIFIER FOR RADAR FREQUENCIES.

Video Amplifiers

In the case of video amplifiers, component values suitable for maximum gain must be modified, in order to obtain a satisfactory response over the much wider range of frequencies required.

Two problems are involved in modifying an R.C. amplifier for video frequencies :—

(*a*) The high frequencies are limited by the output capacity (inter-electrode mainly) C_o, the distributed capacity C_d, and input capacity of the following stage C_i, all acting in parallel to shunt the load resistance R_L.

(*b*) The lowest frequency is limited by the time constant of the coupling condenser and input grid resistance. This time-constant must be low compared with the lowest frequency to be amplified.

In order to improve response at ultra-high frequencies, valves with low inter-electrode capacity must be used and care must be taken to keep leads short and properly spaced.

In video amplifiers for ultra-high frequencies, and where the time constant of the grid circuit is short, an inductance may be substituted in place of the usual resistance for the grid leak. In consequence charges on the coupling condenser can leak rapidly away but a high signal potential can be maintained across the grid-cathode by reason of the inductive reaction of the leak at high frequencies.

Frequency Compensation

Since the principal cause of reduced gain at high frequencies is the shunting effect on R_L by C_S,* and because an inductive

* C_s is the combined shunting effect of C_o C_d C_i.

load increases with frequency, it is possible to compensate to some extent for the shunting effect of C_S by the introduction of a small amount of inductance as part of the anode load (Fig. 32 (a)). This small inductance may also be placed in series with R_L to boost the gain at the higher frequencies. It has practically no effect upon the lower frequencies. The effectiveness of this method is limited, however, by the highest frequency and the number of stages.

Shunt compensation also can be employed. In this case a small inductance is placed in series with the coupling condenser. At the higher frequencies these form a series resonance circuit. In this condition the current through the condenser is increased. So also is the voltage across it as well as the gain.

The shunt series connection shown in Fig. 32 (b) combines the advantages of both the methods.

Low-Frequency Compensation

The increasing reactance of the coupling condenser at lower frequencies tends to attenuate the input voltage to the ensuing stage from about 200 cycles downwards. The voltage across R_g decreases as the frequency is decreased as a natural consequence.

The capacity of the coupling condenser is limited by the necessity for keeping down stray capacities since these affect the higher frequencies. The optimum capacity is about 0·1 mfd.

Compensation can, however, be provided by the addition of a low-frequency filter in series with the load resistance.

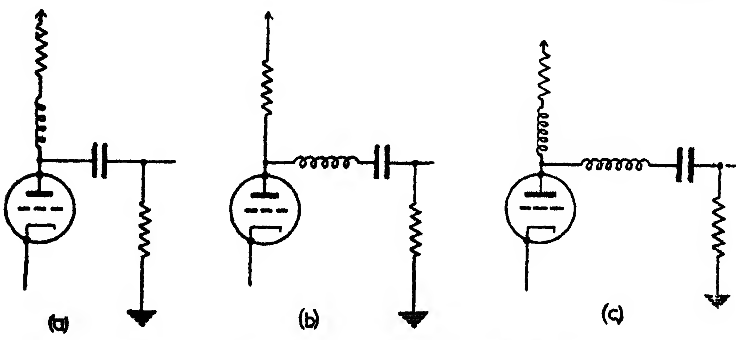


Fig. 32 (a), (b), (c).—HIGH-FREQUENCY COMPENSATION.

(a) Series compensation.
(b) Shunt compensation. All cathode resistances by-passed.
(c) Shunt-series compensation.

Fig. 32 (d) shows a suitable arrangement. A resistance shunted by a condenser to form a filter is joined in series with the load resistance.

The purpose is two-fold :—

(*a*) It introduces a phase shift in the anode circuit to compensate for the phase shift in the coupling circuit.

(*b*) Increases the anode load at low frequencies, thus maintaining the gain.

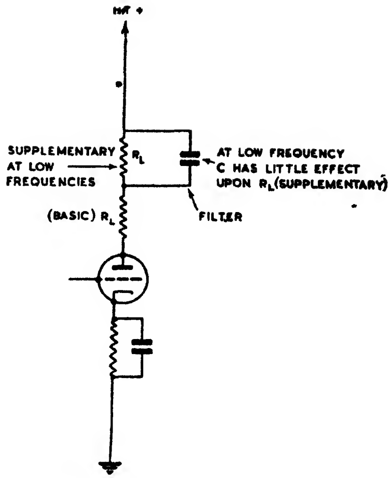


Fig. 32 (d).—LOW-FREQUENCY COMPENSATION.

At medium and high frequencies the reactance X_c will be appreciably less than the value of the compensating resistance and will therefore shunt it effectively.

At the lower frequencies the value of X_c increases above R and therefore the total anode resistance becomes $R_L + R$. In this manner the lower frequency response is extended satisfactorily for video amplification.

When a positive voltage step is applied to the grid of an amplifying valve whose cathode is tied to earth, the rate of fall of the anode volts is slowed up by the effect of stray anode earth capacity.

The voltage change ABC will produce an anode voltage change DEF (Fig. 33 (a) and (b)).

The insertion of inductance in the anode load postpones somewhat the decrease in the rate of rise but tends, if large enough to be effective, to produce new kinds of distortion of type shown in Fig. 33 (c).

Distortion produced by this method is sometimes made use of.

A better solution is the use of a resistance in the cathode load shunted by a condenser of definite capacity (Fig. 33 (d)).

The effect of this arrangement can be seen thus : suppose a voltage step V (ABC) is applied between grid and earth, the cathode would follow the grid, but is prevented partly by the condenser resisting any change in the voltage across it.

The effect on the cathode is shown in Fig. 33 (e), JL/V being the stage gain of the cathode follower, the effect of the anode load being neglected.

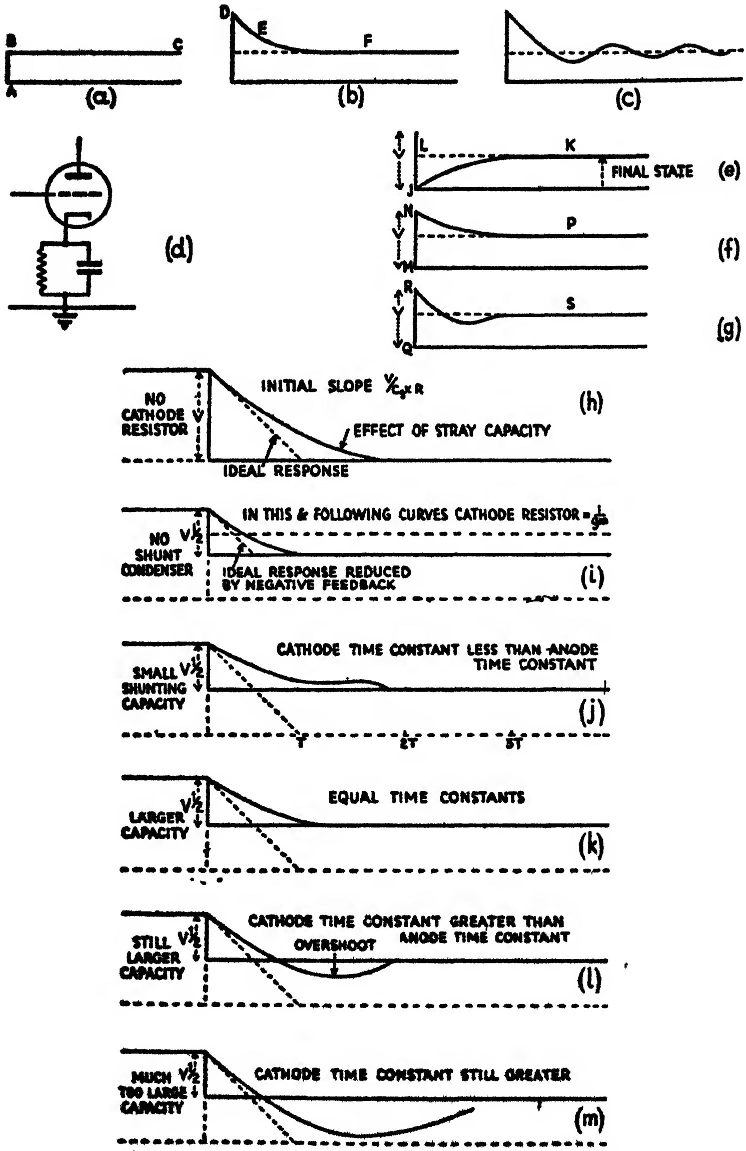


Fig. 33.—EFFECT OF STRAY ANODE-EARTH CAPACITY ON RATE OF FALL OF ANODE VOLTS AND MODIFYING EFFECTS OF CAPACITY SHUNTS ACROSS R_k.

The resulting grid cathode voltage is the difference between ABC and JK as shown at MNP (by Fig. 33 (f)).

It will be noticed that a sudden voltage is applied at first, but this decreases at a rate depending upon the values of C and R. This decrease tends to make the fall of anode volts slow down sooner and even to reverse and so sooner to read a steady value.

The result is shown at QRS (by Fig. 33 (g)) for a particular value of C and R and of the anode circuit time constant.

The modification effected by various values for K is shown in Fig. 33 (h) to (m) to larger scale.

CHAPTER V

SOME IMPORTANT FACTS ABOUT OSCILLATION GENERATORS

IN radar, oscillators are used in transmitter units to generate R.F. at the selected carrier frequency in much the same manner as they are used for a similar purpose in radio communication, with the exception that they are pulsed.

The carrier frequencies employed in radar, range from about 100 megacycles to 10,000 megacycles or more, consequently radar transmitter technique involves standard radio practice, modifications to standard radio practice and new technique.

Other Applications

Oscillators are employed in other radar units for various purposes, over a much wider frequency range—from a few cycles per second upwards.

Oscillators in the lower frequency ranges are frequently employed in pulsing and calibrating systems, particularly for the purpose of developing short, accurately timed pulses. Multivibrators, which are also extensively used, are considered by some technicians to fall within the oscillator classification. They are dealt with here in the chapter on rectangular pulse generators, but that fact does not necessarily alter their fundamental classification.

In the higher frequency ranges, oscillators are used in receiver circuits as local oscillators, in a manner similar to that in which they are normally employed in radio communication receivers. At micro-wave frequencies, however, considerable modifications in circuit elements have to be made, special valves must be employed, and new technique is introduced.

This chapter is therefore devoted to a recapitulation of the general requirements for oscillation and frequency stability, together with a description of the self-pulsing oscillator, which is a simple modification to standard R.F. generator practice.

Oscillators for special purposes, such, for example, as local oscillators for use in receivers for the micro-wave range, ringing oscillators for the production of short pulses, multivibrators,

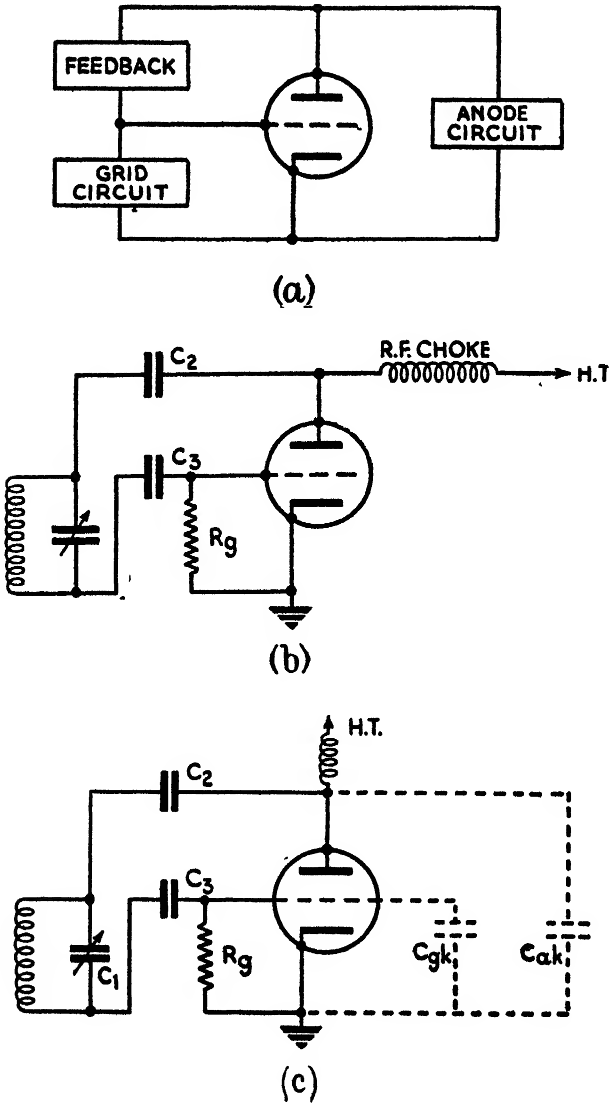


Fig. 34.—ANALYSIS OF THE OSCILLATION GENERATOR.
(a) Schematic oscillator.
(b) Ultra H.F. oscillator.
(c) Colpitts circuit.

for the production of rectangular pulses, etc., are described in later chapters devoted to their special functions in radar systems.

OSCILLATORS

General Requirements for Production of Oscillation

When a valve is connected to function as an oscillator, it becomes a device for converting the D.C. energy of the supply into A.C. energy.

The basic oscillator in Fig. 34 (a) functions because energy from the output is fed back to the grid in phase, and with magnitude sufficient to make good the circuit losses. The energy for these renewals is drawn from the D.C. H.T. supply.

This action constitutes positive feedback or regeneration in contradistinction to negative feedback or degeneration.

When energy is fed back to obtain regeneration, the effect on the circuit so supplied is similar to the effect which would be produced by the presence of a negative resistance or some internal generating element.

The normal E_g — E_a phase shift must be inverted in order to feed back in phase. This operation can be accomplished by the use of A.C. transformers—or L.C. phase shifters and other external devices. Alternatively, the action can be carried out internally by employing the inter-electrode capacities of the valve itself.

Typical Circuit Arrangements

In Fig. 34 (b) the phase inversion is brought about by connecting the anode and grid to opposite sides of a tank circuit.

The grid is automatically maintained at the proper bias for satisfactory operation by the action of C_3 and R_g. The time constant is so proportioned that after C_3 has assumed an average negative charge, small additional charges added at the peak of each oscillation leak off entirely *viâ* R_g. Some form of self bias must be used to enable the oscillator to start up quickly. A fixed negative bias of the correct operating value does not allow the device to break into oscillation immediately it is switched on.

The Colpitts oscillator shown in Fig. 34 (c) utilises the inter-electrode capacities of the valve for regenerative feedback.

The voltage drop across C_{gk} provides grid excitation.

The total capacity of the oscillating circuit is derived from C_1 in parallel with (C_2, C_{ak}, C_{gk} and C_3 all in series). C_2 and C_3

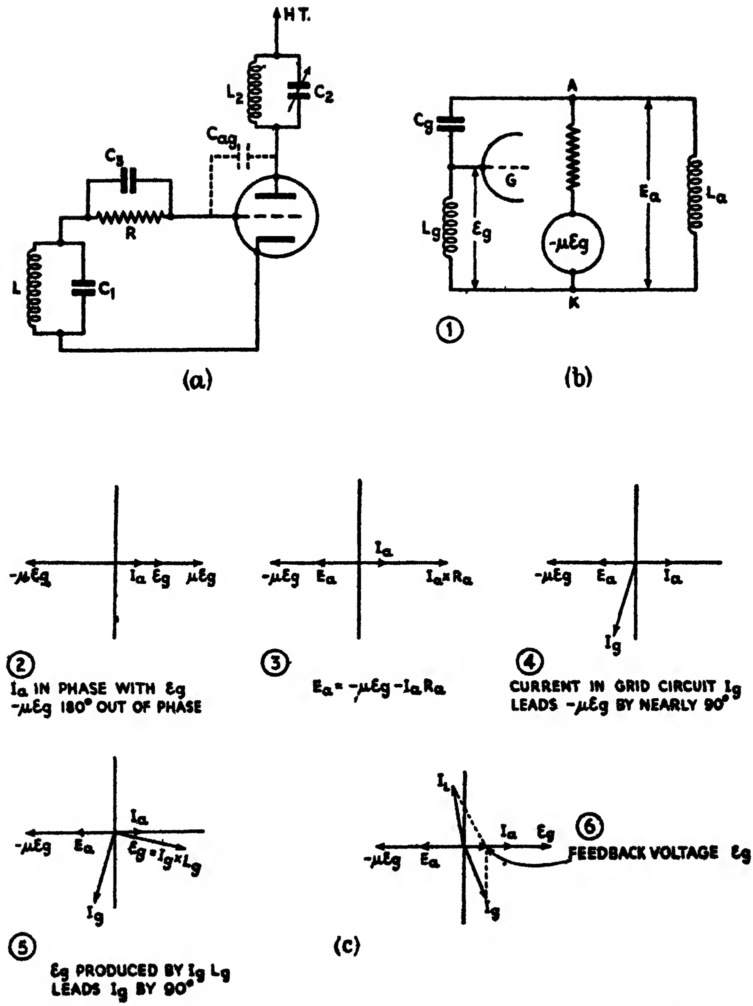


Fig. 35.—TUNED ANODE TUNED GRID CIRCUIT, EQUIVALENT CIRCUIT AND VECTOR DIAGRAMS.

(a) Tuned anode tuned grid oscillator.
(b) Equivalent circuit.
(c) Vector analysis of (a) and (b).

are large, and since they are in series, do not contribute very much to the total capacity.

Fig. 35 (a) illustrates a tuned anode tuned grid circuit.

The operating frequency is less than the natural frequency of either the grid or anode circuit. The equivalent circuit with vector diagrams is also shown in the same figure.

Analysis of Voltages and Currents

The inductance in the anode circuit is not inductively coupled to the inductance in the grid circuit. Feedback is provided *viâ* the inter-electrode capacities.

The operating frequency is lower than the natural frequencies of either the grid or anode circuits. Thus at the operating frequency both circuits offer inductive reactance to I_a.

In the equivalent circuit (Fig. 35 (b)), anode and grid circuits, since they are inductive, are shown as simple inductances, the inter-electrode capacities are represented by the distributed capacities associated with L_a and L_g.

Analysis. Let the oscillator be in oscillation, then a voltage exists at the grid which controls the anode current in such a manner that energy is delivered to the anode circuit in phase to reinforce the oscillations. This is represented by the generator $-\mu E_g$.

In vector diagram 2, I_a is 180° out of phase with $-\mu E_g$.

I_a must be in phase with $\mathcal{E}_g$ since $\mathcal{E}_g$ controls I_a and I_a has maximum value when E_a is minimum.

The vector difference $(-\mu E_g - IR_a)$ is equal to the voltage across the valve.

This voltage is applied across C_{ag} and L_g (the effective grid inductance). The operating requirement is, however, that $X_{C_{ag}}$ must be greater than X_{Lg}, therefore the current in the grid branch as a result of E_a across it tends to lead nearly 90° (see (4).) This current flowing through L_g produces $\mathcal{E}_g$; and $\mathcal{E}_g$, since it is produced by I_g, leads I_g by some angle less than 90°, *e.g.* the voltage fed back is almost in phase therefore with I_a (see (5)), so that oscillations are maintained.

A negative resistance, conveniently considered to be as a generator or reaction effect (positive or negative), is essential for any oscillator to function. If this regenerative action is examined, I_g leads the anode voltage by *more* than 90°, (see (6)), so that the voltage fed back from the anode is exactly in phase with I_a and this fixes the vector I_L.

Summary of Requirements for Oscillation Generator Circuit

The conditions for sustained oscillations in a tuned grid tuned anode circuit are :—

(1) Both grid and anode circuits must be tuned to a frequency somewhat higher than the operating frequency because both circuits must offer inductive reaction.*

(2) X_L of the grid circuit must be less than X_c of the grid-anode inter-electrode capacity at the operating frequency.

(3) A negative resistance (generating action or feedback) must be present, of magnitude sufficient to supply all losses that occur in the circuit. This, in effect, is the joint action of 1 and 2.

Crystal-controlled Oscillators

When greater stability is required than can be obtained from any uncontrolled oscillator, a crystal-controlled oscillator is used.

The frequency of oscillation is determined by the mechanical vibrations set up in a crystal of quartz. This is due to Piezo electric effects as described in standard text-books on radio.

A quartz crystal behaves as a series L.C.R. circuit with a very high Q of the order of 10,000. The crystal holder acts as a small capacity, shunting the series resistance.

A suitable quartz crystal can therefore be used as the tuned circuit of an oscillator, conferring upon the output a very high degree of frequency stability (Fig. 36 (a)).

A circuit of this form has two resonant frequencies at one

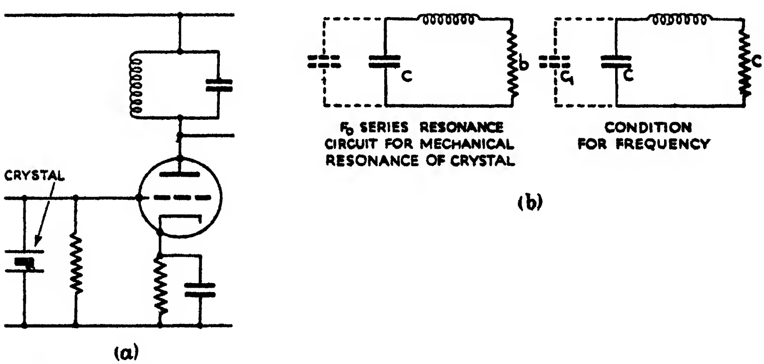


Fig. 36.—CRYSTAL-CONTROLLED OSCILLATOR AND EQUIVALENT CIRCUITS.
(a) Crystal-controlled oscillator.
(b) Equivalent circuits.

* The frequency at which the device oscillates is not the frequency to which anode and grid circuits are tuned, it is above the operating frequency for the device as a whole. It has been shown that when the frequency rises above the resonant frequency the reactance is, in effect, inductive (see Fig. 7, p. 20).

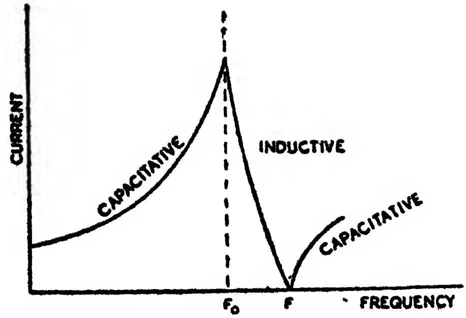


Fig. 37.—PHASE SHIFT CAUSING INCREASE IN FREQUENCY FROM F_0 TO F.

of which it presents a minimum impedance and at the other a maximum.

At the frequency of mechanical resonance for the crystal itself F_0 is equivalent to a series resonance circuit and the circuit as a whole behaves as a resistance in series with capacity (see Fig. 36 (b)).

The circuit is capacitative in this condition, and since R is the minimum value that the crystal can have, the whole circuit has minimum impedance and passes maximum current. At frequencies above F_0 the crystal takes a lagging current and behaves inductively.

Owing to the very high Q, a very small frequency deterioration gives a 90° phase shift (approx.) to the current (Fig. 37).

At some frequency above F_0, therefore, the crystal will tune with C_1 to give parallel resonance circuit presenting maximum impedance (Fig. 36 (b)).

Neglecting R, which is small, since Q is high, this frequency can be found by equating

$$\frac{E\omega C}{\omega^2 LC - 1}$$

through the crystal circuit to the capacitative current $E\omega C_1$ through the holder.

$$(\omega^2 LC - 1)\omega C_1 = \omega C$$

$$\omega = \sqrt{\frac{1}{LC} + \frac{1}{LC_1}}$$

$$\text{Frequency} = \frac{1}{2\pi}\sqrt{\omega + \frac{1}{LC_1}}.$$

It is only between F_0 and F that the circuit behaves inductively and provided that C_1 is large compared with C, which is normally the case with a mounted crystal, F_0 and F are very close together.

The circuit of a crystal oscillator of the tuned grid, tuned anode type in which the crystal provides the grid excitation, is shown in Fig. 36.

Oscillations can only occur in this circuit at a frequency at which both anode and grid circuits behave inductively, from which it is clear that the oscillation frequency must lie between F and F_0. This means that the anode circuit must be tuned above the crystal frequency and that the tolerance is very small. In consequence of this the degree of frequency stability must be very high.

Resistance Capacity or Phase Shift Oscillator

In an ordinary "Hartley" oscillator, the correct phase change is obtained by connecting grid and anode to the opposite sides of an oscillatory circuit, which is oscillating at its own natural frequency.

If a network is built up consisting generally of a number of equal series impedances and equal parallel impedances, an A.C. voltage applied to the input will experience a phase change in passing along the network which will vary with varying frequency. If such a coupling is used between grid and anode of a valve and if, at a certain frequency, a phase change of 180° occurs between input and output, the circuit will oscillate at this frequency, provided that the amplification of the valve will compensate for the attenuation introduced by the network.

The simplest components to choose for the network are resistance and capacity. Inductances have the disadvantage of becoming a source of magnetic interference with other circuits.

A circuit with suitable arrangement of series resistances and paralleled condenser is shown in Fig. 38. A combination of series condensers and parallel resistances sometimes results in a value for the condenser which is impossibly small.

Since each section cannot produce a phase change as great as 90°, it is necessary to have at least three sections. In this case, each section produces a phase change of about 60°, the total phase shift being 180°.

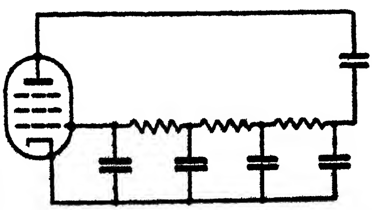
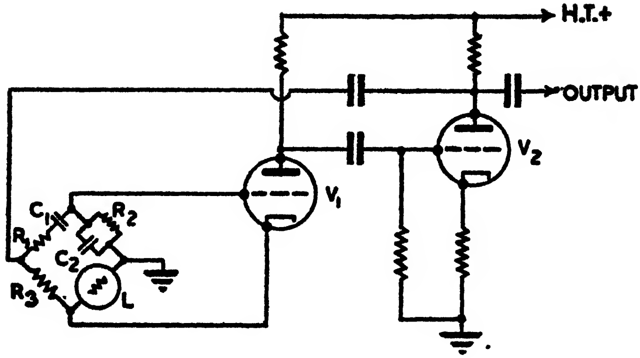
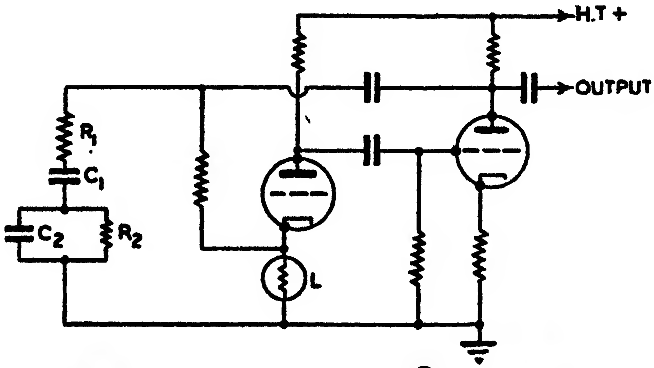


Fig. 38.—Resistance capacity or phase shift oscillator.

(a)



(b)



(c)

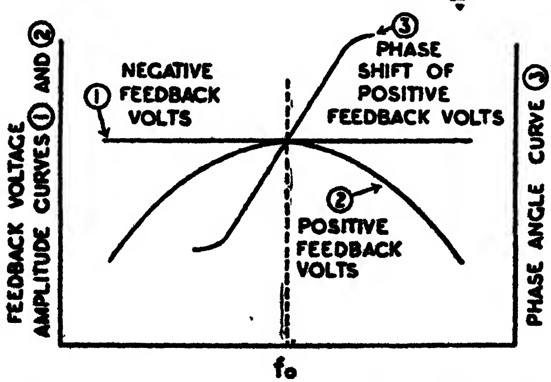


Fig. 39.—WIEN BRIDGE OSCILLATOR AND ANALYSIS OF ITS OPERATION.

(a) Wien bridge oscillator.
(b) Expanded diagram of the Wien bridge oscillator circuit.
(c) Frequency feedback voltages of Wien bridge oscillator.

The loss or attenuation of the network must be compensated for by the gain of the valve.

The Wien Bridge Oscillator

An oscillator which is widely used for generating oscillations of stable frequency, mainly in the low-frequency range, is shown in Fig. 39 (a). These circuits are sometimes employed to generate the Synch. pulse when an independent master timing device for a radar system is to be used.

The equivalent circuit is shown in Fig. 39 (b).

V_1 is an oscillator, and V_2 acts as amplifier and inverter. Thus, without the bridge arrangement for providing feedback the circuit oscillates, since inversion takes place in V_2 in phase with the grid voltage of V_1.

Without the bridge the system is not selective, since it will amplify, invert and feedback in phase oscillation, frequencies over a very wide range. The object of the bridge included in the arrangement for feedback is to eliminate all frequencies other than the desired frequency at which the device is to oscillate.

The selective action of the bridge is due to frequency discrimination in respect of negative feedback and phase shift.

Oscillations can take place only at the frequency F_0 which permits (*a*) the voltage across R_2 to be in phase with the output voltage of V_2, and (*b*) the frequency for which the positive voltage feedback exceeds the negative feedback. Voltages at any other frequency cause (*a*) phase shift between the output of V_2 and the input of V_1 and (*b*) they are attenuated by negative feedback so that they are insufficiently great to maintain oscillations.

Frequency Feedback Voltages

Negative feedback is provided by the voltage divider R_3 and the lamp (L). There is no phase shift across this arm, the resistance is constant for all frequencies, and the amplitude of negative feedback voltages are constant for all frequencies in the output of V_2; see curve of negative feedback voltages in Fig. 39 (c).

Positive feedback voltage is provided by the arm $R_1C_1R_2C_2$. When the frequency is very high X_c is almost zero. In this case R_2 is shunted by a very low resistance, making the voltage between the grid of V_1 and earth almost zero. On the other hand, if the frequency is reduced towards zero, the current that can flow through either C_2 or R_2 approaches zero because of the very high reactance of C_1 and the voltage between the grid of V_1 and earth approaches zero.

At some intermediate frequency the positive feedback voltage is at a maximum, as can be seen by reference to Fig. 39 (c). The change in magnitude of feedback voltage in the vicinity of F_0 is small, but the phase shift that occurs in the positive feedback circuit permits only a single frequency to be generated.

The voltage across R_2 is in phase with the output voltage of V_2 if $R_1 \times C_1 = R_2 \times C_2$. If the frequency of the output of V_2 increases, the voltage across R_2 tends to lag the voltage of the anode of V_2. If the frequency decreases, the voltage across R_2 leads the voltage at the anode of V_2. Curve 3 shows the phase angle between these two voltages as the frequency of the feedback voltage is varied.

The frequency at which the circuit oscillates is

$$F_0 = \frac{1}{2\pi\sqrt{R_1C_1R_2C_2}} = \frac{1}{2\pi R_1C_1}.$$

At this frequency the positive feedback voltage on the grid of V_1 just equals or barely exceeds the negative feedback voltage on the cathode and the positive voltage is of the proper phase to sustain oscillations. At any other frequency the negative feedback is larger than the positive feedback so that the nett result is to suppress these frequencies.

The lamp is used as the cathode resistor of V_1 in order to stabilise the oscillation. If for some reason the amplitude of the oscillation tends to increase, the current through the lamp tends to increase and the resistance of the filament increases. This increased resistance causes a greater negative feedback voltage to be developed. The potential of the grid of V_1 is accordingly reduced and the current output voltage is thus held at a nearly constant amplitude.

Self-pulsing Oscillator

The circuit shown in Fig. 40 is a conventional Hartley oscillator commonly used for the production of R.F. signals. The R.F. frequency is determined by the L.C. constants. Bias is provided by grid current which charges C_1 *viâ* the cathode-grid resistance. R_1 permits C_1 to discharge during the portion of the cycle when the grid is not positive with respect to the cathode. The nett result is a bias on the grid proportional to the amplitude of the R.F. voltage across the grid oscillatory circuit, and a stable oscillatory condition exists.

Self-pulsing Action

If the time constant of the R.C. circuit R_1C_1 is increased greatly, usually by increasing R_1, the charge on C_1 cannot leak off rapidly enough to follow fluctuations of the R.F. voltage, caused by irregularities in the electron stream. As a result, each successive cycle adds a charge to the condenser, until a point is reached where the voltage across C_1 is so high that the amount of feedback provided from the plate circuit is insufficient to maintain oscillations and they cease. The circuit does not break into oscillation again until the voltage across C_1 is low enough to allow the valve to conduct. This combined action results in periods of oscillation and periods of rest. The operation of the oscillator is therefore intermittent and it is said to be self-pulsing.

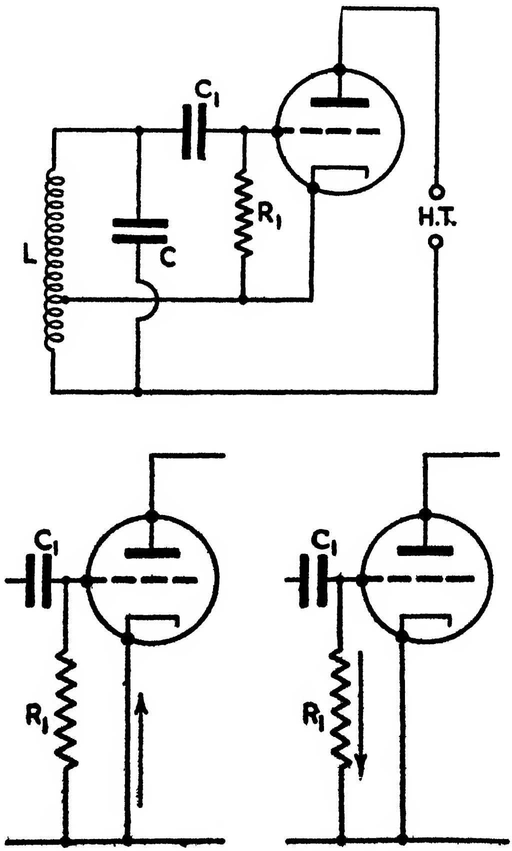


Fig. 40.—HARTLEY OSCILLATOR SHOWING ELECTRON FLOW (C_1R_1 SELECTED TO MAKE A SELF-PULSING OSCILLATOR).

Fig. 41 shows that the circuit starts functioning at the point where the grid condenser is completely discharged. Thus, initially, grid bias is at zero value.

During the initial pulse, which consists of a number of cycles, the grid bias increases until it reaches the point at which oscillations occur. From this point and for each controlling point the grid bias

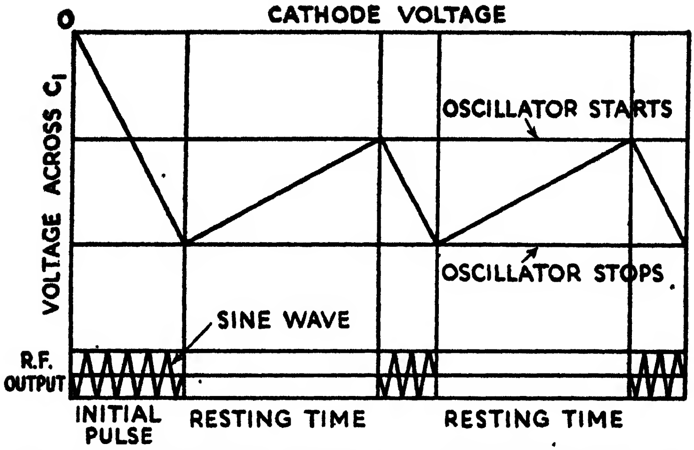


Fig. 41.—INTERMITTENT PULSES OF R.F. ENERGY PRODUCED BY SELF-PULSING OSCILLATOR

varies only between the values at which oscillations start and stop.

Pulse Repetition Frequency

The pulse frequency is determined by the R.C. time constant of the grid circuit and this is the pulse repetition frequency.

The grid of the oscillator swings positive for every positive going portion of the cycle, thus charging C_1. If, say, five cycles are required to provide grid bias to stop oscillations, the grid

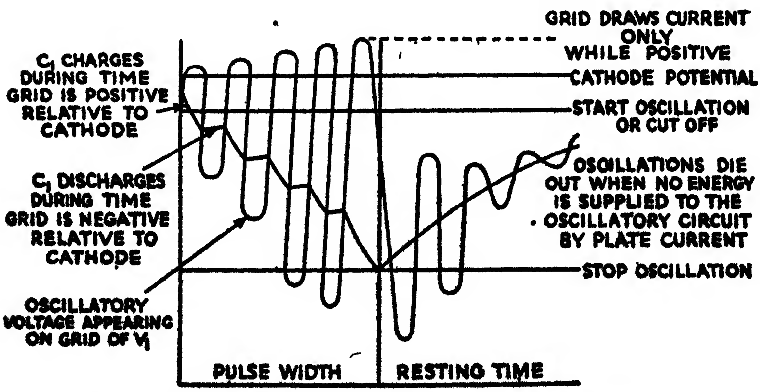


Fig. 42.—CHANGE IN GRID BIAS CAUSED BY SELF-PULSING ACTION.

has to go positive five times for each pulse. During each swing, C_1 receives a nett increase of charge because the grid leak does not completely draw off the charge during the negative part of each cyle. As the charge on the grid condenser builds up, a point is reached at which the grid voltage is sufficiently negative to stop oscillations.

Pulse Width

The pulse width is determined by the size of C_1. If this is small it takes a relatively small time to charge up and the result is a narrow pulse. For a long pulse the capacity must be increased.

Resting Time

The resting time is determined primarily by the R.C. constant. If the time constant is long it takes a long period of time for the charge to leak off and the resting time is comparatively long. Hence the values of C_1R_1 determine the pulse repetition frequency of the oscillator

CHAPTER VI

APPLICATION OF THE CATHODE RAY TUBE TO RADAR

THE cathode ray tube is a special type of vacuum tube in which electrons emitted by a hot cathode are formed into a narrow beam or pencil and accelerated at high velocity towards a specially prepared screen which fluoresces at the point where the electrons strike.

Since the mass of the electron is very small, the beam may be quickly deflected, causing the luminous spot to sweep across the screen, marking its path by a bright trace. The whole length of the trace swept out in this manner can be rendered visible, due to appreciable afterglow of the material with which the screen is coated.

Since the repetition rate is generally far above the maximum limit for the persistence of vision, the nett effect to the observer, is an illusion similar to that of moving pictures as seen on the screen of the cinema. The time base (see Chapter I) appears to be fixed whilst the " pips " representing echo signals returned by targets appear to move along the trace as ranges open or close.

Types

In general, two types of cathode ray tubes are employed in radar, (*a*) the electrostatic type in which the deflecting function is performed by the action of electrostatic fields on the electron beam, (*b*) the magnetic type, in which deflection of the electron beam is obtained by magnetic fields. Both types have their particular applications.

Tubes have been designed in which both forms of deflection are used, but they are not very often employed.

The main outstanding difference between the two types is clearly defined, the electrostatic type being a voltage-operated device whilst the electromagnetic type is essentially current-operated.

Component Arrangement in the Electrostatic Tube

The components and their disposition in an electrostatic tube are shown in Fig. 43.

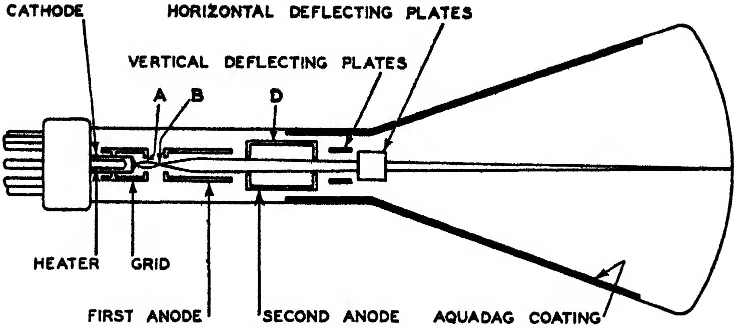


Fig. 43.—ELECTROSTATIC CATHODE RAY TUBE.

The cylindrical cathode with its heater is situated at the base of the tube. It will be noted from Fig. 44 that the heater leads are twisted together in order to render them non-inductive and so to avoid interference with the beam from this source. The cylindrical cathode emits electrons in one direction and from its end only. The end diameter is about $\frac{1}{8}$ in. and carries the oxide coating from which emission takes place.

The control grid takes the form of an outer cylinder surrounding the cathode. A small hole "A" in the end cap permits electrons to pass into the outer space "B" when the control grid is made positive to the cathode.

Functions of Grid and Anodes

The grid performs two functions. It controls the intensity of the electron stream from the cathode and, by its geometry and electric field, causes the electrons, when it allows them to pass through the aperture at "A" to converge and cross over at the point "B."

This preliminary focusing effect is insufficient by itself to prevent the beam from spreading, and it is finally brought to a focus at the screen by the joint actions of the grid-potential,

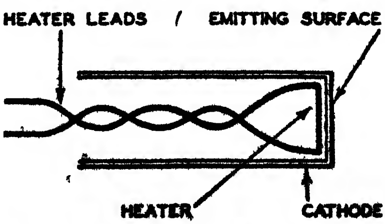


Fig. 44.—DETAILS OF CATHODE AND HEATER.

the aperture " D " and the first and second anodes, marked as such on the diagram, Fig. 43.

Under the influence of the anodes which are always very positive * relative to the cathode, the electrons, formed into a narrow beam, are accelerated towards the screen with a velocity of about 10,000 miles per second or more.

Electrons travelling at this high velocity generate heat at the point of impact with the screen, consequently the coating at this point fluoresces and the brightness of the spot is determined within limits by :—

(1) The velocity attained by the electrons.

(2) The number of electrons that strike a given spot in a given time (electron intensity of the beam).

Brilliance Control

There are, therefore, two methods of controlling the brilliance of the spot, and hence of the trace :—

(1) By increasing the velocity of the electrons.

(2) By increasing the number of electrons in the beam.

The first method is not generally employed. The usual practice followed is to control the beam intensity. This can be done very simply, by altering the potential of the control grid relative to the cathode, by the setting of a potentiometer. In this way the brightness of the trace can be controlled between the limits of zero and saturation. Saturation is reached when the bombardment intensity causes all the material at the point of impact to fluoresce.

" Aquadag "

Secondary emission from the screen takes place. This is caused by displacement of electrons from the coating under the impact of high velocity electrons from the cathode. If the number of new arrivals exactly equals the number of electrons displaced, the potential of the screen relative to earth is constant. If this state of equality is not maintained, the screen will then become positively or negatively charged relative to earth, according to whether the displacements exceed the arrivals, or *vice versâ*. The " Aquadag " † coating, therefore, besides

* The second anode is usually at earth potential but the cathode is made very negative which comes to the same thing as raising the potential of the second anode. This is done for safety in handling.

† A special grade of colloidal graphite for wall coatings has now been developed.

screening the beam from external fields, stabilises the potential of the screen relative to earth, thereby preventing the formation of an accumulated charge, which might interfere with the operation of the tube. In some cases a metal ring surrounding the screen is cemented to the glass inside the tube to intensify the beam and increase brilliance.*

Note. In the case of an E.M. tube, the "Aquadag" is used as an accelerating anode since a second anode is not employed.

Materials for Screen Coating and their Effects

Materials used for screen coating are :—

Willemite (green display) zinc orthosilicate.

Zinc oxide (blue display).

Zinc beryllium silicate (yellow display).

Zinc sulphide and cadmium zinc sulphide or zinc beryllium silicate for nearly white display.

All the above substances exhibit a certain amount of afterglow or persistence, the degree of which can be controlled in manufacture. In the case where changes take place slowly, it is advantageous that the screen coating should exhibit a certain amount of persistence. In other cases where changes of position, for example, take place with great rapidity, the minimum amount of afterglow is desirable.

Focusing in Electrostatic Tubes

Referring to the electrostatic type of tube, Fig. 43, proceeding from control grid to screen, note the following :—

B, the crossover point. This is produced by the joint action of the electrostatic field between grid and cathode and the focusing aperture A.

A then acts as a lens, causing the beam to converge at the point B.

Leaving point B, the beam begins to spread but is brought to a focus on the screen by the action of the electric field between the first and second anodes and the second aperture D.

The grid, the first and second anodes, perform two functions :—

(*a*) A focusing function. (*b*) An accelerating function.

When the grid is made positive to the cathode, the electric field between grid and cathode is shown in Fig. 45.

* In tubes recently developed, the fluorescent coating is sprayed with finely divided aluminium particles in order to reflect light to the face, thereby increasing the brilliance of the display.

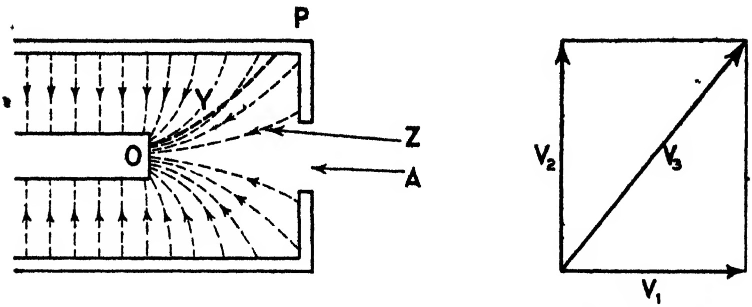


Fig. 45.—Action of grid in concentrating electrons from cathode into beam initially.

Effect of Grid

Electrons enter the electric field from the cathode with initial velocity. If they entered the field with an initial velocity approaching zero they would then be drawn to the positively charged grid along the converging lines of force. Due, however, to the above-mentioned initial velocity which the electron acquires and to the fact that the field is not uniform, the electron, as it recedes from the cathode, follows a curved path. Its velocity at each instant of time can be resolved into two components, both of which change in magnitude and direction from instant to instant.

Briefly, at any instant T the velocity of the electron can be represented by two vectors, V_1 and V_2, V_1 representing the magnitude and direction of its initial velocity and V_2 representing the magnitude and direction of the velocity conferred on the electron by the positively charged electrodes; the resultant, represented by vector V_3, will represent the new instantaneous velocity acquired by the electron.

If at T_2, V_3 is now applied to a new triangle of velocities, another vector, V_4 appears, and so on. If the points obtained in this manner are plotted, the result is a curved flight path.

From the foregoing considerations, it is evident that if the cathode and grid are dimensioned and positioned suitably, most of the electrons emitted by the cathode can be made to pass through the aperture A when the grid is sufficiently positive, relative to the cathode.

When the grid is made sufficiently negative, electrons will be returned to the cathode and none will pass through A. This is the zero intensity condition for the beam.

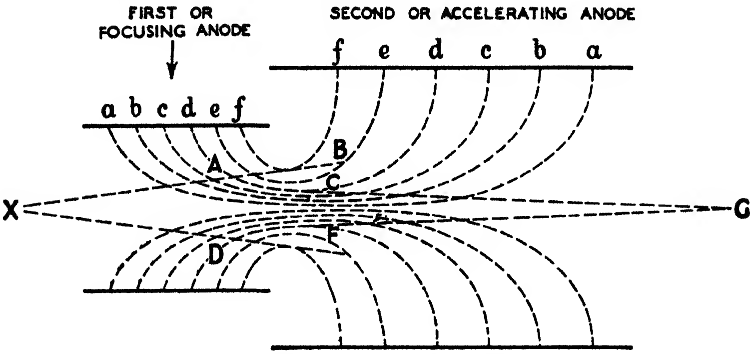


Fig. 46.—FOCUSING AND ACCELERATING EFFECTS OF FIRST AND SECOND ANODES OF AN E.S. TUBE.

The velocity with which electrons pass through A when the grid is positive is determined by the combined fields of the grid and first and second anodes, *i.e.*, by the value of their potentials relative to the cathode.

Effects of First and Second Anodes

In considering the combined focusing and accelerating effects of the first and second anodes, the argument developed for the grid-cathode section can be applied. It may, however, be instructive to look at the matter from another angle.

Consider the movement of an electron entering the field of the first and second anodes along the line XAB from the direction of X, Fig. 46. This electron will have already been accelerated to a considerable velocity. The acceleration will be progressive during the time the electron is passing through the field, as a consequence the electron will attain a very high velocity and pass through quickly. Thus the time during which the field can operate upon the electron to modify the direction of flight is consequently very short indeed.

The nett effect of the electrostatic forces a, b, c and d is to push the electron off its course XAB and modify it to XAC. At C the electron is near the axis where the lines of force are fairly parallel and the field is more uniform. As a consequence, and due to the fact that the lines " dd " and " ee," along which the electron is travelling, accelerate it in the direction of G, the electron will not be forced nearer the axis.

Because of the high velocity attained by the electron before it leaves the field, the electrostatic forces "bb" "aa" will have very little effect upon it.

The total length of path travelled by an electron distant from the axis is greater than for those entering the field on or about the axis. On the other hand, the accelerating forces acting on the distant electrons are greater than those acting on electrons close to the axis, so that the increased length of field to be travelled is offset in time by greater acceleration.

Any tendency to scatter on the part of the electrons forming the beam, due to repulsion between like charges, is counteracted by the high velocity of their flight.

Focusing Control

Sharpness of focus depends upon the strength of the electric field between first and second anode. This is usually controlled by a potentiometer in the circuit of the first anode which can be adjusted by the operator, whilst watching the screen for maximum sharpness of the spot, *i.e.*, with the beam static.*

Horizontal and Vertical Deflection

On leaving the field of the second anode the beam passes between the vertical deflecting plates V_1, V_2, and the horizontal deflecting plates H_1, H_2. If V_1 is made positive to V_2, the beam will be pulled upwards and the spot will move to 2, Fig. 47 (a); conversely, if V_2 is made positive to V_1, the beam will be pulled downwards and the spot will move to 3.

After passing the vertical plates and perhaps being deflected up or down, the beam passes between H_1 and H_2 where H_1 positive to H_2 deflects the beam to the left and the beam as seen from the front moves to 4 (front right). If the polarity of H_1 and H_2 is interchanged the beam moves right (as seen from the deflecting plates) or to left 5, as seen from the front of the screen.

Positioning Voltage and Inputs

In practice, a positioning voltage is applied to H_1, H_2, making H_2 sufficiently positive to H_1 to draw the beam over to the extreme left-hand diameter of the screen, seen from the front of the screen. The horizontal or time base voltage, rising linearly, is applied to H_1 and H_2, generally in *push-pull* and is

* The beam should never be left in a stationary position for any length of time when the intensity is high. This results in damage to the screen by burning at the stationary point.

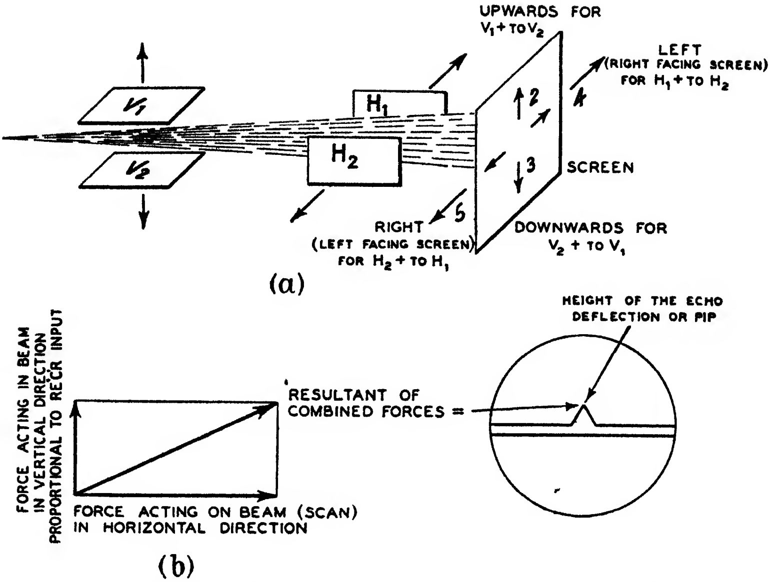


Fig. 47.—HORIZONTAL AND VERTICAL DEFLECTION AND ECHO AMPLITUDE.
(a) Detail of deflecting electrode in E.S. tube.
(b) Pip height proportional to amplitude of received signal.

superimposed upon the biasing or positioning voltage. The *push-pull* method of application is adopted to prevent any defocusing effect due to the distribution of voltages. In this manner the defocusing effect is zero at any instant of time.

Input from Time Base Generator

To do this the time base or deflecting voltage generated by the time base generator is applied to one H plate and inverted before application to the other H plate, in consequence of which a push-pull voltage is applied to H_1 and H_2, causing the potential of H_2 to rise and H_1 to fall by an amount such that when H_2 is fully positive the spot is deflected fully to the extreme right-hand diameter looking at screen front. The average value is zero.

Input from Receiver

The output of the receiver is applied as uni-directional pulses, *viâ* a video amplifier, to the vertical deflecting plates. Each echo signal, as it is received, causes the trace to be deflected vertically for the duration of the pulse, thus producing a "pip" on the trace which marks the *time of receipt of each echo signal, relative to the instant at which the transmitter fired and for each sweep of the time base.* Thus successive pictures superimposed one upon the other present the illusion of motion along the trace, as the ranges from targets by which they are returned open or close.

Echo Amplitude

The height of the "pips" formed on the trace in the above manner is proportional to the amplitude of the received echo signal. This is shown by application of the parallelogram of forces (Fig. 47 (b)).

Relation of Time Base Generator Output to Brightening Pulse

Since the instants at which the transmitter fires are synchronised with the commencement of each successive sweep of the time base, echo signals received and appearing on the trace subsequent to this reference or zero time, indicate by their position in time, the ranges of targets by which they are returned.

Parallel with these activities, a pulse, also synchronised with the commencement of each sweep of the time base, is applied to either the grid or the cathode of the C.R.T. to intensify the beam (which is normally adjusted for low intensity) for the duration of each sweep. This is done in order that the "flyback" may take place in semi-darkness or at zero intensity. The object of this is to avoid the flyback and any echoes received from outside the maximum range for which the time base scale has been set, from appearing on the screen, thereby tending to confuse the observer.

This intensifying pulse may be either a positive pulse to the control grid or a negative pulse to the cathode, whichever is most convenient, having regard to other inputs which may be involved.

At the end of each forward stroke of the sweep the intensifying pulse ceases abruptly and the time base voltage falls quickly to zero. The positioning voltage now draws the beam (almost completely suppressed) back to the left-hand diameter at high speed, and in comparative darkness. On return to the left-hand diameter, special arrangements are made accurately to position it in readiness for the next forward sweep of the time base voltage.

Deflection Sensitivity

If either or both of the deflecting electrodes is moved back along the tube, away from the screen, then, for any measured angle of deflection the length of the resultant seen on the screen increases. Since the vertical deflecting plates are generally further away from the screen than the horizontal deflecting plates, a given deflecting force applied to the vertical plates produces a greater arc of deflection than when the same deflecting force is applied to the horizontal plates. In other words, for a given deflecting force the vertical deflecting plates exercise more control over the beam than the horizontal plates. The sensitivity of an electrostatic tube may therefore be defined in terms of the deflection obtained per volt, *i.e.*, mm./per volt.

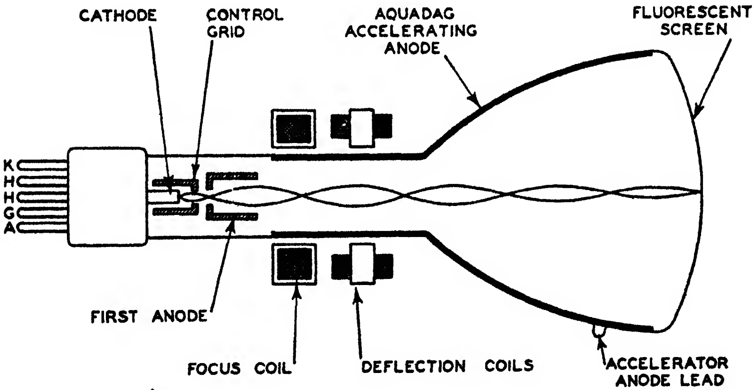


Fig. 48.—ELECTROMAGNETIC CATHODE RAY TUBE.

Electromagnetic Tube

The electromagnetic type of tube is similar in many respects to the electrostatic type.

From Fig. 48 it will be noticed that the shape of the glass envelope is different. More room is required along the neck to accommodate the coils, but the overall length of the tube is less.

General Description of E.M. Tube

The cathode, grid and first anode are similar in arrangement to their corresponding elements in the electrostatic tube, but there is no second anode; the accelerating function of this

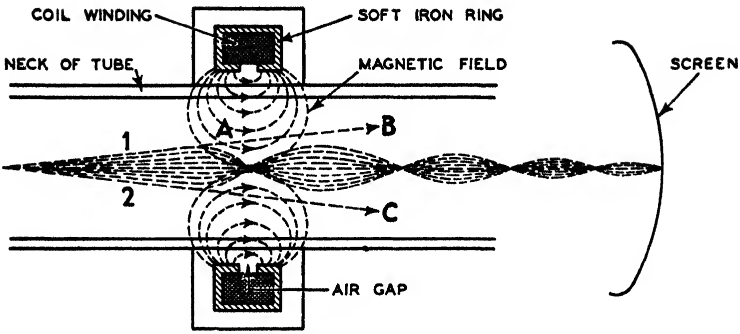


Fig. 49.—MAGNETIC FIELD SET UP IN CATHODE RAY TUBE BY MAGNETIC FOCUS COIL.

element being performed by the "Aquadag." The outstanding differences are:—

(*a*) The anode plays no part in the focusing function; this is performed magnetically by a coil surrounding the neck of the tube, see Fig. 49.

(*b*) The deflecting function of V_1 and V_2, H_1 and H_2 in the electrostatic tube, operated by the output of the receiver and the time base circuit respectively, are performed magnetically when this type of tube is employed to provide an "A" display.

E.M. Tube as Current-operated Device

(*c*) Since the deflecting functions are performed magnetically, the trace is controlled by the *current* in the deflecting coils. This type of tube is then a current-operated device, whereas the electrostatic tube is voltage-operated.

Relation of Time Base Voltage to Time Base Current in E.M. Tube

(*d*) The rise of the time base voltage must be linear for the electrostatic tube. In the case of the electromagnetic type it is the *linear rise of the current* in the deflecting coils that matters. As a consequence, therefore, the output of the time base circuit or, in other words, the wave of voltage, applied to the deflecting coils, must be shaped (having regard to the inductance of the coils) to ensure a *linear rise of current* through them.

(*e*) Positioning of the beam is accomplished by standing current in the horizontal deflection coils.

(*f*) The time base current is supplied to the deflecting coils by a push-pull circuit.

Focusing Control

The focusing coil with its iron core and small air gap is mounted on the neck of the tube as shown in Fig. 49. The coil is wound on an annular ring in which is cut a groove ; an air gap is provided in order that the magnetic path may pass through the tube. The fields at opposite points on the diameter are in the same direction and, therefore, oppose. As a result, the lines of force between poles lie parallel with the axis as shown in Fig. 49. The value of the current and, therefore, the intensity of the field, can be varied by a potentiometer.

Method of E.M. Focusing

Consider the motion of an electron entering the magnetic field along the line AB, with considerable velocity, the resultant velocity at each instant of time can be resolved into two components, *i.e.*, a component of velocity parallel to the axis and magnetic field, and another component of velocity acting in a direction across the field.

The electron in motion is by definition an electric current, consequently the effect of the magnetic field upon it is similar to that which it would have upon a current-carrying conductor situated in the magnetic field and moving with a velocity equal to that of the moving electron.

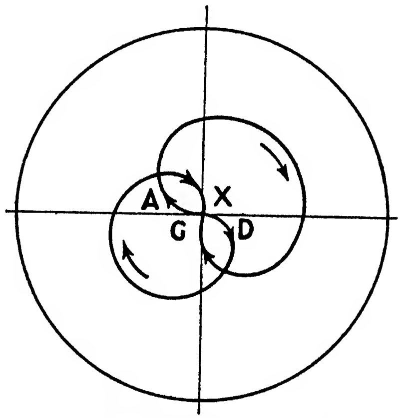


Fig. 50.—MOVEMENT OF ELECTRON (MAGNETIC FOCUSING).

Applying the right-hand rule as though to a current-carrying conductor, it is seen that the motion of the electron in the field at XAD is vertically upwards out of the page (Fig. 50).

Since the first component of velocity is parallel with the axis and is much greater in magnitude, the resultant path of the electron through the field takes the form of a helix. All electrons entering the field will be acted upon by similar forces and will therefore follow a helical path. In this manner, and by suitable adjustment of the field strength, the beam can be brought to focus sharply at the screen as an identical reproduction of the concentration at the initial cross over.

When magnetic deflection is used to produce an "A" display, two pairs of coils are mounted on the neck of the tube, one pair being at right angles to the other.

The Time Base Current

The time base current flows in one pair of coils to produce horizontal deflection, the signal current flows in the other pair to produce vertical deflection. The magnetic field which deflects the beam is the resultant of these two forces, in consequence of which echoes are produced on a trace resulting in a display similar to that of the electrostatic tube.

As in the electrostatic type of tube, the time base generator is started by a pulse which synchronises the commencement of the trace with the firing of the transmitter, and the time base voltage (E.S.) or current (E.M.) grows linearly during each sweep from some fixed positioning value to some other fixed value depending upon the diameter and sensitivity of the tube. In practice both the minimum and maximum values of the time base voltage or current are literally *clamped* by "clamping" circuits to fixed values as a result of which :—

(*a*) Each successive sweep of the time base voltage or current must start from the same level and finish at the same level. The levels referred to are the minimum and maximum values of voltage or current.

(*b*) The beam must commence each successive trace from the same point on the screen and finish at the same point.

(*c*) Thus the bright line which marks the time base is constant in length for any given tube and proportional therefore to some fixed *time*.

(*d*) The *rate* at which each successive sweep of the time base (voltage or current) grows from the starting level to the finishing level determines the velocity of the beam and hence the time value to which the time base is made proportional.

(*e*) This rate of growth must be linear, *i.e.*, the time base voltage or current for each successive sweep must increase by equal amounts in equal time. This enables the trace to be divided into equal lengths each proportional to equal times. Since range (in yards) is also proportional to time in microseconds × 164, the rate of growth of successive sweeps of time base (current or voltage) calibrates the screen for maximum time or range and, because of its linear growth, the time base can also be calibrated for equal sub-units of length.

Brightening Pulse

At the commencement of each successive sweep of the time base (voltage or current) a pulse is applied to synchronise the firing of a brightening pulse generator.

The output from this generator is a square pulse of duration equal to the sweep time of the time base, *i.e.*, it intensifies the beam from the commencement of each sweep until its end. The trace is only illuminated therefore during the forward stroke of the time base and the flyback takes place in comparative darkness.

Flyback Time when Multi-Range Available

At the end of each successive sweep of the time base voltage or current, the maximum value to which it may rise is clamped to a fixed upper level until the flyback takes place. There may be some delay between the instant that each sweep of the time base voltage or current reaches this maximum deflection value, and the moment at which the flyback takes place. This is due to the fact that most display units are provided with a switch, by means of which the time base can be made proportional to perhaps 15,000 yds., 60,000 yds. or 100,000 yds., according to the position selected. This switch controls :—

(*a*) The rate of growth of the time base voltage or current, *i.e.*, the speed of each sweep.

(*b*) The duration of the square pulse which intensifies the beam during each sweep.

Conditions when several Ranges available

It is usual, therefore, to adjust the time base generator so that the flyback takes place almost immediately after the slowest sweep, consequently, because the fast sweep (15,000 yds.) reaches maximum value earlier than the slow sweep (100,000 yds.), there is a period during which the fast sweep is held clamped at its maximum value. This does not show on the screen, however, because the time duration of the brightening pulse is also controlled by the same switch, and the beam is extinguished as soon as the time base voltage or current reaches its maximum value.*

* Although at the end of the fast sweep the beam is clamped to the right-hand side of the tube for some appreciable time before the "flyback" takes place, this does not harm the screen, because the beam is extinguished at the end of the sweep and it cannot, therefore, burn the screen during the time it is held clamped to the right-hand diameter.

These conditions are shown in Fig. 51 and dealt with in greater detail in Chapter XIV, Time Bases and Time Base Circuits.

Shape of Time Base Voltage or Current Wave

From Fig. 51 it can be seen that the deflecting force, whether electrostatic or electromagnetic, must take the form of a sawtooth wave of voltage or current. In the electromagnetic tube, however, a sawtooth voltage applied across the deflection coils will not produce a sawtooth wave of current through the coil. The form which the applied voltage wave must take is discussed in Chapter XIV.

Comparison of E.S. and E.M. Tubes

In making a comparison of electrostatic and electromagnetic tubes the following points emerge :—

(*a*) The electrostatic tube operates on comparatively low power, whilst the magnetic tube requires larger power.

(*b*) On the other hand, the electrostatic tube requires high voltage and is, therefore, subject to all the attendant disadvantages. Its length is greater than the magnetic tube and it is also subject to vibration.

(*c*) The electromagnetic focusing is probably superior to electrostatic focusing.

(*d*) The electromagnetic tube is more easily and cheaply produced than the electrostatic type, one reason being that there is less mechanical difficulty in mounting the electrodes.

The sensitivity of a cathode ray tube may be defined in terms of millimetres deflection per volt.

Factors controlling Forms of Data Presentation

By various combinations of inputs to the vertical and horizontal deflecting plates and to the control grid or to the cathode, a form of data presentation can be selected which is most suitable to any particular application. For example, it is possible to use the vertical and horizontal plates to plot range against azimuth. In this case the time base is applied to the vertical plates and a voltage proportional to the bearing of the aerial is applied to the horizontal plates. The receiver input is then applied to the control grid.

The position of the echo in range and azimuth is indicated on the screen by the appearance of a bright spot.

This bright spot is caused by intensification of the beam,

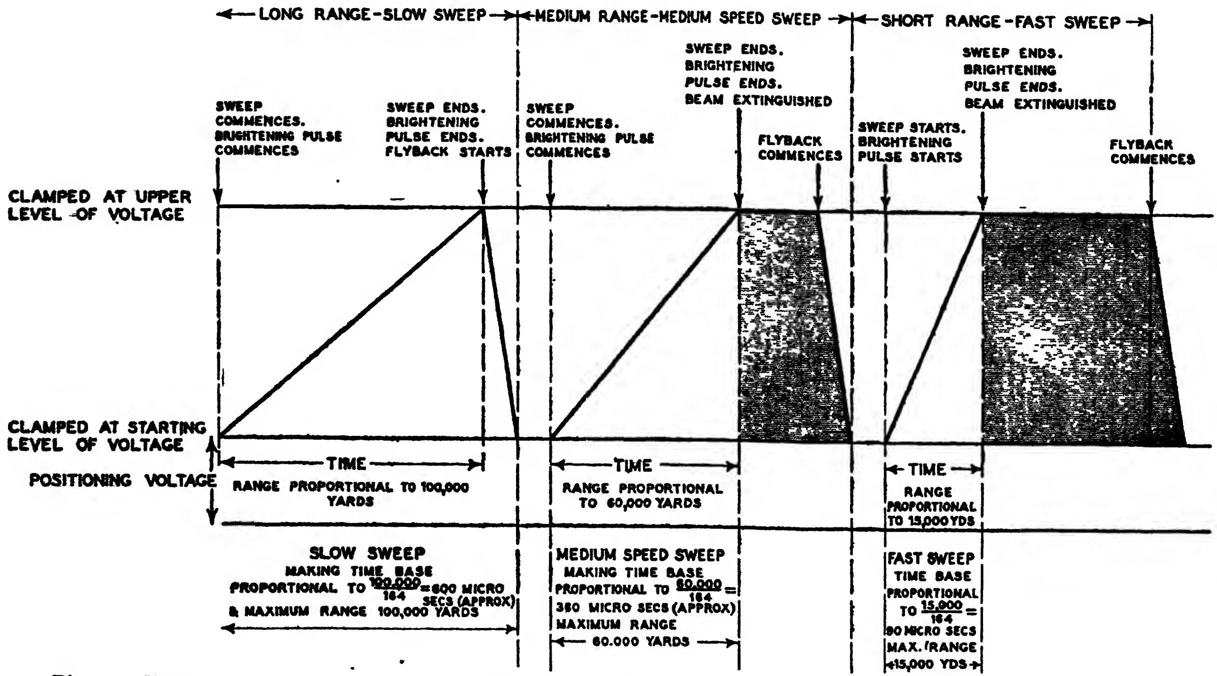


Fig. 51.—TIME RELATIONS AND LEVELS OF TIME BASE, VOLTAGE OR CURRENT, BRIGHTENING PULSE, FLYBACK, AND DURATION OF CLAMPING.

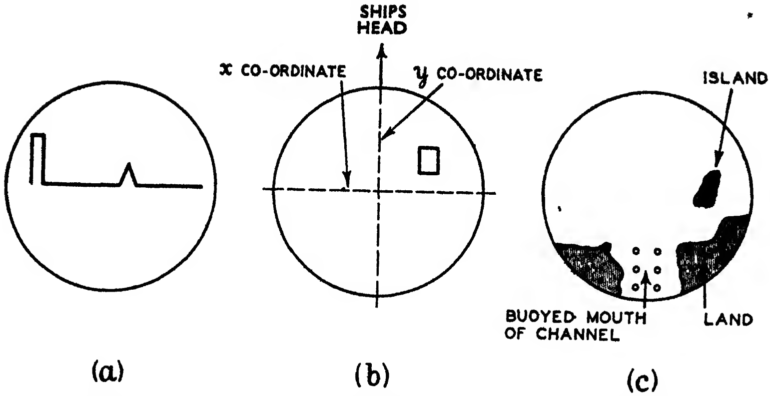


Fig. 52.—APPEARANCE OF SCREEN FOR VARIOUS DISPLAYS.

(a) " A " type display.
(b) " B " type display.
(c) Plan position display or " P.P.I." (plan position indicator).

due to a positive input to the control grid and originated by the received echo.

At the moment of intensification, which coincides with the arrival of the echo, the beam, in semi-darkness, under the influence of the vertical and horizontal deflecting forces, is at a point which in effect corresponds to Y and X co-ordinates,* where range is plotted along Y (vertical) and azimuth along X (horizontal). In these circumstances the beam is said to be intensity modulated.

Alternative Inputs for Intensification Modulation

In this respect it should be noted that a positive input to the grid produces the same effect as a negative input to the cathode. In either case the potential of the grid is raised relative to the cathode and the number of electrons in the beam is thereby increased.

P.P.I. Display

Another useful form of presentation in frequent use is a display which presents a plan of the positions of all targets within a circular area of which the observer's position is at the centre. An electromagnetic tube is usually employed for this display, which is generally obtained by causing the time base to start from the centre of the screen outwards, so that each

* In the j notation this point would be $A + jB$.

trace extends from the centre to the circumference of a circle the radius of which is made proportional to range. The coil carrying the time base current is rotated around the neck of the tube by a small motor, the motor being driven in synchronism with the aerial, which is also continuously rotated.

In this arrangement the line of sight of the aerial continuously scans a circular area of which the aerial is the centre, at perhaps five or six revolutions per minute. The deflecting coil carrying the time base current is rotated at exactly the same speed and exactly in step, thus causing the radial trace to revolve at the same speed. If the output of the receiver is applied to the grid or cathode to give intensity modulation, echoes from all targets in the area will appear on the screen as bright arcs,

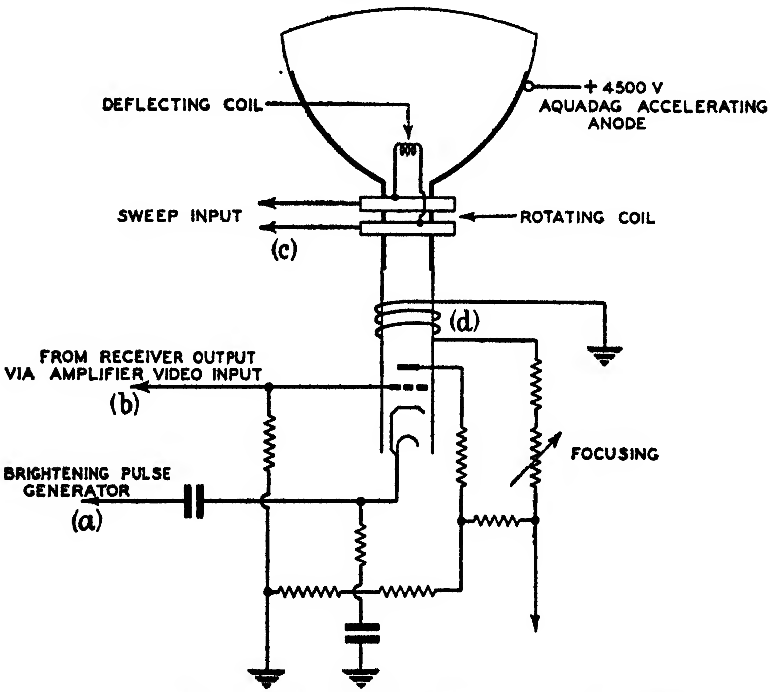


Fig. 53.—ELECTROMAGNETIC TUBE ARRANGED FOR P.P.I. DISPLAY.

(a) Brightening pulse to raise intensity of beam during sweep.
(b) Input from receiver brightening the beam still more when echo signals received.
(c) Rotating deflection coil.
Focusing control.

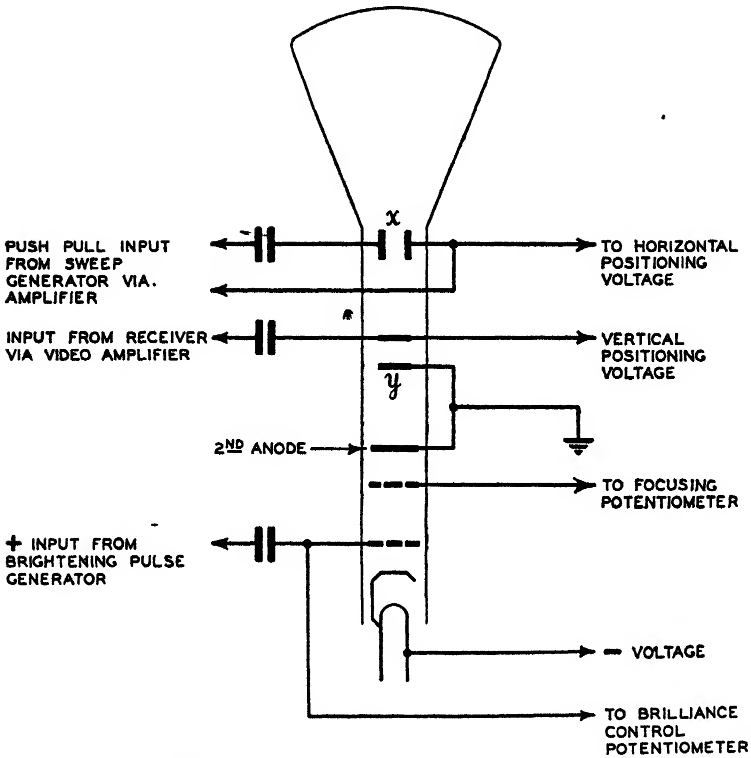


Fig 54.—INPUTS TO ELECTROSTATIC TUBE.

situated at distances from the centre which are proportional to range, and positioned for bearing relative to the ship's head. This type of display is known as "plan position."

The use of cathode ray tubes for various types of display is discussed in the section dealing with indicators, a unit in which the cathode ray tube is the major component.

Fig. 53 shows a typical arrangement of inputs to an electromagnetic tube. Fig. 54 likewise shows a typical input arrangement for an electrostatic tube for an "A" display.

Note. The beam may either be brightened by a brightening pulse or suppressed by a suppressing pulse. In certain instances, a brightening pulse for the duration of the sweep may be replaced by a suppressing pulse for the interval of time between the end of one sweep and the beginning of the next. Occasionally both are employed.

CHAPTER VII

THE BASIC RADAR SYSTEM

EVERY radar system must include units essential to the performance of at least three operational functions plus a co-ordinating or timing-function. To these requirements must be added an aerial system and a power supply unit.

The units performing essential operational functions are :—

(*a*) The transmitter, which includes the R.F. generator and its associated units.

(*b*) The receiver.

(*c*) The display unit or indicator, which includes the cathode ray tube and its associated circuits.

The co-ordinating function may be performed by any of the three essential units or by an entirely independent unit.

The aerial system includes the array and feeder system, together with the supplementary equipment for rotating the array and for indicating bearing to the operating position.

The power supply generally follows standard practice for telecommunication equipment very closely, but with additional precautions for stabilising voltage of the supply when necessary.

Minimum Requirements

Minimum requirements for operating the system are :—

(*a*) The transmitter must be pulsed in short, sharp bursts at regular intervals.

(*b*) It must be possible to aim the R.F. energy generated in this manner by rotating the aerial array to cover any desired target or in any direction for search purposes.

(*c*) The R.F. energy of the echo signals returned by targets must be converted by the receiver to video frequency pulses which are applied (in the simplest form of display) to the "Y" plates of a cathode ray tube.

(*d*) The output of a time base generator is applied to the "X" plates of a cathode ray tube to calibrate the screen in terms of *time*, to which range is made proportional in terms of 164 yds. per microsecond.

(*e*) The output of the time base generator must be a sawtooth waveform of voltage or current (according to whether an E.S. or E.M. tube is employed), the forward stroke of which

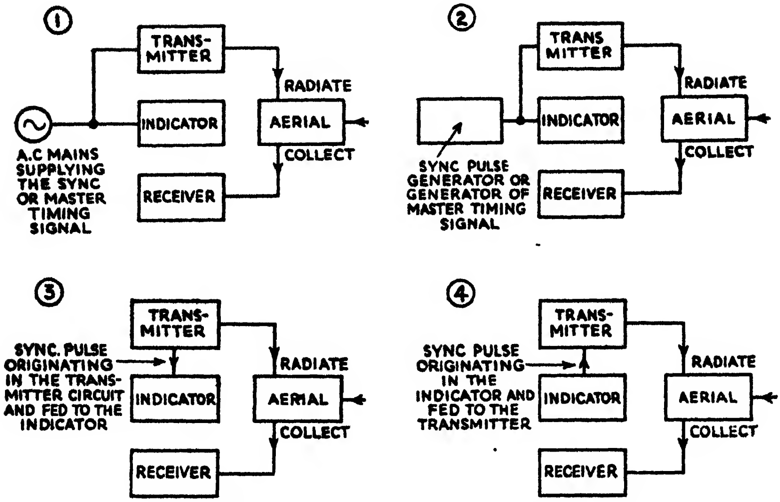


Fig. 55.—Alternative arrangements for synchronising the basic radar system.

Note. It is possible to have combinations of 1 and 3 or 2 and 3; also 1 and 4 or 2 and 4. In such a case, there are two Synch. frequencies each of the same value, independently generated. One, the more stable, is used to stabilise the other. These cases are dealt with in the next chapter.

must grow linearly between a clamped starting value and another clamped maximum value.

(*f*) The operation of the transmitter, the time base generator and the brightening pulse generator must be co-ordinated, that is to say, they must be synchronised to start together at the same instant of time and at the commencement of each successive pulse.

(*g*) Co-ordinating or timing pulses (Synch.) for (*f*) may be supplied from either of the main operating units to the other units, or they may be generated and distributed from an entirely independent unit (Synch. generator). These conditions are shown in Fig. 55.

Fig. 56 shows how the outputs of the major operating units are related to each other in terms of *time* by the synchronising pulse originated and distributed in the manner indicated in (*f*) and (*g*).

Fig. 56 shows how the start of the transmitter pulse, the commencement of each sweep of the time base, the start of

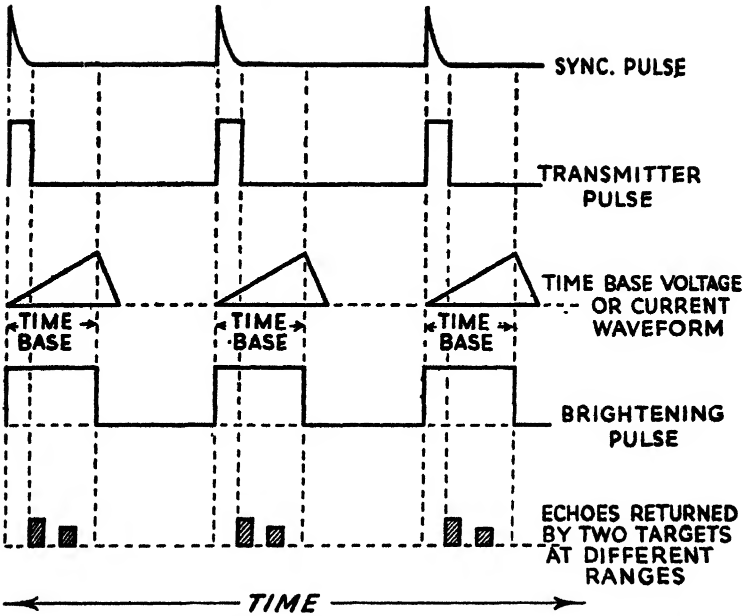


Fig. 56.—SHOWING RELATION IN TIME OF SYNCH. PULSE TO COMMENCEMENT OF TRANSMITTER PULSE, TIME BASE SWEEP AND BRIGHTENING PULSE.

Note (1). Received echoes can be displayed only after the end of each transmitted pulse and for the duration of the brightening pulse. Echoes from nearby targets received during the transmitted pulse are masked on the display.

Note (2). The firing of each transmitter pulse is timed by the Synch. pulse, but the duration of the pulse is not in this case determined by the Synch. pulse. This time may be determined by the constants of the transmitter itself or by some independent control.

Note (3). The time base determined by the sweep time, *i.e.*, the rate of growth of the voltage or current, is determined by the constants of the sweep or time base generator. Duration is controlled by the constants of the brightening pulse generator which is locked to the time base by Synch.

the brightening pulses, are each properly related in *time* by the Synch. pulse.

Constants of the Radar System

Selection of the pulse repetition rate is subject to the following general requirements :—

(*a*) The rate or frequency must be such that sufficient time is allowed between successive transmitted pulses for an echo to return from a target at the maximum range of the equipment. The maximum range for which the set is designed therefore limits the highest frequency that can be used for the repetition rate or pulse repetition frequency (p.r.f.).

(*b*) When an aerial is rotated at constant speed (as for a P.P.I. display), the p.r.f. must be such that, having regard to the speed of rotation of the aerial, a sufficient number of pulses strike the target during each revolution to ensure adequate signal strength. Therefore the speed of rotation of the aerial and the persistence or "after-glow" of the screen of the C.R.T. fix a limit for the *lowest* repetition rate that can be efficiently used.*

(*c*) When the indicator unit is operating for the entire resting time of the transmitter, the repetition rate (frequency) must be very stable to avoid blurring and consequent reduction in accuracy of measurement. There is a special and not very frequent case in which the sweep of the time base, including flyback, extends over nearly the whole of the resting time of the transmitter.

(*d*) In a later paragraph it will be shown that the p.r.f. is one of the factors which determines the maximum peak power that can be safely developed from any given valve. It will also be shown that peak power is inversely proportional to the repetition frequency. In consequence of these opposing factors the selection of pulse repetition frequency is usually a compromise between (*a*), (*b*), (*c*) and (*d*) such that the target definition is sufficient for a particular application.

Pulse width (in terms of time) fixes the *minimum* range at which a target can be detected. If the target is close to the transmitter and the duration of the transmitted pulse is so long that an echo is returned before transmission ceases, the echo will be masked and the target cannot be detected.

Pulse width in conjunction with pulse repetition frequency—the other factor—determines the maximum permissible peak power value that may be developed from any given valve.

The general tendency of short range accurate systems is to make the pulse width as short as possible, *i.e.*, of the order of 1 microsecond or less. In long-range warning sets, pulse widths up to 5 microseconds or more have been employed.

The synchronising signal which controls the pulsing system may be derived from :—

(*a*) The mains supply, if the frequency is suitable.

(*b*) It may be generated at any required frequency by a

* The display may appear to be continuous due to afterglow, etc., but this does not mean that the indicator circuits are all working satisfactorily.

separate unit, which may be called the "Synch. generating unit," or "Synch. generator."

(*c*) It may be originated somewhere in either the transmitter or display systems.

Each of these methods has particular applications, which are discussed in later chapters.

Fig. 55 is a block diagram showing these alternatives applied in their simplest form.* Note that, for the sake of simplicity, power supplies have been omitted. Also it must be understood that each block labelled "transmitter," "indicator," etc., is generally composed of a number of sub-units each of which must also be synchronised by a "Synch." pulse.

As radar systems become more complex to meet wider applications, the number of refinements and sub-units increases, consequently the demand for synchronising pulses not only becomes greater but their duties become more varied.

In some cases a pulse is required to start or stop some particular operation in each sequence, in other cases the pulse is also required to control the duration of the operation as an additional function.

Synch. Pulse and Pulse Generator

In general, any pulse exercises its function by reason of its shape and electrical dimensions, in consequence of which a variety of pulse wave forms are generally required for the operation of a radar system. This means that nearly every control pulse must be generated by its own generator under the control of a Synch. pulse. Thus, in general, the Synch. pulse starts a generator action at the correct moment in each sequence and the electrical constants of the generator circuit shape and dimension the pulse output to suit its particular function in the system. In this way pulse repetition rate is controlled by the frequency of the Synch. pulse distributed from the master source and pulse width is dimensioned in *time* by the electrical constants of the pulse generator. The Synch. pulse may be regarded as a trigger to the pulse generator or pulse shaping circuit.

Average Power and Peak Power

The power relations arising from the fact that the transmitter is pulsed are such that the average resting time of the trans-

* Excluding Synch. pulses to brightening pulse generator, calibration generators, strobe generators, etc., all of which may be regarded as supplementary to the basic units.

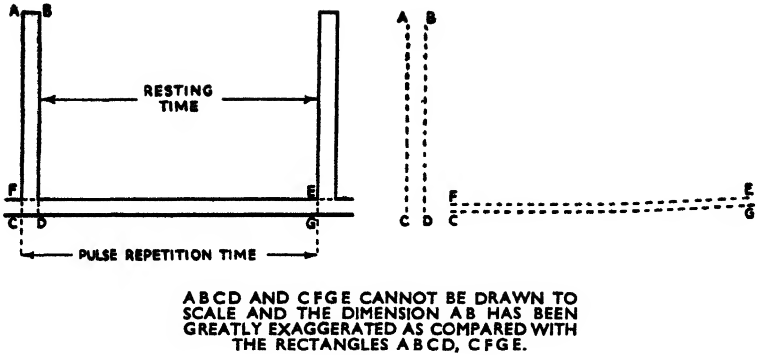


Fig. 57.—RELATION OF PEAK POWER TO AVERAGE POWER.

mitter is usually much longer than its average working time. Consequently, because comparatively long periods of time are available for cooling in between each working period, the valves of the R.F. generator can be safely driven to deliver a much greater peak power for the duration of each short pulse than the average power recommended by the manufacturers for continuous working.

The relationship of average power to peak power is shown in Fig. 57. The energy rectangle A, B, C, D must not exceed the equivalent rectangle C, F, G, E.

In other words, the energy ratio

$$\frac{\text{Average power}}{\text{Peak power}} = \frac{\text{pulse width}}{\text{pulse repetition time period}\left(\dfrac{1}{\text{p.r.f}}\right)}$$

$$= \text{pulse width} \times \text{p.r.f.}$$

Duty Cycle

The operating cycle of the radar system can be described in terms of the total time for which R.F. energy is radiated. This time relationship is termed the "duty cycle." Thus the ratio $\dfrac{\text{pulse width}}{\text{pulse repetition time period}}$ = duty cycle where pulse time period $= \dfrac{1}{\text{pulse repetition frequency (repetition rate)}}$.

Example. A ·1 microsecond pulse repeated 500 times per second represents a duty cycle of ·0005. Similarly, a 2-micro-

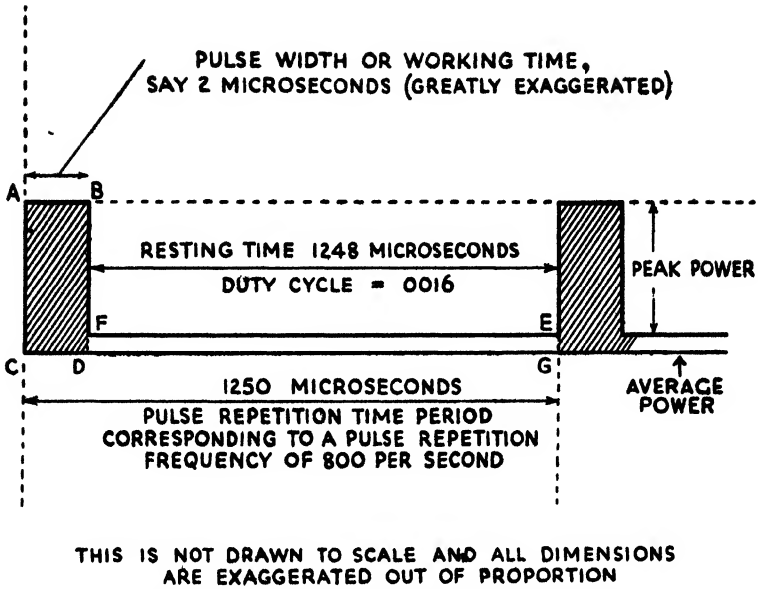


Fig. 58.—GRAPHICAL REPRESENTATION OF RELATIONS IN TIME BETWEEN P.R.F., RESTING TIME, PULSE WIDTH AND DUTY CYCLE.

$$\text{Duty cycle} = \frac{2}{10^6} \times \frac{8 \times 10^2}{1} = \cdot 0016.$$

second pulse repeated 800 times per second gives a duty cycle of ·0016 (Fig. 58).

$$\text{Duty cycle} = \frac{2}{\frac{1}{800}} = \frac{2 \text{ microseconds}}{1{,}250 \text{ microseconds}} = \cdot 0016.$$

Relation of the Power Constants

It now follows that since $\frac{\text{av. power}}{\text{peak power}} = \frac{\text{pulse width}}{\text{pulse time period}}$ and because $\frac{\text{pulse width}}{\text{pulse time period}} = \text{duty cycle}$, then $\frac{\text{av. power}}{\text{peak power}} = \text{duty cycle}$ and average power = peak power × duty cycle.

Example. Let the pulse width be 5 microseconds and the p.r.f. be 500 per second. If the peak power is 15 kW. the duty cycle is ·0025 and the average power is 15,000 × ·0025 = 37·5 watts.

Conclusion. This means that a small valve with an output rated at about 40 watts can deliver 15 kilowatts for 5 microseconds, 500 times per second. The same valve could deliver approximately 23·4 kilowatts for a 2-microsecond pulse repeated 800 times per second.

Peak Power, Pulse Width and P.R.F.

High peak power is of course desirable to produce a strong signal at the desired maximum range. Low average power offers the advantages of small valves and compact circuits. *A low duty cycle is therefore desirable*, consequently, all other conditions being equal, pulse width is generally kept as small as possible and the p.r.f. need not be greater than is necessary to comply with good target definition as specified for the particular application.

When stable 500-cycle mains are available they are often employed as a source from which to derive the "Synch.," thereby conferring on the system a p.r.f. of 500 per second.

Selection of Carrier Frequency

The type of transmitter to be used for a particular radar system depends upon selection of the carrier frequency (radio frequency). Choice of the carrier frequency to be employed is largely dependent upon the particular radar application.

Long-range warning sets may employ carrier frequencies as low as 100 microseconds or less, but short-range accurate sets giving good target definition generally employ micro-waves of from 10 cm. downwards. One of the principal reasons for this latter choice is to obtain the high degree of accuracy which is possible with narrow beams. *Note*, it is only practicable to produce very narrow beams with aerials of small overall dimensions when the frequency approaches the 10-cm. range.

Dispersal of Sub-units

The conditions under which the major functions are performed may involve subdivision of one or more of the main units. Sub-units may be housed remotely from the main unit, either by themselves or with some other main unit or its subdivision.

For example, it may be convenient for technical or other reasons to house the input and initial stages of the receiver in close proximity to the aerial, along with the R.F. generator of the transmitter, and it may also be desirable to incorporate the final stages of the receiver output in the indicator or display unit.

As regards power supply, frequently each unit is self contained in this respect—partly to avoid large central power packs—but mainly, in the case of remote, detached or subdivided units, to avoid long runs of inter-connecting cables carrying high voltage supplies.

In consequence of this, under the terms transmitter, receiver and indicator, it is usual to include all remote sub-units of each of the main units. For example, the term transmitter covers the R.F. generator, together with all the circuits required for pulsing, controlling and supplying it with power, even though it may be housed along with, perhaps, the initial stages of the receiver, near to or as part of the aerial equipment housing.

The Aerial System

The aerial system, which includes the functions of radiation and collection, is treated as a separate unit or units.

The aerial system may be considered to include the transmission line or wave-guide from the transmitter to the aerial array, the array itself and any transmission line or wave-guide between the aerial and the receiver not common to the transmitter. In addition to the foregoing, the aerial, considered as a unit, also includes any switching or protective devices used in switching over from transmission to reception, together with devices for beam switching and equipment for transmitting its bearing relative to the ship's head at any instant of time.

T.-R. Switch

The parallel problems of (*a*) avoiding power losses in the receiving circuits during transmission, and in the transmission circuits during reception, and (*b*) protection of the receiver during transmission, is sometimes solved by using separate aerials for transmission and reception. Since this method is inefficient, uneconomical and bulky, a common aerial is generally employed, together with some form of electronic switch which performs functions (*a*) and (*b*) quite satisfactorily.

The action of the switch is fully described in Chapter XXIV. It is known as a T.-R. switch and depends for its action upon the properties of quarter-wave and half-wave sections of transmission line or wave-guide when used in conjunction with spark gaps and/or resonant cavities (see Appendices I, II and III).

Matching

Efficiency of energy transfer from the R.F. generator to the

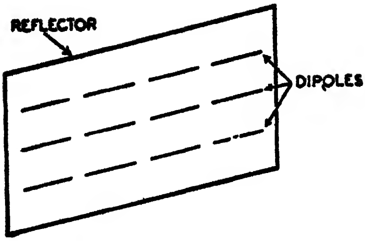


Fig. 59.—DIPOLE ARRAY WITH REFLECTOR.

aerial, and from the aerial to the receiver, must be ensured by matching throughout.

During transmission the transmitter is matched to the transmission line and the transmission line is matched to the aerial. For the whole of this period the receiver circuits are made to present an open circuit or very high impedance to the transmission line. During reception these conditions are reversed.

T.-R. Switching

The problem of switching is greatly simplified by the fact that most transmitters have a different impedance when they are " OFF " to when they are " ON." Consequently, by properly matching to the transmission line during the pulsing period, there will be a mis-match during resting time—looking into the transmitter from the aerial. This fact, together with the action of the T.-R. switch mentioned above, effectively prevents serious loss of energy in the transmitter circuits during reception.

Types of Aerials and their Applications

The possible types of radiators employed in radar systems are :—

(*a*) The multiple dipole array with untuned reflector.

(*b*) The dipole with tuned reflectors and directors (Yagi array).

(*c*) The dipole with parabolic reflector for wavelengths of about 10 cm. (coaxial transmission feeder).

(*d*) Various arrangements of parabolic reflectors used in conjunction with wave-guides.

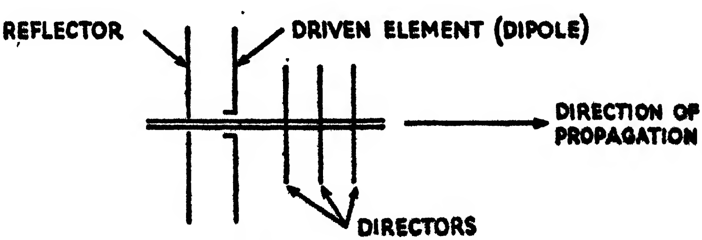


Fig. 60.—YAGI AERIAL.

Fig. 59 shows a typical multiple dipole array. This type may be composed of one or more dipoles. It may be adapted for beam switching (see later paragraph), and can generally be rotated in either azimuth or elevation or both. Its directivity depends upon the number and arrangement of the component dipoles in the array.

Fig. 60 shows a Yagi array. Reflectors and directors pick up energy from the driven element. They are so spaced and proportioned that re-radiation takes place in phase with that of the dipole and the dipole field is reinforced in the forward direction.

Fig. 61 shows a parabolic reflector. This produces a narrow beam when employed with micro-wavelengths. Many points of similarity with the reflection of light can be observed. The

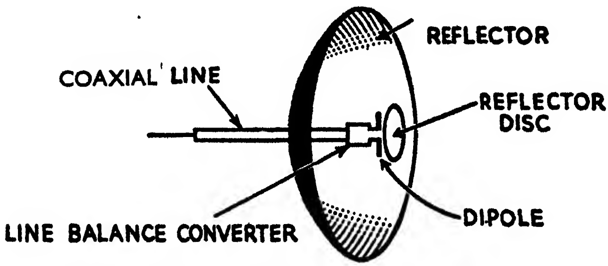


Fig. 61.—DIPOLE WITH PARABOLIC REFLECTOR.

beam is made narrower by increasing within certain limits the radius of the reflector relative to the dimensions of the dipole. R.F. energy is fed to the dipole situated at the focal point of the paraboloid, and the supplementary reflector, placed about a quarter of a wave-length in front of the dipole, reflects back nearly all the energy to the paraboloid from the surface of which it is radiated as a narrow beam.

Requirements for Radar Aerials

The requirements for radar aerials are :—

(*a*) They must be properly designed for the frequency of operation.

(*b*) The impedance of the transmission or feeder system must be matched to the aerial and transmitter output during transmission and to aerial and receiver during reception.

(*c*) Generally the weight of the aerial must be as small as possible, in order that it may be installed and transported easily.

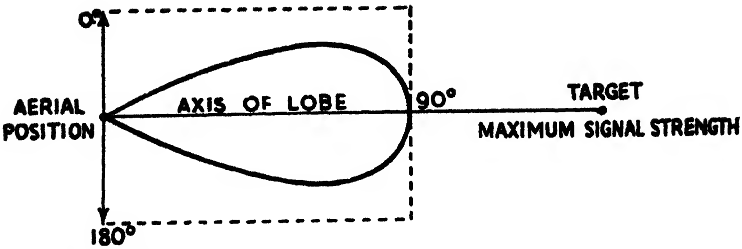


Fig. 62.—LOBE OBTAINED BY PLOTTING SIGNAL STRENGTH AGAINST DISTANCES MEASURED AT ANGLES FROM 0° TO 180° ROUND THE AERIAL.

(*d*) The structure must be strong, mechanically, to withstand vibration and shock.

(*e*) Wind resistance must be made as small as possible.

(*f*) It must be mechanically reliable, capable of rotating in azimuth or elevation, or both.

On account of these conflicting requirements, the final mechanical design, obviously, must be a compromise.

The simplest form of aerial for measuring azimuth is one which produces a single lobe. In this case the aerial is rotated and the beam scans in azimuth until an echo is received. The position of the aerial is then adjusted to give maximum signal strength.

In Fig. 62 relative signal strength is plotted against position of the aerial in respect to a target. The resulting pattern is shown as a lobe.* The maximum signal is received when the axis of the lobe passes through the target.

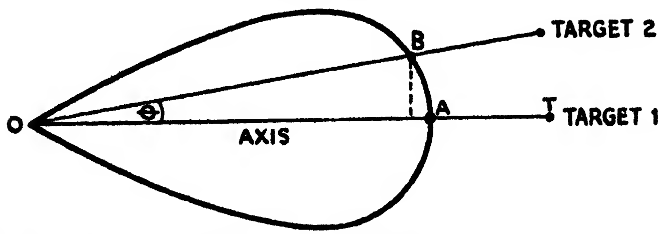


Fig. 63.—OA—OB IS VERY SMALL. CONSEQUENTLY WITH THE AXIS OT PASSING THROUGH TARGET 1, TARGET 2 WOULD RECEIVE NEARLY AS MUCH ENERGY AS TARGET 1, AND THE ECHOES WOULD BE VERY NEARLY EQUAL IN AMPLITUDE.

* The lobe pattern must not be taken as an indication of range. Radii from the centre of the aerial to any point on the lobe boundary bear a fixed ratio to the radius measured along the axis. Consequently the shape of the pattern remains constant because the ratios of all radii to the radius measured along the axis are constant.

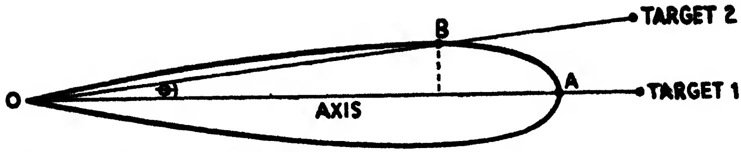


Fig. 64.—OA—OB IS VERY MUCH GREATER THAN IN FIG. 63, AND TARGET 2 WOULD ONLY RECEIVE ENERGY PROPORTIONAL TO $\frac{OB}{OA}$ RECEIVED BY TARGET 1. THE ANGLE T_2, O, T_1 IS SIMILAR TO THAT SHOWN IN FIG. 63.

Beam Switching

When using a single beam, particularly if it is wide, accuracy of adjustment depends upon the operator's ability to distinguish small changes in signal strength for a small degree of rotation about the point where the axis of the lobe passes through the target.

Examination of a typical lobe as shown in Fig. 63 discloses the fact that when the angular width of the lobe is considerable, as in wide beams, the relative signal strengths at A and B are very little different. When the beam is narrow as in Fig. 64, the rate of change of signal strength for the same number of degrees rotation of the aerial is considerable.

Fig. 65 shows that the variation of signal strength per degree of rotation of the aerial is greater near the side of any lobe than near the axis. The greatest rate of change is between the angle which gives 50 to 85 per cent. of the maximum power (measured along the axis).

When wide beams are used and precise measurement of bearing is required, a comparative method based on this fact is employed in order to determine accurately the bearing upon which the centre of the aerial faces target.

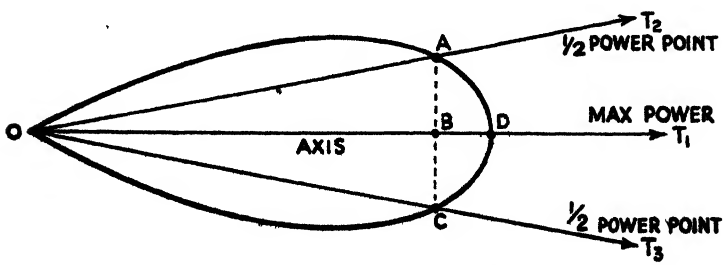


Fig. 65.—OB IS THE PROPORTION OF SIGNAL STRENGTH FOR T_2 AND T_3 AS COMPARED WITH T_1. OD—OB IS AN OPTIMUM CHANGE PER DEGREE.

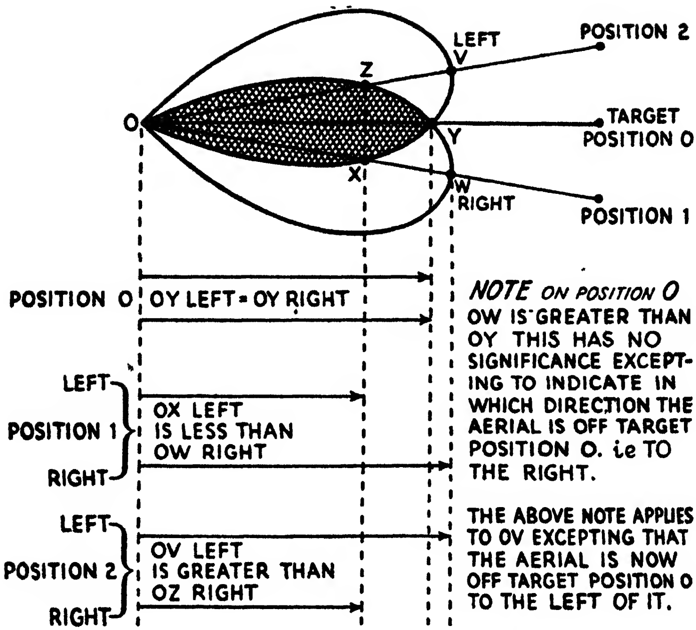


Fig. 66.—GRAPHICAL COMPARISON OF ENERGY RECEIVED AT O WHEN BEAM IS SWITCHED TO LEFT AND RIGHT OF THE TARGET.

This method (known as beam switching), employs electronic means for switching or displacing the axis of the beam alternatively to right and left of the normal axis, so that if the centre of the aerial is held fixed to face target, the beam itself sweeps right and left across the target and a double lobe is produced.*

The success of this device depends upon the fact that the signals received during the right displacement are *equal in magnitude to those received during the left displacement only when the centre of the aerial is facing the target.* It is important to realise that when beam switching is employed a *maximum* signal is *not* received when the aerial faces target. *The on-target position is obtained by comparing the magnitudes of the right- and left-hand signals* for *equality*, and not maximum amplitude. The aerial is rotated until the right- and left-hand signals as displayed on the indicator are equal in height, *i.e.*, in amplitude.

* Beam switching is continuous and quite independent of the instantaneous direction of the aerial, also it is independent of the mechanism which rotates the aerial in azimuth.

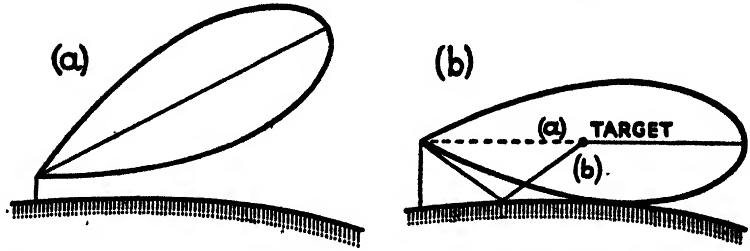


Fig. 67.—HOW THE SPACE PATTERN OF AN ARRAY IS MODIFIED BY GROUND REFLECTIONS.
(a) Aerial elevated. (b) Aerial lowered.

If the centre of the aerial is moved out of line with the target to face position (1), Fig. 66, the right-hand signal proportional to OZ approaches zero and the left-hand signal proportional to OX increases approximately to OV. In this case, in order to make both signals equal (on-target position) the aerial must be rotated to the left until its axis OY coincides with the target as shown in Fig. 66.

Measurement of Elevation and Height

Elevation is expressed as an angle, height as an altitude. The free space pattern of an array is based on the arrangement of the individual elements of the system. When placed near the ground, however, this pattern may be considerably modified by ground reflections (Fig. 67, (a) and (b)).

If the aerial is elevated so that energy does not strike the earth's surface (Fig. 67 (a)), the target receives a direct ray only, and returns an echo in the normal manner.

If the aerial is lowered so that some of the transmitted energy strikes the earth's surface, reflection takes place as shown in Fig 67 (b). In this case the target receives energy from two sources, (*a*) as a direct ray, and (*b*) by reflection from the earth's surface.

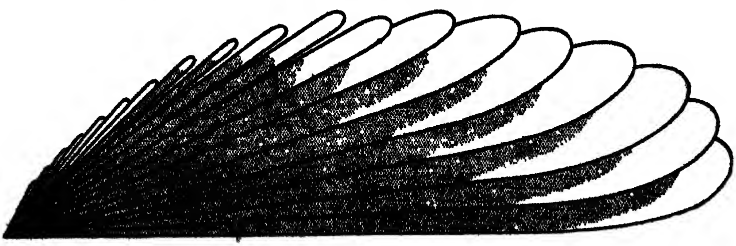


Fig. 68.—VERTICAL AVERAGE PATTERN.

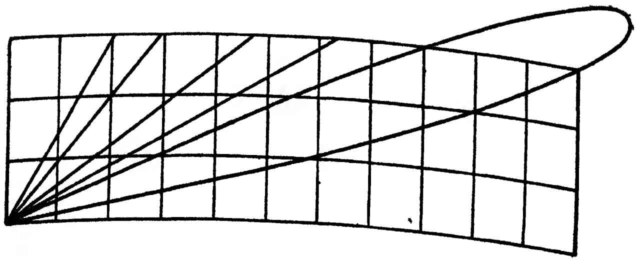


Fig. 69.—Typical calibration chart for estimating aircraft altitude.

Modification to Free Space Pattern

Clearly the paths of the direct and indirect rays have different lengths, and therefore, according to the position of the target relative to the transmitter and to the frequency employed, the direct waves and indirect waves must meet either in or out of phase. This means that the free-space lobe pattern is modified, since, inevitably, there must be areas where the direct and indirect rays reinforce each other, and also gaps where they either partially or wholly cancel one another. These conditions are illustrated in Fig. 68.

Elevation may only be determined directly by using a vertically narrow beam and tilting the aerial, thus avoiding the ground wave. When these conditions do not apply, altitude must be obtained from the known vertical coverage pattern of the aerial (Fig. 69).

A standard coverage pattern can be obtained by flying an aeroplane round the aerial at varying heights and plotting the ranges at which minimum usable signals are received.

Height finding from a coverage diagram is not very reliable, because a number of planes flying together return a signal much larger than a single plane. Modifications of this method have been developed which give much more reliable results.

Tilted aerials measure elevation directly in the same way as azimuth is measured. This method, when it can be used without ground reflection, is preferred.

Conical Scanning

A form of beam switching known as "conical scanning" is employed with narrow beams to obtain a high degree of accuracy in the measurement of elevation. This is dealt with in a later chapter.

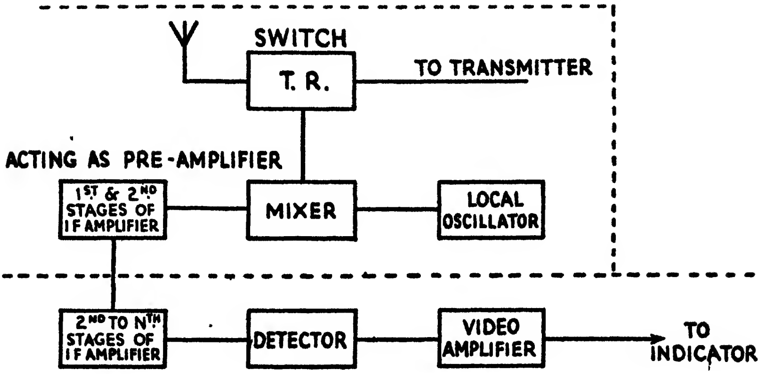


Fig. 70.—TYPICAL DISTRIBUTION OF RECEIVER CIRCUITS FOR RECEPTION OF U.H.F.

Receivers

A receiver of the superheterodyne pattern is usually employed in radar receiving circuits. Automatic gain control and automatic frequency control are frequently incorporated. Automatic gain control is generally applied to a particular echo when it is under close observation. Automatic frequency control is embodied in order to keep the intermediate frequency constant, by adjusting the frequency of the local oscillator, to compensate for any change of frequency of the received echo signal.

In U.H.F. practice the R.F. pre-amplifier is omitted when the " signal to noise " ratio makes its inclusion not worth while. In this case the mixer and local oscillator are generally located as close to the aerial as possible. Usually one or two stages of the I.F. amplifier are also fitted at this position to act as a pre-amplifier, so that a received signal of reasonable magnitude can be passed along the transmission line to the receiver proper. In the absence of some form of pre-amplification, losses along the transmission line would attenuate the signal so much that on arrival at the receiver it would cease to be useful.

Part of the receiver output circuit may be located in the indicator unit, thus the receiver circuits in some systems may be distributed in such a manner that the receiver—as a unit—cannot be identified (see Fig. 70).

Display Unit

The C.R.T. of the indicator unit displays the information and records data in visual form in a manner most suitable to the particular application and to the observer's requirements.

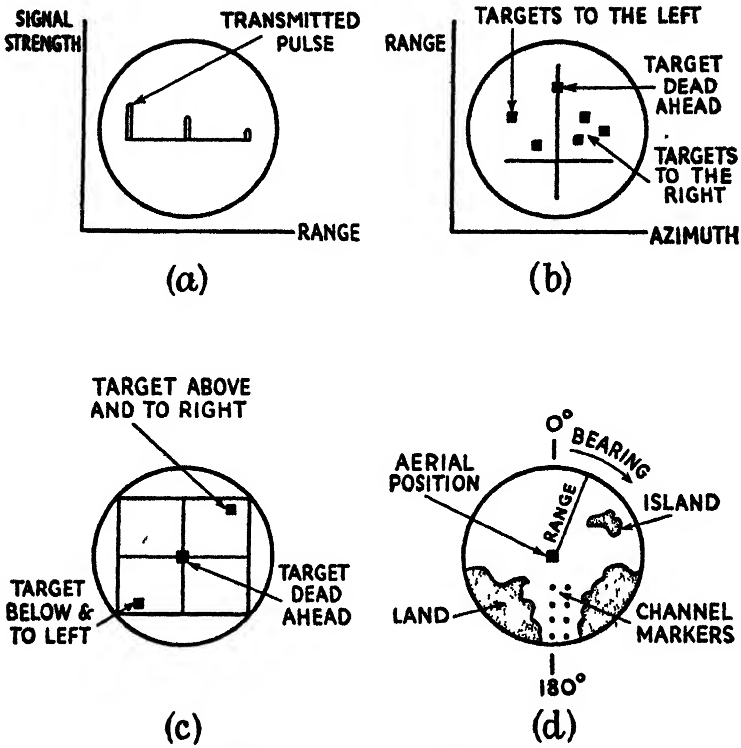


Fig. 71.—TYPES OF DISPLAY.

(a) Type " A " display.
(b) Type " B " display.
(c) Selected range display.
(d) P.P.I. display.

There are therefore many forms in which data can be displayed, and many types of indicators are available.

All indicator units embody one or more cathode ray tubes, together with their associated time base and input circuits. In general they differ only in regard to time bases and input arrangements.

Types

Indicators are classified according to the displays they provide. "A"- and "B"-type displays are shown in Fig. 71 (a) and (b). There are also selected range indicators and various other types, including the P.P.I. display.

The "A" display plots signal amplitude against range (Fig. 71 (a). The "B" display plots range against azimuth (see Fig. 71 (b)). In this case the echo appears as a bright spot somewhere on the screen. A magnetic tube is usually employed, and the position of the bright spot is determined by the time base current flowing through the vertical deflecting coils, whilst a positioning current is applied to the horizontal coils by a potentiometer device operated by the rotation of the aerial. The output of the receiver is applied to the grid or cathode of the C.R.T. for intensity modulation.*

The "B" type of display is generally used when beam switching is employed. The right- and left-hand signals then appear on either side of the "Y" axis taken across the vertical diameter of the tube. "C" is a selected range indicator.

The P.P.I. display (Fig. 71 (d)) has already been outlined. It presents in polar co-ordinates a map of the area being scanned by the aerial, the latter occupying the centre of the screen. The tube, which is usually of the electromagnetic type, is intensity modulated, with the time base moving radially outwards from centre to circumference.

Indicator Unit Components and Sub-functions

The basic components of an indicator unit are :—

(*a*) Cathode ray tube.

(*b*) Time base generator.

(*c*) Circuits for providing the brightening pulse for the duration of each sweep.

(*d*) Calibration and fixed range mark generator.†

(*e*) Other circuits, such for example as positioning and clamping circuits to ensure that the time base shall always start

* Note on displays "B" and "C." Under the joint action of the time base voltage or current applied to the "Y" plates and a voltage proportional to the instantaneous position of the aerial, applied to the "X" plates, the beam at low intensity scans the entire surface of the C.R.T. screen when the aerial is revolved. This scanning takes place in darkness until the arrival of an echo. The received signal after amplification is applied to the grid of the C.R.T. and intensifies the beam, thereby causing a bright spot to appear on the screen at a point which represents the range and bearing of the target relative to the ship's head. In a "C" display, when used in an aeroplane, inputs can also be arranged to show whether the target is above or below the observer.

† Although not strictly included in the basic radar circuit, calibrators are provided with nearly all display units. Chapter XV is devoted to calibration, but in the meantime it is sufficient to state that a calibration generator started by a Synch. pulse at the same time as the time base generator generates "pips" which appear on the trace as bright spots and thus calibrate the trace, generally in lengths of 2,000 yds., to facilitate range measurement by interpolation, or to check against a transparent scale the accuracy of the time base circuits (sweep generator).

and stop at the same positions on the screen, also special forms of circuits to apply the receiver output in push-pull to the appropriate deflecting plates.

Various additions and refinements to the above are added to suit special applications. Also several methods of data presentation requiring time bases of varying degrees of complexity may be employed.

It will be appreciated that this chapter is intended merely to give an outline of the basic radar system. Details of the functioning of the various units and sub-units are contained in subsequent chapters, but before dealing further with systems as a whole, it is first necessary to study the various types of special circuits used for synchronising, pulse forming and general control.

CHAPTER VIII

THE BASIC PRINCIPLES OF PULSING SYSTEMS

THIS chapter is written as a foreword to the study of the various special circuits used in radar systems; many of these circuits may appear to be alternative methods of performing the same function, and to a certain extent this is true.

Outputs may be similar, but it will be found, generally, upon examination that important differences exist in regard to input conditions, control, loading and accuracy.

Because of the large number of supplementary circuits which can be added to the basic radar circuits in the way of refinements and special facilities, the pulsing system for radar systems can become complex.

The Synchronising Function

The main factor which determines the pulsing system layout is, however, the source selected for the master control or synchronising function. It may be chosen from any of the following alternatives :—

(*a*) The A.C. mains supply.

(*b*) The output from an independent oscillator having very stable characteristics.

(*c*) A pulse taken from the transmitter circuit each time it fires. In this case all the other circuits in the system are synchronised to the transmitter.

(*d*) A signal taken from the calibration generator circuit—a sub-unit of the display or indicator unit.

(*e*) A signal taken from the time base generator circuit at the commencement of each sweep of the time base. In this case all the other units are synchronised to the time base generator, *i.e.*, to the commencement of the sweep.

Any one of the above may be selected to perform the synchronising function and to establish the pulse repetition frequency for the whole system.

If the Synch. pulse is taken from either of the sources (*a*) or (*b*), it will, in general, be sinusoidal in shape. A Synch. pulse taken from (*c*) is, however, likely to be rectangular in shape, with steep vertical leading edge and a time duration (width) related directly to that of the R.F. envelope.

Pulses taken from sources (*d*) or (*e*), however, are likely to be square, but related in time to the sweep time, which is also the same as the pulse width of the brightening pulse.

All these pulses are quite suitable for performing the purely synchronising function, but because they all have definite time duration (width) and, therefore, could exercise a time duration control, four cases arise :—

(*a*) The time duration control may be undesirable. In this case the Synch. pulse must act as a " trigger " pulse only, the output from the circuit which it " triggers " being determined by the electrical constants of the triggered circuit itself.

(*b*) The duration of the Synch. pulse may be suitable and desirable, but its magnitude may be too great or too small to perform the Synch. function properly.

(*c*) The duration of the Synch. pulse may be too short and/or of unsuitable amplitude.

(*d*) The duration of the Synch. pulse may be too long and/or of unsuitable amplitude.

To the above must be added the general case in which the source must not be overloaded or distorted by the nature of the load when the shape of the Synch. pulse is rigorously to be preserved.

Considering the general case first, if the load on the output of the Synch. source is likely to distort the input or generated pulse shape, either because of its magnitude or because of its composition, two courses are open :—

(*a*) To generate from the leading edge of the Synch. pulse another pulse whose leading edge shall coincide in time with that of the Synch. pulse itself.

(*b*) To introduce a buffer stage between the Synch. source output and the load (see cathode follower, Chapter XI).

Consideration of the special cases (*a*) to (*d*) leads directly to a distinction between true " trigger " pulses and " trigger-timing " or time-control pulses, the latter being capable of performing a double function.

It is therefore convenient, if not conventional, at this juncture to think of all such pulses as falling under the classifications of Trigger pulses and Trigger-Timing pulses.

The " Trigger " Pulse and Trigger-Timing Pulse

The " trigger " pulse, as the name implies, is employed to start or stop some operation. It is sharp pointed and exercises

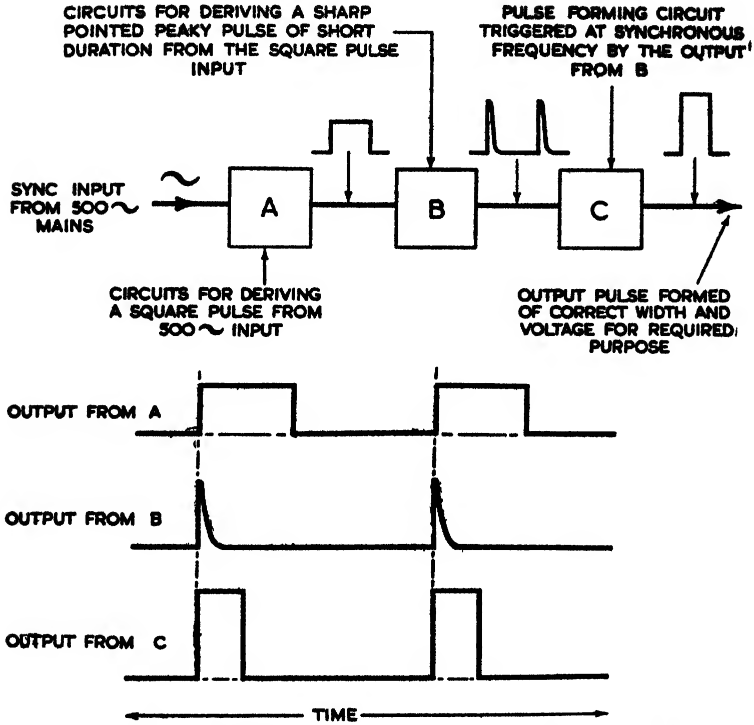


Fig. 72.—DIAGRAM TO ILLUSTRATE DERIVATION OF A PULSE OF REQUIRED WIDTH AND VOLTAGE AT SYNCHRONOUS FREQUENCY, WHEN A SINUSOIDAL WAVE OF LONG TIME PERIOD IS USED FOR SYNCHRONISING PURPOSES.

no time duration control over the operation of the circuit to which it is applied.

The trigger-timing or control pulse is generally flat topped, rectangular and with width proportional to some definite period of time. Such a pulse can be used to start or stop an operation, by its leading edge, and to stop or start the operation again by its trailing edge. In this case the time interval between the leading edge and the trailing edge determines the duration of the effect which the pulse exercises upon the circuit to which it is applied.

Both trigger and trigger-timing pulses have steep vertical leading edges. *It is the leading edge that performs the Synch. function in the majority of cases; delay circuits excepted.*

Thus when a Synch. pulse taken from some Synch. source is unsuitable in duration (width) or amplitude, a new and more suitable pulse must be developed from its leading edge. Since the new pulse can be developed almost instantaneously, for all practical purposes the leading edge of the new pulse coincides with that of the Synch. pulse from which it is developed—the time lapse is negligible.

If the Synch. pulse in cases (*a*) and (*b*) is sinusoidal, the wave is rectified in a half-wave rectifier and the half cycle thus obtained is then roughly squared possibly by an operation of amplitude limiting. If necessary, the semi-square pulse thus obtained may be further squared by a succeeding circuit.

If this pulse has been derived from, say, 500-cycle A.C. mains, it will be very wide and its time period will be $\frac{1}{2000}$th second or 2,000 microseconds.

If, for example, a trigger pulse only is required, it can be developed from the leading edge of this long square pulse by a process of differentiation (see Chapter IX), and amplified, if necessary, after this operation has been performed.

If, on the other hand, a trigger-time control pulse is required of, say, 2 microseconds, the long square Synch. pulse may first be differentiated as above and the sharp-peaked pulse resulting from this operation is then used as a trigger pulse to start up or fire a generator, the output of which is a pulse of 2 microseconds, at the required amplitude. The output of this generator is determined by the electrical constants of the generator circuit. Both these operations are shown in diagrammatic form in Fig. 72.

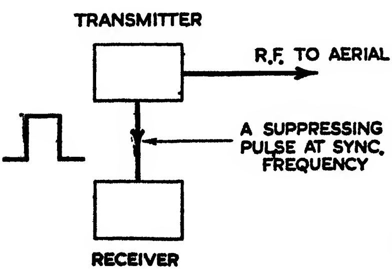


Fig. 73.—ILLUSTRATING A CASE IN WHICH THE SYSTEM IS SYNCHRONISED BY A PULSE FROM THE TRANSMITTER, AN OUTPUT PULSE BEING TAKEN TO THE RECEIVER TO RENDER IT INOPERATIVE DURING THE PERIOD OF TRANSMISSION.

If the pulse taken from the Synch. source is rectangular, the leading edge of this pulse may be differentiated for a trigger pulse, and this may subsequently be used to fire a generator, the output of which may be another square pulse of different magnitude or amplitude, or both, as in the last case. The trigger pulse may, however, be used equally well to fire a generator having a sawtooth or trapezoidal output, since the shape of the output is determined entirely by the constants of the generator.

The Synch. pulse might equally well be a sawtooth or a trapezoidal wave. In either case a new pulse must generally be developed from the leading edge of the Synch. pulse, and in the case of the sawtooth wave this can also be performed by two consecutive operations.

When a sawtooth wave is applied to a differentiating circuit (see Chapter IX) it can be shown that the output is a square wave.

If a trigger pulse is now required, the square wave obtained by the first process of differentiation is differentiated again, and a sharp peaked pulse is developed from its leading edge to coincide in time (for all practical purposes) with the leading edge of the original Synch. pulse.

If a trigger-timing pulse of the same duration as the sawtooth wave is required, the square wave obtained from the first differentiation can be used; or if of insufficient amplitude, it can be amplified and trimmed up.

Alternatively, if this square wave obtained by the first operation is not of the required duration, a new pulse of the right amplitude and duration is generated by the following sequence: (1) differentiation of the sawtooth wave to produce a square wave; (2) differentiation of the square wave to produce a sharp peaked wave; (3) application of the trigger pulse thus obtained to fire a generator arranged to give the required trigger-timing pulse.

Differentiation of a trapezoidal wave produces a sharp peaked pulse for the vertical part of the wave positive going and another peak at the end of the sawtooth part of the trapezoid which would be negative going.

This introduces the question of polarity of Synch. pulses in general.

Polarity of Pulses

The required polarity of a pulse depends upon the input conditions of the circuit to which it is to be applied. For example, a trigger pulse is sometimes required to raise the grid of a valve to a conducting potential, at other times it is required to depress the grid towards cut off.

When a Synch. pulse is of the right shape and amplitude, but of the wrong polarity, it can be inverted. This however involves an additional operation which may be avoided perhaps by modifying the layout of the pulsing system as a whole.

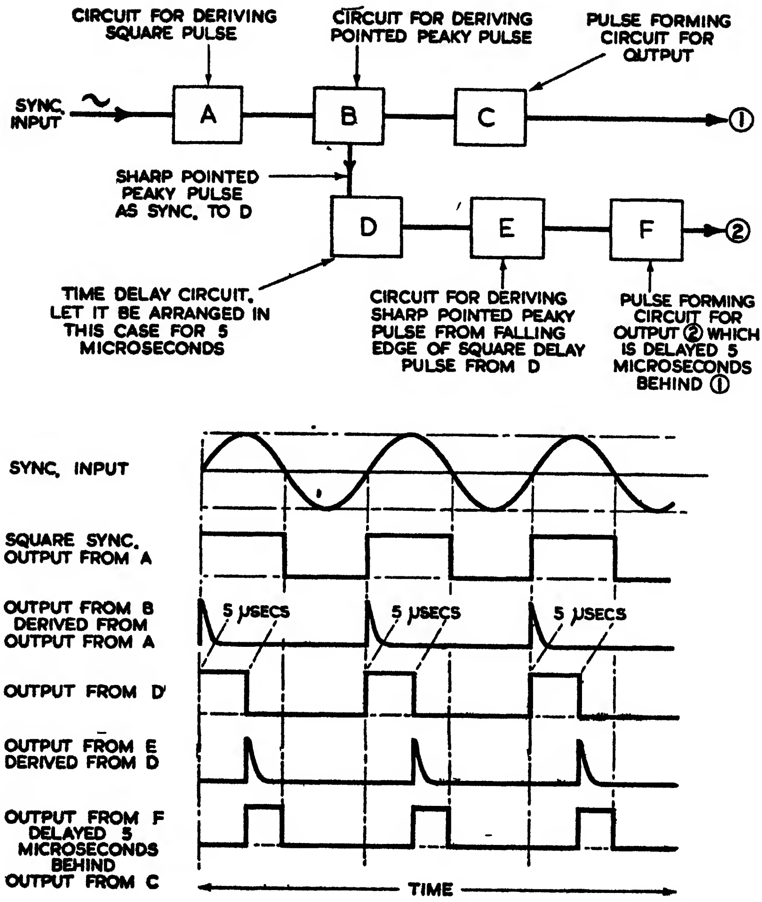


Fig. 74.—ILLUSTRATING USE OF SQUARE PULSE AS A MEASURE OF TIME DELAY. Output 2 from D E F is delayed 5 microseconds behind Output 1.

Layout of Pulsing System

The layout of a pulsing system is determined primarily by the source from which the Synch. pulse is taken, and this in turn is determined by other considerations. The source having been decided by the type and arrangement of the major units to be employed, the number and variety of functions to be performed is the next important factor. This will determine the number of Synch. pulses (trigger and trigger-timing) required, their magnitude, and in the latter case their duration. With this information, a pulsing system can be laid out to provide the necessary pulses, having due regard to their shape, magnitude and polarity, and utilising the fewest number of circuits consistent with efficiency and reliability.

Time Delay

In radar systems it is often essential to bring some particular circuit into operation at some time intermediate between successive Synch. pulses. This means that the commencing operation of this particular circuit must be *delayed in time relative to each synchronising pulse* and therefore to the operation of all other units in the system.

In order to accomplish this, a rectangular pulse suitably dimensioned in width (*time*) is generated such that the leading edge of the generated pulse coincides with the leading edge of the Synch. pulse, and the *time interval* or width between the leading and trailing edges of the rectangular pulse is made equal to the required *time delay*.

The rectangular pulse is now operated upon. Its leading edge is ignored but a sharp-pointed pulse is developed from the trailing edge and to coincide with it in *time*. The sharp-pointed pulse developed from the trailing edge of the rectangular delay pulse is now used to trigger the *delayed* circuit, which comes into operation at the required time interval after each successive Synch. pulse (see Fig. 74).

CHAPTER IX

VOLTAGE DIVIDERS

PROPERTIES OF R.C. AND L/R CIRCUITS, INCLUDING DIFFERENTIATION AND THE DIFFERENTIATOR AMPLIFIER

It is convenient to examine the properties of R.C. and L/R circuits, as applied to radar, from three different angles :—

(1) Voltage division.
(2) Phase-shifting properties.
(3) Differentiating action.

Any R.C. or L/R circuit may be viewed from any one of the above angles, the results obtained in each case depending upon the applied frequency and the proportions of capacity and resistance, or inductance and resistance in the circuit.

In radar, pulses of many different and complex shapes are used, particularly square, sawtooth and peaked, and since these pulses may be regarded as parts of waves, all of which repeat themselves at definite time intervals, they must therefore be composed of sine waves of different frequencies and amplitudes added together, as shown in Fig. 17, p. 42.

If a circuit is to pass a non-sinusoidal pulse without changing its shape, it must pass all component frequencies without shifting their relative phases and without changing any of their amplitudes.

The perpendicular front, which occurs with a perfectly square pulse, is composed of a fundamental and an infinite number of frequencies. In theory therefore it is impossible to pass a pulse of this form without some distortion.

Nevertheless circuits can be designed with an even response from a few cycles to several megacycles. Such a circuit will pass a square pulse with very little distortion. While the leading and falling edges of the pulse are always slightly sloped or rounded (thereby indicating some time of rise and fall), this rise and fall time can be made very small *.

Consequently when analysing the R.C. and L/R pulse

* A wave with perfectly square corners cannot exist since an instantaneous change of 90° implies an infinitely great rate of change which may only be approximated.

circuits used in radar it is usually convenient to consider voltage distribution and phase relations.

Pulse-coupling circuits are generally composed of resistance, or combinations of resistance, and capacity, or elements, resistance and inductance in series, consequently it is essential to have a clear understanding as to how such circuits function under each of the three headings given above, in order that the factors which control the fidelity or distortion of the output in relation to the input may be suitably proportioned to reproduce at the output fidelity or distortion of the input at will.

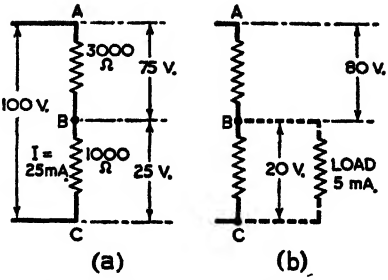


Fig. 75.—VOLTAGE DIVIDERS OR POTENTIOMETERS.

(a) Simple series voltage divider (unloaded).
(b) As in (a) but loaded.

Voltage Dividers, D.C. and A.C.

In the case of the D.C. fixed potentiometers shown in Fig. 75 (a),

$$I = \frac{E}{R_1 + R_2}$$

$$E = E_1 + E_2 = IR_1 + IR_2$$

$$E_1 = E\frac{R_1}{R_1 + R_2} \text{ and } E_2 = \frac{R_2}{R_1 + R_2}.$$

In Fig. 75 (a) the current through the chain ABC is 25 mA., and

Voltage across AB = 75 volts
,, ,, BC = 25 volts.

If a load requiring 5 mA. is placed across BC, the voltage distribution alters (Fig 75 (b)).

Voltage across AB = 80 volts approx.
Voltage across BC = 20 volts approx.

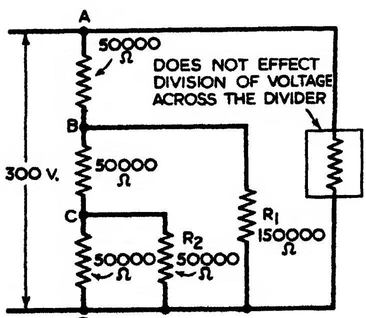


Fig. 76.—VOLTAGE DISTRIBUTION IN RESISTANCE NETWORKS OF POTENTIAL DIVIDERS.

In Fig. 76 :—

$$\begin{cases} R_{CD} = \dfrac{50\text{ k.} \times 50\text{ k.}}{50\text{ k.} + 50\text{ k.}} = 25\text{ k.} \\ + R_{BC} = 25\text{ k.} + 50\text{ k.} = 75\text{ k.} \end{cases}$$

$$(R_{CD} + R_{BD}) = R_{BD} = \frac{75\text{ k.} \times 150\text{ k.}}{75\text{ k.} + 150\text{ k.}} = 50\text{ k.}$$

Total resistance AD $= 50$ k. $+ 50$ k. $= 100$ k. $=$ $[AB + (R_{CD} + R_{BD})]$

Total current $= 3$ mA.

The voltage distribution is therefore :—

Across AB $= 3$ mA. $\times$ 50 k. $= 150$ volts.
BC $= (3$ mA. $- 1$ mA.$)$ 50 k. $= 100$ volts.
CD $= 300$ v. $- (150$ v. $+ 100$ v.$) = 50$ volts.

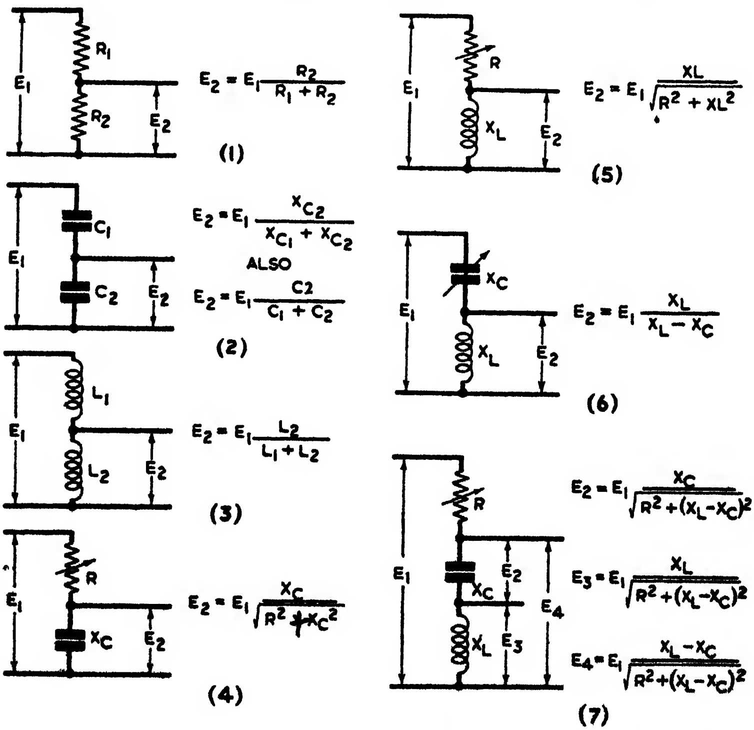


Fig. 77.—VOLTAGE DISTRIBUTION IN RESISTANCE, CAPACITY, INDUCTANCE AND MIXED FORMS OF VOLTAGE DIVIDERS.

The current distribution is :—

Through AB $= 3$ mA.
,, R_1 $= 1$ mA.
,, BC $= (3$ mA. $- 1$ mA.$) = 2$ mA.
,, R_2 $= 1$ mA.
,, BD $= 1$ mA.

A.C. Voltage Dividers

In an A.C. circuit, voltage division can be effected by any of the following arrangements (Fig. 77 (1–7)). The voltage ratios are given for each combination, but these are only accurate when the load impedance is large. The input is assumed to be a pure sine wave.

It will be noted that circuits 1, 2 and 3 may be classed as simple A.C. dividers. One type of component only is used in each circuit, consequently voltage division can be obtained without phase shift between either of the outputs. Circuits 4, 5, 6 and 7 are complex. It is possible with these circuits to vary the voltage division and at the same time obtain a phase shift.

The magnitude of the output in all cases is proportional to the ratio of the impedance across which the output is taken to the total impedance of the dividing circuit.

Phase-shifting Properties

Series circuits containing capacity and resistance and/or inductance can be employed as phase-shifting devices when contingent effects of voltage division are of minor importance.

Phase-shifting Properties of R.C. and L/R Circuits

The phase-shifting performance of such circuits can be calculated to give tolerable accuracy, but when precise measurement is required other means must be adopted.

When the input is a pulse or part of a wave which repeats itself at definite time intervals, the times at which the condenser must charge and discharge are determined by the time period for one half cycle—in other words, by the frequency of the input. Consequently the time constant for any given circuit, as determined by its C.R. or L/R components, may be shorter, or longer, or of equal value to the time period of the frequency of the input.

The *ratio* of the *time constant* of the circuit to the *time period* of the input has important effects upon the output.

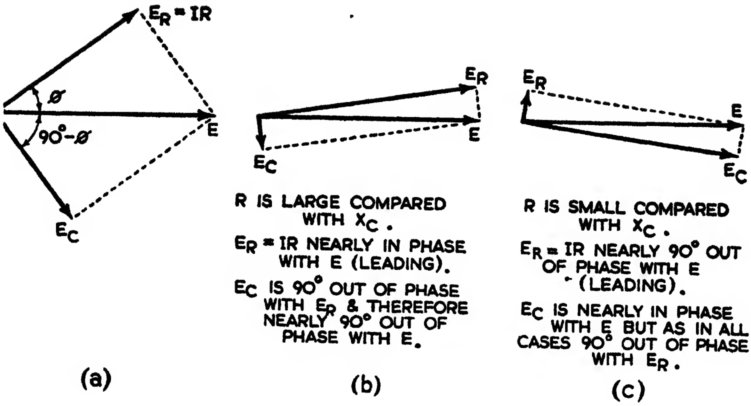


Fig. 78.—VALUES AND PHASES OF COMPONENT E.M.F. IN A SIMPLE RESISTANCE CAPACITY POTENTIAL DIVIDER.

(a) E resolved into component vectors. E_C lagging and E_R leading.
(b) R large and X_c small.
(c) X_c large and R small.

When the voltage input is a pure sine wave, the ratio time constant/time period affects the amplitude and phase of the voltage output. Since, however, there is only one frequency involved, viz., the fundamental frequency, the voltage output waveform is not changed. In this case it is another sine wave of smaller amplitude, displaced in phase from the input by some angle, determined by the relative values of the components, the component across which the output is taken and the frequency of the input.

The paragraphs which immediately follow are therefore devoted to the effects on the phase and amplitude of the output voltage wave when a pure sine wave of voltage is applied as an input to R.C. and L/R circuits.

Phase Shift in Resistance Capacity Series Circuits

When resistance and capacity are in series with an A.C. voltage, having the form of a *pure sine wave*, the voltage drop across R is in phase with the current I.

The voltage across the condenser E_c is, however, 90° out of phase with the voltage across R.

The phase of the current flowing through R is dependent upon the values of the resistance (R) and capacity reactance (X_c) combined to form the total impedance Z.

In such a circuit there are therefore three e.m.f.s to consider, viz. :—

E the applied e.m.f.

$E_R = IR$, and bearing the same phase relation to E as the phase of I does to E, viz. ($\theta = \tan^{-1} \omega CR$), and

E_c which is related in phase to E by the angle θ, $\tan^{-1} \omega CR - \frac{\pi}{2}$

This can be shown graphically, as in Fig. 78 (a).

Vector E is the vector sum of E_R and E_c.

,, $E_R = IR$ with a phase angle to E of $\theta = \tan^{-1} \omega CR$.

,, E_c with a phase lagging E_R by 90°.

θ is determined by the ratio of R to $\frac{1}{\omega C}$; therefore, if R is made large relative to X_c, the current will be nearly in phase with E and E_R will also have a similar phase relationship. See Fig. 78 (b).

Conversely, if R is made small as compared with X_c the phase difference between E and I will be large, and consequently the phase difference between E_R and E will be large and the phase of E_c will approach E, see Fig. 78 (c).

In both cases E_c must be 90° out of phase with E_R. The phase relation of E_c to E depends upon the phase angle between E and E_R. Consequently when X_c is large compared with R, E_c is nearly in phase with E; conversely, when X_c is small compared with R, E_c swings further out of phase with E.

When inductance and resistance are in series, the current flowing lags the applied voltage E by $\theta = \tan^{-1} \frac{\omega L}{R}$.

The drop across the resistance R is in phase with this current, therefore the IR drop E_R lags E by the angle θ.

The voltage across the inductance E_L leads E_R by 90°, and E is the vector sum of E_R and E_L, see Fig. 80 (a).

If R is small compared with X_L the current through R will lag E nearly 90° and E_R will lag E by a corresponding amount. E_L, however, will approach E, so that if E_R is nearly 90° lagging E, then E_L will be almost in phase with E, see Fig. 80 (b).

If R is large relative to X_L, the phase of E_R approaches E more nearly and E_L leads E by an angle approaching 90°, see Fig. 80 (c).

Summary and Conclusions

Analysis of Fig. 78 (a), (b) and (c) yields the following results :—

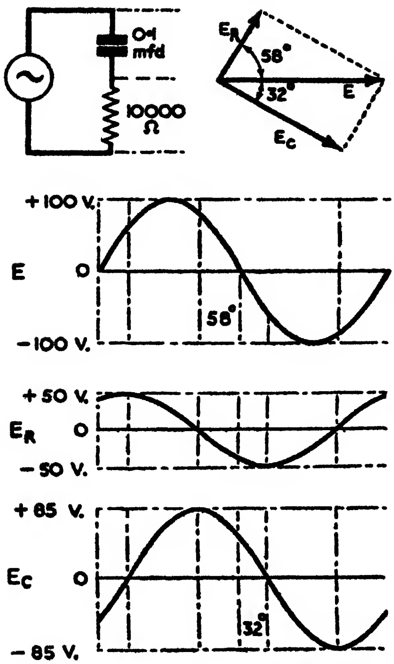


Fig. 79.—VECTOR DIAGRAM AND CURVES OF VOLTAGE DISTRIBUTION ACROSS CAPACITY AND RESISTANCE SERIES.

(1) When the input voltage is a pure *sine wave*, the output voltage taken across either component can be changed in amplitude and phase, but not in *shape*.

(2) If E is a pure *sine wave*, E_R and E_c are both pure sine waves displaced in phase, relative to E by angles the vector sum of which is constant, *i.e.*, 90°.

The respective angular displacements of E_R and E_c relative to E are determined by the values X_c and R.

(3) Since X_c varies indirectly with frequency, an increase of frequency applied to a circuit having fixed values of C and R causes the output across R to increase in magnitude and the phase of E_R to approach nearer to E. Parallel with this the output across $C = E_c$ diminishes in amplitude and its phase displacement from E increases, maintaining the vector sum of E_c and E_R equal to 90° (see Fig. 78).

(4) If the frequency is fixed and the values of C or R are varied, the results are as shown in the vector diagrams, Fig. 78. Fig. 79 shows the amplitudes and phases of E_R and E_c relative to a *sine wave input E* when the frequency is 1,000 cycles and a resistance of 10 k. is in series with a capacity of 0·1 mfd. The time constant is 1,000 microseconds and the time period of the applied frequency is 1/1000th sec.*

(*a*) E_R has smaller amplitude than E and leads E by 58° (approx.).

(*b*) E_c is nearly equal to E in amplitude.

(*c*) E_c lags E by about 32°.

(*d*) The vector sum E_R and E_c equals 90°.

When the phase shift of E_R is such that it approaches very

* R.C. time constant $= 10^4 \times 10^{-1} = 10^3$ micro-seconds.

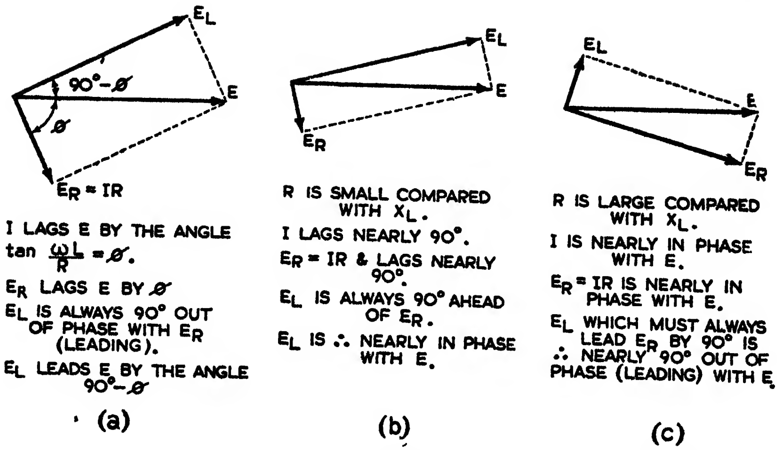


Fig. 80.—VALUES AND PHASES OF COMPONENT E.M.F. IN A SIMPLE RESISTANCE POTENTIOMETER.

(a) E resolved into vectors E_L leading and E_R lagging. (b) R small and X_L large. (c) R large, X_L small.

nearly to 90°, the output taken across R will be the *cosine wave* (approx.). Similarly when a *sine* wave of voltage is applied to a series R.C. circuit, the time constant of which can be varied in relation to the time period of the applied frequency, the voltage output taken across R is also a sine wave proportional to the rate of change, of the input voltage across R but shifted in phase relative to the input. This result is similar to the results obtained by the mathematical process of differentiation. A series R.C. circuit having a time constant that is short compared with the time period of the applied frequency is termed a differentiating circuit, and the output is said to be the differentiation of the input.

L/R Circuit

When an L/R circuit is employed the differentiated output is taken across L. Reference to Fig. 80 shows that an increase in applied frequency causes I to lag E by a greater amount, and to decrease in magnitude, wherefore E_L, which leads I by 90°, comes further into phase with E but is still leading relative to E. The voltage across L is, therefore, proportional to the rate of change of the input voltage.

When the input to an R.C. circuit is non-sinusoidal, such, for example, as a square,* sawtooth or peaked wave, there are

* Sometimes termed a voltage step.

other frequencies present besides the fundamental; consequently, any phase-shifting action which is applied to the fundamental frequency is also applied to the harmonics, *but in a greater or less degree than to the fundamental* according to their relative frequencies. Since the phase shift applied to a harmonic is proportional to its frequency, *all the harmonics* must be shifted in phase relative to the fundamental, and conditions for distortion will be present.

Thus, the time constant of a circuit must be *long* compared with the time period of the applied frequency if a *non-sinusoidal wave* is to be passed without serious alteration in shape, *i.e., if it is not to be differentiated.* If the voltage wave input is differentiated, the shape of the output voltage wave is changed.* Thus distortion to shape by *differentiation* can be caused and controlled by using a short time constant, the ratio of which to the time period of the applied frequency will determine the amount of distortion introduced and therefore the shape of the output wave.

DIFFERENTIATION

Before examining differentiating circuits, it is desirable to analyse the distribution of voltages and currents in R.C. and L/R circuits.

This examination is facilitated by assuming that the circuit is connected and disconnected from a D.C. source by switches. Both switching on and switching off times are in this way under control *of the respective switches, thereby avoiding any complications due to the ratio time constant/time period.*

In making an analysis, where the inputs are *non-sinusoidal,* A.C. technique does not apply; it will be noticed also that the conditions are such that it is the *transients* caused by the sudden switching on and switching off of the D.C. supply that are under examination, and that it is these transients which produce an effect similar to a square wave input.

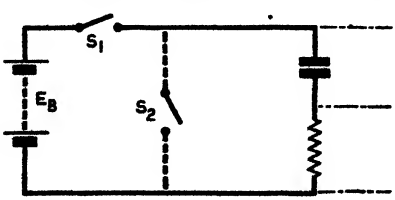


Fig. 81.—CHARGE AND DISCHARGE OF CONDENSER IN A CAPACITY RESISTANCE SERIES CIRCUIT.

When a condenser is connected to a source of constant e.m.f. and the circuit contains resistance, the condenser does not instantaneously acquire a p.d.

* This must follow since the output waveform reflects the rate of change of the input wave.

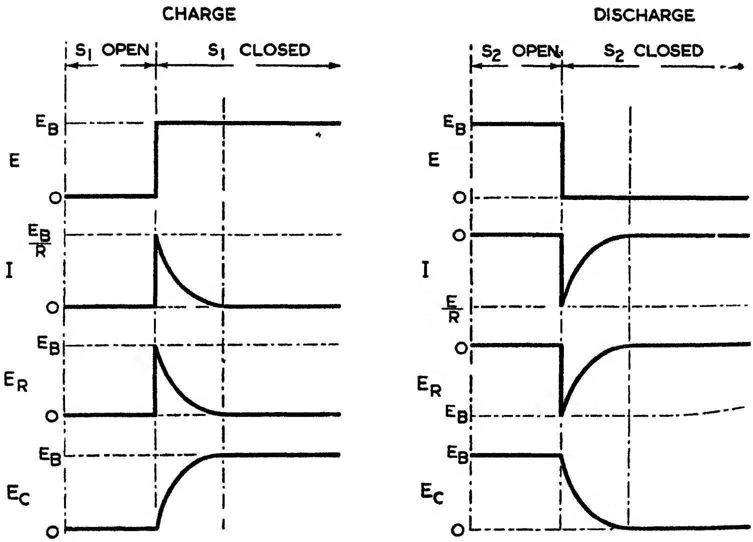


Fig. 82.—CHARGE AND DISCHARGE OF AN R.C. CIRCUIT (S_1 OPEN).

equal to that of the charging source. At the instant the e.m.f. is applied there is no charge in the condenser and, therefore, no p.d. across it at all. By Kirchoff's second law a drop must appear somewhere in the circuit equal to the charging voltage. This can only take place at the instant in question across the resistance, consequently at the instant the charging switch S_1 is closed, see Fig. 81, the charging voltage E_B appears in full across R and a current $\frac{E_B}{R}$ commences to flow.

As the charge Q in the condenser grows a back e.m.f., $E_c = \frac{Q}{C}$ grows with it. This growing back e.m.f. (E_c) opposes the applied e.m.f. (E_B), consequently the charging current decays and the charge Q grows less quickly. The voltage (E_c) across the condenser thus grows towards the value of E_B in an exponential manner.

In fact, it can be shown that $E_c = E_B (1 - \epsilon^{-\frac{t}{RC}})$ where ϵ = the basis of the Naperian logarithm (2·718) and t is time in seconds.

Now the value of the constants R and C must obviously determine the charging time t. C determines the magnitude of the charge Q required to make $E_c = E_B$, and R determines the instantaneous charging current $\frac{Q}{t}$, which is $\frac{E_B - E_c}{R}$. Therefore, C is proportional to t, and because I is inversely proportional, R must also be proportional to t.

In the equation $E_c = E_B\,(1 - \epsilon^{-\frac{t}{RC}})$ we can put $t = R \times C$,

therefore
$$\begin{aligned} E_c &= E_B\,(1 - \epsilon^{-\frac{RC}{RC}}) \\ &= E_B\,(1 - \epsilon^{-1}) \\ &= E_B\left(1 - \frac{1}{2{\cdot}718}\right) \\ &= {\cdot}632\,E_B. \end{aligned}$$

This means that in any series R.C. circuit E_c will reach 63·2 per cent. of E_B in RC seconds where R is in ohms and C in farads, Q growing proportionally.

At the moment the charging current commences to flow through the resistance, it is at maximum $= \frac{E_B}{R}$, and because it then decreases exponentially, Q (and therefore E_c) must rise exponentially in order that Kirchoff's law may be satisfied. Thus at the end of RC seconds, when E_c and Q have reached 63·2 per cent. of their full value ($= E_B$), the current through R must have fallen to 36·8 per cent. of its original $\frac{E_B}{R}$ value, and the drop across R is 36·8 per cent. of E_B in RC seconds after the application of E_B.

Thus at time RC seconds after commencement of the charging operation, Kirchoff's law is satisfied because

$E_B =$		100	
$E_c =$	63·2		lagging E_R 90°.
$E_R =$	36·8		in phase with I.
	100·0	100·0	

Voltage waveforms and phases are shown in Figs. 82 and 84.

The product of R × C where R is in ohms and C in farads is the time constant of the circuit (in seconds) and represents the time required to charge the condenser to 63·2 per cent. of an applied voltage E_B.*

* If C is in microfarads and R is in ohms, the RC time constant is given in microseconds. If R is in megohms and C in microfarads, the RC time constant is in seconds.

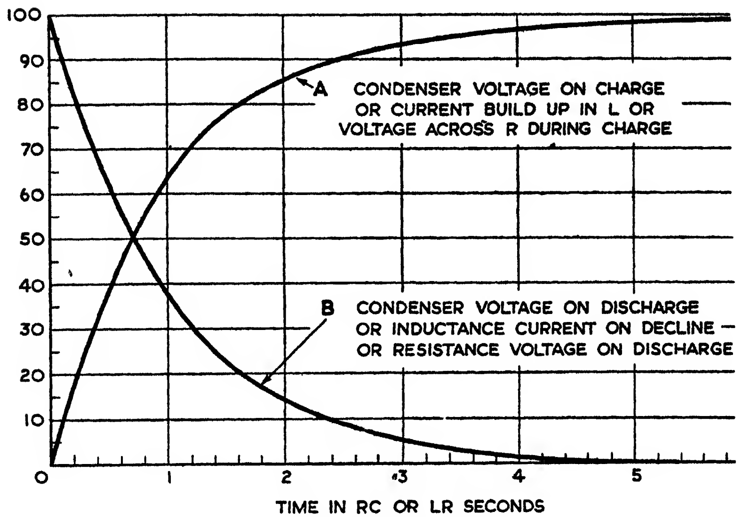


Fig. 83.—GENERAL FORM OF RC OR L/R TIME CONSTANTS.

Theoretically the condenser can never be really charged to the full applied voltage E_B but for all practical purposes at time (RC seconds $\times$ 10) the difference between the condenser charge and the applied voltage is so small that it can be disregarded, wherefore 10 RC seconds may be regarded as the time required for the condenser to charge fully.

Discharge Conditions

At the instant that discharge commences, $-E_c = E_B$ and at any time t therefore, $E_c = E_B\,(\epsilon - {}^{-\frac{t}{RC}})$,

$$E_c = E_B\left(\frac{1}{\epsilon}\right)$$

$$= \cdot 368\ E_B.$$

This means that in RC seconds after discharge commences the voltage (E_c) across the condenser falls to 36·8 per cent. of the voltage which appeared across it when fully charged (E_B). Since this is the only e.m.f. in the circuit during discharge this will also be the drop across R, wherefore the current at this instant of time (RC) is $\frac{36{\cdot}8 \text{ per cent. of } E_B}{R}$.

The RC time constant during charge represents the time required for the condenser to charge to 63·2 per cent. of E_B.

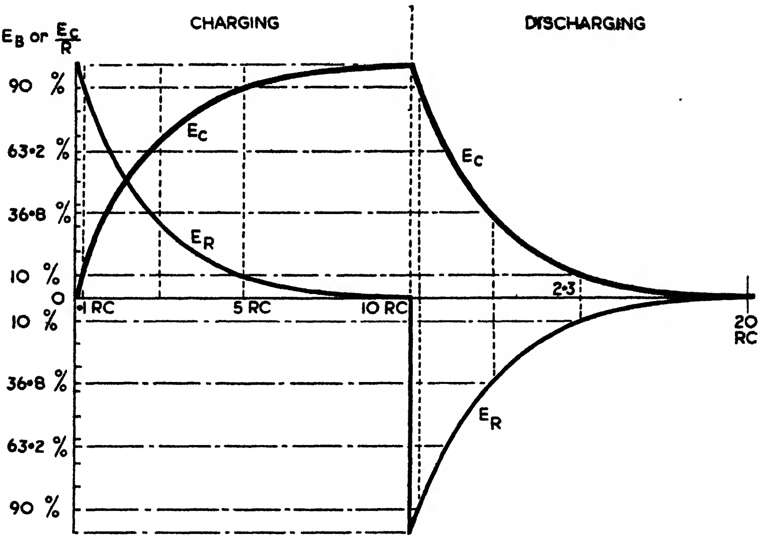


Fig. 84.—CHARGING AND DISCHARGING CURVES FOR E_c AND E_R PLOTTED AGAINST TIME IN R.C. CIRCUITS.

In this case the times for commencing the charge and discharge and the duration of each operation are controlled by switches, charging being performed from a D.C. source.

Note. (*a*) The heavy curve reflects the shape of the voltage wave taken across the condenser. It is a back-to-back exponential curve or sawtooth. (*b*) The fainter line reflects the shape of the voltage wave output across R. It has a sharp positive going peak and a sharp negative going peak, very suitable in shape for pure trigger pulses, since the time duration of the peak is indeterminate.

At discharge this constant represents the time required for the condenser voltage E_c to fall to 36·8 per cent. E_B (see Fig. 84).

Universal R.C. Graph

Fig. 83 is a graph of the voltage rise or fall across condenser and resistance in a series R.C. circuit.

The time scale is graduated in RC seconds and the voltage scale is graduated in percentages of full voltage (E_B).

If the time constant and the initial or final voltage for a circuit is known, the voltage across the components of the circuit can be obtained from curves A or B for any time after the switch is closed for charge or discharge.

Example of Use. The condenser in an R.C. circuit must charge to $\frac{1}{8}$th of the charging voltage (E_B) in 100 microseconds.

A resistance of 10 k. is to be used. Find the capacity of the condenser required.

From curve A it is seen that the RC time necessary to give $\frac{1}{5}$th or 20 per cent. of E_B is 0·22 RC.

In the example the charge must reach 20 per cent. of E_B in 100 microseconds, therefore 100 microseconds must be put proportionate to 0·22 RC seconds and the full RC time for the circuit is therefore $\frac{100}{0{\cdot}22} = 455$ microseconds.

Thus $C \times R$ for this circuit $= 455 \times 10^{-6}$ seconds.

$$\text{or} \qquad C = \frac{455}{\text{R (in ohms)}}$$
$$= 0{\cdot}0455 \text{ microfarads.}$$

Waveform Graph of Charge and Discharge

Fig. 84 is similar to Fig. 83 but shows the conditions at discharge plotted as an extension to the charge. Hence the complete voltage waveforms taken across C and R respectively are revealed.

L/R CIRCUITS

Inductance in Series with Resistance

Inductance can be used in conjunction with resistance to produce effects similar to those of an R.C. combination. There are, however, important differences which determine the conditions under which one or the other is employed.

Analysis of the R.C. circuit showed that the output voltage across the resistance is the IR drop due to the condenser charging current.

In an L/R circuit the current changes due to variations in voltage input and causes proportional changes in voltage across the inductance. The output from an L/R circuit is, therefore, taken across L. The difference in action between R.C. and L/R circuits may be stated as follows :—

(*a*) A constant *current* flowing into a condenser causes the *voltage* across it to rise at a constant rate.

(*b*) A constant *voltage* applied across an inductance causes the *current* flowing through it to rise at a constant rate.

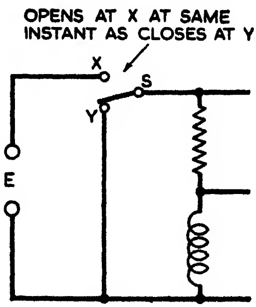


Fig. 85.—DISTRIBUTION OF E.M.F. ACROSS RESISTANCE AND INDUCTANCE IN SERIES L/R CIRCUIT.

From (*b*) then, it is clear that in an L/R circuit it is the changes of current through the inductance for each change of input voltage that is important *i.e.*, rate of change $I_L \propto$ rate of change V input.

(1) When the switch (Fig. 85) is closed at $T = 0$ there will be no instantaneous flow of current. There is, therefore, no drop across R and the full value of E_B appears across L (Kirchoff's law).

(2) Current now rises through the circuit at the rate which will produce ultimately a counter e.m.f. in L equal to E_B.

(3) When this current (2) begins to flow a voltage drop occurs across $R = IR$, therefore E_L must decrease so that the sum of the vectors E_L and E_R is equal to E_B (Kirchoff's law).

(4) The current continues to grow, but at a reduced rate, and this sequence goes on progressively until I reaches the final value $\frac{E_B}{R}$.

(5) During this period of current growth the voltage across R rises exponentially. I increases in like manner.

The relationship in time of I to E_B, R and L is :—

(1) I at any time $= \frac{E_B}{R}\left(1 - E^{-\frac{Rt}{L}}\right)$

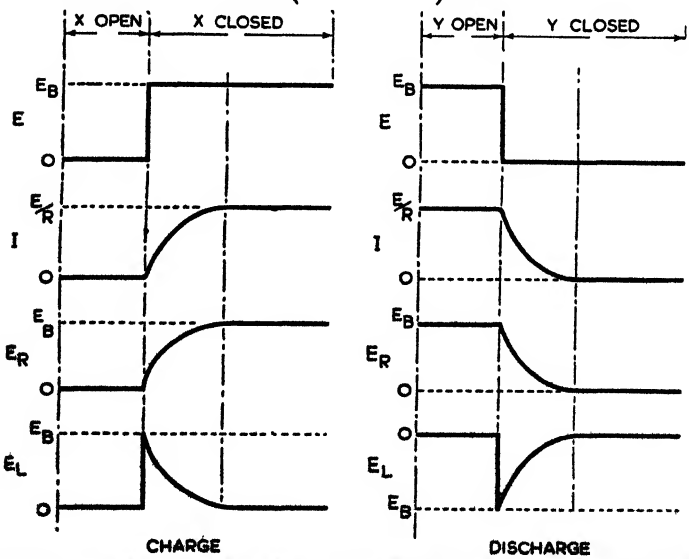


Fig. 86.—OUTPUT ACROSS R AND L AT "SWITCH ON" AND "SWITCH OFF" FOR A SERIES L/R CIRCUIT.

(2) Putting t $= L/R$,
the above reduces to :—

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

$$= \frac{1}{2\cdot718}$$

This means that in L/R seconds I rises to a value of 63·2 per cent. of its $\frac{E_B}{R}$ value.

Therefore in L/R seconds E_R will be 63·2 per cent. of E_B,
E_L ,, ,, 36·8 per cent. of E_B.

Increasing L increases the time constant because I requires a longer time to rise.

Increasing R decreases the time constant because $\frac{E_B}{R}$ is a smaller maximum value for I to attain.

It may be assumed for all practical purposes that in L/R × 10 seconds the current has reached its $\frac{E_B}{R}$ value.

If the input voltage is now reduced instantly to zero, the current $\frac{E_B}{R}$ will not immediately change, since L generates an instantaneous e.m.f. $= E_B$ (Lenz's law). The current will, therefore, fall exponentially and will be 36·8 per cent. $\frac{E_B}{R}$ in L/R seconds after the input voltage has been made zero.

The vector relations and phase differences between E_R, E_L and E, are shown in Fig. 80 (a), (b) and (c). The output waveforms are shown graphically in Fig. 86.

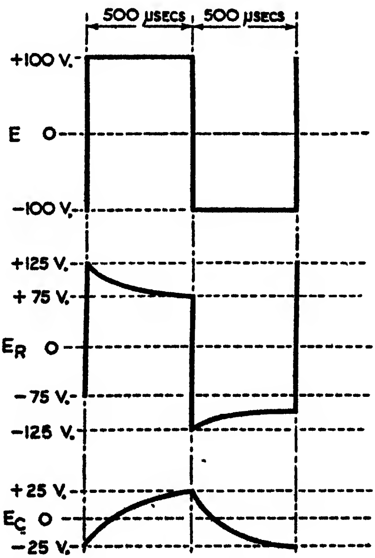


Fig. 87.—RATE OF CHARGE OF POTENTIAL OR COMPONENT E.M.F. ACROSS RESISTANCE OR INDUCTANCE IN AN R.C. SERIES CIRCUIT WHERE $\frac{2}{T}$ = R.C.

Square Wave Inputs. Analysis when Time Constant = Time Period

If the square wave shown in Fig. 87 is applied as an input to the circuit shown in Fig. 79, the output waveforms of voltage across E_R and E_c will also be as shown. The frequency of the square wave is 1,000 cycles per second.

The time constant of the circuit is assumed *equal* to the time period of the input and some distortion therefore occurs, because all the harmonics that matter of the square wave are not shifted in phase by the same amount as that of the fundamental frequency.

The square wave output taken across E_R varies about an average value of zero, and the leading edge of the positive-going pulse rises to 125 volts (25 volts above E). In 500 microseconds (the time between the leading and falling edges), the voltage drops 50 volts, so that the trailing edge is at + 75 volts at the moment when it commences to fall. Similar conditions are exhibited by the negative-going half cycle, the negative signs taking the place of the positive signs but the values corresponding.

When E_R is at + 125 volts, E_c is at — 25 volts. Also E_c has risen 50 volts to + 25 volts when the trailing edge has fallen below the leading edge by the same amount.

Thus the sum of the vectors for E_c and E_R, taken at any instant of time is equal to E.

At time T = 0, when the leading edge of the positive pulse is at + 125 volts and the value E_c is — 25 the charge on C is of such polarity that the resulting e.m.f. E_c reinforces E_B. The condenser discharges from — 25 volts during the time interval between the leading and trailing edges and becomes charged to + 25 volts in time to reinforce E_B again as the negative-going half cycle as the trailing edge begins to fall.

Analysis for Square Wave Input when Time Constant/Time Period Ratios (a) Long, (b) Short

Referring to Fig. 88, if the frequency of the input is constant, a long time constant will give outputs E_c and E_R as shown in "a." A short time constant will cause the output pulses to take the forms shown in "b."

A detailed comparison should now be made of the results shown in Fig. 88, comparing conditions "a" with conditions "b."

The inputs applied to "a" and "b" are identical. "a"

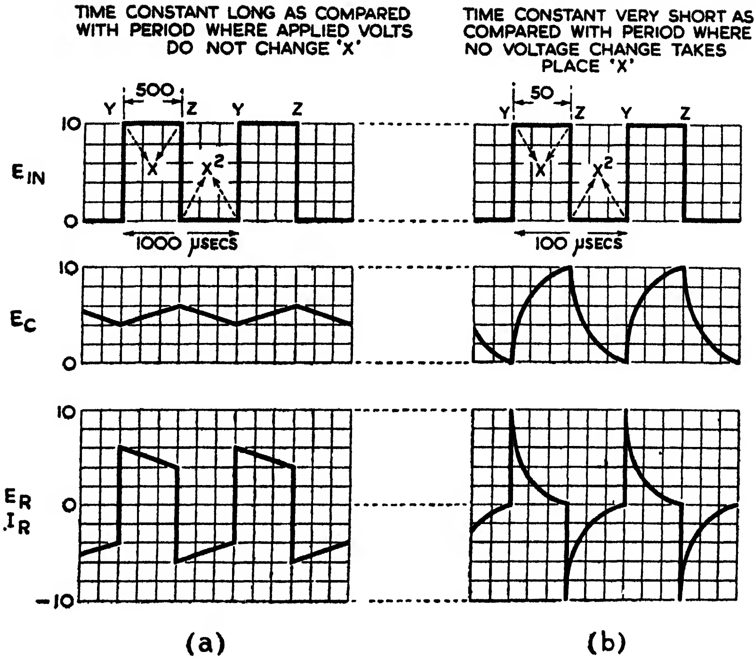


Fig. 88.—Comparison of output waveforms for circuits having (a) long and (b) short time constants compared to time period of the input.

has a time constant much longer than the time period of the input rectangular pulse. The time constant for "b" is much shorter than the time period for a similar input. Supposing the input frequency to be say 1,000 cycles per second, its time period is $\frac{1}{1,000}$ second = 1,000 microseconds. The time constant to give output "a" might, therefore, be about 5,000 microseconds. The time constant for "b" to give the "b" output would be of the order of about 50 microseconds.

It is not, therefore, the length of the time constant itself that makes it a long or short time constant; its classification is determined by the length of the time constant of the circuit relative to the time period of the input, *e.g.*, in the above case 5,000 microseconds is a long time constant for an applied frequency of 1,000 cycles per second. (Time period $\frac{1}{1,000}$

second = 1,000 microseconds.) It would, however, be short for an input frequency of say 100,000 cycles per second.

Analysis of the results in "a" and "b," Fig. 88, shows :—

(*a*) Long R.C. Time Constant.

During the first few cycles of the input (not shown) an average value for the charge Q has been established. This follows from inspection of the curve for E_c which shows E_c varying about an average value of 5. Since $E_c = \frac{Q}{C}$, and since C is a constant, Q must also have been established at some average value in order that E_c may vary about it in the manner shown.

The maximum value of C, although not small, is generally limited by other considerations, wherefore in order that the product RC shall be sufficiently large, R must also have a fairly large value. Thus during the first few cycles and prior to the establishment of an average value for Q, the increments contributed towards this average value at each rise of the input voltage, by the current flowing into the condenser for the *time* marked "X" Fig. 88, are comparatively small, and a comparatively large condenser is therefore only partially charged by each cycle.

For similar reasons the reverse current flowing at the discharge during X^2 (see Fig. 88) is unable fully to reduce Q to zero. Consequently Q builds up over a number of cycles until it reaches a stable value determined by C and R.

When a stable state is established, each subsequent rise and fall of the input voltage causes charging and discharging currents to flow proportional to the rate of change of the input voltage. E_R will, therefore, follow these changes and the input wave will be reproduced with reasonable fidelity across R. The voltage output is therefore proportional to the rate of change of the voltage input. In other words, $V_o = \frac{dV_i}{dt}$ where V_o = voltage output and V_i = voltage input.

The condenser commences to charge when E_R is at its maximum value. E_R falling to zero as Q, and therefore E_c, simultaneously reach their maxima. E_R and E_c are 90° out of phase but their vector sum at any instant = E_B which satisfies Kirchoff's law.

Thus the general conditions for passing an undistorted pulse from input to output of an R.C. coupling require a long RC time constant, so that the condenser charge Q can attain an average value and vary as little as possible with changes of the input voltages.

(*b*) **Short Time Constant**

In the case of " b," the time constant is short (one-tenth of the time period of the input), consequently C and R must be small; this means that the comparatively large current flowing into the smaller condenser at each rise of the input voltage charges the condenser exponentially well within the time limit set by YZ.

If the time constant of the circuit is very short, relative to the time period of the input wave, the condenser may charge long before the end of the period YZ is reached, but in any case as long as the charge is completed *within* the time YZ, conditions will be as follows :—

At the rise of the input volts E_B will appear across R. The initial current through R will be $\frac{E_B}{R}$. This will fall exponentially to zero as E_c rises. Consequently, if the condenser charges very quickly the current through R and therefore E_R will be a very short, sharp pulse, the sharpness of which is determined by the time taken for the condenser to charge and for the charging current to fall to zero.

If the charge is completed before the input volts fall at Z and since there is no change of input volts between Y and Z, the condenser will hold its charge, and E_c will have a constant value from the time the charge is completed until the input volts commence to fall. The output waveforms across the resistance and condenser will then be as shown in Fig. 88 (b).

Clearly, therefore, when the time constant of the circuit is shorter than the time period of the input, the output wave across the resistance is again proportional to the rate of change of the input volts (differentiation) but the rate of change of the input volts across the R.C. elements is greater, and this is reflected in the output waveform.

The conditions for differentiation are therefore a short time constant such that the condenser may be fully charged or discharged between successive variations of the input voltage. In other words, Q must be able to rise to maximum and to fall to zero during each half wave of the voltage input so that I, and consequently E_R, may be proportional to the rate of change of E_B which depends upon the values of C and R.

Comparison with L/R Circuits

In an R.C. circuit the average charge on the condenser and, therefore, the p.d. across it, builds up cycle by cycle until its charging and discharging currents are equal.

In an L/R circuit the current through L will build up from cycle to cycle until it reaches a value where the current through L increases and decreases with equal intensity. If the L/R time constant of the circuit is being compared with the time period of the input for a long time constant, the energy stored around the inductance magnetically ($= \frac{1}{2}LI^2$) must not change very much from cycle to cycle. This condition is similar to the state when the average charge in the condenser of an R.C. circuit reaches stability and the energy is $\frac{1}{2}CV^2$. Since the current changes through the inductance are proportional to the voltage input, the voltages across the inductance are likewise proportional.

If the L/R time constant is made short relative to the time elapsing between successive input voltage changes, the current must rise to its $\frac{E_B}{R}$ value very quickly and fall to zero in a similar manner, thus the voltage across L will follow changes of the input voltage. In other words, E_L will be proportional to the rate of change of E_B and E_B will be differentiated.

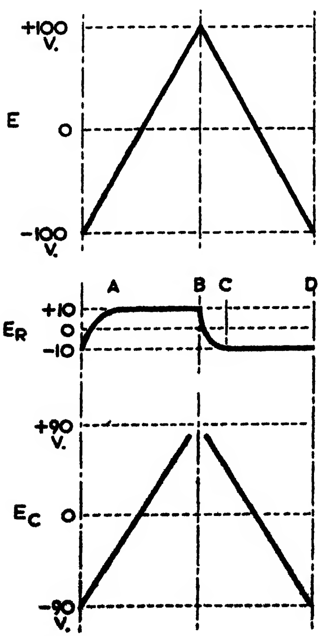


Fig. 89.—DIFFERENTIATING AND INTEGRATING ACTION ON A TRIANGULAR OR SAWTOOTH INPUT.

Application of L/R Circuit

An inductance is sometimes used in the anode circuit of a valve so that it forms with the anode resistance a long L/R series circuit in the anode. This is done to ensure that the waveform of the anode voltage shall follow the input waveform to the grid, the anode voltage then varies about the supply voltage.

Average Value of Q in R.C. Circuit

In an R.C. circuit E_c is proportional to the frequency and amplitude of the input pulses. This is because the average value of Q builds up from cycle to cycle.

The average value of Q is, therefore, directly proportional to the input frequency, and inversely proportional to the time during which the input voltage does not change.

R.C. Integration

All other things being equal, the current I also varies directly with E_B; thus an output taken across the condenser E_c being proportional to Q is the integration of the input.

The current variations across the resistance, and hence E_R, is considered to be a purely alternating component, the charge Q being regarded as the direct current component.

L/R Integration

In an L/R circuit integrated voltage waves may be taken across the resistance. When $T = 0$, $E_R = 0$, $E_L = E_B$. When current flows, E_L drops and E_R rises, consequently if a square wave is applied to the input, E_R rising as E_L falls, produces what is in effect a sawtooth wave across R for a long L/R and a complete exponential wave for a short L/R.

Sawtooth Input to Differentiating Circuit

If a sawtooth voltage is applied to an R.C. circuit with a time constant approximately one-sixth of the period of the input voltage, the waveform outputs for E_R and E_c are as shown in Fig. 89.

If the time constant is short, the output taken across E_c will be very similar to the input E, since there is only a small phase difference (lag), also there will be very little change in amplitude.

The charging and discharging currents are constant, because E_c increases at a constant rate, consequently the differentiated output taken across R_1 will take the form shown in Fig. 89. After the initial exponential rise of the charging current to A there is no change in its value, since $\dfrac{E_B - E_c}{R}$ is constant until the point B is reached. The current then falls to zero and to its maximum

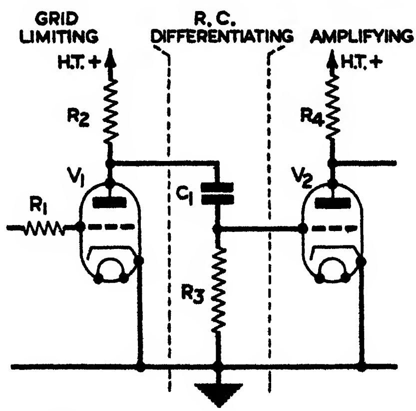


Fig. 90.—R.C. DIFFERENTIATING AMPLIFIER.

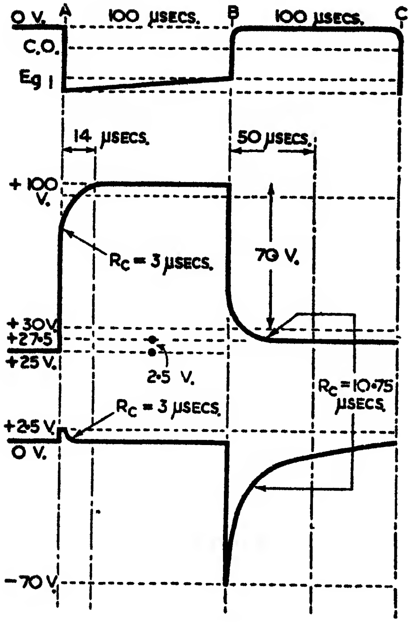


Fig. 91.—WAVEFORMS OF DIFFERENTIATING AMPLIFIER CIRCUIT SHOWN IN FIG. 90. E_gV_1, E_aV_1, E_{R3}.

negative value exponentially at C, where $\frac{E_B - E_c}{R}$ is again constant. Wherefore E_R also remains constant until it reaches point D.

THE DIFFERENTIATING AMPLIFIER

Fig. 90 shows the circuit of an R.C. differentiating amplifier.

The input voltage is assumed to have a half period of 100 microseconds which drives the grid of V_1 behind cut off during the negative half-cycle, and maintains it at approximately earth potential, because of grid limiting * during the positive half cycle. The grid and plate voltage waveforms for V_1 and that applied to the grid of V_2 is as shown in Fig. 91. The dotted waveform indicates the shape of E_{R2} with differentiator disconnected. The maximum value of E_{R2} is 100 volts and the minimum value as determined by the valve characteristics is 25 volts.

The anode circuit of V_1 is in effect an open circuit when V_1 is cut off. When conducting the anode circuit resistance approximates to R_a and $R_a = 7{,}500$ ohms $R_L = 29{,}000$, (R_s on diagram), $R_3 = 100$ k.

During the interval preceding time A, see Fig. 91, V_1 is conducting and C_1 charges to 25 volts. Since E_{a1} is steady at this level during the conducting period, $E_{R3} = 0$. The voltage distribution around the circuit at the instant V_1 is cut off is shown in Fig. 91 and it must be remembered in making an analysis that C_1 cannot change its voltage instantaneously.

* Any tendency for the grid of V_2 to go positive is checked by the drop across R_1 when grid current flows as it must do when the grid of V_1 is positive to the cathode.

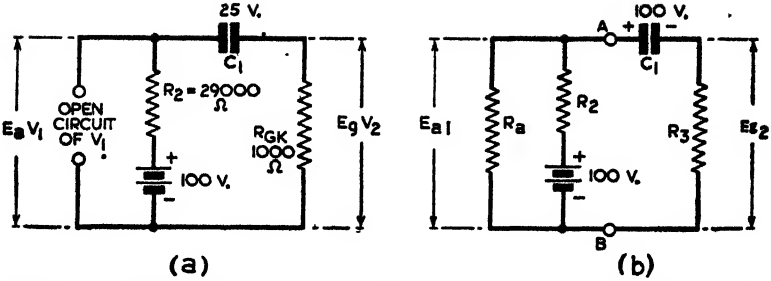


Fig. 92.—EQUIVALENT CIRCUITS FOR FIG. 90 DURING "CUT-OFF" AND CONDUCTION PERIODS RESPECTIVELY.

(a) Equivalent circuit during time V_1 is cut off.
(b) Equivalent circuit during time V_1 is conducting.

Development of Short Time Constant Path

At cut off C_1 charges from 25 volts (initial) to 100 volts (max.). V_2 is at zero bias, therefore a positive signal to E_g V_2 causes *grid current to flow,* placing the equivalent of about 1,000 ohms in parallel with R_3 which would be about 100 k. Consequently by comparison R_3 can be neglected (Fig. 90).

The charging time constant $= (R_2 + R_{Gk} \times C_1) = (29 + 1)$ $10^3 \times \cdot 0001 = 3$ μsec.(Fig. 92 (*a*)).

The initial charging current flowing through the circuit is the *nett voltage* around the circuit divided by total resistance.

$$I \text{ (initial)} = \frac{100 - 25}{(29+1)\ 10^3} = 2{\cdot}5 \text{ mA.}$$

This current flows through the grid-cathode resistance (equivalent) of V_2 and causes an initial drop from grid to cathode of 2·5 volts.

The voltage across the condenser at this instant is 25 volts. But the initial value of E_a $V_1 = 27{\cdot}5$ volts, *i.e.*, 25 volts + 2·5 volts as above to which value E_c rises.

The condenser now charges to 100 volts in a time determined by the 3 μsecs. time constant.

Simultaneously, E_{g2} drops from 2·5 volts to zero.

The variations of E_{a1} and E_{g2} are shown in Fig. 91 between times A and B.

At the end of the cut-off period of V_1 the potential across C_1 is 100 volts. Thus at time B when V_1 starts to conduct again the voltage across C_1 cannot change at once and remains momentarily at 100 volts, but the instantaneous change of E_{a1} from 100 volts to 25 volts appears across R_3 as E_{g2}.

Equivalent Circuits for the Discharge of C_1

The equivalent circuit for the time V_1 conducts is shown in Fig. 92 (b).

During this time C_1 discharges from 100 volts to its final value of 25 volts.

$E_g V_2$ is normally zero but $E_a V_1$ has dropped, and $E_g V_2$ becomes negative. No grid current can flow during this half cycle. Therefore, $R_{Gk} V_2$ is an open circuit and the discharge to earth of C_1 is *viâ* R_3 and $R_a V_1$ in series.

Long Time Constant

The time constant for the discharge is therefore :—

$(R_a + R_3) C_1$.

Taking the value of R_a as 7,500 ohms for this instant of time:—

$= (7{,}500 + 100{,}000)\ {\cdot}0001.$

$= 10{\cdot}75\ \mu$secs.

The initial discharge current is :—

$$\frac{E_a - E_c}{R_a + R_3}$$

$$= \frac{25 - 100}{107{,}500} = \frac{-75}{107{,}500}$$

$= -0{\cdot}7$ mA.

This current flows in opposite direction to the charging current, hence the minus sign.

This current flows through R_3 and causes an initial drop of $-0{\cdot}7 \times 10^{-3} \times 10^5 = -70$ volts. But the voltage across C at this instant is 100 volts, therefore the nett value of the voltage $ER_3 = 100 - 70 = 30$ volts.

$E_a V_1$ cannot drop its volts instantaneously, because the condenser discharges exponentially to 25 volts in a time determined by the $10{\cdot}75\ \mu$sec. time constant. $E_g V_2$ rises exponentially from -70 volts to zero in equal time. These relationships are shown in Fig. 91 from B to C.

Thus the shape of the anode voltage E_{a1} is modified by the R.C. circuit into which it is feeding.

$E_a V_1$ with the differentiator disconnected is rectangular in shape.

The rounding of the corners in the actual case is due to loading. Thus grid current through V_2 during the charging period when V_1 is cut off, places in effect, a low grid to cathode

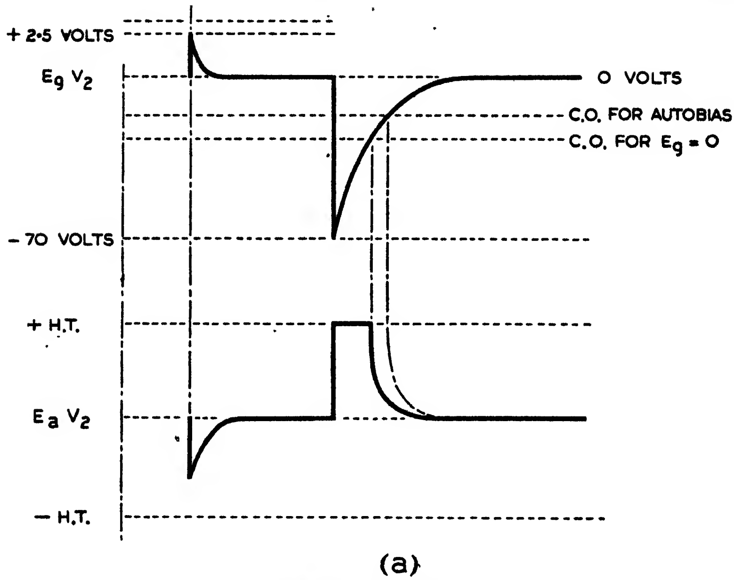


Fig. 93 (a).—OUTPUT WAVEFORMS FOR E_g V_2 AND E_a V_3 OF CIRCUIT IN FIG. 90.

resistance in parallel with R_3. This produces a short charging time constant, wherefore a higher current flows and the loading is greater than for the longer time constant of the discharge.

V_1 can be regarded as a square wave generator with internal resistance. Its terminal voltage alters as the load varies.

In general then a sharply peaked output can be obtained by using an R.C. differentiator having a time constant very short compared with the duration of the input pulse.

The loading effect of the differentiator may be reduced when a low impedance output is required by replacing V_3 with a cathode follower.

Note. The only change in this circuit which it is necessary to make in order to pass a square pulse without distortion is to adjust the values of C and R so that their product results in a long time constant compared with the duration of the pulse. In this case C acquires an average value which does not change appreciably from one part of the operating cycle to another. Thus the loading effect is negligible and there is practically no distortion.

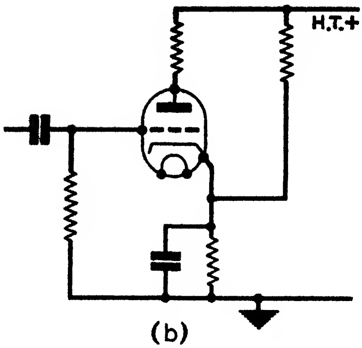


Fig. 93 (b).—V_2 WITH AUTOMATIC BIAS AND POSITIVELY BIASED CATHODE.

The output waveform from V_2 for the circuit and input arrangements shown in Fig. 90 are shown in Fig. 93 (a).

The small positive-going peak at the grid of V_2 derived from the leading edge of the square wave input produces a sharp negative-going pulse at the anode of V_2. The duration of this pulse in the anode circuit of V_2 is 3 μsecs.

The large negative-going peak at the grid of V_2 derived from the trailing edge of the square wave input drives this valve below cut off, consequently the pulse appearing at the anode of V_2 will be a positive-going pulse, flattened for the duration of time that the valve is held at cut off by the peak of the negative-going pulse at its grid. Since the valve is cut off, this pulse will rise to the H.T. supply voltage.

If instead of limiting the positive half cycle of the square wave input the negative half cycle is eliminated by biasing V_1 (Fig. 90) to cut off, the positive pulse at the anode of V_2 would appear at the same time as the *leading* edge of the input square

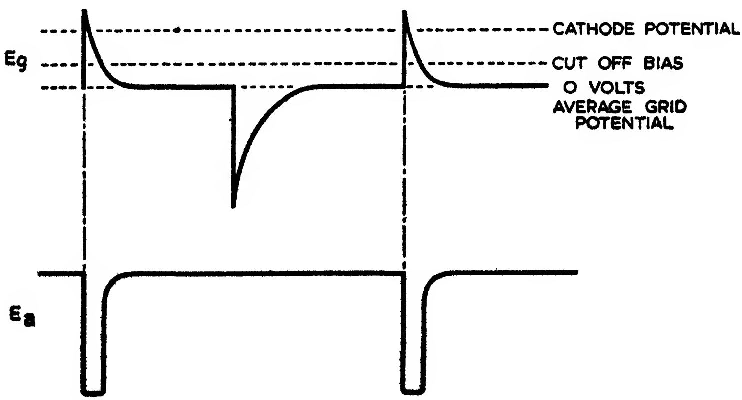


Fig. 93 (c).—WAVEFORMS E_g V_2 AND E_a V_2 WHEN ARRANGED WITH CUT-OFF BIAS AND POSITIVELY BIASED CATHODE.

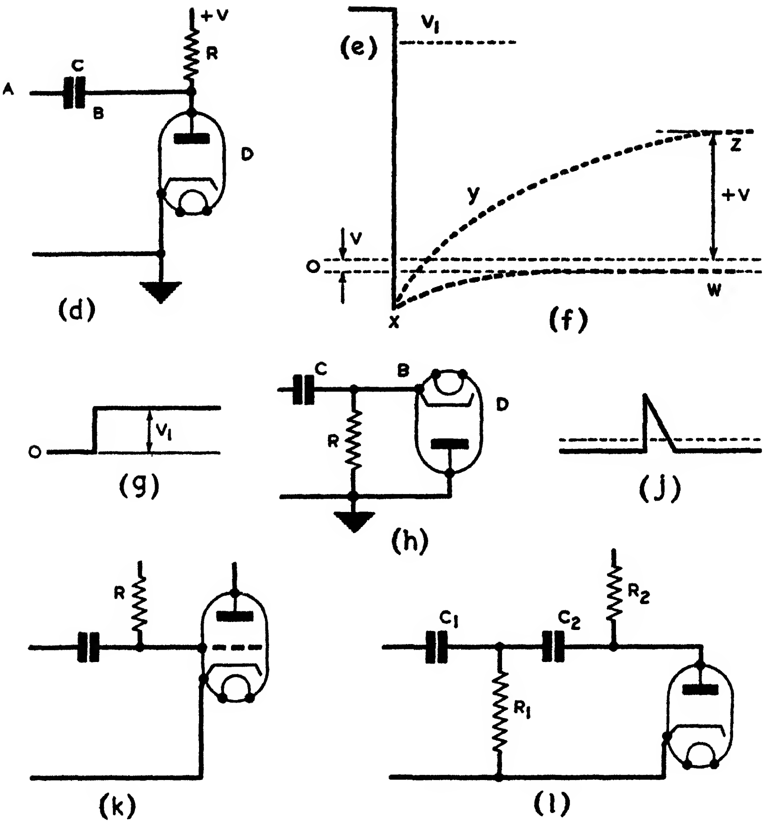


Fig. 93 (d)–(l).—EFFECT OF BIAS ON OUTPUT WAVEFORM.

wave and the negative pulse would coincide with the *trailing* edge. In other words, the output for the second case is the mirror effect of the output for the first case.

The effect of applying automatic bias to V_2 is mainly to cause the negative-going pulse on the grid to cut off V_2 earlier. As a consequence of this, the positive pulse appearing at the anode is broadened, because the peak of the negative-going pulse on the grid holds the valve at cut off for a longer time.

If in addition to the automatic bias applied to V_2 as above the cathode is further positively biased from the H.T. source, see Fig. 93 (b) (*for positive limiting only at V_2*), the output from the anode of V_2 will be as shown in Fig. 93 (c).

In these conditions V_2 is cut off and cannot conduct until the grid, which is at earth potential, rises above the positively biased cathode.

The input to the grid of V_2 from the differentiator circuit will be somewhat similar to that shown in Fig. 93 (a), but since the valve cannot conduct until the grid of V_2 rises above the cathode, only the peaks of the positive-going pulses are effective. These drive the grid very positive and a long negative pulse appears at the anode of V_2 with duration approximately equal to the time constant of the R.C. differentiator circuit.

Narrowing the pulse base

When the base of the pulse derived from differentiating a square wave is too broad its width can be reduced by application of a positive or negative bias to the grid. This is simply explained by first considering the case of the circuit shown in Fig. 93 (d) in which a diode is shown in place of the triode. The action is, however, similar in both cases.

Under normal conditions before the application of any voltage at point A the potential at point B is settled by the conducting path D. Point B can be considered as being near the end of a voltage divider consisting of R and the diode. Point B therefore takes up a voltage V slightly positive.

If a voltage V_1 as shown at Fig. 93 (e) is applied to A, the voltage at B will drop by V_1 volts as shown at X in Fig. 93 (f).

If the voltage at A remains steady, the voltage at the point B will start rising, at a rate depending upon the values of C and R, as if it would follow the curve Y Z to a voltage V.

As soon, however, as the voltage at B becomes positive, D starts conducting and the point B quickly comes back to its original voltage V.

If V is many times the value of V_1, only a small part of the exponential curve XYZ is made use of, this part of XY being nearly straight.

The result is that the normal waveform shown at XW is turned into the steeper waveform shown at XYZ.

The same effect can be produced in the case of a positive pulse, by connecting the diode as at Fig. 93 (h) and applying a negative pulse bias to R. The result is a positive triangular pulse (Fig. 93 (j)).

The diode in Fig. 93 (d) may be replaced by a triode as in Fig. 93 (k). The same arguments apply as regards the potential at B (the control grid of the triode) is concerned. The amplified output is taken from the anode of the valve.

Repeated Differentiation

The circuit shown in Fig. 93 (l) is sometimes used for presenting a short triangular pulse.

This is seen to be a combination of a normal differentiating circuit and a biased circuit. This circuit is used when a short pulse is required, because the shortness of a pulse depends upon small values of C and R. In the above circuit, R_2 cannot be small as, in this case, the standing current would be too large, the lower limit to C_2 is settled by stray capacities. However, by using another circuit consisting of C_1R_1, R_1 can be as small as required.

If the amplitude of a long pulse is P it can be shown that the rate of change of the output voltage is equal to the sums of the following rates :—

A circuit C_1R_1 with an input voltage P.
A circuit C_2R_2 with an input voltage P plus V.
A circuit C_1R_2 with an input voltage P plus V.

Therefore the output voltage changes at a higher rate than in a simple R.C. circuit.

CHAPTER X

SOME APPLICATIONS OF THE DIODE TO RADAR CIRCUITS, INCLUDING CLAMPING AND RESTORATION D.C.

THE characteristics of the diode are :—

(*a*) Unilateral conductivity.
(*b*) Low resistance when conducting.
(*c*) $R_a = \frac{E_a}{I_a}$.

Because of these characteristics, diodes, in conjunction with various circuit arrangements, can be used to perform a variety of different functions :—

(*a*) Rectification.
(*b*) Frequency conversion (detection).
(*c*) Limiting.
(*d*) Clamping and restoration D.C.
(*e*) Counting circuits.
(*f*) Voltage doubling.

Limiting

The function of a limiter circuit is to limit or clip either or both the positive or negative amplitudes of a wave to a predetermined level. A diode may therefore be employed with the primary object of exercising amplitude control, or if amplitude is unimportant, perhaps because of subsequent amplification, it may be used with the primary object of shaping. For example, to produce a partly squared wave from a sine wave input.

Clamping

A clamping circuit is employed when some point in a circuit must be free to vary within certain limits *only*. It may be applied to control the upper limit, or the lower limit, or both upper and lower limits. Clamping circuits are employed, for instance, in connection with the start and end of the time base in display units.

Another form of clamping is known to television engineers as restoration D.C. This usually applies to the case where the output from the anode of one valve is coupled to the grid of the next valve by means of a condenser. In passing through the condenser the pulse loses its D.C. component, which is

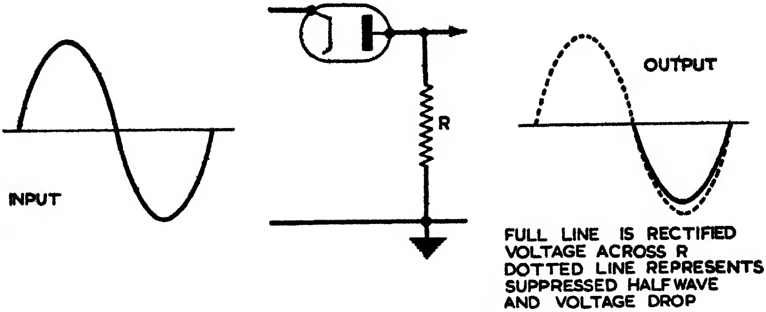


Fig. 94.—SERIES DIODE FOR SUPPRESSION OF POSITIVE HALF WAVE AND RECTIFICATION OF NEGATIVE HALF.

to be expected, since the condenser does not pass D.C. Thus its level is altered, the input being A.C. superimposed upon D.C., and the output pure A.C. A diode can be used to restore the D.C. component of the signal, or to raise the A.C. output to any desired level.

Counting

Counting circuits are employed in radar, mainly when it is desired to generate from an input frequency some other frequency which is to be a sub-multiple of the applied frequency.

Voltage Doubling

Voltage doubling is employed when a higher voltage is required from the H.T. supply. Its use is limited, however, to cases in which the load is very small because the regulation of all voltage doublers is inherently poor.

It has been assumed that the reader is familiar with the other applications which are well known to established practice in radio communication, *i.e.*, rectification and detection.

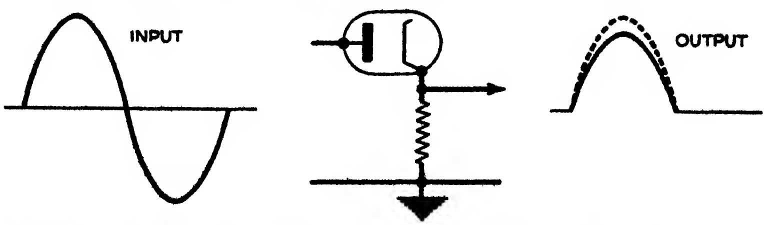


Fig. 95.—SERIES DIODE FOR SUPPRESSING NEGATIVE HALF WAVE AND RECTIFICATION OF THE POSITIVE HALF. FULL AND DOTTED LINES AS IN FIG. 94.

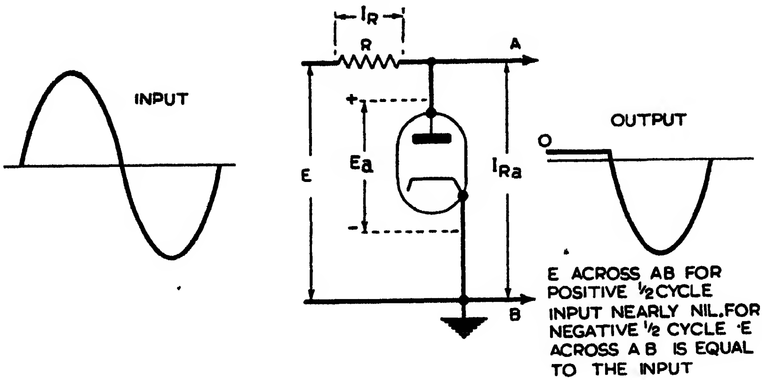


Fig. 96.—PARALLEL DIODE. NEGATIVE RECTIFICATION.

Diode limiting

Diode limiting may be carried out by either a series or parallel arrangement. Figs. 94 and 95 show series arrangements for eliminating either the positive or negative half cycle. In both cases the diode acts as a rectifier. When the anode is driven or held positive to the cathode the diode conducts, for values for E_a/I_a (its resistance R_a) may fall as low as to approach a short circuit. For reversed polarity of anode and cathode the diode does not conduct at all and its resistance is infinity. Consequently in both the above series arrangements the device is only justified, in this form, as long as the value of R is large compared with R_a.

In the parallel arrangements shown in Figs. 96 and 97, R (now in series) is still very large compared with R_a of the diode when conducting. When the diode is not conducting the full

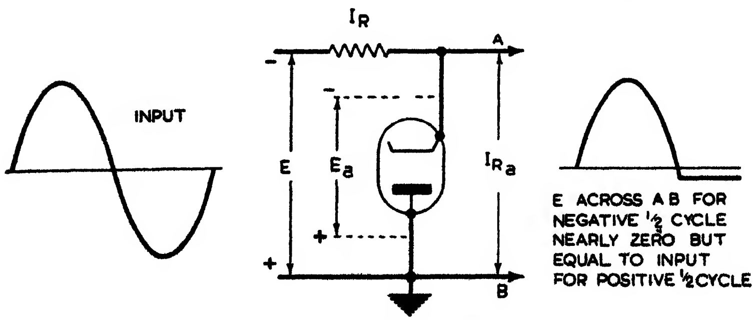


Fig. 97.—PARALLEL DIODE. POSITIVE RECTIFICATION.

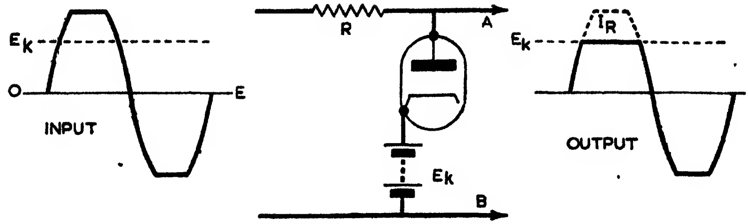


Fig. 98.—POSITIVE LIMITING BY MAKING E_k POSITIVE TO EARTH.

voltage of the input appears across AB (the output). When the diode *conducts*, the *current through it* produces a drop IR across R (in series) and only the very small and almost negligible voltage IR_a appears across the output terminals. Thus $E_{out} = E_{in} \frac{R_a}{R_a + R}$. Since R is very large compared with R_a, IR_a is very small. These conditions are shown in the waveforms of Figs. 96 and 97.

Positive Limiting

An input voltage can be limited or clamped to any desirable positive or negative value by holding the proper diode electrode at that voltage by the application of a biasing voltage.

In Fig. 98 the cathode is held at a higher positive potential than the anode by positive bias.

The diode cannot therefore conduct on the positive half cycle until the input raises the anode above the potential of the cathode. During this non-conducting period the output voltage at AB is equal to the input voltage, but when the diode conducts the *output* falls practically to zero. The amplitude of the positive half cycle can therefore be limited

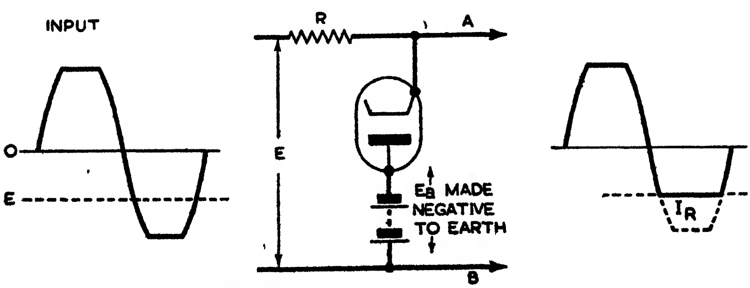


Fig. 99.—NEGATIVE LIMITING BY MAKING THE ANODE NEGATIVE TO EARTH.

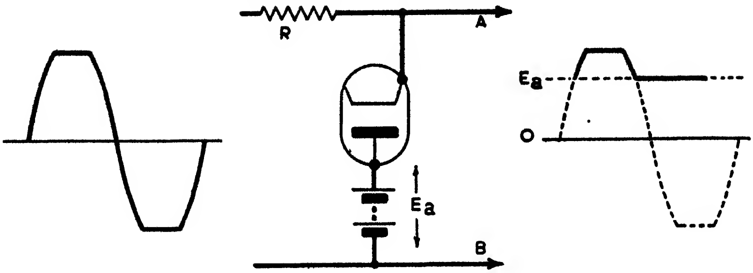


Fig. 100.—ALL SUPPRESSED EXCEPTING POSITIVE PEAK.

to, or *clamped* to, any desired value by suitable adjustment of the bias on the cathode.

Negative Limiting

Fig. 99 shows that in order to limit or clamp the negative amplitude, it is necessary to reverse the diode and to bias the anode negatively with respect to earth. The diode does *not conduct during the positive cycle,* so that the output voltage across AB for the positive cycle is equal to the input voltage. During the negative half cycle the diode *still does not conduct* until the cathode becomes more negative than the anode.

Extremity Limiting

When it is desired to pass to a succeeding stage the positive or negative extremities of the waveform *only,* this can be accomplished by the circuit arrangements shown in Figs. 100 and 101.

The arrangement shown in Fig. 100 permits the diode to conduct for the whole period of the waveform until the cathode potential on the positive half cycle rises above the positively

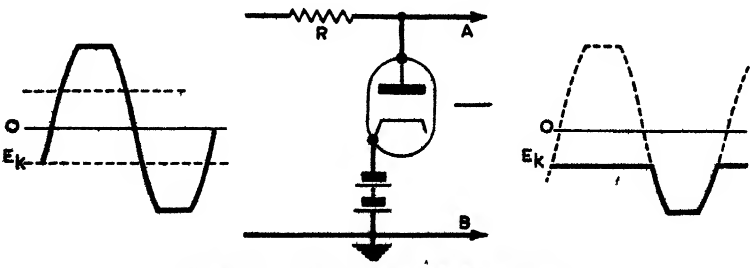


Fig. 101.—NEGATIVE PEAK RETAINED.

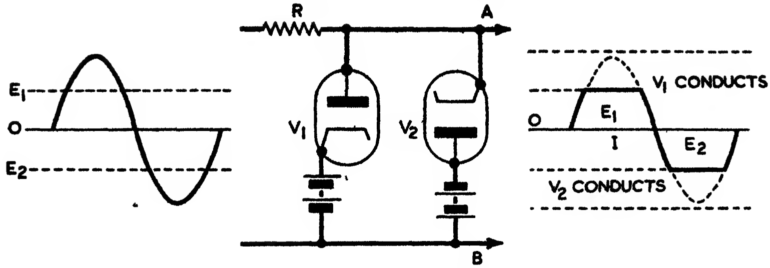


Fig. 102.—SYMMETRICAL CLIPPING.

biased anode. Thus the output at AB is the tip of the positive half cycle of voltage.

By reversing the diode and by applying negative bias to the cathode, the diode is made to conduct over the whole waveform, until the anode (on the negative half cycle) becomes more negative than the negatively biased cathode. In this case the output at AB is the tip of the negative waveform.

In both arrangements the output is zero (practically) during the time the diode *conducts* and only becomes equal to the input when the diode is shut down. By suitably adjusting the bias, therefore, any desired fraction *only* of either half cycle of the waveform may be passed to the output.

The arrangement shown in Fig. 102 limits both amplitudes by adjustment of the bias to any desired values for the positive and negative half cycles—double limiting. Thus the output from the arrangement shown in Fig. 102 is a rectangular-shaped wave.

Triode as Limiter

The input conditions to the grid of a triode can also be adjusted to produce grid limiting and squaring of the output wave.

The positive portion of the output wave can be limited by the use of a grid stopper or suitable resistance in series with the grid (see Fig. 103). When the grid is made positive, current flows in the cathode-grid circuit. The drop I_gR is in opposition to positive inputs and therefore produces almost complete cancellation. The effect on the output is shown by the dotted line in Fig. 103.

The negative output can of course be limited by employing a negative bias such that the valve cuts off when the required limiting value is reached.

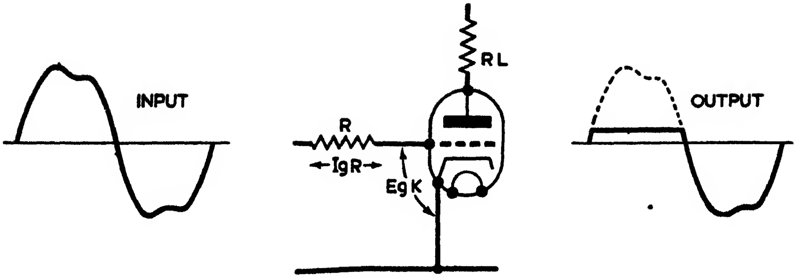


Fig. 103.—POSITIVE GRID LIMITING FOR TRIODE BY GRID STOPPER.

It has been shown that a modification to the shape of an input wave must of necessity result from any limiting operation, consequently limiting or clipping operations are frequently used as part of some shaping process. Reduction in amplitude is of course inseparable from any limiting operation. When this is not of secondary importance the loss can be compensated for by subsequent amplification.

Diode Clamping

In the very large majority of cases, amplifying stages, used in radar, are coupled together by a condenser; the result is that only the varying or A.C. components of the anode potential of the first stage are coupled into the grid-cathode circuit of the next stage.

The potential at which the grid is held (with or without bias) is the value to which all positive and negative grid voltage swings are referred for nett evaluation and comparison, thus the varying potentials passed to the grid by the coupling condenser, swing about a reference value which is zero when no bias is applied to the grid; alternatively, they take for pivot the reference value conferred on the grid by some bias voltage.

These are the ideal conditions for class " A " amplification, for the voltage variations of the anode of the first valve constitute alternating e.m.f.s which vary above and below the reference voltage of the grid of the next valve.

RESTORATION D.C.

When operating under conditions other than class " A," the conditions just described are not always desirable. Output

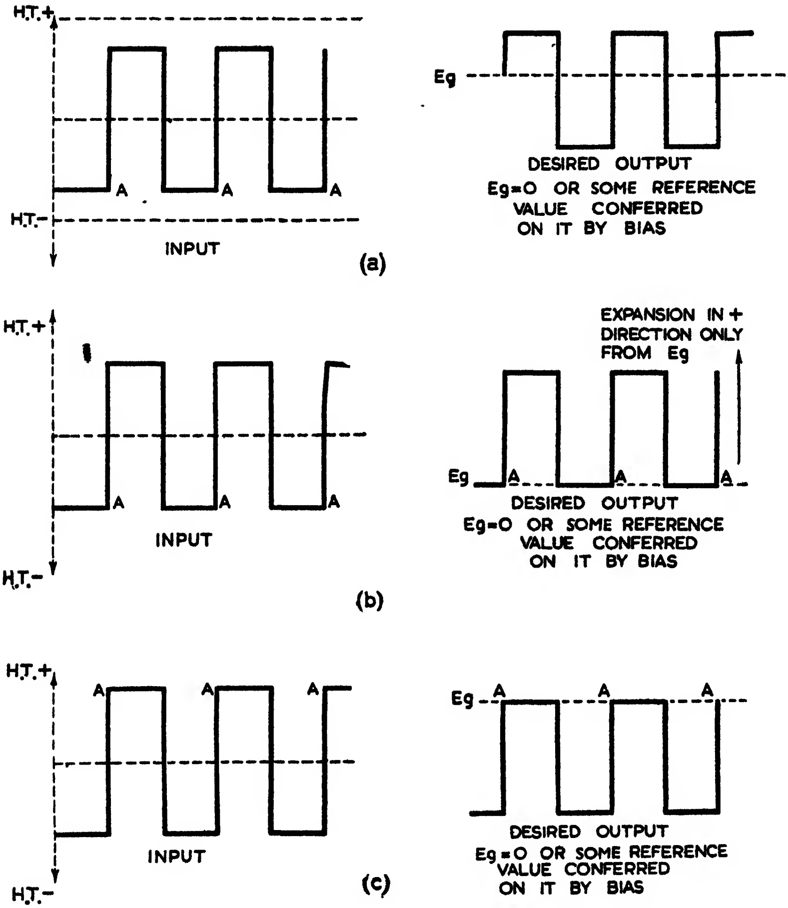


Fig. 104.—INPUT CONDITIONS IDENTICAL FOR (a), (b) AND (c).
Value of E_g is also the same for (a), (b) and (c).
No diode used for output (a).
Outputs (b) and (c) produced by diode action and according to arrangement of diode.

requirements frequently demand that the input variations shall take place entirely above or below the reference voltage of the grid. This means that the variations must all take place from and above the reference voltage in a positive direction, or from and below it in a negative direction (see Fig. 104 (a), (b) and (c)).

In order to bring this about, the coupling condenser must be made to assume, and hold, a charge of value equal to either the minimum applied voltage from the anode V_1 or to the maximum, according to whether variations in the grid circuit V_2 are to take place entirely above, or entirely below, the grid reference voltage.

This means that if the grid coupling condenser is made, by diode action, to *acquire and hold* a charge which produces across the condenser a voltage equal to the *minimum* voltage applied by the input (point A in Fig. 104 (b)), all variations of the condenser potential must take place *from* this value in a positive direction, because the potential across the condenser, by the action of the diode, is maintained at point A and it is never permitted to fall progressively below this value. In consequence of this diode action all voltage variations applied to the grid must take place in the positive direction starting from the grid reference voltage (as determined by the presence or absence of bias).

Alternatively, if under similar conditions the condenser acquires and holds a charge producing a voltage across it equal to the *maximum* voltage applied to it from the input (point A, Fig. 104 (c)), all variations of the wave must then take place in a negative direction because point A is now the highest value which the voltage wave can reach. All grid voltage swings also take place in the negative direction, starting from the grid reference voltage. The diode action in this case is such that the condenser potential is never allowed to rise progressively above point A, Fig. 104 (c).

These conditions are shown in Fig. 105 (a) and (b).

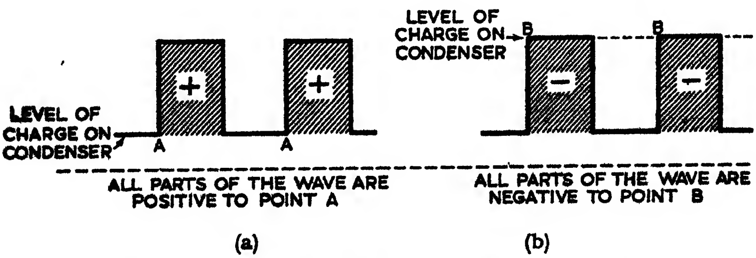


Fig. 105.—VARIATIONS IN POSITIVE AND NEGATIVE DIRECTIONS FROM A PREDETERMINED POSITIVE REFERENCE.

(a) Variations in the positive direction.
(b) Variations entirely in the negative direction towards "cut off."

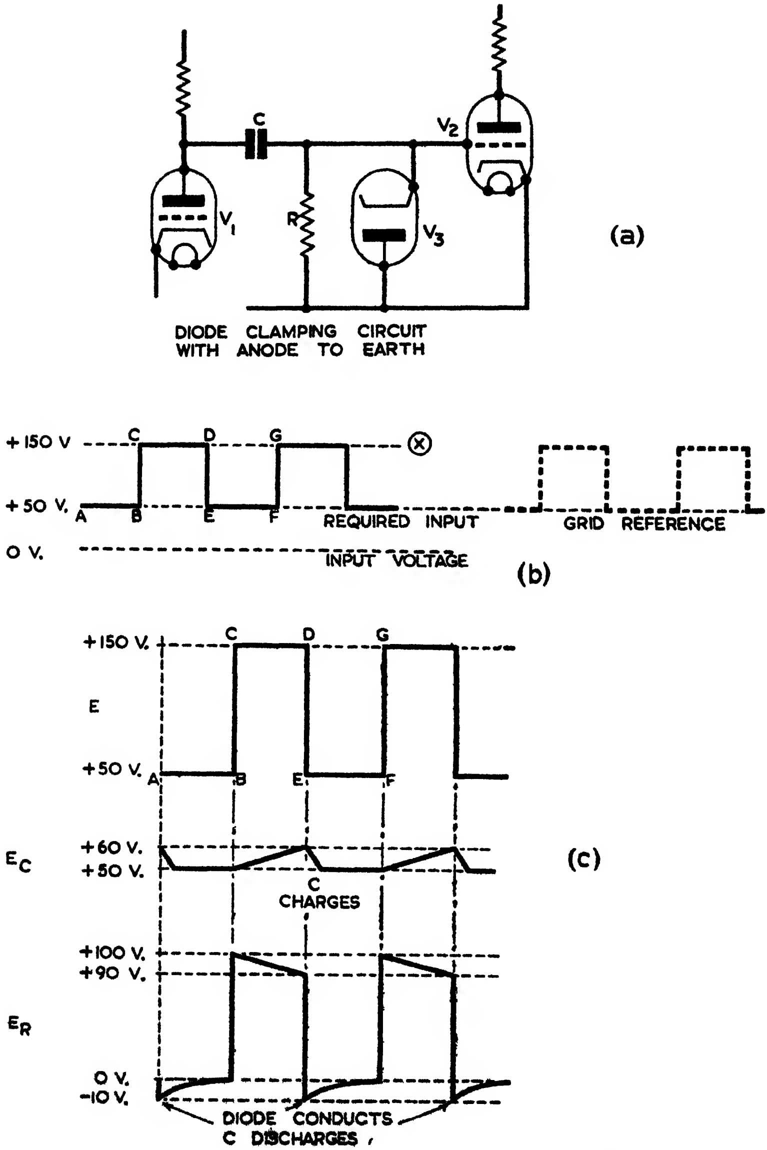


Fig. 106.—WAVEFORMS FOR E_C AND E_R DIODE CLAMPING CIRCUIT AS SHOWN ABOVE.

Lower Extremity of Wave Clamped to Zero

Suppose that the square wave of voltage, shown in Fig. 106 (c) developed in the anode circuit of V_1, is to be R.C. coupled to the grid-cathode circuit of V_2 (see Fig. 106 (a)), and let the RC time constant of the coupling circuit be long as compared with the frequency of the square wave input ; let it also be a requirement that the output voltage wave applied to the grid-cathode circuit of V_2 is to swing between zero and maximum in the positive direction only, as shown by the dotted curve marked "required input." A diode V_3 must be added to the normal R.C. circuit, connected as shown in Fig. 106 (a), in order to obtain this result.

For the initial period of application of the square wave, *i.e.*, from A to D in this particular case, the diode can be neglected. It has no effect on the output during the A—D period because its anode is at earth potential and its cathode is positive to earth during this time.

Thus the conditions at A are that in the first few cycles previous to A (not shown) the condenser has charged to an average value of + 50 volts. This is the base of the input waveform and the value at which the condenser must be held in order that all parts of the wave may be relatively positive to it. In a normal coupling circuit it would probably charge to an average value, during succeeding cycles, of say 100 volts. In the absence of the diode, 100 volts would then be the average value of E_c, and E_g would alternate between the values of plus and minus 50 volts.

The voltage applied to E_g would thus be a simple alternating e.m.f. of 50 volts amplitude above and below the grid reference voltage (either zero or as adjusted by bias). This condition is the one to be avoided in this case, consequently the tendency of the condenser to go on charging by incremental gains, cycle by cycle, from 50 volts to its normal average value of say 100 volts must be checked at + 50 volts. This is the function which the diode has to perform.

At time "A" it is assumed that the voltage has been applied for some time and that the condenser has been charged up cycle by cycle to + 50 volts.

Between times A and B the charge will remain constant and $E_c = E_B$.

At time "B," E_B rises 100 volts, *i.e.*, from + 50 volts to + 150 volts. The condenser cannot, as called upon to do, change its charge immediately, so the condenser voltage at time " C " is still + 50 volts. At this instant therefore 100 volts must appear across the output R. This drop of 100 volts across R means that a current must be flowing through R. In these circumstances the charge in the condenser must begin to increase, but only by a small amount, due to the long time constant of the circuit. Let this increase in charge due to the current through R, flowing for the period C D, be say —10 volts.

If during period C D the condenser gains 10 volts, the voltage across R, which was initially 100 volts, must fall simultaneously to 90 volts. At point " D " the applied voltage E_B suddenly drops to + 50 volts. *But the condenser is at* + 60 *volts, which makes the anode of the diode* 10 *volts positive to earth.* Accordingly, the diode conducts momentarily and discharges the condenser from + 60 volts to + 50 volts—the value at which it is to be held.

The sequence is repeated at each cycle, consequently the small increment to the average charge of the condenser which is normally gained during the rise of the input wave is discharged through the diode at time D (*i.e.*, at the fall).

Due to the action of the diode the average voltage of C is never allowed to rise *progressively* above + 50 volts, which means that the current flowing through R can only vary from *zero* towards the *positive* direction, excepting for the very brief moment, after positive D, when the condenser discharges very rapidly through the parallel path formed by the high resistance R and the low resistance of the diode. The effect of this is shown in Fig. 106 (c) by the slight distortion appearing at the base of the E_R waveform.

Alternative Input

Now, suppose the input in Fig. 107 * is applied to the circuit of Fig. 106. In this case it will be convenient to consider what happens at the instant the input voltage is first applied. At time A the input E_B is 0 and the charge in the condenser is 0, *i.e.*, $Q = 0$ and $E_c = 0$. From A to B there is no variation and therefore no change.

At B C, E_B drops suddenly to — 100 volts. Since the condenser cannot change its state suddenly, the instantaneous voltage across it—E—is still zero, therefore, — 100 volts must appear across R—the output (Kirchoff's law).

The cathode of the diode is now 100 volts negative to the

* This input swings from ·0 volts to — 100 volts in the negative direction, but the *diode has not been reversed.* The grid swings 0 — 100 positively.

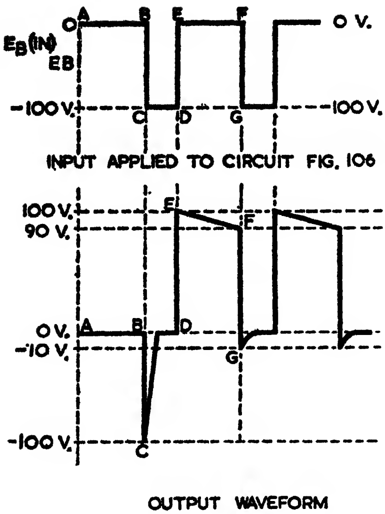


Fig. 107.—EFFECT ON OUTPUT WAVEFORM ON APPLYING ABOVE INPUT TO CIRCUIT OF FIG. 106.

Input swings from 0 to — 100. Output swings from 0 to + 100.

anode and a very high current must flow through the diode, charging the condenser very quickly, so that $E_c = E_B$.

$IR(E_g)$ is now zero, and C remains charged to — 100 volts until, at D, E_B rises to zero, and 100 volts positive appears across R, because the condenser cannot change its charge instantaneously.

C now discharges through the high resistance R very slowly—the diode being in a non-conducting state.

If the RC time constant is long compared with the frequency of the input, the condenser will only lose a comparatively small portion of its — 100-volt charge during the time EF, say — 10 volts, and E_c across the condenser will now be 90 volts.

At F the input voltage again drops to — 100 volts, but E_c is only 90 volts, therefore the anode of the diode is — 90 volts—(— 100 volts) or + 10 volts relative to its cathode.

The diode must again conduct and a current flows sufficient to recharge the condenser again from — 90 volts to — 100 volts and the output is restored to zero. But for the action of the diode the charge on C would rise towards zero and attain an average value at which E_c would be — 50 volts. Variations of E_B would cause E_R to alternate between + 50 volts and — 50 volts which, in this case, has been avoided.

The action of the diode makes C charge initially to —100 volts and maintains it at that value. Thus a variation in E_B which would normally cause E_c to drop — 50 volts is ineffective, since the condenser is already charged to — 100 volts. E_B, in these circumstances, can only cause E_R to vary between 0 and + 100 volts in the positive direction, and this is the voltage wave that appears at the input to the grid-cathode circuit of V_2.

If the diode is reversed it will cause the output waveform to vary between some negative value and the zero reference voltage.

Let an input as shown in Fig. 108 be applied. The sequence is as follows :—

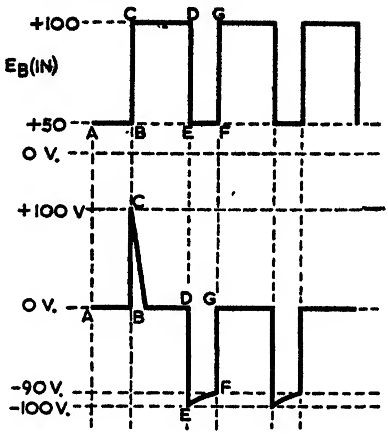


Fig. 108.—INPUT AND OUTPUT WAVEFORMS FOR CIRCUIT WITH ANODE OF DIODE TO EARTH.

At time "A" the anode of the diode is positive. Diode conducts and charges C quickly to — 50 volts. E_R drops to zero.

,, ,, "B" E_B rises to 100 volts. C cannot change instantaneously. 100 volts rise appears across output E_R.

,, ,, "B—C." Anode of diode now very positive. C charges up still further very rapidly. E_R falls to zero.

,, ,, "D." C is at + 100 volts. E_R drops 50 volts.

,, ,, "D to E." C cannot change instantaneously. Therefore E_R drops 100 — 50 = 50 volts.

,, ,, "E to F." Condenser discharges a small amount *viâ* R. Diode being open circuited.

,, ,, "G" Loss during E—F replaced as anode of diode is again momentarily positive.

This sequence repeats.

Effects of Grid Current viewed as clamping Action

In the case of triodes, tetrodes, pentodes, etc., when E_g is made positive grid current flows from cathode to grid. This is sometimes used to derive bias by proportioning C and R so that a positive charge accumulates on C as described in Chapter III, p. 30.

The effect under these conditions may be regarded as resulting from either bias action or a clamping action. For looking at the effect from a clamping point of view, it is seen that the grid of a triode, for example, acts as an anode to the cathode when it draws current and the grid-cathode combination is therefore comparable to a diode.

The current drawn by the grid when it is positive to the

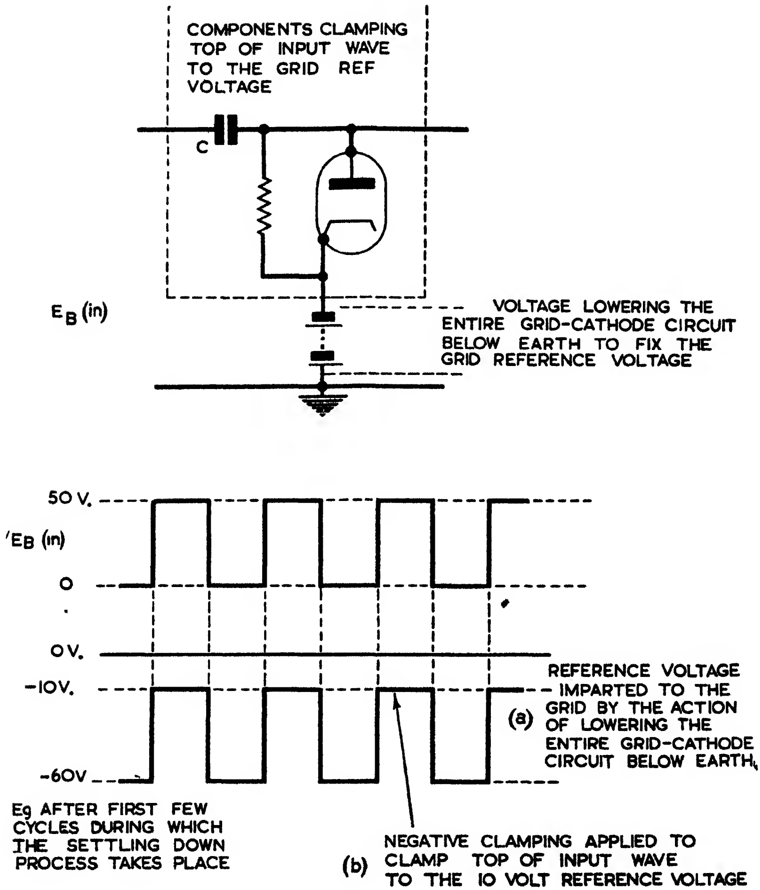


Fig. 109.—INPUT AND OUTPUT WAVEFORMS WHEN NEGATIVE CLAMPING APPLIED TO A NEGATIVE REFERENCE VOLTAGE.

cathode charges C to the value of E_B and $E_c = E_B$. The effect is therefore that E_g can only swing in the negative direction, since it is clamped to zero by the diode action of the grid-cathode of the triode itself.

Fig. 109 shows an extension of the clamping principle.

Reference Positive or Negative to Earth Potential

If the potential of the entire cathode-grid circuit is raised or lowered with respect to earth, the grid reference is no longer directly relative to earth potential. The grid reference in Fig. 109 is the potential of the cathode. Since the cathode is 10 volts negative to earth, the grid reference is also 10 volts negative to earth. E_g can therefore be clamped to this reference instead of to zero. If positive clamping is also applied so that the inputs to the grid are always positive (for example, variations of say 50 volts, all in the positive direction), E_g will swing between -10 and $+40$. Alternatively, if negative clamping of 50 volts is applied when the E_g reference is -10, E_g will swing from -10 to -60 volts. Thus, either or both extremes of the input from a preceding stage may be clamped to a reference potential for E_g, the value of which can also be adjusted by raising or lowering the entire grid-cathode circuit with respect to earth potential.

Synchronised Clamping

In radar practice it is sometimes required that signals generated in a separate free-running circuit and applied continuously to the control grid of a controlled valve shall be neutralised or suppressed at the grid during the intervals between Synch. pulses. A further requirement is that on receipt of a Synch. pulse neutralisation or suppression must cease abruptly and the grid of the controlled valve must be left free to follow the input from the free-running generator for the duration of the Synch. pulse.

These conditions are met by synchronous clamping. During the intervals between consecutive Synch. pulses the grid of the controlled valve is held at a fixed potential by the action of clamping circuits. On arrival of the Synch. pulse, and for its *duration*, the clamping circuits are rendered inoperative, and the input from the free-running generator is then reproduced in the output circuit of the controlled valve.

Analysis of Operation of the Synchronous Clamping Circuit shown in Fig. 110

(*a*) The free-running generator input to V_3 is *viâ* C_1.

(*b*) The potential of the grid of V_3 is controlled by the potential of the anode of V_2.

(*c*) The grid potential of V_1 is controlled by the anode current of V_2.

(*d*) The anode potential of V_2 is controlled by the current flowing through V_1 and V_2 in series.

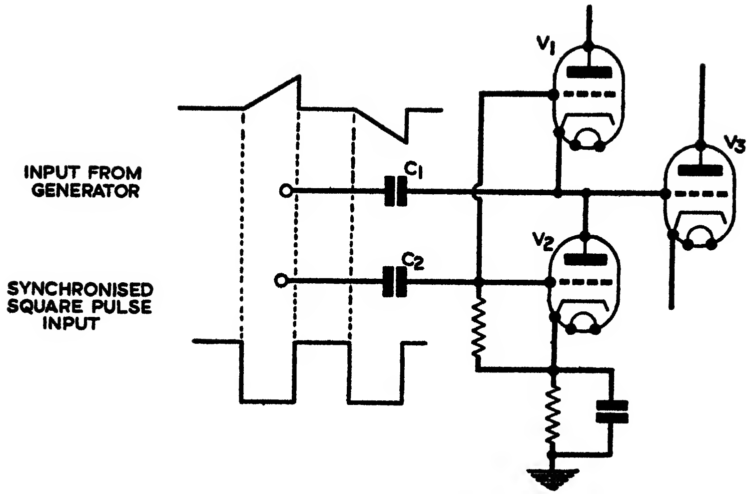


Fig 110.—SYNCHRONISED CLAMPING CIRCUIT.

(*e*) In view of (*b*), the potential of V_3 is therefore controlled by (*d*).

(*f*) The negative Synch. pulse on arrival cuts off V_2 and therefore V_1.

(*g*) The grid of V_3 is free to follow the input wave and to reproduce it at its anode.

Example. A variation of the input *viâ* C_1 in the positive direction makes the anode of V_2 more positive, therefore I_aV_2 increases; E_gV_1 increases; the series current through V_1/V_2 increases; E_aV_2 is restored to normal by the drop due to increased series current and any tendency of E_gV_3 to rise is immediately counteracted.

A similar form of reasoning can be applied to a negative-going input signal.

Example of Clamping Application

One outstanding application of clamping, amongst many, is its use in time base circuits to ensure that each sweep of the time base shall always start from the same value of positioning voltage or current. If this is not done, the trace may not begin at the same point on the screen for each sweep of the time base.

The circuit shown in Fig. 111 is an instance of the application of clamping to ensure that the time base sweep of a magnetically

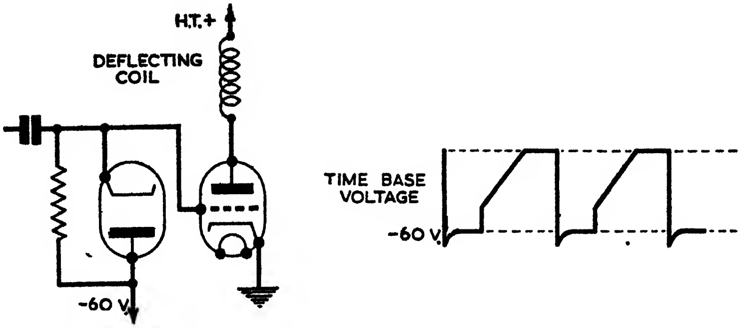


Fig. 111.—DIODE EMPLOYED IN ELECTROMAGNETIC TIME BASE CIRCUIT.

controlled cathode ray tube always starts at the same value of the positioning or static current.

The static current flowing in the deflecting coil before the commencement of the sweep positions the beam at the left-hand side of the screen. Thus the anode voltage at the beginning of the sweep clamped to — 60 volts, as shown in the diagram, ensures constant standing current.

Counting Circuits for Frequency Dividing

Frequency-dividing circuits may be used to detect slow changes in frequency or to trigger other circuits at some multiple of the input frequency or to count.

The positive pulse input to the circuit shown in Fig. 112 produces a voltage across R which is proportional to the input frequency.

Essential requirements to the operation of this device are that the amplitude of the input and its frequency must be constant. If the amplitude varies it will be necessary to carry out a limiting operation on the input, reducing all amplitudes to that of the smallest, before application to the frequency-divider circuit.

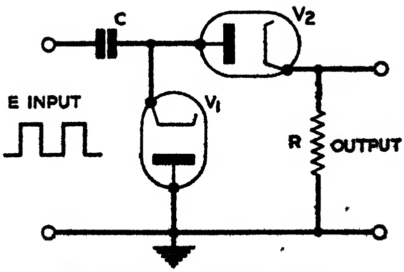


Fig. 112.—POSITIVE COUNTING.

Analysis (Fig. 112).

(*a*) The leading edge of each positive pulse makes the cathode of V_1 positive.

(*b*) Current flows up through R and the condenser charges *viâ* V_2.

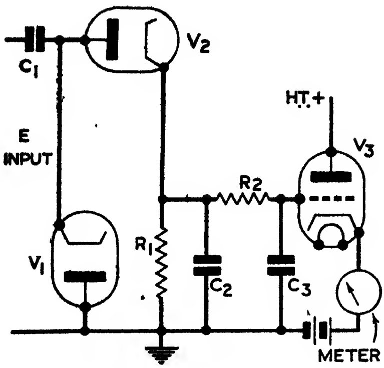


Fig. 113.—MEASURING CIRCUIT FOR POSITIVE COUNTING.

(*c*) At the trailing edge V_1 is open circuited, V_2 conducts and discharges C.

Conclusion

Unilateral currents flow through R each time a pulse is applied; therefore, over a period, an average current flows, which increases with frequency. The drop across R varies with frequency and a change in voltage indicates a change in frequency.

A circuit for measuring slow changes in frequency is shown in Fig. 113. The maximum rate of change measurable is limited by the response of the indicating device. If the diodes are reversed, as in Fig. 114, negative pulses operate the device.

Fig. 115 shows a modification of Fig. 114. In this case positive pulses input charge C_2, which takes the place of R. The charge on C_2 grows step by step as each positive pulse is applied. When C_1 discharges at the trailing edge of the pulse, V_2 conducts and clears the anode of the input stage in readiness to pass the leading edge of the next positive pulse.

Clearly the output voltage E_c across the condenser can be used to trigger some other circuit at any sub-multiple of the input frequency.

Example. Let the input frequency be, say, 2,000 cycles per second, and let it be required to start up a valve (normally at "cut off") 500 times per second, *i.e.*, we require to count down the frequency from 2,000 to 500 cycles. The bias on the output valve must be adjusted so that every fourth input pulse raises the voltage E_c by an amount sufficient to lift the potential on the grid above the "cut off" potential.

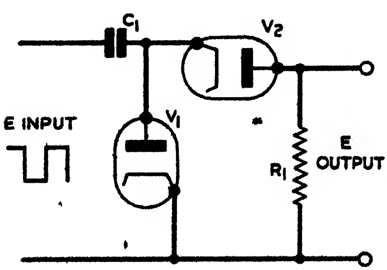


Fig. 114.—NEGATIVE COUNTING.

In this case frequency dividing 2,000/4 is performed by the counting function which accumulated charges on C_2 exercise on the grid of the output valve.

Voltage Doublers

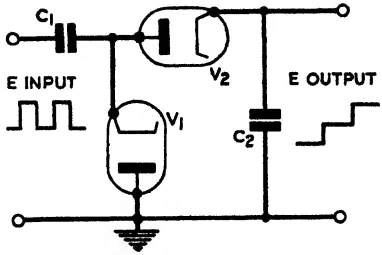


Fig. 115.—STEP-BY-STEP COUNTING.

Radio and radar requirements often call upon power supply circuits for a very high voltage and a small current load. In these cases it is usual to employ a voltage doubler, where possible, since it is much more economical in cost than a special high voltage transformer. Fig. 116 is a schematic diagram of a common form of voltage doubler.

Circuit Analysis

(1) Assume that the upper end of the transformer is positive.

(2) The cathode of V_1 is made negative and the lower plate of condenser C_1 is negatively charged *viâ* V_1 to the peak voltage of the supply.

(3) The lower end of the transformer is positive for the next half cycle.

(4) The cathode of V_2 is made negative by this half cycle causing it to conduct.

(5) Condenser C_2 charges to peak voltage *viâ* V_2 and the transformer so that the lower plate is negative.

(6) Condensers C_1 and C_2 both charged to peak voltage are in series across the load.

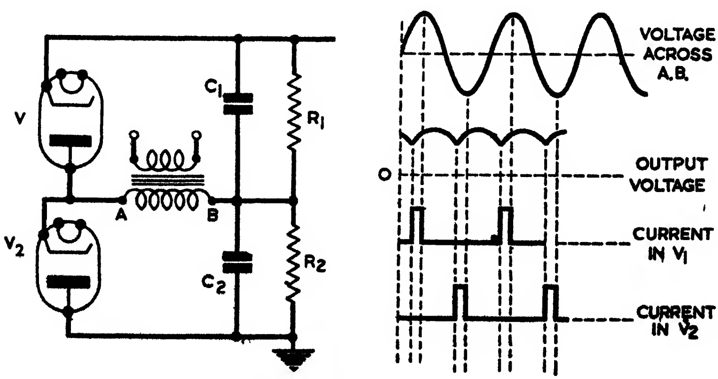


Fig. 116.—VOLTAGE DOUBLER CIRCUIT AND WAVEFORMS.

(7) The output voltage will therefore be approximately twice the transformer voltage.

The voltage regulation of such an arrangement is poor in any case, but may be tolerated if the condensers have large capacity compared to the load.

The cases discussed in this Chapter by no means exhaust all the applications of the diode to radar circuits. One notable exception is that in which a diode is incorporated in the input circuit of a magnetron. In this case, the undirectional conductivity of the diode may be employed to isolate the source of high-tension supply during the period of its discharge through the magnetron. The foregoing is, therefore, to be regarded as a brief description of some of the effects that can be obtained by the use of diodes in combination with various other components.

CHAPTER XI

THE CATHODE FOLLOWER

It is frequently necessary, in radar, to couple a high impedance output to a low impedance input without distortion, such as usually occurs when transformer coupling is employed.

The impedance of the output of the anode circuit of a valve is usually very high and therefore unsuited for coupling directly into a low impedance of the order of 200 to 1,000 ohms.

Such a mismatch would attenuate the output at the far end of the low impedance to a useless value.

A valve, (generally a triode), alternatively a pentode, can be connected to function as a cathode follower, having a high impedance input and a low impedance output. The low impedance output is undistorted and has the same phase as the input.

A cathode follower does not unduly load the source from which it derives its input, and, because of its power gain, it can supply a reasonable current to a low impedance load.

The cathode follower may be connected directly to the anode of the preceding stage since no coupling condenser or grid leak is required. Also positive pulses, within the limit of the bias, may be applied to the grid without drawing grid current.

The amplification factor is, however, less than unity. This is the price that must be paid for all the advantages.

A radar system generally comprises a number of units coupled together with coaxial cable varying in length according to location of the units.

Since the characteristic impedance of a coaxial cable is essentially low, to be effective, inputs must be matched to the characteristic impedance, at the same time having due regard to the load which the latter is required to feed.

Triode Connected as a Cathode Follower

Fig. 117 shows the circuit of a triode connected as a cathode follower. The anode is unloaded and the output is taken across

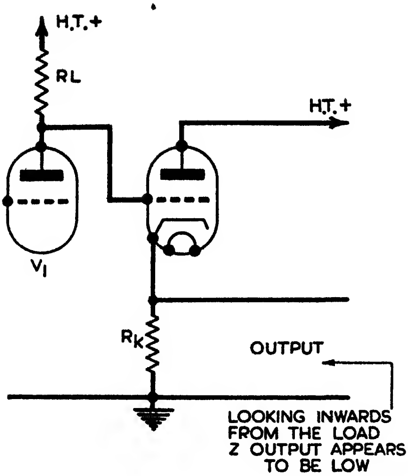


Fig. 117.—CATHODE FOLLOWER BASIC CIRCUIT.

R_k. *Note.* There is a considerable resemblance between this circuit and the normal arrangement for securing automatic bias. In this case, however, no portion of R_k has been by-passed by a condenser, consequently all the requirements for negative feedback are present. In some respects, therefore, the cathode follower has the characteristics of a single-stage amplifier with negative feedback.

Analysis

If a positive signal is applied to the grid, I_a increases and therefore I_aR_k increases so that the cathode is raised to a higher potential above earth. Since the potential of the grid is determined by its input, the cathode potential has moved up towards the grid potential or, in other words, *followed it.*

Since there is no dropping resistance in the anode (although one can be used if required, and provided it is decoupled), the current flowing through the valve for $E_g = 0$ for any given value of H.T. is determined by the value of R_k.

If grid current tends to flow as a result of making E_g positive, I_a increases and the cathode potential rises.

The change of voltage of the grid is E_g input $- E_k$ *output.* This is small, therefore comparatively large positive inputs can be handled without drawing grid current or loading the input impedance.

Within fairly wide limits then the cathode potential rises and falls with the grid potential and by nearly the same relative amount.

For example, if a sine wave of e.m.f. is applied to the grid the cathode potential varies sinusoidally.

If I_{a1} increases to I_{a2}, $I_{a1}R_k$ increases to $I_{a2}R_k$, the grid potential must change from $E_g - I_{a1}R_k$ to $E_g - I_{a2}R_k$. The cathode must now follow this change in grid potential, the nett result of which is that the grid-cathode relations are restored and the load tends to remain constant for varying values of I_a. For this reason R_k appears to the load as a low impedance where $Z = \frac{1}{g_m}$.

When a triode is operated as an ordinary amplifier the impedance of the anode load must be high in order to obtain high gain. Any low impedance in parallel with this load reduces the gain, particularly when the following stage is driven into grid current.

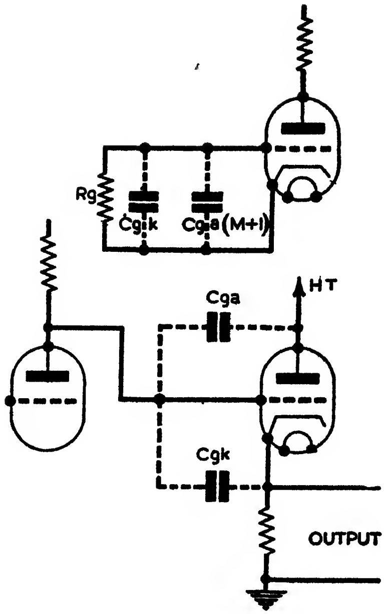


Fig. 118.—CATHODE FOLLOWER AND INTER-ELECTRODE CAPACITIES.

The Miller Effect

Fig. 118 illustrates this condition, C_{ga} and C_{gk} are in parallel across the grid resistance, C_{ga} having capacity due to the Miller effect * of $C_{ga}(M + 1)$ where M is the stage gain.

The shunting effect of these capacities obviously increases with frequency, consequently at high frequencies there will be serious loss of gain and phase distortion, particularly if the grid draws grid current.

Results when a Triode is Connected

Fig. 118 shows that when a triode is connected as a cathode follower the results are :—

(*a*) There is no anode load and the Miller effect is therefore eliminated.

(*b*) C_{ga} is now in series with R_g.

(*c*) Grid bias and grid leak can be dispensed with and direct coupling can be substituted.

These changes bring about a high impedance input which is constant in value over very wide limits.

* The Miller effect is fully described in standard textbooks on radio. Briefly the effect of a capacity connected internally or externally between anode and grid is equal to that of a capacity $C(M + 1)$ connected from grid to earth. C = capacity of the condenser between anode and grid, and M is the stage gain.

If $R_a = \dfrac{E_a}{I_a}$

$g_m = \dfrac{1}{R_a}$

$G_m = \dfrac{I_a}{E_g}$ mutual conductance.

μ = amplification factor.

R_k = cathode load resistance.

$G_k - G_a = \dfrac{1}{R_k}$ cathode load conductance.

E_g = change of grid/cathode volts.

$E_g - E_i$ = change of grid/earth volts or input.

E_a = change of anode volts.

E_k = change in cathode earth voltage.

I_a = change in anode current produced by E_g.

Z = grid-cathode impedance of the valve.

$$\mu E_g = I(R_k + R_a)$$
$$E_g = E_i - E_k$$
$$E_k = I_a R_k$$
$$\mu(E_i - E_k) = \frac{E_k}{R_k}(R_k + R_a)$$
$$\mu E_i = E_k\left(\mu + 1 + \frac{R_a}{R_k}\right)$$
$$\text{stage gain} = \frac{E_k}{E_i}$$
$$= \frac{\mu}{\mu + 1 + \dfrac{R_a}{R_k}}$$

but since $\mu = G_m R_a$

$$= \frac{G_m}{g_m + g_a + g_k.}$$

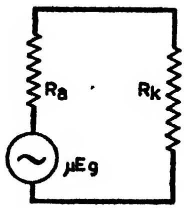


Fig. 119.—EQUIVALENT CIRCUIT FOR DERIVATION OF AMPLIFICATION FACTOR.

This accurate expression for stage gain is obviously always less than unity (the denominator must always be larger than the numerator in the above expression).

A close approximation is therefore :—

$$\text{Stage gain} = \frac{G_m}{G_m + G_k}$$

provided that R_k is much greater than $\dfrac{1}{g_m}$ this value approaches unity.

A valve in which $G_m = 5$ mA./volt has $\frac{1}{g_m} = 200$ ohms, the condition for a stage gain of near to unity in this case is therefore that R_k must be much greater than 200 ohms.

The output impedance of the cathode follower looking across R_k from the output terminals can be examined by applying an external e.m.f. across the output terminals and finding the resulting change in I_a. The grid is held at earth potential and no grid input is applied.

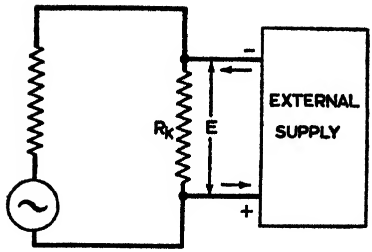


Fig. 120.—EQUIVALENT CIRCUIT FOR DERIVING OUTPUT IMPEDANCE.

The equivalent circuit is shown in Figs. 119 and 120.

A change in voltage applied across the cathode is similar in effect to a corresponding change of grid-cathode volts. The valve generates a voltage change of μE.

Deriving Output Impedance

In the equivalent circuit (Fig. 120), the voltage change across R_a is therefore :—

$$\mu E + E = \frac{E(\mu + 1)}{R_a}$$

also the application of E across R_k gives :—

$$IR_k = \frac{E}{R_k}$$

$$\therefore \text{ the supply current } = I_a + IR_k$$

$$= \frac{E(\mu + 1)}{R_a} + \frac{E}{R_k}$$

$$= E\left(\frac{\mu}{R_k} + \frac{1}{R_a} + \frac{1}{R_k}\right)$$

$$\therefore \text{ the output impedance } = \frac{E}{I}$$

$$= \frac{\mu}{R_a} + \frac{1}{R_a} + \frac{1}{R_k}$$

$$= \frac{1}{G_m + G_a + G_k}.$$

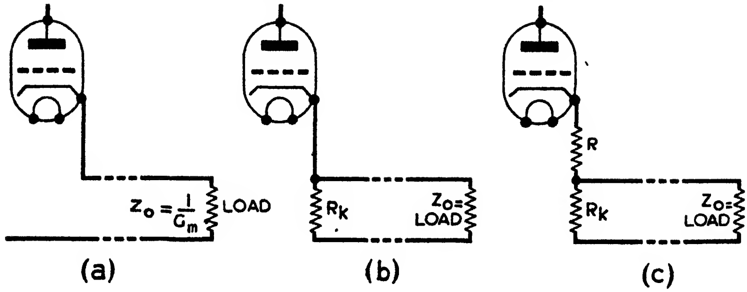


Fig. 121.—THREE ALTERNATIVE METHODS OF MATCHING CATHODE FOLLOWER OUTPUT INTO A COAXIAL CABLE TO SUIT DIFFERENT CONDITIONS OF LOAD.

The output impedance $= \dfrac{1}{G_m + G_k}$

provided that R_k is much greater than $\dfrac{1}{G_m}$.

The output impedance $Z_{out} = \dfrac{1}{G_m}$.

Taking G_m as 6 mA./volt, the following values for output impedance and gain can be calculated :—

R_k	Output Impedance	Gain
1,000	141 ohms	·646
10k.	161 ,,	·966
100k.	164 ,,	·984

Matching

Coaxial cables are electrically *long* or electrically *short*, according to whether the length is greater or less than $\frac{\lambda}{4}$ of the highest frequency involved. If the length of the coaxial cable is greater than $\frac{\lambda}{4}$ of the highest involved frequency it is said to be electrically *long*, when it is less than this standard it is electrically *short*.

In matching to a coaxial cable that is electrically *long* three cases arise :—

(*a*) When $Z_o = \dfrac{1}{G_m}$ the cable is terminated in a load equal to its characteristic impedance Z_o, (see Fig. 121 (a)). This provides the load on the valve and replaces R_k.

(*b*) When Z_o is less than $\frac{1}{G_m}$ (Fig. 121 (b)), a resistance R_k must be included such that $\frac{1}{G_m + G_k} = Z_o$.

This means that $R_k = \frac{1}{G_k} = \frac{Z_o}{1 - G_m Z_o}$.

Whilst this is a perfect match the stage gain will be considerably less than unity if R_k is less than $\frac{1}{G_m}$.

Example. To match a cable of $Z_o = 80$ ohms to a valve with G_m value 6 mA. per volt,

$$R_k = \frac{80}{1 - \cdot 006 \times 80} = 154 \text{ ohms.}$$

The termination at the remote end of the cable would be 80 ohms, and this in parallel with R_k would give an effective load :—

$$\frac{80 \times 154}{80 + 154} = 52\cdot 6 \text{ ohms.}$$

The amplification of $\dfrac{\cdot 006}{\cdot 006 + \dfrac{1}{52\cdot 6}} = \dfrac{1}{\mu}$ approx.

(*c*) When Z_o is greater than $\frac{1}{G_m}$ the output of the circuit can be raised by the addition of the resistance R shown in Fig. 121 (c).

In the normal cathode follower, the impedance "looking in" from the output terminals across R_k is $\frac{1}{G_m}$. The added R is in series with G_m so that the new output impedance is $\frac{1}{G_m} + R$.

$\therefore$ to match to Z_o R must $= Z_o - \frac{1}{G_m}$.

When the Cable is Electrically Short

If the cable is electrically *short*—say less than a quarter of a wavelength of the highest important frequency involved, the question of characteristic impedance does not arise and the load on the valve is simply R_k in parallel with the load at the far end of the cable.

The meaning of the term electrically *short* is made clear in the following example :—

A coaxial cable is carrying a 1·5 microsecond pulse, the lowest frequency involved is therefore $\frac{10}{3}$ c/s. If th efifteenth harmonic of this is considered to be the highest important frequency nothing above 5,000 kc/s. need then be considered. This frequency is proportional to a wavelength of 60 metres, and therefore as long as the cable is less than 12 metres in length it may be considered to be electrically short.*

In order to find the input impedance it is necessary to find the change of input current I produced by change in the voltage E :—

$$\begin{aligned} E_g &= IZ \\ &= R_a I_a + R_k(1 + I_a). \end{aligned}$$

Eliminating E_g and I_a,

$$\frac{E}{I} = \frac{R_a R_k}{R_a + R_k} + Z\left[\frac{R_k(1 + \mu) + R_a}{R_a + R_k}\right]$$

If R_k is large compared with R_a and μ is much greater than 1, this simplifies to :—

$$Z \text{ input} = R_a + \mu Z,$$

which is considerably greater than Z for the same valve when functioning as a normal amplifier.

The high input impedance of the cathode follower, almost entirely capacitative, makes it suitable for connecting to circuits in which damping must be avoided. If it follows a normal stage of amplification, the shunt capacity will be small and satisfactory amplification of frequencies higher than usual can be obtained. It follows therefore that amplifiers can be given a larger load and more amplification can be

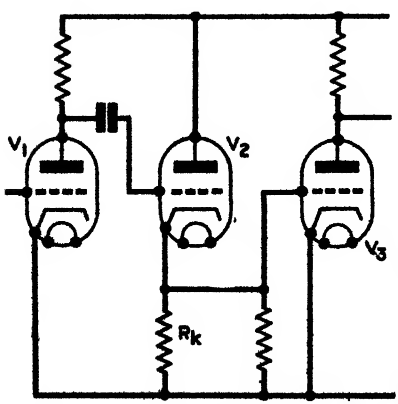


Fig. 122.—Cathode follower coupling V_1 to V_3.

* Note 15 metres would be $\frac{\lambda}{4}$ for a 60-metre wave but it is advisable to make the acceptable maximum about 20% less than the optimum.

obtained without the CR values of the circuit becoming too large.

The absence of the grid coupling resistance puts grid and cathode at the same potential as the anode of V_1.

In Fig. 122 the comparatively large input condenser of V_3 is associated with low output impedance V_2 and will therefore have less effect than if V_3 were coupled direct to V_1. V_2 acts as a buffer stage between V_1 and V_3.

A cathode follower can be used between any components when the resultant loading, if directly coupled, would impair operation.

Applications in a Buffer Stage

In a linear time base, for example, where a condenser is charged through a constant current device, any shunt resistance across the condenser would adversely affect the linearity. In order to avoid this a cathode follower is incorporated as a buffer stage between time base and amplifier, thus avoiding distortion to the sawtooth waveform.

A second example is that of the linear potentiometer used in certain range transmission units. The voltage of the slider must vary linearly with the setting. To avoid any shunt resistance between slider and earth, the slider is taken to the grid of a direct-coupled cathode follower.

Low output impedance makes it suitable for use with valves that are to be driven well into grid current or non-linear circuits in general.

For example, the gas triode of a modulator and large modulating valves of some transmitters where the peak grid current may run to 2 amps. or more. In order to minimise distortion due to this cause, a source with impedance as low as possible must be used.

A cathode follower may even be applied to improve the regulation of some H.T. supplies.

Use as Impedance Transformer

The combination of high impedance input and low impedance output enables it to be used as an impedance transformer when it is required to switch a low impedance load to a high impedance source, as, for example, in matching a moving-coil loud speaker to an output valve.

The cathode follower may be matched to any desired output. It is therefore used for feeding into pulse-carrying cables between various units of a system.

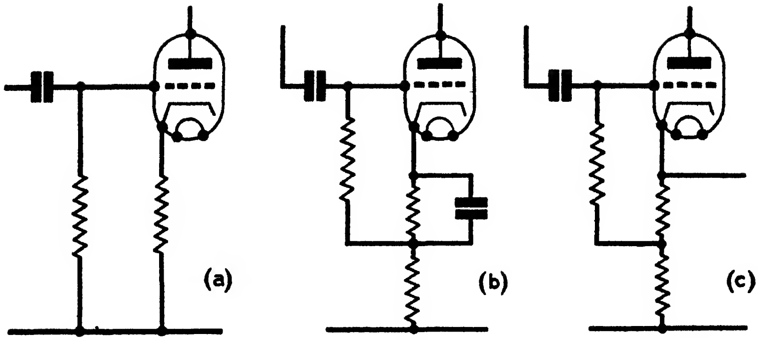


Fig. 123.—ALTERNATIVE METHODS OF OBTAINING BIAS WITH A CATHODE FOLLOWER.

(a) Parallel with the Output Cathode resistance.
(b) Auto bias in series with the Output Cathode resistance.
(c) Auto bias derived directly from part of the Output Cathode resistance.

If the cable is electrically long the cathode follower is matched as described. If it is *short*, matching is not necessary. In the latter case, however, the cathode follower is still used, as its low output impedance avoids the shunting effect which the capacity of the cable would otherwise cause.

There is only slight distortion at the input circuit when it consists of a C.R. circuit of short time constant, the output being across the condenser.

For example, the resistance in the anode load of a preceding stage and the condenser formed by the small input capacity of the valve itself.

As regards the output circuit, account must be taken of the stray capacities across the load. This is again a C.R. circuit with the output across the capacity. Provided the valve is operated within its grid bias, the time constant is short, particularly if $\frac{I}{G_m}$ is small.

For a positive input pulse the valve does operate within its grid bias owing to the following action of the cathode, and it will take an input of considerable magnitude without running into grid current. In fact, the output is only limited by the input applied to the valve.

Thus a cathode follower gives less distortion with positive pulses than it does when the time constant is the load resistance × output stray capacity.

For negative pulses of reasonable amplitude, however, the valve is driven to cut off when the output impedance ceases to be $\frac{1}{G_m}$ and rises to the load R. The time constant then rises to $C \times R$, which is much longer than before, and causes a correspondingly greater distortion of the pulse.

For positive A.C. input pulses the valve requires bias almost to cut off.

Methods of Obtaining Bias

Fig. 123 shows three methods of obtaining a small negative bias such as is required for positive input pulses. The grid lead is taken to the earthy end of R (value 200 ohms). The drop across it produces the necessary bias.

A pentode can be connected as a cathode follower. The advantage obtained in using this valve is that its input capacity is only about one-third that of a triode when similarly connected. Unless the object is to keep down input capacity, a triode is nearly always used.

The Pentode Cathode Follower

The pentode cathode follower has the following disadvantages :—

As regards input impedance, the screen is effectively in parallel with the cathode load and therefore limits its maximum value.

Because of decoupling at the screen, the A.C. component of the screen circuit does not flow through the cathode load so that the effective gain of the valve is reduced.

A pentode gives a considerably decreased signal to noise ratio.

Sometimes pentodes are used as triode cathode followers. In this case screen, suppressor and anode are strapped together. In these circumstances the screen picks up the majority of the valve current and the valve constants are modified accordingly.

The following table is an indication of the effect produced :—

	E_a	E_s	G_m	R_a	μ
As pentode	200	100	4·4 mA/v.	$\frac{1}{2}$–1 megohm	3,000/4,000
When screen suppressor and anode are strapped together					
	200		6·4 mA/v.	10,000 ohms.	

CHAPTER XII

RECTANGULAR PULSES HAVING A DEFINITE TIME DURATION AND THE DEVELOPMENT OF HIGH-VOLTAGE HIGH-POWER TIME CONTROL PULSES AT LOW AND HIGH LEVELS

IN general, a rectangular pulse is employed when one of the primary objects is to exercise an operating time control over the circuit which it must also trigger.

The circuit to which the pulse is applied, therefore, generally loads the pulse generator, but there is also the special case of loading in which the pulse, besides acting as a trigger-timing pulse, is also required to act as a source of power supply for the duration of the pulse. Then the load on the pulse generator itself is necessarily comparatively heavy.

In order to obtain accurate timing for synchronisation and duration, a rectangular pulse should have a waveform with vertical leading and trailing edges and a top as flat as possible.

General Requirements of Trigger Time Control Pulses and Pulse Generators

If the pulse is to function purely as a trigger time control pulse, its amplitude is generally determined by the magnitude of the voltage that must be applied, either to drive a valve to increased conduction (or into conduction if initially cut off), alternatively to reduce normal conduction or to depress the grid to " cut off."

The duration of the applied pulse determines the duration of the output in all cases when the electrical characteristics of the circuit which it triggers do not themselves exercise any operating time control function.

When, in addition to triggering, the pulse exercises a time control function, because it is also the source of power supply, the general requirements are more exacting.

This arises from the facts that (*a*) the pulse generator is more or less heavily loaded for the duration of the pulse, (*b*) the voltage amplitude of the pulse required for power supply to a magnetron, for example, may be of the order of several thousand volts, perhaps 20,000 or more.

It is obvious that when the pulse generator is heavily loaded by reason of the power it is called upon to supply for its duration of operation the regulation of this source must be reasonably good. In other words, the terminal voltage of the generator must not drop very much during the time the load is applied (pulse duration). This means that the internal resistance of the pulse generator must be low and the load power factor high, this latter implying good matching and high overall efficiency.

The maximum voltage which a valve can stand without inter-electrode arcing taking place is in the order of 27,000 volts, consequently when voltages of this magnitude or greater are required, the valve can no longer be employed and some more robust device (electrically) must take its place.

Primarily, therefore, two entirely different methods of generating high-voltage high-power rectangular pulses arise :—

(*a*) With a valve when the voltage and power output are within its handling capacity.

(*b*) By the use of a spark gap, which is capable of handling very large power at high voltage.

These distinctions apply only to power-time control pulses for application to R.F. generators and the like.

When very accurate time control pulses of short duration and comparatively small amplitude are required, it is usual to generate, under the control of a Synch. triggering pulse, a damped oscillation, the frequency being such that the time period of one half cycle is equal to the required time duration for the pulse. Either the positive or negative half cycle is then isolated (choice of polarity depending upon input requirements of the circuit to which it is to be applied). The selected half cycle generated in this manner is then clipped, squared and amplified if necessary.

In the case of control pulses of small amplitude, if there is any danger of the load influencing or distorting, the shape of the rectangular pulse unduly, because of its phase or magnitude, a cathode follower can be introduced as a buffer stage between the pulse generator and its load.

Convenient Sub-classifications and Requirements

In order to facilitate examination of the various types of rectangular pulse generators commonly used in radar, it is convenient to consider them under two main headings :—

(*a*) Circuits for generating power-time control pulses such as those employed in transmitter circuits.

(*b*) Circuits for generating similar pulses of smaller voltage amplitude, when power supply is not involved as a major function.

The requirements common to (*a*) and (*b*), apart from their rectangular shape, are related to (*a*) repetition rate and (*b*) pulse duration.

Generators for (*a*) and (*b*) must both be synchronised to the repetition rate established for the system, and whilst the pulse durations for some transmitter circuits may differ from those employed in other parts of the radar system certain common components (differing perhaps in value) may be used in similar circuit arrangements for pulse-width-forming equally well in both cases. This refers to artificial lines.

The Artificial Discharge Line

This device is a particularly good example of a pulse-forming component commonly employed to function as a pulse-forming device in various positions in radar systems. It serves to store energy for the pulse and at the same time controls the rate at which the energy is released.

Thus the requirements for the production of a rectangular pulse may be summarised as *means for storing energy during the interval between consecutive pulses and for releasing it at some constant rate which controls pulse time and average power. Thus*

$$\frac{\textit{Energy stored}}{\textit{Release time}} = \textit{pulse width.}$$

The fundamental difference between requirements of transmitter pulses and pulses employed in other parts of the radar system is, therefore, *the amount of energy to be stored per pulse.*

Thus the considerations in connection with the formation of a power-time control pulse are :—

(1) Means for storing energy.

(2) Means for releasing the energy at the right moments (synchronisation with the p.r.f.).

(3) Means for controlling the rate of release.

(4) Means for cutting off the supply of energy when the proper pulse width (time duration) has been established.

As a point of interest, it may be mentioned here that the artificial discharge line performs functions (1), (3) and because the amount of energy which it stores can be related directly to the rate of delivery, it performs function (4) automatically. For this reason it is used extensively in rectangular pulse generators to produce accurately timed pulses of low or high power for many purposes.

Alternative Systems

As an alternative to the artificial line, a condenser may be employed in which the energy can be stored, but in this case the rate at which it gives up its energy and the duration of its discharge (which constitutes the pulse width) must be controlled by a modulator.

The modulator, which may be a valve or a spark gap, acts as a switch for turning on and off stored energy. In this way the commencement of each discharge and its duration, *i.e.*, the repetition rate and the pulse width, are both controlled by the modulating device.

When a valve is used as a modulator, it is generally controlled for Synch. and duration by a sub-modulator. The sub-modulator may or may not be the shaping circuit according to circumstance, but it is generally controlled for Synch. by a Synch. pulse supplied either by some form of independent Synch. pulse generator or, in some cases, it may be self-generated.

As stated above, the shaping of the pulse applied by the sub-modulator to the modulator, for width (time duration), may be performed sometimes in the sub-modulator itself, but when the modulator has to control large voltages the sub-modulator must also supply high-voltage pulses to the modulator to enable it to function in this manner, consequently the shaping operation is usually performed in special shaping circuits, at low level, and amplified.

An alternative system to the above is to charge an artificial discharge line so that it stores the necessary amount of energy required for each pulse, and to discharge it by means of some form of spark gap or other discharging device. The instant of each discharge (p.r.f.) is then regulated by the discharging device, mechanical or otherwise, but the artificial line controls the rate at which the energy is given up and consequently the duration of the resulting pulse.

When this system is employed the power-time control pulse is said to have been generated at high level, in contra-distinction to the method previously described, when it is generated at low level *viâ* shaping circuits, sub-modulator, and modulator or similar device.

It will be noted that when the pulse is generated at high level successive discharges across the spark gap establish a repetition rate and, therefore, it is usual to use this effect to synchronise the entire radar system by a pulse taken from the transmitter

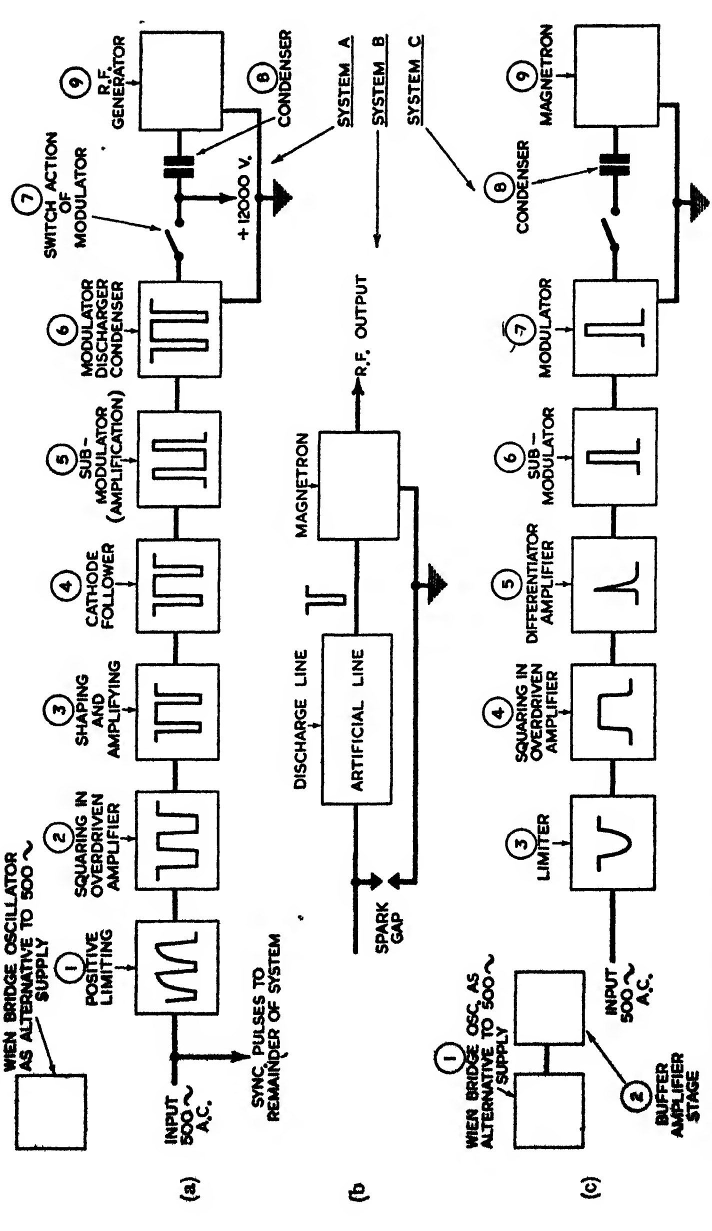


Fig. 124.—Comparison of methods of generating rectangular power pulses for application to magnetron at low and high levels circuit when high-level generating methods are employed. Figs. 124 (a) and (b) contrast the two systems.

Another alternative system which has been used for the production of small and medium power-time control pulses at low level is very similar in principle to that just described. The main difference lies in the use of a high-gain amplifier as a sub-modulator. In this latter case, too, the sub-modulator circuit is the shaping circuit and therefore the differentiator amplifier is arranged to provide a sharp-peaked wave of large voltage amplitude to trigger the sub-modulator circuit. In system A, however, the output from the over-driven amplifier is applied to a shaping circuit which develops an output wave of the required pulse width by differentiating the grid input and using this in conjunction with bias action to produce a rectangular output of the required duration.

The two low-level systems of generating the power pulse for the magnetron and the high-level system are compared in Fig. 124 (a), (b) and (c).

Thus in the following pages the circuits involved in the two low-level systems are first discussed, followed by a similar discussion of the circuits employed for the high-level system.

Details relating to spark-gap dischargers and similar devices are fully dealt with in the chapter on transmitters.

Accordingly, let it be assumed that an R.F. generator is required to produce a 2-microsecond pulse of R.F. energy repeated at say 500 times per second. Let the Synch. pulse be taken for this repetition frequency from a 500-cycle A.C. main supply,* and let the pulsing system selected be as that shown in Fig. 124 (a).

The wave of 500 cycles per second obtained from the main supply has been roughly squared by a limiting or clipping process and further squared, so as to make the leading and trailing edges more vertical, by an over-driven amplifier.

A square pulse is now available of duration 1,000 microseconds with a recurrence frequency or repetition rate of 500 per second.

* In case the frequency of the main supply is unsuitable for the repetition rate which has been chosen for the system and where the synchronising function is to be performed by an independent master timing control, a Wien bridge oscillator circuit can be used as a Synch. pulse generator, adjusted to give a p.r.f. at any given rate. This oscillator is described in Chapter V. When a Wien bridge oscillator is used as a master synchronising device a buffer stage is inserted between it and the shaping circuits in order to prevent deterioration of shape or frequency by the loading conditions imposed by succeeding circuits.

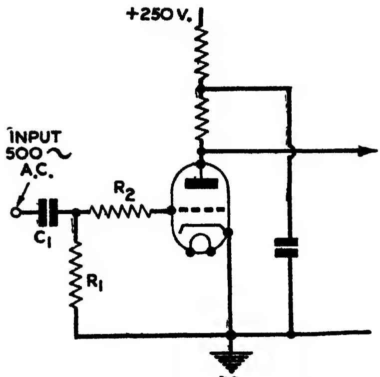


Fig. 125.—Positive limiting by use of grid stopper.

From this square pulse a high-amplitude negative-going pulse of 2 microseconds' duration is developed in the shaping circuit (3 in Fig. 124 (a)) at a frequency of 500 per sec.

This pulse is fed to the grid of the sub-modulator (5) *viâ* a cathode follower (4) in order to avoid the distortion in the shaping circuit which would be caused by heavy loading.

The voltage output from the sub-modulator is a positive-going amplified version of the input from the shaping circuit, and this is applied to the grid of the modulator valve (6) in order to drive it into heavy conduction from its normal cut-off condition.

The condenser (8) which has charged up to the value of the

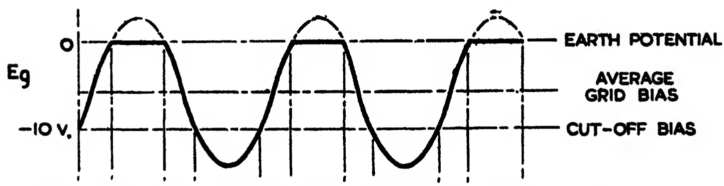


Fig. 126.—Grid voltage waveform due to action of stopper. Circuit fig. 125.

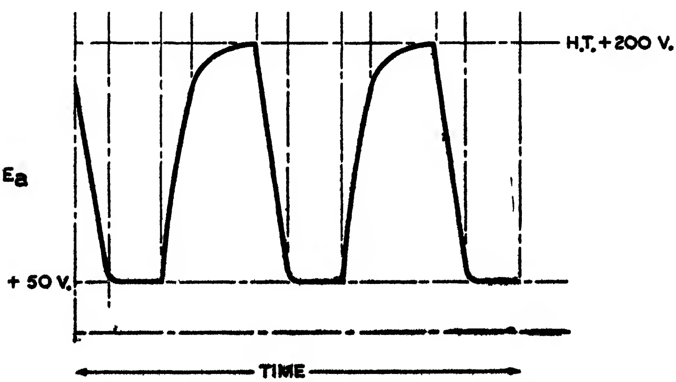


Fig. 127.—Waveform at anode for circuit fig. 125.

supply voltage during the shut-down time of the modulator, on arrival of each positive pulse from the sub-modulator, now discharges through the modulator and the magnetron for the whole time during which the modulator is conducting, the discharge being cut off at the end of each 2-microsecond positive pulse from the sub-modulator, as the modulator reverts to the " cut-off " state.

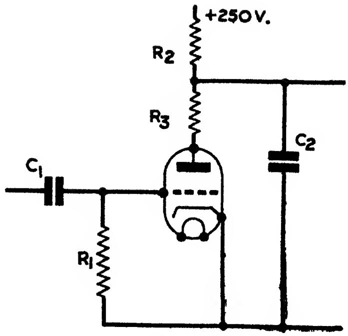


Fig. 128.—CIRCUIT ARRANGEMENT FOR OVER-DRIVEN AMPLIFIER.

Each discharge of the condenser *viâ* the modulator constitutes a 2-microsecond negative-going pulse to the cathode of the magnetron.

The circuit arrangements for stages 3, 4, 5, 6 and 7 in the block diagram, Fig. 124 (a), can now be examined.

Positive Limiting by use of Grid Stopper

The limiting operation for initial squaring purposes can be performed, for example, by the circuit shown in Fig. 125. The grid voltage waveform due to the action of the grid stopper R_2 is as shown in Fig. 126. The output waveform is shown in Fig. 127. Bias is obtained from the residual charge in C_1 resulting from the difference between the time constant for charging on the positive swing $\left(i.e., C_1 \dfrac{R_{gk}R_1}{R_{gk}+R_1}\right)$ and the time constant for discharging which is C_1R_1 (very much longer).

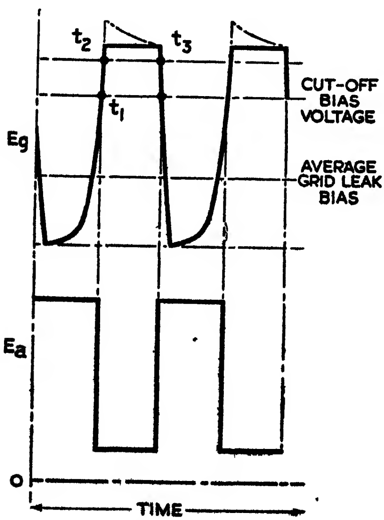


Fig. 129.—WAVEFORM FOR FIG. 128.

The average grid potential about which the sine wave input swings is represented by this residual bias. This is shown in Fig. 126. When the grid is driven positive R_2

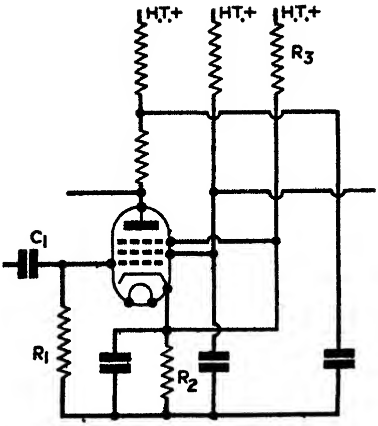


Fig. 130.—SHAPING CIRCUIT.

still further limits the signal by the voltage drop across it due to grid current, *i.e.*, R_2I_g. The effect on the positive peak in the output is shown in Fig. 127. The least positive portion of the anode waveform corresponds to the period of grid current flow and is flattened by grid limiting. As the grid signal swings negative, anode current is reduced and ceases to flow when the combination of signal and bias reaches the cut-off value.

The input to stage 2 (Fig. 128) is the voltage waveform shown for the output of stage 1 (Fig. 127). This is a pulse of large amplitude, consequently the anode current of stage 2 is cut off early in the negative alternation and is driven to a maximum (saturation) early in the positive alternation.

A high negative bias is developed by the charge on C_1 resulting from large grid current drawn during the positive half cycle of the input.

The anode waveform is distorted somewhat by the charging and discharging of C_1 (Fig. 129).

The average grid leak bias developed fixes the reference potential of the grid at a value below cut off, consequently the valve is not driven into conduction until the potential of C_1 on the rising input curve is reached. The valve then conducts heavily until time t_1 is reached, *i.e.*, saturation condition attained. The anode current then remains constant until time t_3 is reached, when it rapidly falls from saturation value to cut off.

The drop across the anode load caused by the large anode currents which flow between t_1 and t_3 results in a negative-going square pulse of large amplitude. The duration of this pulse is however very much longer than the 2 microseconds required; consequently the next operation in stage 3 (Fig. 130) is to reduce the pulse width to the required value.

When a Narrow Pulse of High Voltage is to be Developed

Fig. 130 shows a circuit sometimes employed when a narrow pulse of high voltage is to be developed by two or more stages

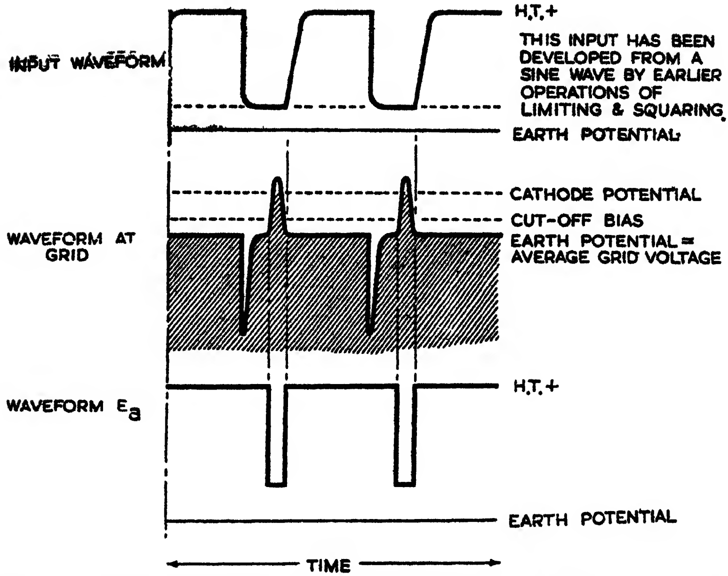


Fig. 131.—INPUT AND OUTPUT WAVEFORMS FOR CIRCUIT SHOWN IN FIG. 130.

of amplification from a broad and rectangular wave input. This circuit is suitable for the shaping operation.

If the output is to be a negative-going pulse of say 40 volts with a duration period or width of 2 microseconds, this means that a long narrow pulse is required to be formed from the broad rectangular wave which forms the input. From the E_g waveform, Fig. 131, it can be seen that the waveform of voltage at the grid of the pentode in Fig. 130 is the differentiation of the broad square wave output of the previous stage, *i.e.* the input waveform, Fig. 131. The bias is so arranged that all the shaded portion of the peaky wave derived from differentiation of the leading edge of the broad square wave is lost with exception of the positive extremities shown above the cathode potential reference. These positive " pips " at comparatively high level cause the valve to conduct heavily and so produce the negative-going square wave by the drop across the anode load (Fig. 131).

The first operating requirement is that the time constant of the input circuit shall be short. Because of this, C_1 charges and discharges completely, consequently the average bias which it produces is negligible. The grid may therefore be considered as at earth potential.

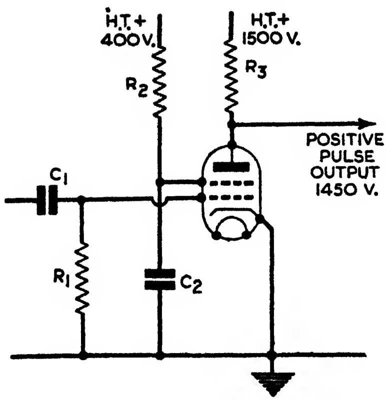


Fig. 132.—SUB-MODULATOR.

The *time constant* of the grid-coupling circuit, C_1R_1, is made approximately equal to 2 microseconds for a negative voltage swing, but it would be less than 2 microseconds for a positive voltage swing at the grid, because of the low cathode-grid shunting resistance when grid current flows.

When the positively increasing pulse is applied, E_g rises instantaneously to maximum value and falls to zero quickly. Similarly, when the applied signal swings the grid negative, E_g is driven below earth potential but returns rapidly to zero.

The valve is, however, biased beyond cut off to eliminate the negative swings, and the cathode is raised above earth potential by + bias *viâ* R_2 and R_3. Since the required negative-going pulse at the anode must be produced by a positive voltage applied to the grid, this bias is such that the valve does not conduct until the grid is + 6 volts above earth, hence Fig. 131 shows both negative pulses and the broad lower base portion of the positive pulse are lost due to the bias on the valve.

The initial portion of the grid signal which tends to drive the grid more positive than the cathode is also lost because of the limiting action of grid current and C_1.

The peaks of the positive pulses, however, tend to drive the grid very positive relative to the cathode and therefore cause a large anode current to flow for the duration of the pulse tip, which appears at the grid for 2 microseconds. E_a therefore drops to a low value during this period and the output is a high amplitude negative-going pulse of 2 microseconds.

Fig. 132 shows the circuit of an amplifier suitable for amplifying the 40-volt 2-microsecond pulse output of the circuit just described. A 1,450-volt positive-going pulse of 2 microseconds is developed in this stage.

The control grid is at zero bias so that the valve normally draws a current of about 150 mA. This causes the anode volts to fall to about 50 volts (a drop of about 1,450 volts). In the absence of a signal on the grid, R_2 limits the screen current to a safe value. C_2 maintains the screen volts constant.

Under the conditions of operation the cut-off bias is — 38 volts. The output from the previous stage is 40 or more volts negative. The application of this pulse cuts off the valve and the anode voltage rises to 1,500 volts.

Thus the output is a positive-going pulse of 1,450 volts.

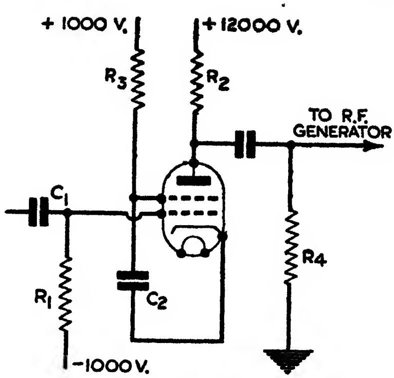


Fig. 133.—THE MODULATOR.

The Modulator (6) and (7) (Fig. 133)

Fig. 133 shows the circuit arrangements for the modulator and energy-storing condenser (8), Fig. 124.

Normally the modulator valve is made non-conducting by the — 1,000 volts bias on the grid. In this condition the anode condenser charges to + 12,000 volts.

However, the large positive-going 2-microsecond pulse from the output of the sub-modulator causes the modulator to conduct heavily and the condenser discharges through the modulator in series with the parallel combination of R_4 and the R.F. generator (magnetron).

The capacity of the coupling or storage condenser is large and such that a drop of approximately 5 per cent. only occurs during the oscillation of the magnetron. This is important because any excessive drop in the pulse voltage on load causes the frequency * and power output of the magnetron to change during the pulse.

In the alternative system for generating a power-time control pulse at low level, shown in Fig. 124, the Synch. pulse is taken from the A.C. supply or from an independent Synch. pulse generator of the Wien bridge type. (In the latter case a buffer stage follows for reasons already explained.) The positive half cycle can be suppressed by a series rectifier and the negative half cycle is applied to an over-driven amplifier, the output of which is a broad square pulse. The leading edge of

* Unlike most R.F. generators, the frequency generated by a magnetron does not remain constant if the voltage of the H.T. pulse rises or falls, instability occurs when the rectangular high voltage pulse is not flat topped.

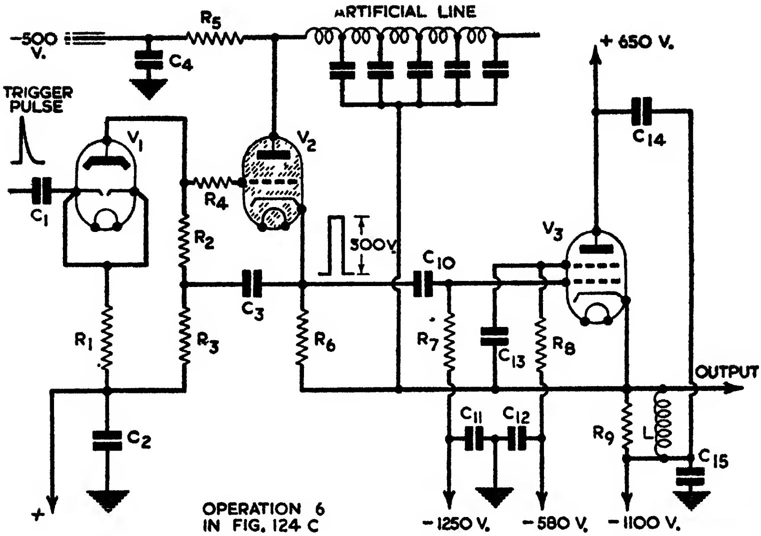


Fig. 134.—HIGH-GAIN VOLTAGE AMPLIFIER OR " BOOT-STRAP " CIRCUIT (U.S.A.), USED IN CONJUNCTION WITH AN ARTIFICIAL LINE. THE COMBINATION SHAPES AND AMPLIFIES.

the square pulse, which is now too broad, is differentiated and amplified in a differentiator amplifier, the circuit components and time constants of which are arranged to produce a sharp-pointed trigger pulse of large amplitude at the output.

All these processes and the circuits involved have already been described either in this chapter or in preceding chapters.

The trigger pulse of large amplitude obtained from the output of (5), Fig. 124 (c), is applied to trigger the sub-modulator circuit, Fig. 124 (c) (6).

The particular form of sub-modulator circuit employed in this case is an interesting example of an amplifier circuit arrangement for high gain.

Sub-modulator for Position (6) in Fig. 124 (c)

High Gain Amplifier Operated by Output from Artificial Line

In the circuit shown in Fig. 134 a sharp positive-going trigger pulse causes the artificial line to discharge *viâ* the thyratron. The rectangular pulse thus formed is then amplified by the high-gain amplifier. The artificial line performs the shaping function for pulse width.

The artificial line * is connected between — 500 volts and — 1,100 volts; it is therefore charged to 600 volts *viâ* R_5. A positive-going Synch. pulse with steep leading edge is applied *viâ* C_1 and the diode V_1 which causes V_2 to ionise and conduct. The current that now flows through V_2 must come from the artificial line, because R_5 is too large to allow a flow of current sufficiently large to maintain ionisation of the thyratron. The discharge of the artificial line through V_2 and R_6 produces a — 300-volt pulse across R_6 during the discharge.

V_3 is a double beam power valve (shown as a single valve). It is normally cut off since its cathode is connected to — 1,100 volts and its grid is connected to — 1,250 volts.

The nett bias is therefore — 150 volts, which holds the valve at cut off.

The — 300-volt pulse developed across R_6 is coupled to the grid of V_3 *viâ* C_{10}, causing V_3 to conduct strongly.

The cathode of V_3 *rises from* 1,100 *volts below earth to a positive potential with respect to earth because of the drop across* R_9.

Note. The life of the thyratron is shortened by positive ion bombardment of the cathode if the grid is driven sufficiently negative to cause acceleration of ions towards the cathode. This condition is avoided by the use of V_1 and C_3. A positive swing makes V_1 to conduct and a signal is produced across R_2 and R_3. The drop across R_2 is applied between grid and cathode of V_2 *viâ* C_3. When the pulse is removed, C_1 discharges and the diode prevents the negative swing appearing at the grid of V_2. The pulse developed in the cathode of V_2 is coupled back to the grid *viâ* C_3, causing the grid to rise with the cathode. This prevents a negative voltage on the grid from this source. R_4 limits grid current through V_2.

This amplifier contains elements which cause the voltage at the grid of V_3 to rise with the cathode voltage, *thus maintaining a constant signal voltage from grid to cathode.*

These elements are :—

(*a*) The source of anode voltage for the preceding valve which is *not tied directly to earth* (the artificial line discharge).

* An artificial line is, as the name denotes, a device to simulate the properties of a real line. One of these properties is that it can be charged to any given potential and when discharged it gives up the energy stored in the charge at a definite rate determined by the length of the line. This device is used extensively for pulse-forming purposes, since by adjusting its length (electrically), accurate pulses of any desired duration may be formed. The artificial line is fully described later in this chapter.

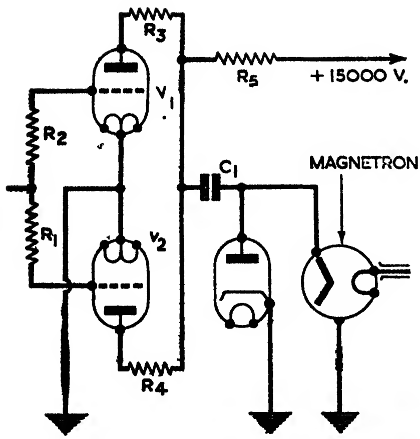


Fig. 135.—MODULATOR OPERATED BY THE OUTPUT FROM THE CIRCUIT OF FIG. 134.

(*b*) A means for coupling the rise of potential of the cathode to this anode supply.

(*a*) The artificial line is the source of anode voltage for V_2 during the pulse, and it is partly isolated from earth by R_5. The only other path to earth is *viâ* R_9 in the cathode of V_3. Thus the cathode of V_3 rises during the pulse and the entire circuit of V_3 is also raised by the same amount above earth.

The drop across R_6 is maintained by the discharge of the line and is applied directly between grid and cathode of V_3. *Thus the amplifier raises its grid circuit voltage in order to maintain a constant grid signal.*

When the artificial line is completely discharged the p.d. across R_6 disappears and the potential of the grid of V_3 at once returns to — 150 volts with respect to the cathode and the valve is again cut off.

The large positive pulse that appears across R_9 during the above sequence of operations is the output. This form of amplifier is known in the United States as a " boot-strap " for reasons which are obvious.

The Modulator (No. 7 in Fig. 124 (c))

Initially C_1, Fig. 135, is charged from the 15,000 volts supply *viâ* the diode and R_5, and V_1 and V_2 in parallel are " cut off." The arrival of the large positive-going rectangular pulse from the sub-modulator, drives V_1 and V_2 into heavy conduction. C_1 now discharges to earth through V_1 and V_2.

Since V_1 and V_2 are in parallel, the resistances R_1, R_2, R_3 and R_4 are provided to prevent parasitic oscillations.

The diode is employed to permit charging to take place but to prevent loss of energy during the pulse, since none of the discharge current can flow through the diode.

Two possible alternative methods of producing power-time

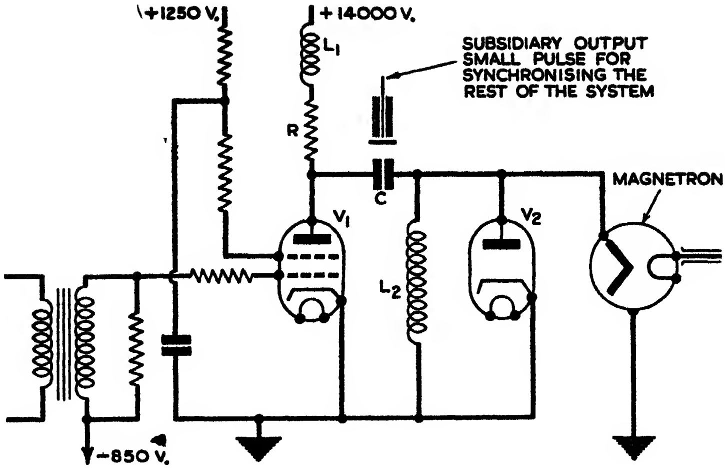


Fig. 136A.—ALTERNATIVE CIRCUIT FOR MODULATOR DRIVEN BY A FORM OF SELF-PULSING SUB-MODULATOR.

control pulses of large voltage amplitude are shown in Figs. 136A and 137 (c).

The system shown in Fig. 136A employs the controlled discharge of a condenser to supply the energy and to form the high-voltage power pulse.

The system shown in Fig. 137 (c) is, however, unique, and is really an example of *generating* as distinct from developing a power-time control pulse of high voltage at high level.

Alternative Circuit for Modulator Driven by a Self-pulsing Sub-modulator

Fig. 136A shows one of the many alternative circuits available for modulators.

In this case the sub-modulator is a self-pulsing oscillator, the output of which is a trigger-control pulse having a peak voltage of about + 1,000 volts.

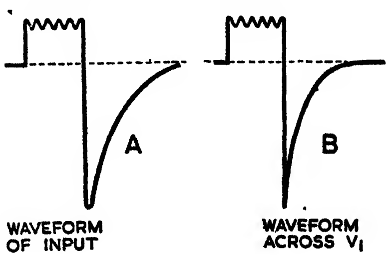


Fig. 136B.—CURVES A AND B SHOW THE EFFECT OF L_2 ON THE RATE OF DECAY OF THE PULSE APPLIED TO THE MAGNETRON. A IS WITHOUT COMPENSATING EFFECT OF L_2 AND B IS TAKEN WITH L_2.

V_1 is normally cut off by the — 850-volt bias applied

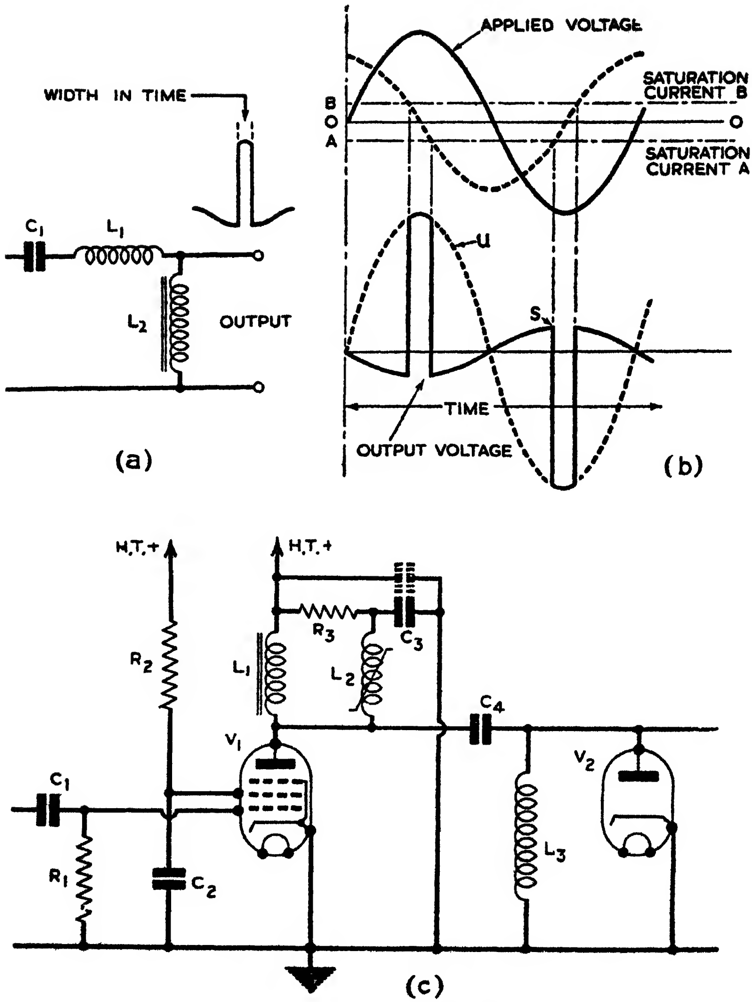


Fig. 137.—METHOD OF GENERATING HIGH-VOLTAGE HIGH-POWER PULSE OF SHORT DURATION FROM A SINE WAVE INPUT.

(a) The series resonant circuit $C_1L_1L_2$ (C_1).
(b) Development of peak voltage across L_1 during period L_2 is unsaturated.
(c) Square wave generator with saturated coil as switch (L_2) 3 μ sec. pulse.

to its control grid. C is charged from the 14,000-volt supply through V_2, R and L_1.

When the 1,000-volt trigger pulse from the sub-modulator is applied to the control grid of V_1, C discharges through the magnetron and V_1 and this excites the magnetron to generate an R.F. pulse.

V_1 cuts off again when the pulse from the sub-modulator falls and oscillations are set up in L_2 and the distributed capacity between the magnetron filament and earth. L_2 is included deliberately to cause these oscillations in order to give the pulse applied to the magnetron filament a sharp falling edge. Fig. 136B, curve A, shows the effect of distributed capacity in prolonging the decay of voltage at the magnetron filament. Curve B shows the steeper fall caused by the action of L_2.* The negative halves of the oscillations could cause the magnetron itself to oscillate in an undesirable manner, but they are removed or damped out by the action of the diode.

Alternative Method of Developing a High-voltage Pulse of Short Duration Direct from a Sine Wave Input

When the current value is between A and B, see Fig. 137 (b), the effective inductance of the coil L_2 is high. When the current is raised beyond B in the positive direction, or becomes lower than A in the negative direction, the effective inductance of L_2 drops to a very low value. L_2 acts as a switch.

The curve (U) shows how the impedance of L_2 falls when the total current through it rises above the maximum value of the magnetising current required to produce saturation of the core † (curve S).

The series circuit C_1, L_1, L_2 is made nearly resonant at the *input* frequency. L_1 is chosen with a value such that the circuit is slightly inductive when L_2 is unsaturated and slightly capacitative when L_2 is saturated. Thus when L_2 is saturated the current leads the voltage by nearly 90°. But since the voltage across L_2 leads the current through it by nearly 90° the voltage across L_2 saturated is approximately 180° out of phase *leading* the applied voltage.

Also, because the saturation inductance component of the

* The steeper fall is very desirable since it causes the trailing edge of the R.F. pulse to become more vertical—a factor which affects accuracy of reading on the C.R.T. display.

† L_2 is in fact a choke with a core of high permeability, designed to saturate with a comparatively small magnetising current.

current I_2 of L_2 is almost zero, the voltage across the inductance is of very small amplitude.

If L_2 were replaced by a normal coil whose inductance equalled the inductance of L_2 (unsaturated), the circuit would become inductive and the current would lag the voltage by 90°. In this case the voltage across L_2 is in phase with the applied voltage.

Because the circuit is near series resonance, and because the reaction of L_2 is now large, a large amplitude sine wave appears across L_2, see Fig. 137 (b).

L_2 is, however, saturated during most of the cycle and the voltage across the inductance, in this condition, is essentially the small amplitude sine wave " S."

Whilst the current passes from " A " through zero to " B " in the opposite direction, see Fig. 137 (b), the core is unsaturated.

During the short unsaturated interval the voltage across L_2 changes from curve "S" to curve "U," producing a large pulse. When the coil again becomes saturated the voltage across L_2 falls to curve " S."

The duration of the pulse is made short by the use of the series resonance circuit, the effect of which is to cause large currents to pass through L_2, so that only a short time is required for the magnetising current to pass from saturation value in one direction to saturation value in the other.

Fig. 137 (c) shows the use of this device in a transmitter assembly when a trigger pulse is applied to start the action of the circuit

Analysis

V_1 is normally cut off by grid bias developed by C_1R_1.

When the grid is driven positive, currents I_1 and I_2 flow in L_1 and L_2.

I_2 builds up quickly to a value in excess of saturation value so that L_2 becomes a low impedance.

C_3 quickly discharges *viâ* L_2 to the voltage of the anode of V_1.

The magnitude of I_2 is mainly limited by the value of R_3.

L_1 is consequently shunted by a high impedance whether L_2 is saturated or not. Therefore most of I_aV_1 must flow through L_2.

As I_1 increases through L_1, energy is stored $= \frac{1}{2}LI_1^2$.

When V_1 is cut off by the negative swing of the input pulse, a high voltage is induced in L_1.

This is momentarily short circuited by the low impedance of L_2C_3 and the filter across the power supply, all tending to reverse the direction of flow of I_2.

I_2 decreases to saturation value and L_2 becomes a high impedance again.

Simultaneously with this action the stored energy $\frac{1}{2}LI_1^2$ is being transferred to C_4 and a very high voltage builds up across it.

In the case shown the voltage is approximately 15,000 volts, and the time for build up is about 10 microseconds, which is equal to a quarter of the period of the oscillation of the resonant circuit made up chiefly by L_1 and C_4.

The charging path of C_4 is *viâ* the diode V_2 and the filter capacity across the power supply.

Whilst C_4 is charging, I_2 decreases from saturation " A " to zero and builds up to saturation " B " in the opposite direction.

Again, the time required for this reversal to take place is 10 microseconds, approximately.

When L_2 is again saturated it again becomes a low impedance earthing the positive sides of C_4 *viâ* C_3. C_3 is about fifteen times larger than C_4, and L_2 is at that moment a low impedance. Consequently 14,000 volts approximately appear across the output. L_3 shunts this after about 3 μ secs.

Development of High-voltage Power-time Control Pulses at High Level

Before discussing the development of high-voltage power-time control pulses at high level, it is necessary to examine the properties of an artificial line when used as a discharge line. Accordingly we proceed to a discussion of this device and to the development of high-voltage power pulses. Comparison of this method with the low-level systems shown in Fig. 124 follows immediately afterwards.

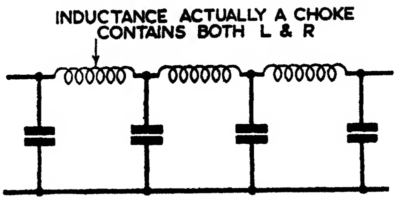


Fig. 138.—ARTIFICIAL TRANSMISSION LINE.

Artificial Line

An artificial transmission line * can be thought of as an electrical network having input and output terminals and composed of lumped elements of inductance, capacity and

* This is a brief interim description of the theory and action of an artificial line. For further details, the Appendix No. 1 should be consulted, " Theory of Transmission Lines."

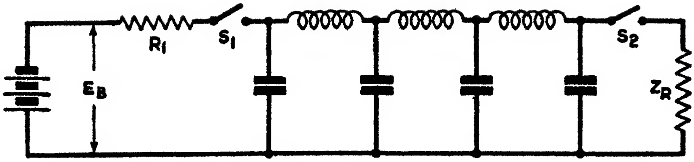


Fig. 139.—ARTIFICIAL LINE AS AN ENERGY-STORING DEVICE. S_1 AND S_2 SHOWN OPEN.

resistance, thus imparting to the device characteristics similar to those possessed by an actual transmission line. It derives its name from the fact that in order to avoid the bulk of an actual transmission line an artificial line can be built of inductance and capacity, having approximately the same total values as the distributed inductance and capacity of the line which it is desired to imitate.

In an actual transmission line the total inductance and capacity is spaced uniformly along its length, consequently the more sections that are used to build up the artificial line the nearer its performance will approximate to the actual transmission line it is intended to imitate. In practice, the conductance or leakage factor which is very small can be neglected. Three to eight sections are sometimes employed, but when the performance of an actual transmission line has to be more closely adhered to, more sections are required (Fig. 138).

The primary purpose of a transmission line is to guide energy from point to point, but it also exhibits certain characteristics some of which are made use of in radar circuits.

Functions

By suitable arrangements an artificial line can be made to perform any of the following functions :—

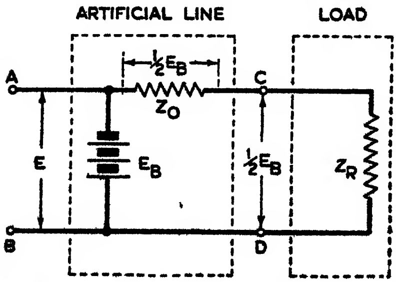


Fig. 140.—DISTRIBUTION OF VOLTAGE AT MOMENT OF DISCHARGE WHEN $z_R = z_0$.

(*a*) To store energy and subsequently deliver it at a predetermined rate to form pulses of given duration.

(*b*) To establish a given delay time between any two consecutive operations.

(*c*) To act as a low pass filter, as exemplified in power supply circuits (smoothing), after rectification has taken place.

If a voltage is applied to

one pair of terminals of a transmission line, the far end being open circuited, a surge of voltage accompanied by current travels along the line towards the open end in a manner resembling that of a tidal wave.

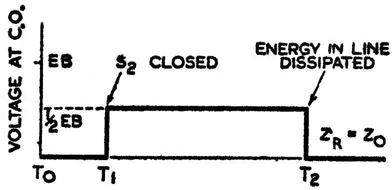


Fig. 141.—WAVEFORM OF THE DISCHARGE.

Although the surge may be travelling along the line with a velocity of approximately 186,000 miles per second, it must take a definite time to reach the open end.

On arrival at the open end a change occurs (*i.e.*, there is a discontinuity of the conducting path), and both voltage and current are reflected back to the source*—a voltage appearing across the input terminals equal to the applied voltage.

Providing no change is made to the circuit and its supply values, and assuming that there is no leakage, the line will remain charged to the voltage of the supply indefinitely.

It follows from the above that if a pulse is applied across the terminals of an artificial line open circuited at the far end, some delay occurs between the moment of application of the pulse and the moment at which it reappears at the input terminals.

This time delay can be controlled by adjusting the values of L, C and R to simulate lines of different lengths and characteristics.

When used as an Energy-storing Device

When an artificial line is used to store energy and to re-deliver it at some predetermined rate, the charging conditions are similar to those already described. The charging current may be obtained from a suitable direct current source or with certain modifications the output of a transformer may be used for this purpose.

The discharge however does not, as might be expected, follow the exponential curve of a condenser discharge.

The conditions during charge and discharge can be understood by consideration of Fig. 139.

Let switch S_2 be opened and S_1 closed. A surge of voltage accompanied by current will travel to the free end of the line and be reflected back to the source (the input terminals), where

* An action analogous to the reflection of sound waves from impact with a hard surface, *i.e.*, a change of state in the medium at the moment of impact being comparable with a discontinuity met with at the terminals of an open line.

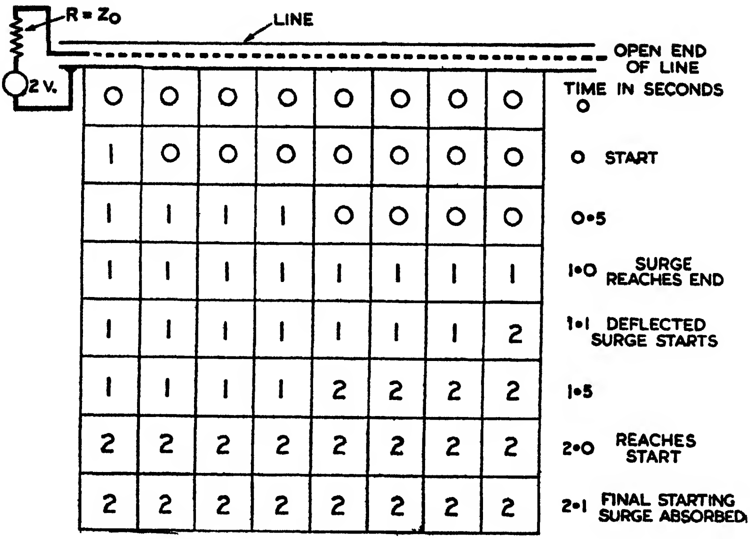


Fig. 142.—VOLTAGE DISTRIBUTION (INSTANTANEOUS VALUES) ALONG AN OPEN LINE FROM MOMENT OF CONTACT, *i.e.*, FROM COMMENCEMENT OF FORWARD SURGE UNTIL REFLECTED SURGE RETURNS TO E_B.

after a delay dependent upon the constants of the line it will re-appear across the input terminals. The voltage value of this reflected pulse will be equal to that at which the line was charged in the first place, *i.e.*, E_B.

If now S_1 is opened and S_2 is closed, the line will discharge through Z_R.

At the instant of discharge the voltage appearing across the terminals is not E_B as might be expected ; furthermore, the voltage does not decrease along some curve like an R.C. curve. The line instead acts as though the voltage across the line capacity (equal to E_B) had divided itself between Z_0, the characteristic impedance * of the line and the load Z_R.

When Z_R is made equal to Z_0 the conditions during discharge are :—

(*a*) The voltage to which the line has been charged—E_B—divides equally between Z_0 and Z_R, consequently the voltage across Z_R is $= \frac{1}{2} E_B$, see Fig. 140.

(*b*) When $Z_R = Z_0$ the voltage across Z_R is maintained at

* The characteristic impedance of a transmission line $= Z_0$ is a constant for the line depending upon its *geometry* and not upon its length. (See Appendix I.)

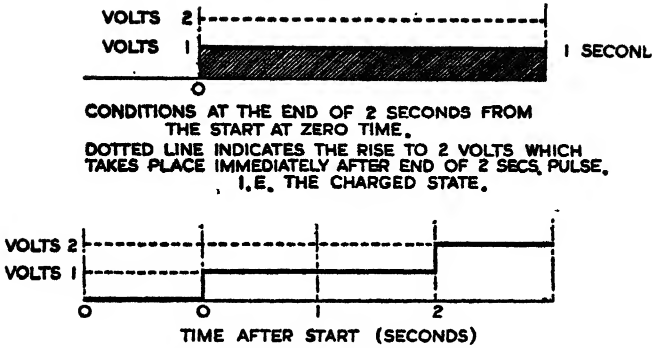


Fig. 143.—VARIATION WITH TIME OF VOLTAGE AT NEAR END OF THE OPEN LINE IN FIG. 142.

the constant value of $\frac{1}{2}$ E_B as long as there is any energy left in the line.

(*c*) The output at CD will therefore be a rectangular pulse, the width of which depends upon the time required for the line completely to discharge, or, in other words, upon the length of the line.

A mental picture of the effects which take place during the periods of charge and discharge can be obtained from a consideration of the following example :—

If a line (uncharged initially) is connected to a source of, say, 2 volts (internal resistance $= Z_0$), a surge of 1 volt ($\frac{1}{2}$ E_B) travels along the line and is reflected when it reaches the open end.

Let this wave take 1 second to travel to the end of the line. If it were possible during this period to measure the voltage

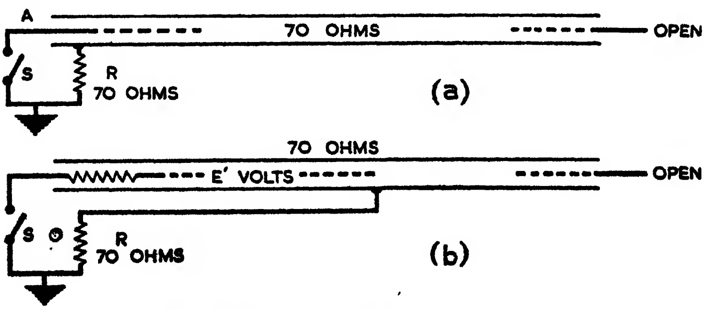


Fig. 144.—VOLTAGE DISTRIBUTION ACROSS R.
(a) Actual connection, switch open. (b) Equivalent arrangement, when S (shown open) is closed.

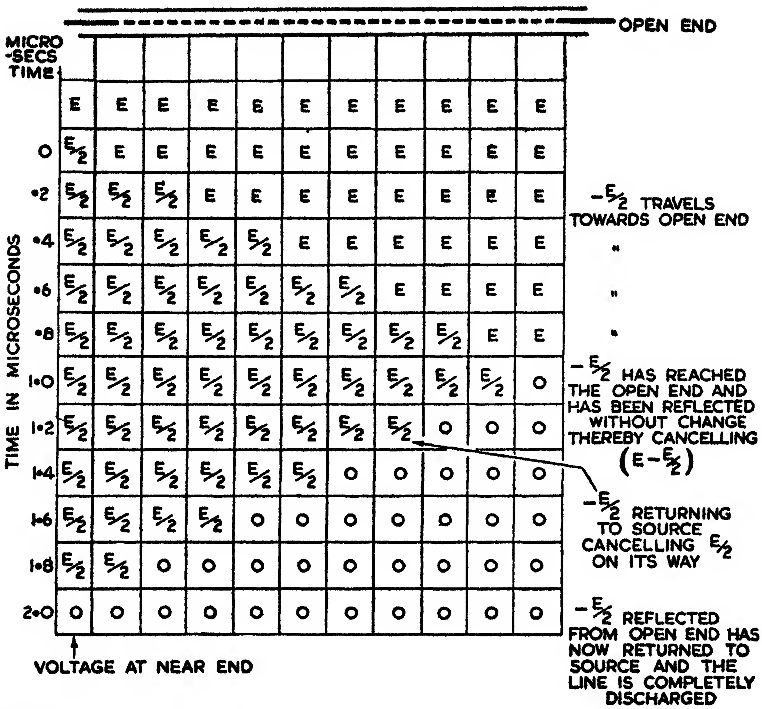


Fig. 145.—VOLTAGE ACROSS R DURING THE TIME SURGE TAKES TO TRAVEL TO OPEN END AND BACK. THIS ALSO SHOWS INSTANTANEOUS VALUES OF VOLTAGE ALONG THE LINE DURING THE SURGE.

at a number of points along the line with a meter that responded instantaneously and passed no current, the readings at various successive intervals of time would look somewhat like Fig. 142.

If the impedance across the E_B end is $= Z_0$ the surge will be completely absorbed by this terminating impedance and the line will be left permanently charged to 2 volts.

The voltage of the near end plotted against time would be as in Fig. 143.

The line has thus produced a pulse of 1 volt lasting for *two* seconds (1 second to travel to the end of the line plus 1 second for the reflected surge to travel back to the near end). Completion of this pulse is followed by a sudden rise to *two* volts—the final steady charged state when the charge of the line has become equal to E_B.

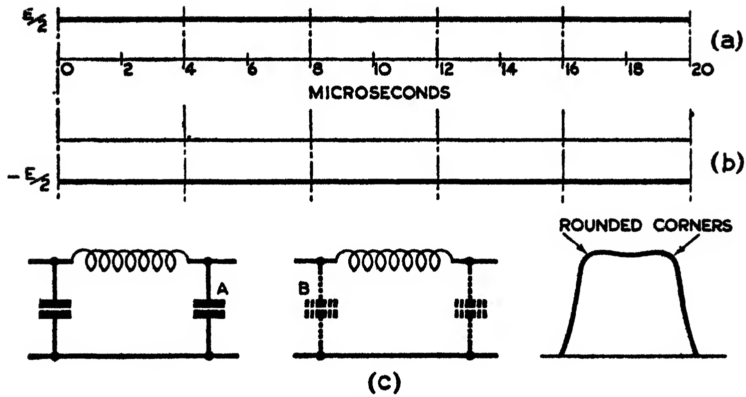


Fig. 146.—VOLTAGE DISTRIBUTION, IMPEDANCE AND OUTPUT CHARACTERISTIC OF CABLE EXAMPLES, FIGS. 144 AND 145.

(a) and (b) Conditions across R during surge.
(c) Rounding of corners of pulse.

A Practical Example

Another and more practical example is—when the time interval of the surge along a line is, say, 1 microsecond.

In this case an open switch is connected to the centre conductor of a coaxial cable whose characteristic impedance, Z_0, is, say, 70 ohms. *Initially*, in this case the cable is already *charged* to E volts (Fig. 144).

If at zero time the switch is closed the cable discharges through S and R in series. See Fig. 144 (a) and (b).

Immediately after the switch is closed the cable acts as a source of E volts in series with the characteristic impedance Z_0 as in the equivalent circuit, Fig. 144 (b).

Since the load is $= Z_0$, at time $= 0$ a surge of $-E/2$ is applied to the near end. This travels along the cable reaching the far end after 1 microsecond. Here it is reflected without change so that a reflected surge travelling back towards the source reduces the $-E/2$ charge progressively to zero as it returns. Thus the state of zero volts created by cancellation at the far end now spreads back to the start, reaching it after a period of another microsecond (Fig. 145).

The voltage across the load resistance is $E/2$, which is equal to the voltage across the cable end.

Since the initial change of the charged cable was E volts, and because the voltage at the near end changes to $E/2$ immediately after closing the discharging switch, the change produced by the closing of the switch is equivalent to applying a surge, E/2, to an uncharged cable.

Since the outer sheath of the cable is connected to earth *viâ* the switch, the voltage across R will be as shown in Figs. 145 and 146 (a), consequently there must be a voltage of —E/2 existing across R during the time taken by the surge to travel to the end of the cable and back.

To produce a pulse lasting for 1 microsecond, 195 yds. of cable would be required. (The velocity of the wave in this particular cable being 195 yds. per microsecond.) A cable of this length, capable of standing, say, 10 kV., is not convenient, hence the use of an artificial line to perform a similar function.

Since the surge is rectangular in shape, it can be regarded as composed of a number of sine waves of many different frequencies, consequently the artificial line does not pass all frequencies equally well, also the impedance of the line varies with frequency. The effect of this is that the sharp corners of the rectangular pulse are rounded off. These undesired effects can be reduced by careful choice of the values of L and C. By suitable design of the artificial line reasonably square pulses can be obtained.

In one arrangement shown in Fig. 134, one conductor 0.6 kV. is initially charged to — ·6 kV., and during the discharge the other part of the line (represented by the

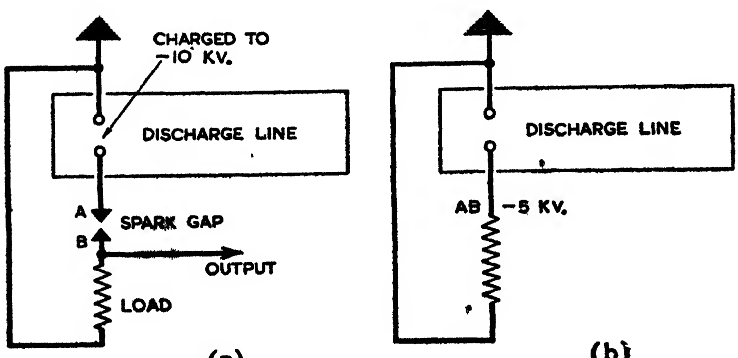


Fig. 147.—GENERAL PRINCIPLE GOVERNING ARRANGEMENT OF CIRCUIT WHEN A SPARK GAP IS EMPLOYED.

(a) Voltage distribution before spark takes place.
(b) Circuit conditions and voltage distributed during spark action.

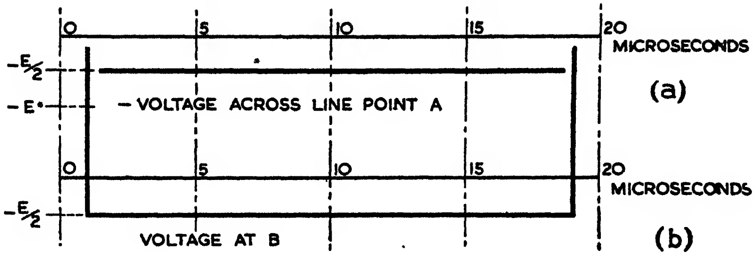


Fig. 148.—VOLTAGE DISTRIBUTION ACROSS LINE AND LOAD FOR CONDITIONS SHOWN IN FIG. 147.

(a) Voltage distribution across the discharge line.
(b) Voltage distribution across the load.

outer conductor) is ·5 kV. below earth. Both sides of the line must therefore be very well insulated from earth. This system is used when a thyratron is used as a switching or modulating device.

When a Spark Gap is Used

When used with a spark gap taking the place of the thyratron, the circuit arrangement is shown in Fig. 147.

At the start, electrode " B " is at earth potential, when the gap sparks over and behaves as a low resistance half the 10 kV. to which the line is initially charged appears across the load and the other half across the line itself. This condition is dependent upon the assumption that the load impedance is equal to Z_0.

The circuit conditions and voltage distribution are now as shown in Fig. 147 (b). Thus a surge of $+5$ kV. will travel along the line, be reflected at the far end, returning as a -5 kV. surge.

Fig. 148 (a) and (b), shows the voltage across the discharge line and across the load. These should be compared with Figs. 145 and 146.

If the load as shown in Fig. 147 represents the cable leading to the cathode of a magnetron (see section U.H.F.), a negative

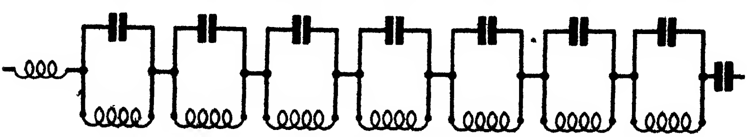


Fig. 149.—GUILLEMIN LINE.

pulse lasting for the time of travel along the line and back is applied to the magnetron cathode.

It must be realised that one side of the discharge network is always at earth potential. Due to this fact, it is unnecessary to insulate both sides of the earth network, and a much more compact arrangement of the network results, compared with that shown in Fig. 147.

The artificial line, then, can deliver a constant voltage to a load for a definite period of time and is similar in action to that of a battery when switched on and off. Its advantages lie in the accuracy with which magnitude and duration of a pulse of energy can be delivered and also in the extreme rapidity of its action.

The Guillemin Line

An alternative construction for the artificial line is shown in Fig. 149. This is known as a Guillemin line.

If an artificial line is to meet all the requirements for simulation of performance of a real line as closely as possible (this is very desirable when a good rectangular pulse is to be produced), it must have many sections, and if all the condensers are required to stand a high voltage the device would be very bulky. The Guillemin line offers an adequate means for improving the required performance and at the same time it keeps the bulk within reasonable dimensions.

In the Guillemin arrangement the series condenser is the only one that need be insulated for high voltages, since it is the only one that remains charged. The other condensers are paralleled by inductances, and hence voltage is applied across them only during the periods of charge and discharge, when they divide the applied voltage amongst their number.

Development of Negative Pulses of Medium or High Power Suitable for Pulsing a Magnetron

Fig. 150 (a) is a circuit used to derive a suitable negative-going pulse for application to the cathode of a magnetron from a D.C. source.

The artificial line may be charged from either a D.C. or A.C. source to a high voltage.* As the electrodes carried by the

* The voltage to which the condenser or artificial line is charged may be approximately twice that of the charging source. The load being capacitative, an inductance may be placed in series with it to produce a resonance effect with the D.C. pulses. When the input is A.C., the circuit is not resonant in this respect, but the frequency of the supply can be when circuit constants are suitably selected. In this condition the voltage across the capacity builds up to approximately twice that of the supply voltage as demonstrated in standard textbooks on radio. Also see chapter on transmitters.

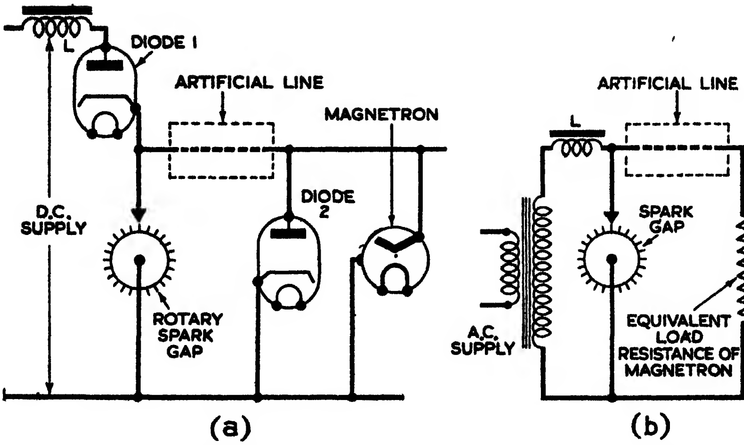


Fig. 150.—D.C. CHARGING AND EQUIVALENT CIRCUIT FOR A.C. CHARGING THROUGH INDUCTANCE.

(a) Circuit for charging artificial line from a D.C. source (high voltage) and for discharging the stored energy in the line through a magnetron, thereby applying to the cathode of the magnetron a negative pulse of high voltage of any desired duration.

(b) As above but using an A.C. transformer operating on a suitable A.C. supply. (Equivalent circuit.)

rotating disc revolve past the fixed electrode the gap breaks down and the artificial line discharges.

One half of the voltage to which the artificial line was charged is applied to the cathode of the magnetron * as a negative-going square pulse, and because of the properties of the artificial line, regulation is good over the whole period. In other words, the peak voltage does not drop very much, due to the magnetron load for the period of each pulse.

Alternative arrangements are the use of fixed spark gaps either in air or some suitable gas, or a gas-filled valve operating on the spark gap principle, such as the trigatron. Details of

* To obtain this effect, Z_R (the load) must be adjusted to Z_0, the characteristic impedance of the line. In the section dealing with transmitters it is pointed out that pulse transformers are frequently used.

In this case the artificial line feeds the pulse into a coaxial cable at a comparatively low voltage level for transmission to the primary of the pulse transformer. The coaxial cable must be matched at the input and also to the load.

For the sake of simplicity, the pulse transformer has been omitted in this section, since the component units are again considered together in detail as a unit in the chapter on transmitters.

the discharger or modulator are considered further in Chapter XVIII.

The foregoing are typical cases of pulse generators suitable for high voltage rectangular pulse output for application to R.F. generators. These will be referred to in detail later in the chapter on transmitters.

In Fig. 150 (a), diode 2 performs the function of damping out any oscillations which may take place at the end of each pulse, the negative half of which might cause prolonged or unwanted R.F. outputs from the magnetron. This is dealt with more fully in the chapter on transmitters. A damping diode is included in the A.C. arrangement but is not shown in the equivalent diagram.

Diode 1 is used in the charging circuit when the source is D.C. It is shown in the chapter on transmitters that the energy of each D.C. charge can be made to oscillate at a low frequency, with the important result that the capacity of the artificial line can be charged to twice the voltage of the supply source. Diode 1 is used to prevent any loss to the charge should the spark gap not discharge at the exact moment when the charge in the capacity has reached its peak value.

Comparison of the High- and Low-level Systems

Comparing the high- and low-level systems for developing high-voltage time control power pulses, it is obvious that the high-level system is more simple, takes up less space, and is more economical and efficient than the low-level system. It does, however, suffer from two disadvantages :—

(*a*) No opportunity is afforded for modifying or improving the shape of the rectangular wave generated by the discharge of the artificial line. The quality of the wave generated depends solely upon the "goodness" of the artificial line as a rectangular pulse former.

In the low-level system, however, opportunity is afforded to improve the pulse shape by adjustment to the output circuit of the sub-modulator or by the input characteristic of the modulator.

(*b*) The repetition rate of the transmitter depends upon the regularity of discharge of the spark gap.

The efficiency of the breakdown performance of the spark gap depends essentially upon ionisation of the gases in the spark gap itself, and since there is a factor of irregularity about this, stability of the repetition rate may be adversely affected.

The principal factors are the nature of the gas in the gap and the pressure. At high altitudes this is very noticeable.

Means, described in Chapter XVIII on transmitters, can be provided to improve the regularity of the spark gap discharge, but when extreme precision of measurement is required, the stability of the p.r.f. is better when an independent and very stable source of Synch. is employed, which necessitates the use of the low-level system for development of the high-voltage power pulse. Since, however, extreme accuracy is generally confined to short-range operation, the peak power required for efficient working is also very much smaller.

Development of Low-Power Time Control Pulses

Chapter XIII deals with circuits for the development of low-power time control pulses for use in units other than the transmitter. The term "low power" is used in a relative sense mainly as a distinguishing label or line of demarcation, since in the majority of these cases the power requirements are not of major importance.

There are, however, some borderline cases. For example, when it is essential to receive echoes from targets which are very close to the transmitter, extreme precautions must be taken to avoid any temporary paralysis in the receiver circuits due to strong signals from the transmitter, and this requirement may make it necessary to provide synchronised pulses of high voltage as H.T. for the receiving valves.

With the best possible arrangements for switching, some of the transmitter output does enter the receiver branch of the feeder system where it is subsequently amplified and may produce large residual charges on condensers and in circuit capacities.

Since the time interval between the end of the transmitter pulse and receipt of an echo from a very near target may be fractional, there is insufficient time for residual charges to leak away sufficiently to permit the receiver circuits to function immediately in the normal manner. This leads to masking or non-recording of echoes returned by nearby targets.

In order to avoid this, various devices are employed, such, for example, as increased bias applied during the transmission period and the still more satisfactory method of removing the H.T. supply entirely from the anodes and screens of certain valves in the I.F. section of the receiver.

When the last-mentioned method is employed, a considerable amount of amplification is required for the pulse which re-establishes the H.T., and since an appreciable load must be supplied, this pulse may be looked upon as a power pulse, although the voltage is not likely to exceed about 300 volts.

The circuit arrangements discussed in this chapter and in Chapter XIII are not necessarily part of any equipment actually in service at the present time. The main purpose is to convey a general understanding of what can be done with pulses and how they may be shaped, controlled and engineered to achieve a desired purpose.

CHAPTER XIII

DEVELOPMENT OF LOW-VOLTAGE LOW-POWER TIME-CONTROL RECTANGULAR PULSES

THE circuits now to be described are typical of some of those used to develop rectangular pulses directly by virtue of the electrical characteristics of the circuits themselves or by a process of shaping and amplifying.

When the pulse width becomes very narrow—say at some fraction of a microsecond, it becomes necessary to develop it from a damped oscillation and to square and amplify this by means of subsequent stages.

When accuracy is of primary importance the pulses should in general be as rectangular as possible, consequently the output waveform of the various types of generators to be described is of considerable importance.

In the case of multivibrators, of which there is a considerable variety, it has been necessary to distinguish them by the polarity and shape of the input they require and also by their general performance and outputs.

Circuits using artificial lines as pulse-forming networks and as delay lines are largely used in units other than transmitters. The peak voltages required are very much smaller, otherwise the manner in which they operate resembles closely their use in transmitter circuits.

There is no essential difference between (*a*) rectangular pulses generated to control the time of operation of some other generator and (*b*) between similar pulses generated for a delaying function.

The pulses in group (*a*) are generally designed to start some generator from the leading edge and to stop it with the falling edge. In case (*b*) the rectangular pulse is generally subjected to a differentiating operation, in which the sharp trigger pulse which would normally result from differentiating the leading edge of the rectangular pulse is deliberately suppressed. The trigger pulse resulting from the falling edge is used to start some generator, which is thus delayed in operation by a time corresponding to the width of the rectangular delay pulse.

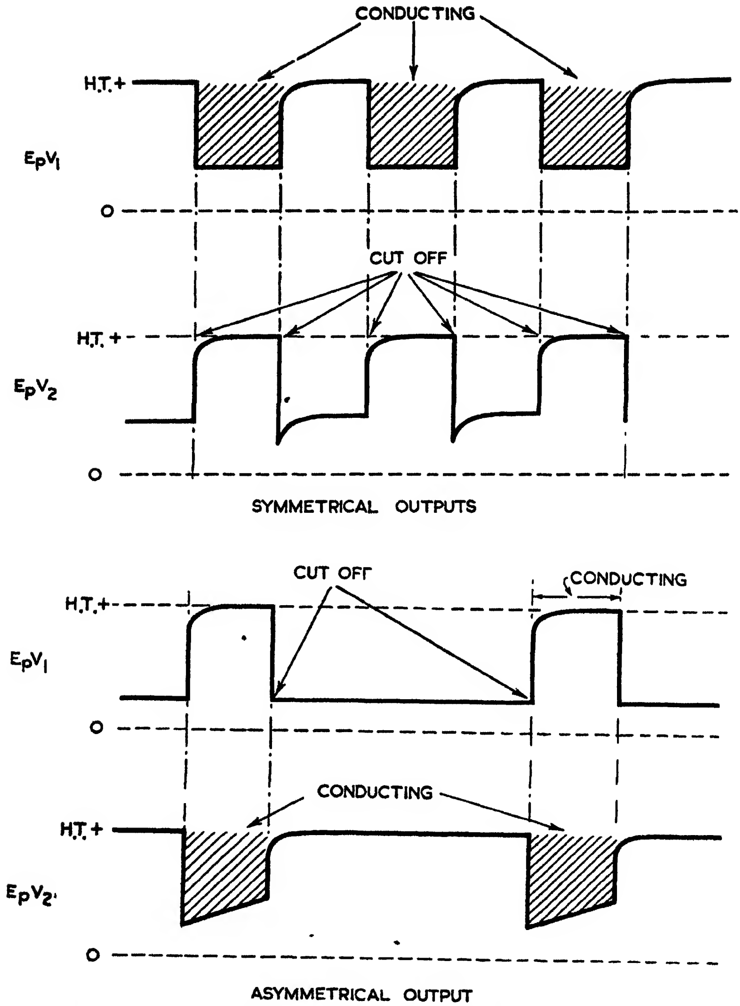


Fig. 151.—TYPICAL SYMMETRICAL AND ASYMMETRICAL OUTPUTS.
Note that when the output of V_1 is a positive-going pulse, that of V_2 is negative going.

In the case of differentiator amplifiers which are largely used, the differentiated input to the grid may be made to give different outputs by adjusting the bias and operating conditions of the valve. Limiting, clamping and D.C. restoration principles may all be used in conjunction with grid and cathode bias to eliminate the whole or any portion of either the positive or negative differential. Similarly, an input may be applied to a grid in any desired relation to the reference voltage of the grid.

Thus, from a trigger pulse input, the output from an R.C. coupled amplifier may be either an amplified replica of the input or a wave of some other form.

MULTIVIBRATORS

A multivibrator, frequently called a " flip-flop " because of its action, consists generally of a pair of valves coupled together to repeat a sequence of operations, either continuously, or at intervals determined by Synch. trigger signals.

The sequence is a sudden change of state from the quiescent condition to a working condition in which one valve becomes quiescent and the other comes into operation alternately.

The circuit may be so arranged that each valve functions alternately for the same length of time, in which case their alternate quiescent periods are also equal. A multivibrator of this type is said to be balanced or symmetrical. Alternatively, the components associated with each valve may have different values, in which case one valve will remain in the working condition longer than the other and, consequently, the quiescent period of one valve will be shorter. This type of multivibrator is said to be unbalanced or asymmetrical (see Fig. 151).*

Multivibrators may also be classified as either continuous (free running) or controlled (performing a sequence only when triggered).

Thus the continuous or free-running type repeats its sequence continuously from the time it is switched on until it is switched off again. Whereas the controlled type, as the name implies, functions under the supervision of a trigger pulse which itself may or may not be synchronised to the system, Synch. frequency or p.r.f.

In both classes of multivibrator the output from the anode of each valve is a rectangular wave. They are in antiphase, as might be expected (see Fig. 151).

* The reader will find that anode voltage has been denoted sometimes by E_p (p = plate) and sometimes by E_a (a = anode); both suffixes, however, denote the same quantity.

The continuous or free-running type of multivibrator is tolerably consistent in its performance, but when a high degree of accuracy is required it may be synchronised by a Synch. pulse so that during each sequence each valve must perform each of its functions at the allotted time.

Prior to receipt of a Synch. or trigger signal, the controlled type of multivibrator is found in the " ready-to-work " state, *i.e.*, one valve biased beyond " cut off " and the other passing a steady anode current.

On receipt of a Synch. or trigger signal, the " ready-to-work " state of equilibrium is suddenly upset by :—

(*a*) A cut-off valve being suddenly forced to conduct, or

(*b*) A conducting valve being suddenly cut off.

The bias on the grid of the valve at " cut off " rises rapidly, simultaneously the bias on the other valve drops towards " cut off." If these valves are distinguished by " A " and " B " respectively, the sequence on receipt of a Synch. or trigger signal is as follows :—

(*a*) The grid potential of A rises from " cut off." A passes maximum anode current and B cuts off altogether.

(*b*) After a time determined by the circuit constants, the grid potential of A falls to " cut off " and the grid potential of B rises. B passes maximum anode current again and A is cut off.

The conditions prevailing prior to receipt of the Synch. signal are now re-established, and equilibrium having been again restored, the multivibrator, if of the controlled type, becomes quiescent until the next Synch. or trigger signal is received.* In the free-running type the sequence goes on repeating itself until the device is switched off.

Multivibrator Characteristics

Any type of multivibrator can generally be arranged to produce a balanced or an unbalanced output, *i.e.*, a working period of long duration and a " cut off " time of short duration for one valve, and a working time of short duration with a long period of " cut off " for the other valve.

There are several types of controlled multivibrators, each having different input and output characteristics, according to the method of coupling employed.

Synchronisation can be applied to the operation of controlled

* Certain types of controlled multivibrators execute only one alternation per ½ cycle per Synch. pulse.

multivibrators for the same reasons as it can be applied to the continuous or free-running type. This is discussed in later paragraphs.

In general, choice of type depends upon the functions to be performed.

Multivibrators are generally designed to operate at some definite frequency which can be selected by a switch. The range of frequencies that can be covered by multivibrators is from 1 cycle per minute to about 100 kc/s.

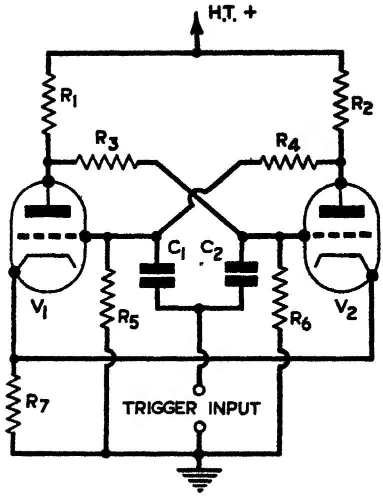


Fig. 152.—ECCLES-JORDAN TRIGGERED, TRIGGER CIRCUIT OR MULTIVIBRATOR.

Circuits embodying multivibrator principles are largely employed in radar. They may be used to perform any of the following functions:—

(*a*) For generating a synchronising signal to control an entire radar system (continuous type).

(*b*) To generate a square wave either for direct application as a source, or from which trigger pulses can be developed.

(*c*) To establish a definite delay time.

(*d*) To operate an electronic switch.

The Eccles-Jordan Circuit

The Eccles-Jordan circuit shown in Fig. 152 is a simple and particular case of the controlled type.

A positive or negative trigger pulse may be used to operate this device. One pulse must be applied to the paralleled grids *viâ* C_1 and C_2 each time the device is required to operate.

The output at each anode is a rectangular-shaped pulse of the form shown in Fig. 153, their phases being in opposition. In the example shown, the periods during which each valve is cut off are equal, so are the conducting periods. The output is therefore balanced or symmetrical in this case.

It will be noted that the anode of V_1 is directly coupled to the grid of V_2 and that the anode of V_2 is directly cross coupled back to the grid of V_1.

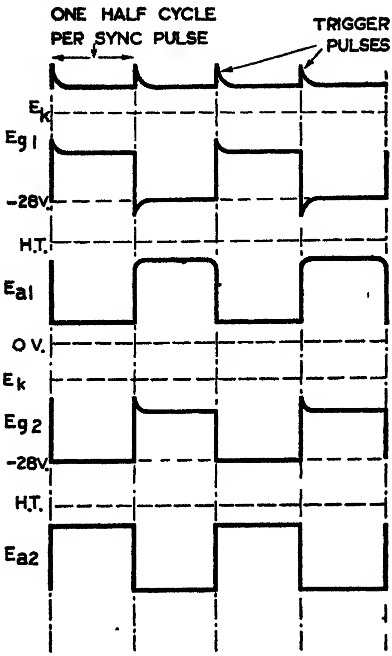


Fig. 153. — ECCLES-JORDAN WAVE-FORMS.

Analysis. 1st Part. (*a*) Assume that H.T. is applied before the filament volts are switched on.

(*b*) The parallel high resistances R_1, R_3, R_6 and R_2, R_4, R_5 will carry a small current and cause a drop, but it will be small and the voltage between anodes and cathodes will be nearly equal to the "no load" value of the applied H.T.

(*c*) If the filament volts are applied, conduction commences but small differences in manufacture of the valves themselves or their circuit components will inevitably cause one of the valves to conduct before the other.

(*d*) Anode currents flow through R_7.

(*e*) Grid bias is established for both valves.

2nd Part. Assume both valves start to :—

(*a*) Conduct about the same time but that V_1 conducts more heavily than V_2.

(*b*) E_a V_1 is lower than E_a V_2.

(*c*) E_a V_1 *viâ* R_3 reduces E_g V_2.

(*d*) I_a V_2 decreases. E_a V_2 rises.

(*e*) Increased E_a V_2 coupled to E_g V_1 causing E_g V_1 to rise.

(*f*) I_a V_1 increases, E_a V_1 falls.

(*g*) E_g V_2 further depressed.

(*h*) This cumulative feedback continues progressively until E_g V_2 is driven to cut off.

(*i*) V_1 now conducting heavily and V_2 is "cut off."

It must be understood that this series of actions takes place very quickly indeed.

The circuit will now remain in the state (*i*) until the next trigger pulse is applied, as a result of which V_2 will rise and go to maximum conductance whilst V_1 falls and is depressed to "cut off."

Eccles-Jordan Waveforms

Fig. 153 shows the waveforms when a series of positive trigger pulses are applied.

3rd Part. On arrival of positive trigger pulse :—

(*a*) V_1 conducting heavily. V_2 cut off.

(*b*) Since V_1 conducting heavily, positive pulse has little effect on this valve but will remove the cut off potential on V_2.

(*c*) V_2 driven into conduction.

(*d*) E_a V_2 drops sharply, so does E_g V_1.

(*e*) I_a V_1 falls, E_a V_1 rises and drives E_g V_2 more positive.

This cumulative action goes on again progressively until V_1 is driven to "cut off" and V_2 conducts heavily. The circuit remains in this state until the next trigger pulse arrives, and so forth.

If negative trigger pulses are used, they become effective on the conducting valve, and by reducing its current cause the rising anode volts to make the grid of the cut off valve sufficiently positive to start an action similar in effect to that described above.

The time for the valves to switch over is very small compared with the time elapsing between pulses. The vertical edges of the waveform indicate that it is almost instantaneous; thus one trigger pulse is required for the Eccles-Jordan device to complete a half cycle; two pulses must be applied to complete one cycle.

In practice the resistances are by-passed by condensers to improve the coupling between stages, this reduces the switching time by overcoming the attenuation caused by grid cathode inter-electrode capacities.

The drop across R_7 has an average value, which is reasonably constant since V_1 and V_2 conduct alternately

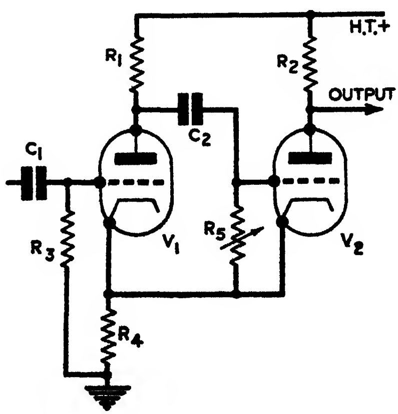


Fig. 154.—"ONE CYCLE PER TRIGGER PULSE" MULTIVIBRATOR.

and therefore the increase in one valve tends to balance the decrease in the other.

"One Cycle per Trigger Pulse" Multivibrator

The multivibrator shown in Fig. 154 differs from the previous arrangement in as much as it accomplishes one complete cycle when triggered by a positive pulse.

Note. The input must be a positive pulse for this multivibrator, because it is not common to V_1 and V_2 as in the Eccles-Jordan arrangement, also the trigger pulse should always find V_1 in the "cut off" state and V_2 conducting.

When V_1 is cut off and V_2 conducting, the circuit is in a state of equilibrium. Examination shows that V_1 is held at cut off by the anode current of V_2 flowing through R_4, consequently any change in the value of this current will render V_1, and therefore the circuit as a whole, unstable until V_1 is again permanently cut off, but V_1 can only be cut off permanently when V_2 conducts fully. V_1 cannot cut itself off by its auto bias alone.

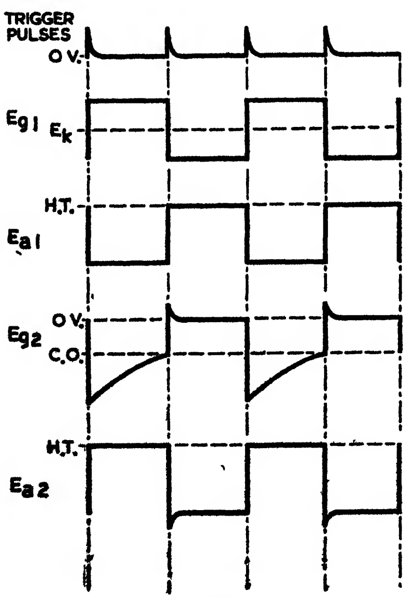


Fig. 155.—Waveforms of "one cycle per trigger pulse" multivibrator.

Analysis. (*a*) Positive pulse *viâ* C_1 raises E_g V_1 above C.O. (cut off).

(*b*) V_1 begins to conduct and E_a V_1 decreases.

(*c*) Since C_2 cannot charge instantaneously, E_g V_2 is reduced by the change of potential which must thus occur across R_5 and I_a V_2 diminishes.

Simultaneously the voltage drop across R_4 decreases and the consequent fall in potential of the cathode of V_1 is equivalent to making E_g V_1 more positive.

(*d*) I_a V_1 therefore increases, and E_a V_1 falls still more; the sequence is repeated until V_2 is cut off and V_1 conducts fully. The negative charge on C_2 leaks

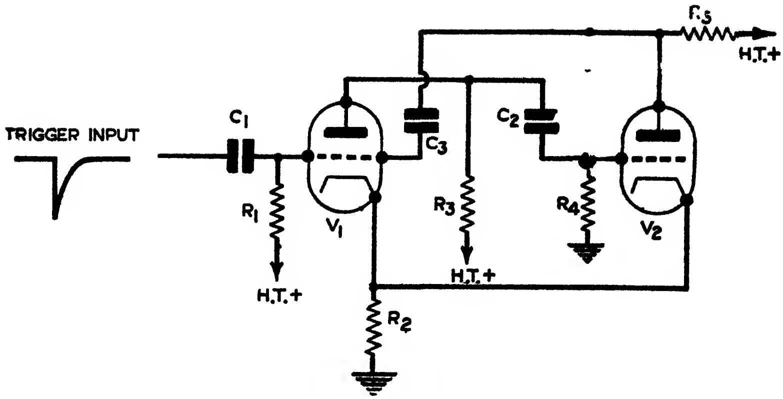


Fig. 156.—AS FIG. 154, BUT WITH POSITIVE GRID RETURN.

away via R_5, R_4 permitting V_2 to conduct again. The anode current of V_2 through R_4 raises the potential of the cathode of V_1 and $E_g V_1$ is thereby made more negative. $I_a V_1$ now falls, $E_a V_1$ rises, so does $E_g V_2$. Clearly this series is cumulative in effect and continues until V_2 conducts fully, thereby shutting down V_1.

The multivibrator circuits have now returned to their original state of equilibrium and will remain in this condition until the next positive pulse is received.

A large positive pulse will therefore be developed at the output of V_2 for each positive pulse input to the grid of V_1. The length of this positive output pulse depends upon the time constant $C_2 R_5$, which controls the length of time that C_2 holds $E_g V_2$ below cut off. (Waveforms are shown in Fig. 155.)

Input. Positive trigger pulse.

Output. Large rectangular positive square pulse from anode of V_2.

One cycle per trigger pulse.

Delay time between leading edge and trailing edge can be varied by adjusting R_5.

Arrangement with Positive Grid Return

Fig. 156 shows a multivibrator circuit which is a modification of Fig. 154.

Analysis. The main differences are :—

(*a*) R_1 is taken to the positive side of voltage supply instead of to earth (see R_3 in Fig. 154).

(*b*) The anode of V_2 is coupled back to the grid of V_1 by a condenser (C_3).

(*c*) R_4 is earthed instead of going to the cathode of V_2 (see R_5, Fig. 154).

Change (*a*) alters the conditions for equilibrium ; V_1 now conducts slightly and V_2 is cut off by the drop I_a V_1 across R_2.

Change (*b*) is necessary in order to couple back changes in E_a V_2 to E_g V_1 and to hold down V_1 to cut off for the discharge period of C_3.

Sequence of Operations

The sequence of operations is :—

(*a*) *Negative* trigger to E_g V_1.

(*b*) Decrease I_a V_1. Increase E_a V_1. Increase E_g V_2.

(*c*) V_2 begins to conduct. I_a V_2 increases. E_a V_2 falls. E_g V_1 depressed still further.

This sequence is repeated until V_1 is cut off and V_2 conducting.

The circuit remains in this condition with the negative charge on C_3 leaking off *viâ* R_1 sufficiently to allow E_g V_1 to rise to cut off and commence conducting again.

(*d*) I_a V_1 increases. E_a V_1 decreases. E_g V_2 depressed.

(*e*) I_a V_2 decreases, E_a V_2 rises. E_g V_1 rises, causing V_1 to increase its conductance progressively.

During the time V_1 is cut off C_2 is charged to the positive voltage applied to R_3 *viâ* R_2—E_{gk} of V_2 and consequently, when V_1 conducts, the anode volts of V_1 drop so sharply that the large negative pulse applied to the grid of V_2 (due to the drop in anode volts of V_1 in series with the voltage developed across C_2 by its charge), cuts off V_2. V_2 is then held in this condition by the drop across R_2 due to the anode current of V_1, until arrival of the next trigger pulse. (C.O. V_2 accelerated.)

The differences between the arrangement shown in Fig. 154 and Fig. 156 have already been discussed, the difference in effect is, primarily, consistency in pulse width or time interval between the leading edge and trailing edge of the output wave. This depends upon the time taken for C_3 to discharge to a value sufficiently low to allow V_1 to conduct again after it has been cut off.

Variations in this time produce varying pulse widths (in time), and therefore inconsistent and inaccurate results.

Small variations of the supply voltage or other changes taking place during the time C_3 discharges may cause inconsistencies in the discharge times ; these changes are, however,

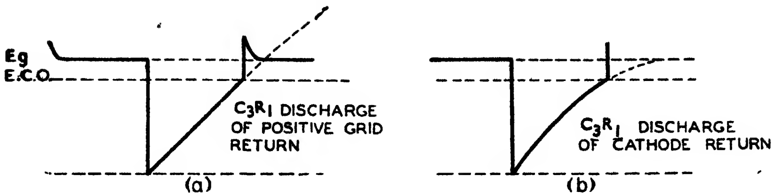


Fig. 157.—COMPARISON OF GRID WAVEFORMS.
For accuracy, the waveform in (a) is preferable to that in (b).

likely to have less effect when the rate of change of E_c during the discharge is made as large as possible.

For any given discharge time this *rate of change* will be a maximum when the discharge curve is linear, wherefore the object of causing C_3 to discharge towards positive H.T. becomes clear. The discharge of the condenser is *accelerated* along the early part of its exponential discharge curve, which is now tolerably linear. This effect is shown in Fig. 157 (a) and (b).

For accuracy the waveform in Fig. 157 (a) is preferable to that in Fig. 157 (b).

Characteristics. Input. Negative trigger.

Output. Rectangular wave. One cycle per trigger pulse. Accurate interval time and pulse width for any fixed value between trailing and leading edges of the output waveform.

Cathode-Coupled Multivibrator using Direct Coupling

Fig. 154 is a diagram of a cathode-coupled multivibrator using direct coupling, and Fig. 158 is the same circuit modified for condenser feedback.

A comparison between Fig. 154 and Fig. 158 shows that the only difference is that in Fig. 154 the grid leak of V_2 is connected to the cathode of that valve, whereas in Fig. 158 the grid leak of V_2 is taken direct to earth. Although the circuit change is small, the effect is to change the multivibrator of Fig. 154 from a controlled type to a continuous or free-running type (Fig. 158).

Similarly the small change shown in Fig. 160, *i.e.*, the addition of condenser C_1, is small in itself, but the effect is to change the output of the multivibrator of Fig. 158 from a balanced or symmetrical output to an unbalanced or asymmetrical output.

The multivibrator shown in Fig. 154 has become continuous or free running in Fig. 158 because it is not possible for the anode current of V_2 to maintain a large enough bias by drop across R_k, completely to hold V_1 at cut off.

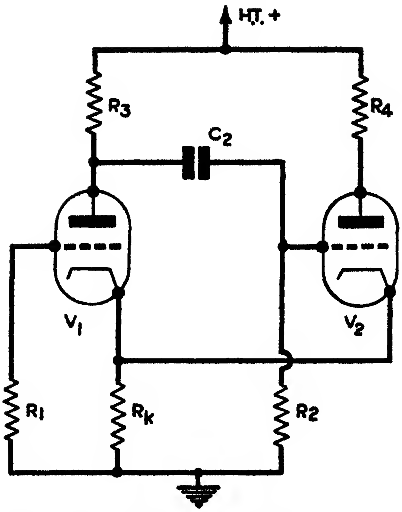


Fig. 158.—CATHODE-COUPLED MULTIVIBRATOR USING DIRECT COUPLING (FIG. 154 MODIFIED).

Analysis. In Fig. 154. Since the grid is directly connected to cathode the bias on $V_2 = 0$ and a large anode current can flow. The drop across R_k is therefore large, due to I_a V_2.

In Fig. 158. The grid of V_2 is returned to earth, so that when C_2 is fully charged E_g $V_2 = 0$, but since the cathode of V_2 is positive by an amount equal to the drop produced by the anode current of V_2 flowing through R_k, E_g V_2 is therefore relatively negative to the cathode of V_2 by this amount. Thus the bias on V_2 is equal to this voltage drop. Since this bias is produced by the anode current, V_2 cannot be cut off by this bias voltage which it produces. If V_1 and V_2 are the same type of valve, it follows that V_2 (I_a R_k) cannot hold V_1 cut off, consequently the multivibrator must function continuously from the time it is switched on until it is switched off again, and no trigger pulses are required.

Sequence. Before H.T. is applied, there will be no charge on C_2 and the grids of V_1 V_2 will be at earth potential. When H.T. is switched on anode current flows in both valves and a bias is jointly developed across R_k.

(*a*) E_a V_1 falls due to drop across R_3.
(*b*) E_g V_2 depressed. I_a V_2 decreased. E_{R_k} decreased.
(*c*) I_a V_1 increases and E_a V_1 drops further.
(*d*) I_a V_2 reduced still more.

This action is cumulative until I_a $V_2 = 0$ and I_a $V_1 = \text{max}$.

V_2 is now held at C.O. while C_2 discharges *viâ* R_2, V_1 and R_k. The discharge current of C_2 *viâ* R_k makes E_g V_2 negative relative to earth, but this negative potential on the grid of V_2 decreases exponentially as C_2 discharges.

V_2 begins to conduct again and I_a V_2 flowing through R_k makes the bias on the grid of V_1 more negative, decreasing I_a V_1 and increasing E_a V_1.

At time B see Fig. 159.

$E_g V_2$ is driven more positive, increasing $I_a V_2$ and making the bias on the grid of V_1 still more negative. This sequence is repeated until $I_a V_1 = 0$, $I_a V_2 = \max$.

Since $E_g V_2$ is positive, grid current is drawn and C_2 charges quickly *viâ* R_k, R_{gk} of V_2 and R_3 (time BC), Fig. 159.

At C, the charge on C_2 is such that $E_g V_2$ is reduced to cathode potential leaving $E_g V_2$ still positive to earth, consequently C_2 continues to charge but much more slowly, because E_g is now no longer positive to the cathode and therefore draws no grid current. The continued charging path of C_2 is therefore shifted from the comparatively low resistance path *viâ* R_{gk} of V_2 to R_2 and R_3 (high resistance). The time constant is therefore increased and the charging rate is slowed down.

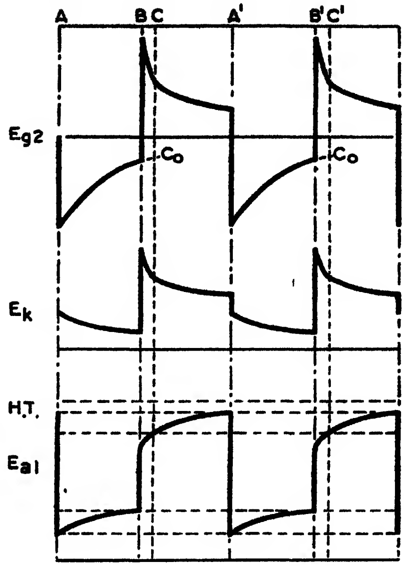


Fig. 159.—WAVEFORMS IN CATHODE-COUPLED MULTIVIBRATOR USING DIRECT COUPLING.

The continued charging of C_2 *viâ* the new path causes the bias on V_2 to become more negative and this makes $I_a V_2$ decrease, in consequence of which E_k decreases.

The grid of V_1 is held at earth potential and V_1 is held at cut off only just as long as E_k is positive relative to earth by more than the cut off voltage. When E_k drops to the cut off voltage, due to decrease of $I_a V_2$, V_1 conducts and cuts off V_2.

Analysis of the foregoing sequence shows that $I_a V_2$ is unable to hold V_1 at cut off permanently. This multivibrator must therefore pass through one sequence after another as long as it remains switched into circuit and no trigger pulses are therefore required to start or maintain it in operation.

Cathodes, Capacity Coupled

In Fig. 160 the cathodes are capacity coupled. The cut off time for V_2 is controlled by the discharge of C_2 through R_2,

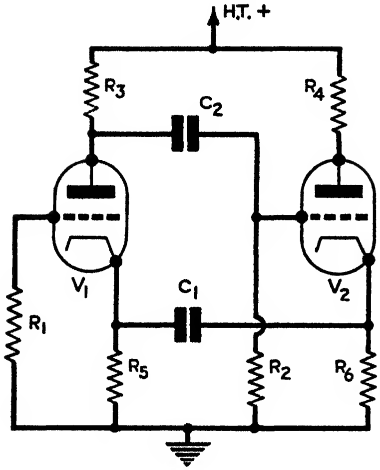


Fig. 160.—CATHODE-COUPLED MULTIVIBRATOR WITH FEEDBACK THROUGH A CONDENSER.

but the cut off time for V_1 is controlled by C_1 discharging through R_6.

Analysis of Sequence. Assume V_1 conducting and V_2 cut off. C_2 discharges *viâ* R_2, V_1 and R_5. As the discharge current through R_2 decreases V_2 rises. · I_a V_2 rises and flows *viâ* R_6. The drop across R_6 is coupled *viâ* C_1 to the cathode on V_1. The drop adds to the existing drop across R_5, I_a V_1 decreases, E_a V_1 increases, E_g V_2 becomes more positive. I_a V_1 progressively reduces to 0, and I_a V_2 goes to max. C_1 now charges to the voltage across R_6, *viâ* R_5 V_2, and R_4 and the H.T. supply. The charging current holds V_1 beyond cut off, but as C_1 charges the current through R_5 decreases until the drop which it causes across R_5 is too small to hold V_1 at cut off.

Simultaneously I_a commences to flow in V_1. E_a V_1 decreases and is coupled *viâ* C_2 to E_g V_2 causing I_a V_2 to decrease and the voltage across R_6 to drop. Simultaneously the voltage across R_5 is increased by I_a V_1.

C_1 now discharges in an endeavour to adjust its charge to the voltage across R_5.

The combined effect of the discharge of C_1 *viâ* R_6 and the negative voltage applied to the grid of V_2 causes this valve to cut off almost immediately.

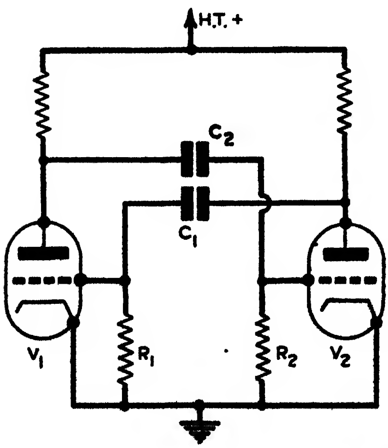


Fig. 161.—CONTINUOUS OR FREE-RUNNING MULTIVIBRATOR CIRCUIT.

The Continuous Multivibrator

The continuous or free-running multivibrator shown in Fig. 161 is widely used in radar systems. When switched on, condensers C_1 and C_2

charge as anode current commences to flow in both valves. Some small difference in the values of these currents is bound to exist and this difference provokes a further unbalancing action which is cumulative, thus the unbalanced state increases progressively until the anode current in one valve is zero and the other maximum. They remain in this state of equilibrium until that negatively charged condenser, which holds its valve at cut off, is discharged sufficiently to raise the grid of the cut off valve to a conducting potential once more.

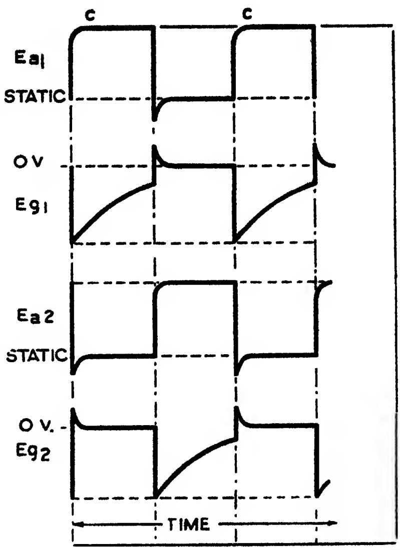


Fig. 162.—TYPICAL MULTIVIBRATOR WAVEFORMS.

Equilibrium is again upset and the sequence consequent upon unbalance is repeated but this time in a reverse direction. These conditions are typical of the action of all continuous or free-running multivibrators, and may be summarised as follows :—

At switching on, unbalance leads to the equilibrium state "A" (say V_1 cut off and V_2 conducting). This is maintained for time "T" depending on the rate of discharge of a condenser to some given value. At this critical point unbalance again occurs because V_1 commences to conduct again, V_2 is driven to cut off, and V_1 rises to maximum.

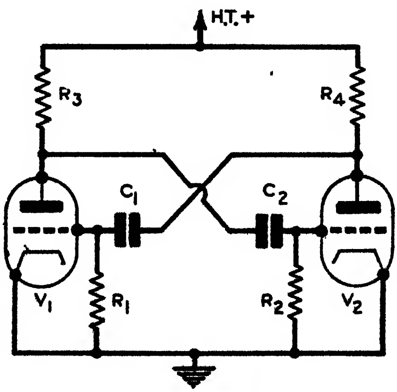


Fig. 163.—BASIC MULTIVIBRATOR.

Assume V_2 has just been cut off and V_1 is conducting heavily, see Figs. 163 and 164. C_2 is negatively charged

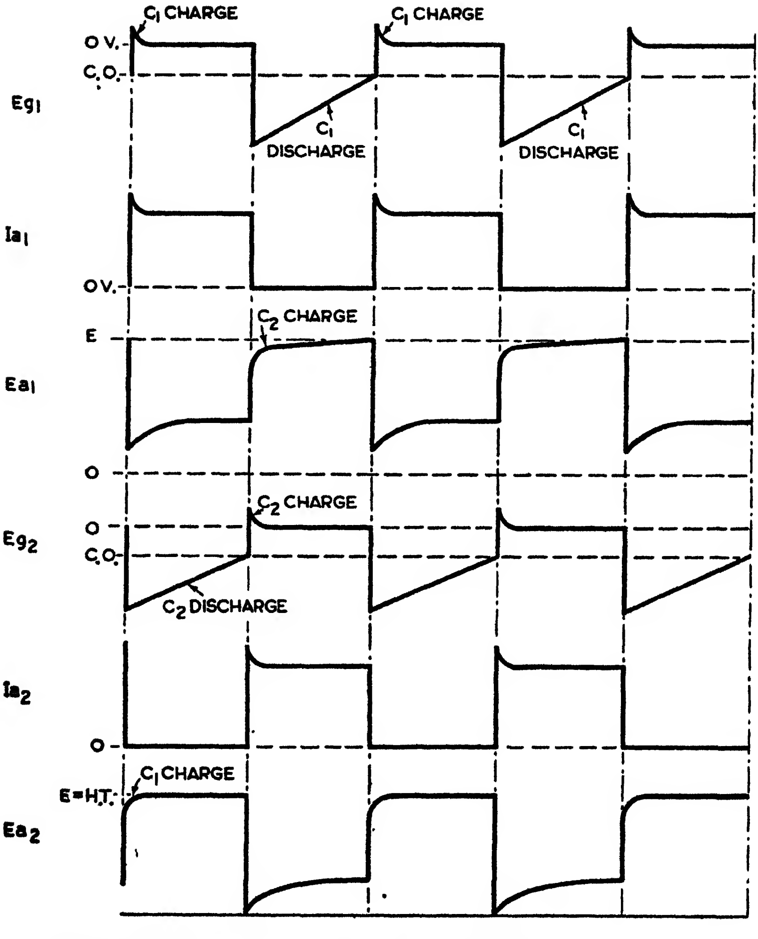


Fig. 164.—WAVEFORMS IN A BALANCED CONTINUOUS MULTIVIBRATOR.

and is holding V_2 down to cut off. The conditions are as follows :—

V_1 (*a*) I_a V_1 is maximum. E_a V_1 is minimum. E_g V_1 is positive.

C_1 (*b*) Charging *viâ* R_4, and the cathode-grid circuit of V_1 thereby producing a drop across R_4 so that when V_2 is cut off E_a V_2 equals $E - IR_4$, also E_g V_1 falls towards earth potential due to the drop across R_4.

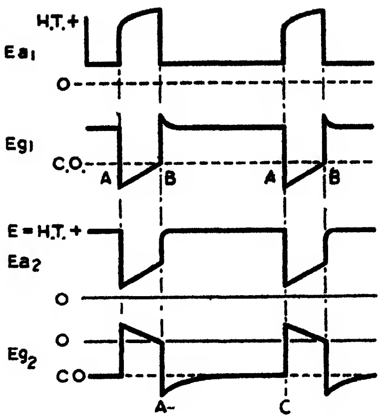


Fig. 165.—UNBALANCED MULTI-VIBRATOR WAVEFORMS.

(*c*) When C_1 is charged, E_a V_2, which was lower than E as a result of the action described in (*b*), rises to E at the same time as E_g V_1 reaches earth potential.

(*d*) Simultaneously with action (*c*), C_2 discharges *viâ* R_2 and the cathode-anode path of V_1. The time constant of this path determines the conducting time of V_1 and the cut off time of V_2. E_g V_2 is rising from extreme negative bias towards cut off during this period.

V_2 (*e*) This is the point at which equilibrium is disturbed. Since E_g V_1 has fallen due to the drop caused by the charging current of C_1, I_a V_1 decreases, E_a V_1 rises, also due to the simultaneous action taking place in (*d*), E_g V_2 rises.

(*f*) V_2 now goes progressively to maximum conduction and V_1 is cut off. Although the processes by which these reversals take place appear to be involved, the exchange of functions takes place almost instantaneously.

In some radar applications, pulses of short duration with a relatively long time between them are required. To obtain this effect the time constant of the grid circuit of one valve must be made longer than that of the other, so that one valve remains cut off for only a small portion of the cycle.

Fig. 165 shows the output when the above conditions obtain. The time constant R_1C_1 is short compared with R_2C_2 and V_1 remains cut off for only a small fraction of the total cycle. The time AB is much less than the time AC.

Electron-Coupled Multivibrator

Valves, other than triodes, can be connected as multivibrators. Pentodes may be used for example. In this case the screen

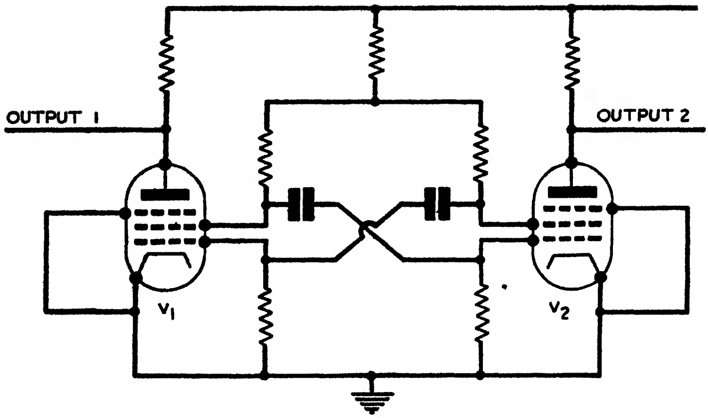


Fig. 166.—Circuit of electron-coupled multivibrator.

grids are cross connected just as though they were the anodes of triodes. The load is taken from the anode which shares the space current with the screen grids. That portion of the space current reaching the anode is modulated in passing through the cathode screen grid space to which it is electronically coupled (Fig. 166).

In such a circuit, output load reactions are isolated from the input and timing circuits, and the frequency of the multivibrator section is reasonably independent of the load.

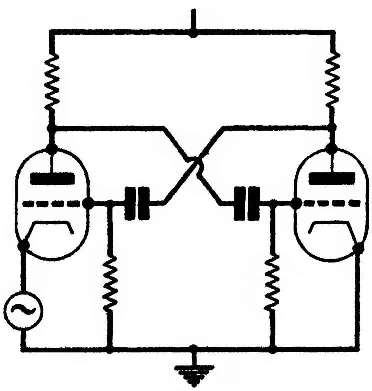


Fig. 167A.—Multivibrator with sine wave of synch. voltage applied to cathode of V_1.

Synchronising a Multivibrator

The frequency stability of continuous or free-running multivibrators is not very good, consequently when good frequency stability is required, the frequency can be controlled by a synchronising signal. This synchronising signal may take the form of a wave or pulse of almost any shape, and it may be injected into the control grid or to the cathode of one of the valves forming the multivibrator.

In practice, a sine wave or

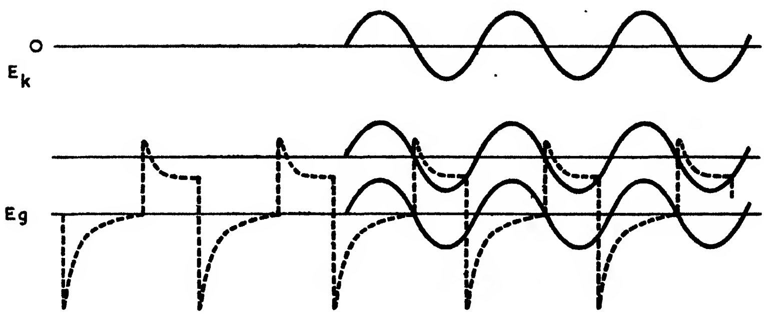


Fig. 167B.—WAVEFORMS FOR MULTIVIBRATOR SYNCHRONISED BY SINE WAVE ON CATHODE.

pulse is generally used for synchronising purposes. The frequency of the multivibrator is then locked to the frequency of the synchronising signal. Sine waves are preferable to pulses, but positive or negative pulses may be used for this purpose.

Multivibrators driven at synchronous frequency may have a natural frequency higher or lower than the synchronising signal. In these circumstances the difference between the natural frequency of the multivibrator and the synchronising signal must not be too great, otherwise the multivibrator will lock on to a multiple of the synchronising frequency. Frequency division can be obtained in this way.

When a sine wave is injected into the cathode of, say, V_1 for synchronising purposes, the internal impedance of the sine wave source should be low in order to avoid excessive drop on load and consequent distortion of the synchronising wave.

The actual grid to cathode voltage is: E_g to earth—E_k to earth (Fig. 167B). Thus the voltage on the grid is not affected by the synchronising voltage injected into the cathode, but the grid to cathode potential now contains this sinusoidal component, since the potential of the cathode relative to earth must vary sinusoidally. In this manner the cut off voltage also varies sinusoidally about the normal value, but in phase with the synchronising voltage on the cathode. This means that the cut off voltage must occur at the same point on each cycle, in which case the frequency of the multivibrator is effectively locked to the synchronising frequency. This locking on takes place after the first few initial cycles have been performed during which the multivibrator is pulled into step by the action of the synchronising sine wave (see Fig. 167B).

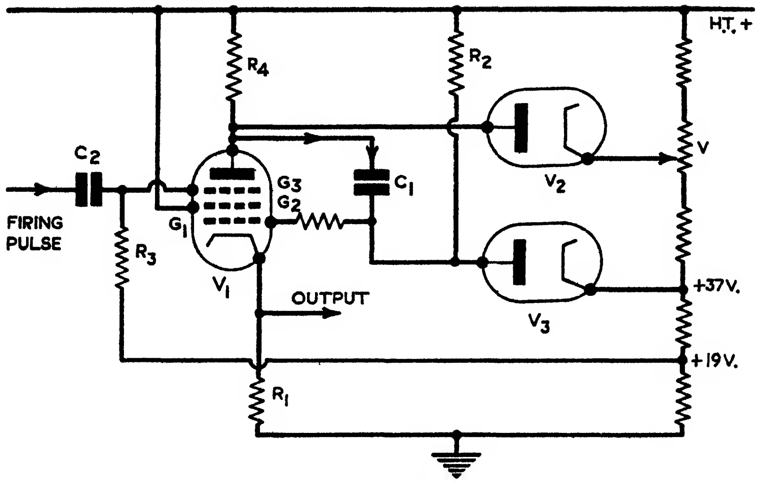


Fig. 168.—PHANTASTRON CIRCUIT.

When the synchronising voltage is applied to the grid, it adds directly to the grid voltage and, consequently, alters the grid voltage waveform.

The modified waveform indicates that the grid voltage is changing at a different rate and therefore the cut off point is established to suit the synchronising frequency and maintained in a constant relative position by it. If the synchronising frequency is higher than the natural frequency of the multivibrator, the grid voltage waveform is distorted in such a manner that it reaches cut off sooner.

Alternatively, short trigger pulses may be used for synchronising, although they are not as satisfactory as sine waves for this purpose. The Synch. pulses may have either positive or negative polarity. If a positive pulse is injected into a conducting valve it has practically no effect since it only causes a momentary increase in anode current. If, on the other hand, it is applied to a cut off valve it may raise it to conduction.

For satisfactory synchronising with pulses, the period of the multivibrator must be greater than the interval between pulses. Each synchronous pulse applied then triggers the cut off valve into conductivity at equally spaced intervals of time.

Several cycles may elapse before the multivibrator becomes locked to the synchronising pulses. These pulses are superimposed on the grid voltage, which they use as a pedestal.

A multivibrator may synchronise to a sub-multiple of the trigger pulse frequency.

PHANTASTRON

This circuit device may perform either of two functions :—

(*a*) To enable a short sharp pulse to be developed at an accurately timed interval after receipt of a Synch. pulse (delay function).

(*b*) Production of pulses with a repetition rate $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, etc., of that of Synch. pulses. This is called "counting down," and is not dealt with here.

The voltage divider in Fig. 168 clamps the potentials of the grid and anode of the pentode.

A positive Synch. pulse is applied to the suppressor.

Stage 1. The stable state before the Synch. pulse arrives is as follows :—

Volts above earth of grid G_1 set by clamping divider to + 37.

Cathode a little more positive than G_1. (Since the valve is acting as a cathode follower), say 40 +.

Screen at H.T.

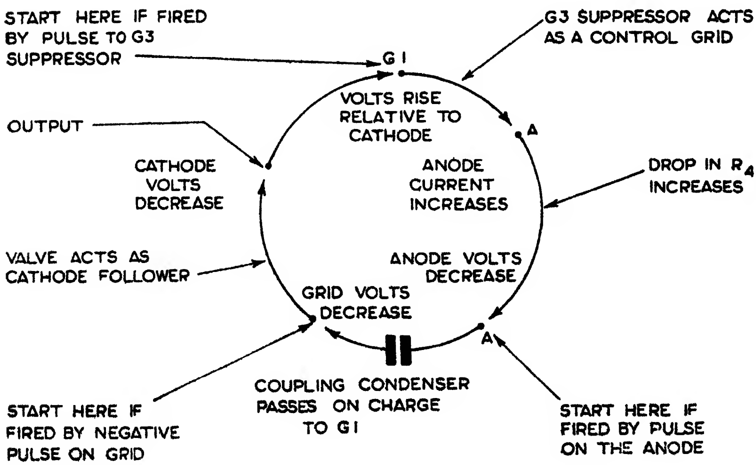


Fig. 169.—PHANTASTRON.

Chart showing cause and effect. Spots represent events and curved lines the links. The sequence is determined by choice of firing point. Three alternative firing points are shown.

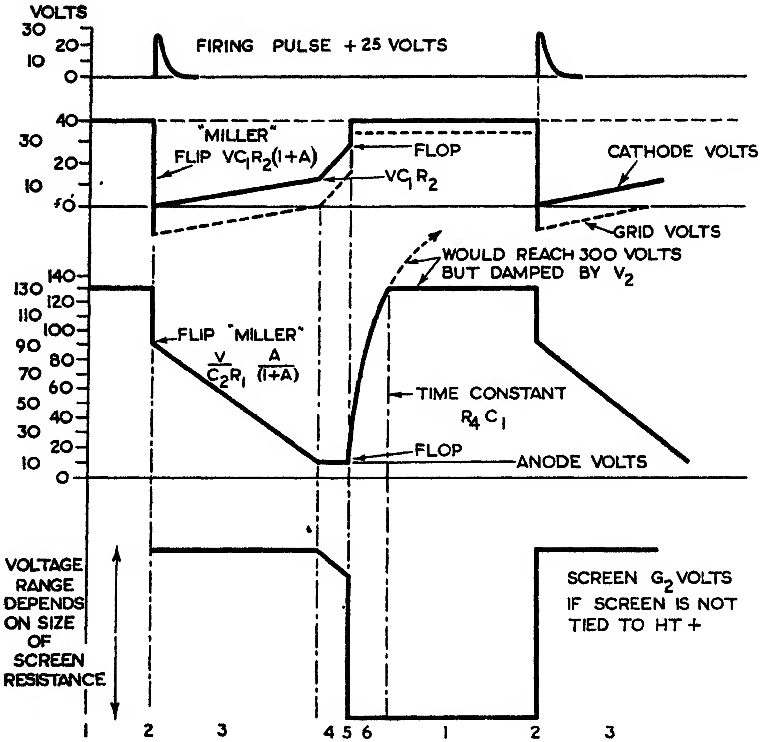


Fig. 170.—GRID CATHODE, ANODE AND SCREEN WAVEFORMS OF PHANTASTRON.

Suppressor held at 19 + therefore 21 volts more positive than the cathode.

Anode clamped at some voltage V by the action of V_2.

In this condition the anode current is practically cut off and since the suppressor is here 21 volts below cathode, no current flows.

Screen current does flow, and it is this current through R_1 which makes the cathode 40 volts above earth.

Stage 2. A short Synch. pulse, 25 volts positive going, is applied to G_3 *viâ* C_2. When 10 volts below the cathode anode current begins to flow.

E_a drops and the following actions occur :—

E_a drop drives the grid negative (*viâ* C_1).

E_k *follows* the grid and goes negative.

The suppressor is therefore made more positive resulting in an increase in I_a and consequent further drop of E_a.

The result of this action is :—

The valve soon gets into an unstable condition because the effect of the change of suppressor voltage on the anode current is greater than the change in anode current which (indirectly) causes it.

The valve now flips over to a new stable state in which the chain of events stated above no longer fulfils the condition for instability, due to the fact that as the grid goes progressively more negative, the efficiency of the cathode follower action decreases (see Fig. 169).

The spots represent the events and the curved lines the links which connect each event to its neighbour.

Where the effect is " thermionic " (could be absent if the valve were turned off) the lines are – – – –. Where the effect depends on the circuit, the lines are ————.

The arrows show the direction " Cause to effect."

Three possible firing points are shown.

Stage 3. The reverse action is started by the grid going positive and the events in the chain are then the opposite of those given in the diagram (for rise read fall, etc.).

The amount of drop of E_k is nearly independent of anode volts.

It will also be noticed that the grid has dropped nearly the same amount (since the valve is acting as a cathode follower).

The grid drop is equal to the anode drop (because of C_1), so that at the end of the flip the anode has dropped a definite and constant amount below the starting voltage V. Let the value be called V—K where K is a constant.

The valve now works as a Miller time base and the grid voltage rises at a rate $+ HT/C_1R_2\ (a + 1)$. The anode voltage falling at a rate $(HT +) \times A/C\,R_2\,(a + 1)$ where A is the stage gain of the valve, or nearly $(HT +)/C_1R_2$. The rate of rise is nearly constant and continues until the anode voltage is down to a constant level of about 10 volts and the gain of the stage has dropped to zero.

Note. The rate of rise changes a little near the end of this period because the anode voltage becomes so much lower than the screen voltage that conditions change slightly. This effect changes the time of this period by nearly a constant amount and is therefore of little importance if *difference in delay time* only is being studied.

Stage 4. The Miller effect having disappeared, the grid volts rise at the rate HT/C_1R_2 (*i.e.*, much more quickly) until :—

An unstable condition is again reached, the chain of events being as before but in the reverse direction and the flop occurs.

After the flop the anode is again cut off and it rises slowly to the clamped voltage V.

The condenser C_1 and the resistance R_4 determining the time taken to rise.

The voltages are now back to the value at the start. When the next pulse arrives the whole sequence is repeated.

It was stated that the anode voltage at the end of the flip is V—K, the voltage at the end of the constant rate part of the Miller time base is 10, so that the time taken to drop at a constant rate is proportional to V—K—10.

The rise of grid volts at the rate (HT/C_1R_2) to the point at which the flip occurs is constant, so that the delay time (Synch. pulse to flip) is made up of :—

Flip (nearly instantaneous) time for anode voltage to drop V—K—10 at rate HT/C_1R_2 volts per second, *i.e.*, $\frac{C_1R_2}{HT+}(V - K + 10)$ and constant time for grid to rise to start of flop (say) X secs.

The total is $\frac{C_1R_2}{HT+}(V_1 - K + 10) + X$.

If another voltage V_1 is chosen the delay is :—

$$\frac{C_1R_2}{HT+}(V_1 - K + 10) + X.$$

The difference is $\frac{C_1R_2}{HT+}(V - V_1)$, which is proportional to the change of V

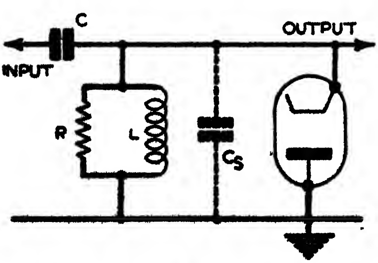


Fig. 171.—Circuit for developing a narrow pulse of positive polarity.

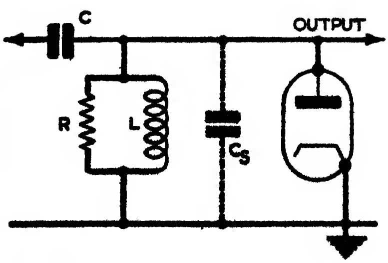


Fig. 172.—As in fig. 171 but for a pulse output of negative polarity.

If the output taken from the cathode is differentiated, a pulse is produced whose delay (after the Synch. pulse) can be accurately set by control of V.

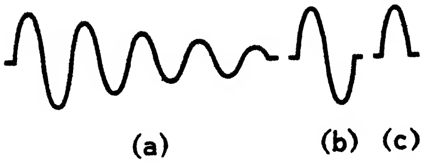


Fig. 173.—RESULTANT WAVEFORMS IN SHORT PULSE OSCILLATORS.

(a) Wave train output from circuit Fig 171 without diode.
(b) Damping of wave train by parallel diode.
(c) Rectification of output by diode.

SHORT PULSE OSCILLATORS

The process of differentiation produces a short pulse which has a sudden rise and an exponential fall; or a sudden rise and nearly linear fall, according to the arrangement of the circuit.

When short pulses with a sudden rise and a sudden fall are required they can be produced from a sudden change in voltage or shock in the following manner:—

When the voltage change is positive going the circuit in Fig. 171 is used.

The input voltage is applied through a condenser C to a parallel combination of resistance and inductance.

A diode is connected across R and L with the cathode connected to the condenser and anode to earth. The stray capacities, including the distributed capacity of the coil, are represented by C_s.

A positive-going voltage change applied through C "kicks" the oscillatory circuit LC_s into oscillation. If R is not connected and the diode out of action, the resulting waveform at the output is as shown in Fig. 173 (a). When R is connected the waveform is as Fig. 173 (b) because it damps out the oscillations more quickly.

When the diode is also in action, the waveform is as Fig. 173 (c) because the diode conducts during the negative part of the cycle and damps out the negative part of the wave still more.

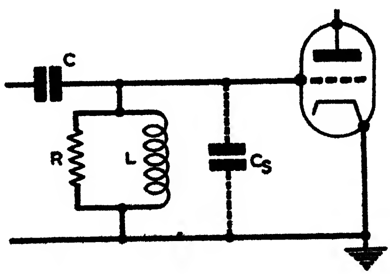


Fig. 174.—CIRCUIT WITH PURPOSE SIMILAR TO FIG. 172, BUT USING A TRIODE WHICH FUNCTIONS IN PLACE OF THE DIODE OF FIG. 172, AND PROVIDES AMPLIFICATION AS AN ADDITIONAL FEATURE.

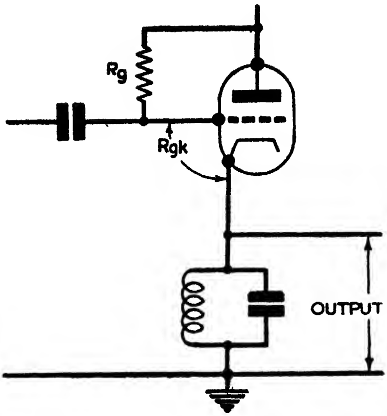


Fig. 175.—RINGING CIRCUIT.

The steepness of rise depends upon the steepness of the voltage change. The sharpness of fall and duration of pulse depends upon the LC_s circuit.

When the voltage change is negative going the circuit in Fig. 172 is employed. The only change made in this figure is that the diode is inverted.

The waveforms are as in Fig. 173 but inverted.

In the case of the negative-going voltage change a triode can be used instead of a diode and the added amplification of the triode can be made use of. The effect of grid current during the positive-going parts of the oscillation is the same as that of the diode in the circuit shown in Fig. 172.

Another form of ringing-circuit or shock-excited oscillator is shown in Fig. 175. R_g is a large resistance and the anode volts are divided between R_a and the very low value of R_{gk}. The grid is a fraction of a volt positive. The valve normally is therefore conducting, and near to saturation anode current. The full anode current flows through L of the tank circuit LC. If now a large negative pulse * is applied to the grid of the valve sufficient to cut off the anode current the resonant tank circuit is shocked into oscillation.

The frequency of oscillation is $\frac{1}{2\pi\sqrt{LC}}$ and the width in time of any half oscillation is $2\pi\sqrt{LC}$.

The number of oscillations per train depends upon the Q of the tank circuit.

When an oscillation is generated to form the basis of a short pulse, any of the oscillators described in Chapter V may be employed, but it is important to note that an oscillator which is crystal controlled cannot be pulsed because it does not immediately break into oscillation.

* The magnitude of the applied negative pulse must be sufficiently great to maintain the valve at cut off, *i.e.*, it must be larger than the maximum-amplitude of the positive-going oscillation generated by shock.

THE SINGLE-VALVE MULTIVIBRATOR—NIEMAN'S OSCILLATOR OR TRANSITRON

The special circuit arrangement of a pentode which causes it to function primarily as a multivibrator and by further modification as an oscillator may be recognised under any of the three above designations.

In its basic form it belongs essentially to the multivibrator family, but by suitable adjustment of components and their values it can be made to generate sustained oscillations.

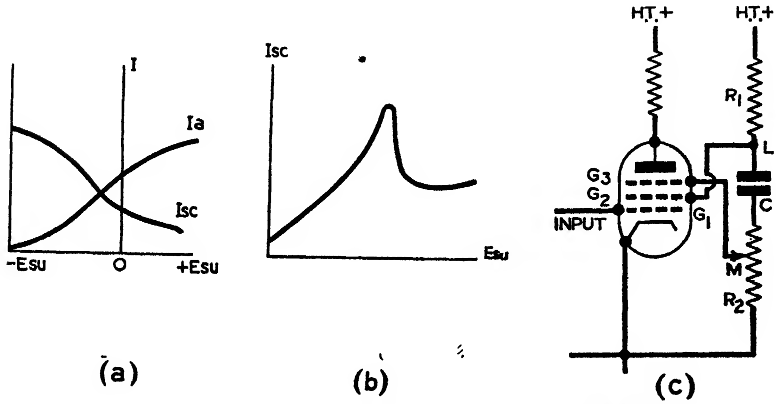


Fig. 176.—CIRCUIT AND CHARACTERISTICS OF THE TRANSITRON.

(a) Anode current suppressor voltage and screen current suppressor voltage characteristics.
(b) Screen current or suppressor voltage characteristic of a pentode with screen suppressor coupling.
(c) Basic single-valve multivibrator or transitron.

" Transitron " is the name by which it is generally known to American technicians.

In common with all multivibrators, it has a stable and an unstable state, the transition from one state to the other being effected by an incoming pulse, which in this case must be positive going.

The potential of the suppressor grid of a pentode has negligible control over the flow of electrons constituting the total space current; also it is effectively shielded by the screen grid and control grid.

It does, however, effect the division of the space current between screen grid and anode.

Fig. 176 (a) shows the relationship between suppressor volts and anode and screen currents.

Fig. 176 (b) shows the variations of screen current with suppressor volts when the screen and suppressor are strapped together.

(*a*) In general, if G_3 (Fig. 176 (c)) is made a little negative, the anode receives a smaller proportion of the space current and, consequently, G_2 receives a larger share. This increase of screen current causes the voltage on the screen to fall (due to drop across R_1). Under normal conditions, the drop in potential of G_2 is greater than the drop in potential of G_3, which causes it. The insertion of the resistance R_2 between the suppressor and its normal connection to the cathode allows G_3 to take up varying potentials and so to control the distribution of space current between anode and screen.

(*b*) The condenser connected between G_2 and G_3 causes G_3 to follow the instantaneous potential changes of G_2.

The effects in paragraphs (*a*) and (*b*) exist together and cannot really be separated ; consequently, when the potentials in a stable condition of (*a*) and (*b*) are disturbed by a short pulse, they seek new conditions, which can only be stable if they comply with the conditions imposed by the above effects.

Two separate cases are met with :—

(*a*) When the drop in potential of G_2 which would be caused (in the absence of C) by the drop in potential of G_3 is less than the drop in potential of G_3.

(*b*) When the drop in potential of G_2 (in absence of C) by the drop in potential of G_3 is greater than the drop in potential of G_3.

In case (*a*) the current will return to its original condition at the end of the disturbing pulse. This is the stable state.

In case (*b*) the circuit will change violently to a completely different condition. This is the unstable state which is made use of. Suppose that a positive-going pulse is applied to the control grid G_1, under condition (*b*).

G_3 and G_2 drop equally in potential, G_3 thus becoming very negative, the anode current is nearly cut off. The potential difference between L and M is the same as at the start.

C now commences to discharge *viâ* R_2 and the potentials of G_2 and G_3 will vary until a state similar to (*a*) is reached.

Simultaneously, the slow change in potentials of G_2 and G_3

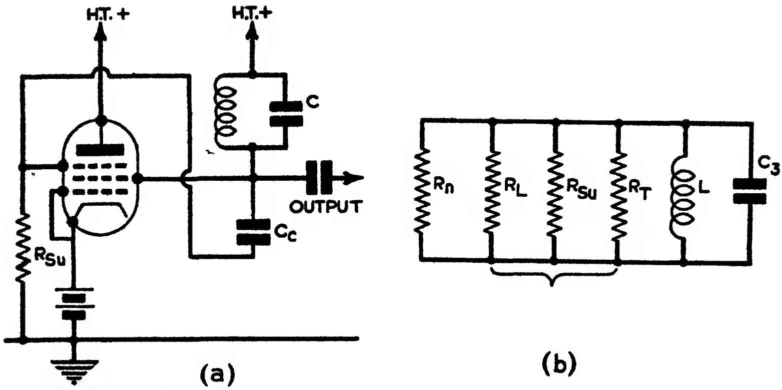


Fig. 177.—ELEMENTARY CIRCUIT OF THE TRANSITRON, WHEN USED AS AN OSCILLATOR, AND EQUIVALENT CIRCUIT.

(a) Elementary transitron oscillator.
(b) Equivalent circuit when acting as an oscillator.

acts as another disturbance and another violent change takes place in the opposite direction, until the original stable state is again reached.

Thus, a positive pulse applied to G_1 produces a sudden rise in anode volts, followed by a delay, the duration of which depends upon the values of C and R. A sudden fall of anode volts then occurs.

Consequently, when this device is connected as a multivibrator, one cycle is performed for each positive pulse received.

When connected as an oscillator with regenerative feedback, provided the magnitude of the energy fed back is in phase and equal to the energy dissipated in circuit losses, oscillations may be sustained at a frequency determined by the LC constants.

The elementary circuit of the transitron, when used as an oscillator and the equivalent circuit, is shown in Fig. 177.

The negative slope of the I_{sc} curve, that is the decrease of I_{sc} with increase of E_{su}, indicates that the transconductance between suppressor grid voltage and screen current is negative.* Since the reactance c of C_c is negligible at oscillation frequency, the alternating component of the suppressor voltage is of the

* It is sometimes *convenient* to think of negative resistance $\left(\frac{1}{\text{conductance}}\right)$ in terms of a generating effect taking place within the valve itself between its elements.

same polarity as that of the screen voltage. Thus, the negative transconductance becomes a negative resistance between screen grid and cathode. An increase of screen voltage causes a corresponding increase in suppressor voltage and hence a decrease in screen current.*

To maintain oscillations the losses in the circuit must be offset by energy supplied by the valve.

In the equivalent circuit R_n represents the negative resistance presented by the valve to the tuned circuit and the tuned circuit losses are represented by R_T paralleling the LC tank circuit, R_L represents the load resistance across the output terminals.

In order to produce oscillations of constant amplitude, the power supplied by R_n must equal the power consumed by all three resistances bracketed together. Thus the current through R_n must be equal and opposite to the sum of the three positive resistance currents through R_L, R_{su} and R_T. If the current through R_n is too small oscillations die out. If it is too large, they increase in amplitude.

* See footnote on p. 245.

CHAPTER XIV

RADAR TIME BASES AND TIME BASE GENERATOR CIRCUITS

THE function of the indicator unit, of which the cathode ray tube is the principal component, is to measure the time required for a transmitted pulse to travel outward to a target, and for the reflected energy to travel back to the transmitter-receiver position.

Measurement is accomplished by causing the electron beam of the cathode ray tube (normally held at the extreme left of the screen, looking at it from the front), to traverse the screen to the extreme right at some uniform speed, thereby sweeping out the trace of a time base in some given time, determined by the velocity of the beam or scan.

The course of the scan is marked by a bright line which presents the illusion of having been painted on the screen, because of the " afterglow " and the frequency of repetition, which is above that at which the eye can discriminate between successive time base sweeps.

This bright line, the length of which is calibrated by the constant velocity of the scan in terms of time, is the time base, and since 1 microsecond is proportional to 164 yds. of range, the time in microseconds represented by full-scale deflection of the time base $\times$ 164 is also proportional to range. Full-scale time base deflection can therefore be made proportional to any desired range by adjustment of the velocity of the scan.

The range of an echo is measured by making some full-scale time base deflection proportional to some convenient range and by causing the time base to commence at the same instant as the transmitter fires. The range of any echo which then appears on the trace is determined by its distance along the time base, measured from the start of the time base to the point at which the echo appears.

The time represented by the full-scale deflection of the time base is always longer than that of the transmitted R.F. pulse, in order that echoes returning after each transmitted

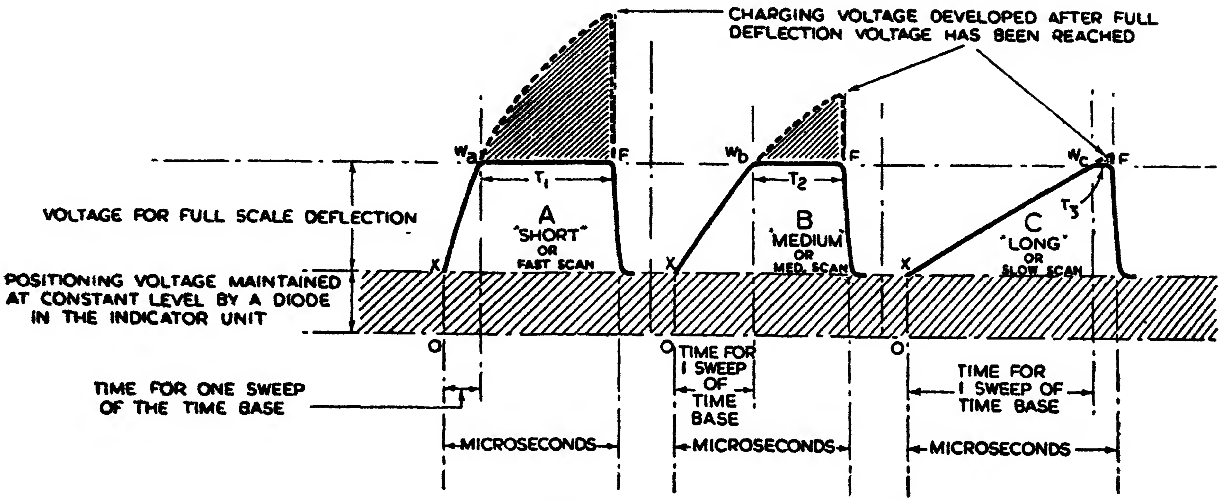


Fig. 178.—Time base requirements for accurate measurement.

A, B and C are all approximately linear. A is fast time base giving short range for full-scale deflection.
B is medium time base giving medium range for full-scale deflection.
C is slow time base giving long range for full-scale deflection.
W_a, full deflection voltage reached.
W_b, diode applies limiting operation.
W_c, the end of the brightening pulse.
F, controlling valve recovers and becomes a low resistance in parallel with the time base generator condensers.

Note.—At time X, controlling valve cuts off and time base condensers start to charge from H.T. supply. This is the start of the time base voltage; also :—

(1) The start of the time base trace.
(2) The instant of commencement of the brightening pulse.

pulse may be registered, but the full-scale deflection must not be greater than the interval between the start of one pulse and the beginning of the next, otherwise echoes will be masked. This means that the highest permissible repetition rate for a given set is limited by the maximum distance from which echoes are to be received.

Requirements for Accurate Measurement

The requirements for accurate measurement are :—

(*a*) Accurate synchronisation of the time at which the transmitter fires with the start of the time base. This means that the repetition rates of the time base and that of the transmitter must be identical and that they must fire simultaneously.

(*b*) The scan must traverse the screen from left to right with constant velocity, *i.e.*, it must move through equal distances in equal times, which is but another way of saying that the time base must be linear.

Another requirement is that a brightening pulse must be provided to intensify the beam from the commencement to the end of each sweep. At the end of each sweep the intensity of the beam must be reduced so that the beam itself may be returned to the extreme left-hand diameter in darkness; furthermore, the return stroke or " flyback " should take place as quickly as possible. These two latter requirements arise because (*a*) echoes received from outside the maximum range for which the full-scale deflection is calibrated would cause confusion in taking readings from the scale, and (*b*) because it is generally desirable to allow as much time as possible between the end of one sweep and the beginning of the next, particularly in the case of electromagnetic tubes.

From the foregoing considerations, it is clear that the force deflecting the beam must grow at a constant rate from the instant it starts until full-scale deflection is reached, the maximum value to which it must grow is determined by the diameter of the tube and is a constant for that tube. The rate at which it is required to grow determines the velocity of the scan and hence the time-range calibration for full-scale deflection (see Fig. 178).

If a deflecting force to fulfil all these conditions is plotted against time it will take the form of a sawtooth. The forward stroke is linear, rising by equal amounts in equal times. The flyback is as nearly vertical as possible. The gradient, or rate of growth of the deflecting force determines the velocity of the scan, the maximum value to which the deflecting force must rise for any given tube is determined by the diameter of the tube and is constant for that tube. In order, therefore, to calibrate the time base for a shorter range, it is necessary to increase the rate of growth of the deflecting force, in consequence of which the gradient of the linear part of the time base becomes steeper.

In practice it is desirable that provision should be made for selection by means of a switch from two or more time base speeds. The reason for this is that as the target approaches, full-scale deflection is conveniently made proportional to a shorter range thereby assisting the operator to follow the movements of the target with greater accuracy.

The deflecting force is an electrostatic field in the case of an electrostatic tube and a magnetic field in the case of an electromagnetic tube. Thus the deflecting force is proportional to the *voltage* between the deflecting plates in the first case and to the *current* in the deflecting coils in the second case, consequently the voltage wave generated by the time base generator must take the form of a sawtooth for the electrostatic tube. In the electromagnetic tube, however, it is the *current* that must be given a sawtooth waveform and the only requirement of the voltage waveform is that it should be such that when applied to the deflecting coils, the current waveform will be a sawtooth.

Since the deflecting coils of an electromagnetic tube are highly inductive, a sawtooth voltage applied across them will not produce a sawtooth wave of current through them ; consequently, the voltage waveform generated by the time base generator for an electromagnetic tube may differ widely from the waveform of current which it causes to flow in the deflecting coils.

Development of Voltage Waveform

The voltage wave shape necessary to accomplish this can be determined in the following manner :—

A coil having resistance and inductance may be considered as two elements in series (Fig. 179 (b)).

If a current through the deflecting coils equivalent electrically to Fig. 179 (b) is to be of sawtooth waveform examination of the facts shows :—

(*a*) A current waveform passes through the resistance component of the deflection coil unchanged in shape.

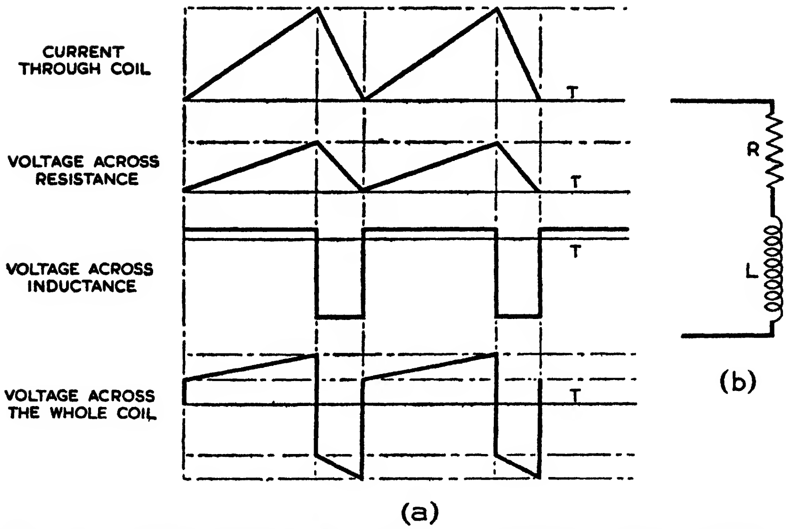


Fig. 179.—VOLTAGE WAVEFORM ACROSS RL TO PRODUCE A SAWTOOTH WAVEFORM OF CURRENT IS A COMBINATION OF A SAWTOOTH AND A RECTANGULAR WAVE.

(a) Development of voltage waveform required to generate sawtooth current wave through an inductance L and having a resistance R.

(b) Equivalent circuit of a coil having a resistance R.

(*b*) If a current wave of sawtooth shape is applied to an inductance a square wave of voltage appears across the inductance.

From (*a*) it follows that a sawtooth wave of voltage across a resistance produces a sawtooth wave of current. From (*b*) it is evident that a square wave of voltage must be applied to an inductance in order to produce a sawtooth wave of current through it. This also follows from results obtained when considering the L/R. circuit in Chapter IX.

Since the resistance of the coil is in series with the inductance of the coil (see equivalent circuit, Fig. 179 (b)), it follows that the waveform necessary to apply to a coil containing some resistance with inductance is obtained by superimposing the resistance e.m.f. waveform upon the E_L waveform. The result is the trapezoidal waveform shown in Fig. 179 (a), the magnitudes of E_R and E_L being related to values of R and L of the coil respectively.

In nearly all cases sawtooth waveforms of voltage are developed from the waveform resulting from the rise of voltage across a condenser during the charging period.

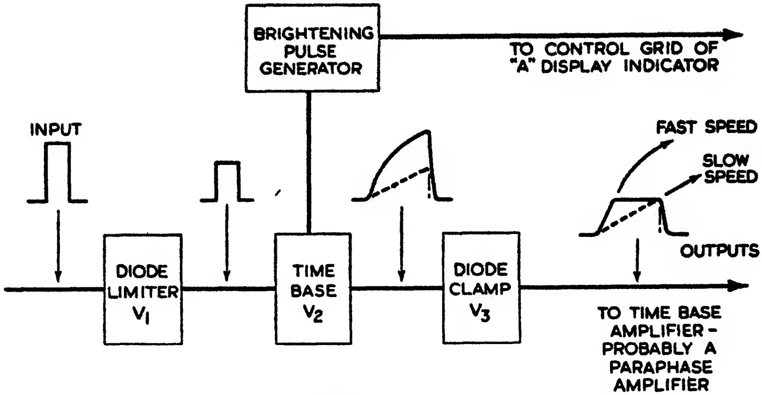


Fig. 180.—Block diagram of sweep generating assembly.

This voltage rises *exponentially* from zero to the value of the supply voltage. If, however, the ultimate value of the voltage across the condenser when fully charged (*i.e.*, E_B) is made very large compared with the maximum voltage required to produce full deflection for the time base, the waveform at its rise from zero to the required fraction may be regarded as tolerably linear and suitable for use as a time base voltage (see Fig. 178).

When a greater degree of accuracy of ranging is required the Miller time base * is used. In this case the small fraction of the charging voltage used is, by an artifice, modified and amplified to produce a time base voltage which is more linear by an amount equal to the stage gain of the amplifier.

Time Base Voltage

Consider in the first place circuits for developing a sawtooth wave of voltage. Circuits for the development of time bases for electromagnetic tubes are described later.

Fig. 180 shows the block diagram of a typical sequence of operations for developing a time base voltage from the rise of voltage across a condenser during the charging period. The circuit shown is merely one of many ways of performing this function (see Fig. 181).

The waveforms depicted in Figs. 180 and 181 show an amplitude for the time base for 20,000 yds. much greater than that for 100,000 yds. This condition arises because

* The Miller time base is fully described later in this chapter.

cut off for the time base generator must be fixed to take place only when the slowest time base voltage has reached the maximum value required to give full-scale deflection. As a consequence of this the fast time base voltage rises to a greater amplitude than the slow one and their waveforms at the output of V_2 (Fig. 180) are as shown.

The joint action of the brightening pulse and V_3 limits at the effective amplitudes of both time base voltages to just the value required to give full-scale deflection.

To Prevent Distortion (Fig. 181)

To prevent distortion, the grid of V_4 must not be allowed to go negative, relative to earth. V_3 clamps the grid of V_4 in this condition. In effect it acts as a one-way grid leak to earth. C_1 is comparatively small and therefore charges very quickly, consequently when grid current flows each increase in the applied positive-going wave causes an increase in grid current, the condenser charges to a higher value, but the potential of the grid relative to earth does not rise further.

In effect C_1 absorbs all increases of positive potential after grid current commences to flow and therefore by its action *limits* the amplitude of the output of V_4. Thus the voltage amplitude at the output of V_4 is controlled by the time at which grid current commences to flow.

The switch S_2 selects R_1 or R_2 and so adjusts the auto bias. Clearly the larger this bias is made the longer time elapses before grid current flows. Thus if S_2 is ganged to S_1, position 1 selects the smaller condenser in the time base generator and S_2 selects the smallest resistance for auto bias.

The fast sweep generated by the time base generator is thus limited for amplitude at the output of V_4 by the instant at which grid current commences to flow. In position 2, S_1 selects the larger condenser in the time base generator circuit and the larger resistance for auto bias of V_4. The grid being now more negative to the cathode a longer time elapses before it rises to the potential at which grid current commences to flow. This gives the slow time base time to rise to the full-scale voltage deflection value at which the faster time base is also limited.

Analysis of Fig. 181

A positive-going rectangular Synch. time-control pulse is applied to the grid of V_2 (the time base generator) *viâ* the limiter V_1. This ensures triggering of the time base generator

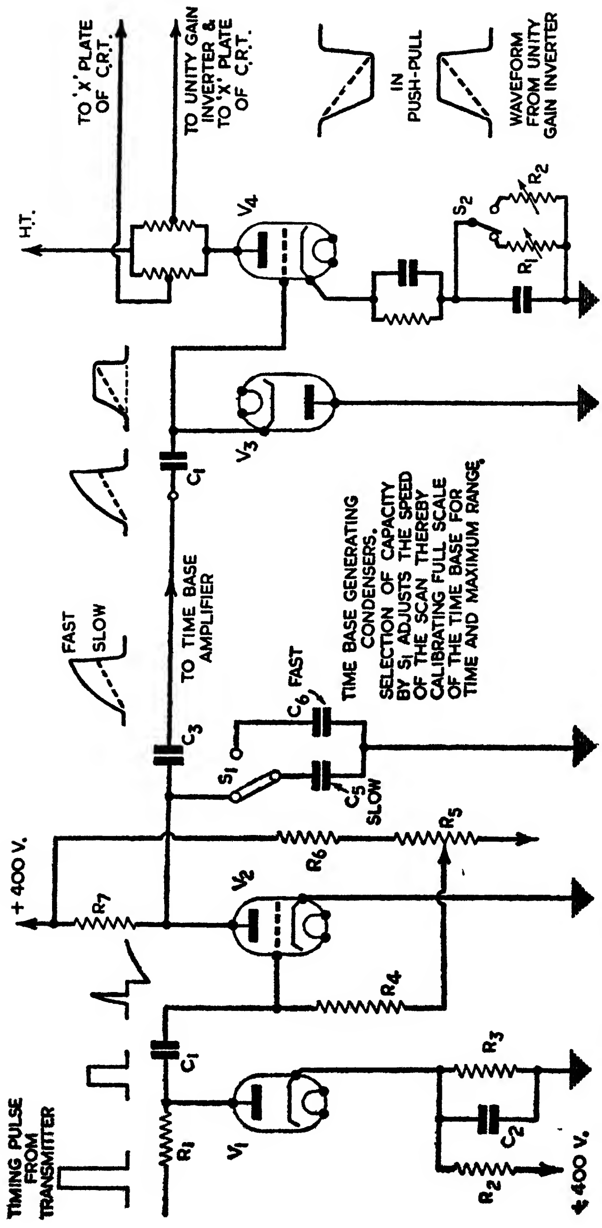


Fig. 181.—SCHEMATIC DIAGRAM SHOWING INPUTS AND OUTPUTS OF CIRCUITS OF TIME BASE GENERATOR AND AMPLIFIER.

NOTE. C_3 normally incorporates functions of C_1V_3.

at the proper operating voltage and synchronises the commencement of each sweep of the time base with the operation of other units of the system. The width of the rectangular pulse (in terms of time) determines the duration of *operation of* V_2.

Normally, V_2 is cut off. The cathode of V_1 is held at a positive potential above earth by positive bias and the voltage divider R_2R_3. V_1 does not therefore conduct until the amplitude of the positive-going rectangular pulse makes the anode of V_1 more positive than the cathode, consequently a voltage is applied to the grid of V_2 *viâ* C_1, which is limited in amplitude by the joint action of R_1 and V_1 as soon as the diode conducts.

The grid of V_2 is normally held slightly positive with respect to its cathode by the voltage divider R_5R_6. A change in the value of R_5 has little effect on the grid potential, because of grid current drawn through R_4, which limits the effective grid voltage supplied by the divider.

When the positive pulse is applied to the grid of V_2 it becomes instantaneously very positive and a large grid current flows to charge the condenser C_1 *viâ* the grid-cathode of V_2.

When C_1 has been charged to its full potential the grid of V_2 is returned to zero potential (approximately). At the end of the rectangular pulse, the falling edge drives the grid very negative and cuts off V_2.

(1) The charge on the grid condenser begins to leak off *viâ* R_4 to the positive supply.

(2) When V_2 is cut off by the discharge of C_1 the time base condenser C_5 or C_6 begins to charge to the H.T. supply *viâ* R_7. This charging action continues until V_2 again becomes a parallel low resistance across the time base condenser.

The duration of rise of the time base voltage (*i.e.*, the charging voltage across R_5 or R_6) depends upon the length of time for the negative grid voltage caused by the discharge of C_1 to reach cut off for V_2.

This time is made greater than the time required for an echo to return from the maximum range.

The waveforms for the time base voltages for Ranges 1 and 2 will therefore be as shown in Fig. 180.

For Range 1—the short range (assume that this range is, say, 20,000 yds. and that Range 2 is for, say, 100,000 yds.)—the time base must rise to maximum deflection value five times quicker than time base 2. This result is obtained by making condenser C_5 five times smaller than condenser C_6.

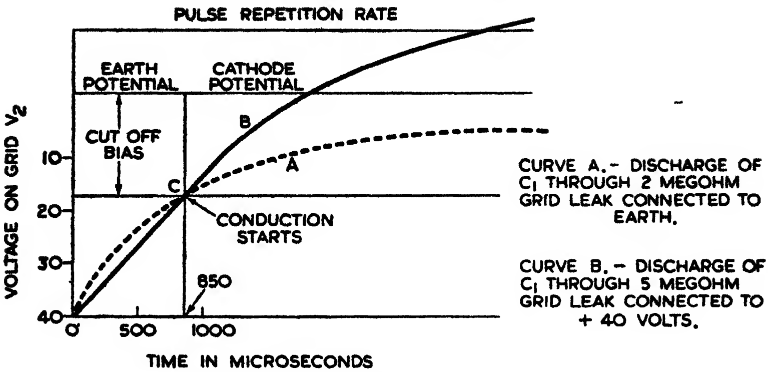


Fig. 182.—COMPARISON OF EFFECT ON GRID OF V_2 WHEN A HIGHER LEAK RESISTANCE CONNECTED TO A HIGH POSITIVE POTENTIAL IS USED.

The *charging time* (the leak-away time for the negative charge on the grid of V_2) is, however, determined by the time required for the time base voltage for the *longest* range (slowest rise) to attain to the full deflection voltage, consequently, in the interval between the time C_5 reaches maximum deflection voltage and cut off (as set for C_6), C_5 will continue to develop its exponential charging curve.

V_3 and V_4 now perform a limiting and amplifying operation so that the final waveform of the output for Range 1 is a steep linear rise of voltage to the maximum deflection value, limited at this value until cut-off time, as set for C_6 and followed by an almost vertical fall.

The waveform output for Range 2 is an approximately linear rise of voltage at a much lower rate to the maximum deflection voltage, and since this is set to be reached only just before cut-off time occurs, C_6 is not able to go on developing its exponential charging curve for very long. In consequence of this very little limiting is required for Range 2 waveform, and hence it has practically no flat top. The fall occurs almost at the instant that full deflection voltage is attained.

Consideration of the time base generator output waveforms, after the limiting operation has been applied, makes it clear that when the beam has been fully deflected to the extreme right it will be held there for the interval between the end of the rise of the time base voltage and the commencement of the fall, *i.e.*, for the duration of time indicated by the flat top of the time base voltage wave. This would cause a bright spot to appear at the extreme right of the time base. Since, however, the trace is only brightened by the brightening pulse for the duration of the forward stroke of each sweep, the undesirable bright spot which would otherwise appear at the beginning and end of the trace is eliminated, the intensity of the beam having been reduced to the point of extinction immediately either limit is attained.

It is important that C_5 or C_6 should discharge completely when V_2 conducts. To ensure this, it is essential that the negative charge left on the grid of V_2 (resulting from the discharge of C_1 through R_4 when V_2 is cut off) should leak completely away before the arrival of the next positive Synch. pulse.

The two curves A and B in Fig. 182 compare the effect on the grid of V_2 when a higher leak resistance connected to a high positive potential is used.

Curves A and B both reach the cut-off point C at the same time, but it is clear from its general shape that Curve A can never, theoretically, reach finality, which means that C_1 can never be fully discharged before arrival of the next positive pulse. Curve B however shows that C_1 will be discharged within the required period of time.

Input Requirements to X Plates of the C.R.T.

When a time base voltage is applied to the horizontal or X plates of a cathode ray tube to deflect the electron beam across the screen, it is generally desirable that the deflecting force acting on the beam should be balanced in order to avoid distortion of the beam and hence defocusing effects.

In order to accomplish this the deflecting voltages applied to the X plates must be connected in push-pull so that the mean value between them is zero at any instant of time.

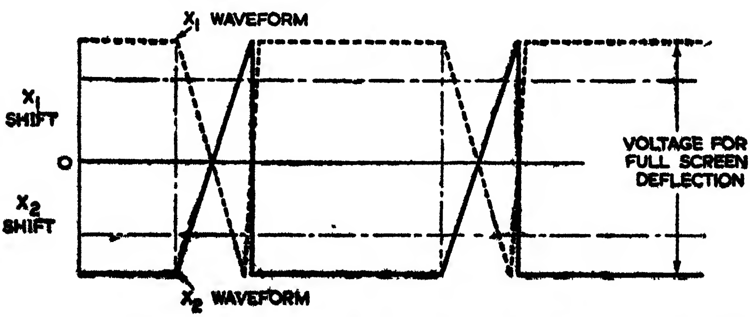


Fig. 183.—Ideal input waveforms in anti-phase for application to the X plates of a C.R.T.

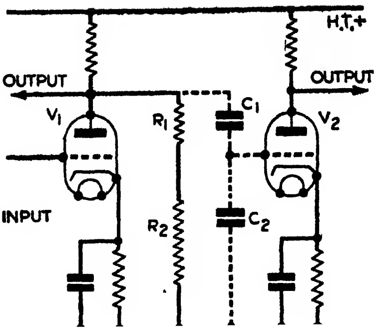


Fig. 184.—CIRCUIT OF A PARAPHASE AMPLIFIER.

Fig. 183 illustrates the conditions which it is desired to establish. This is done by feeding the time base voltage to the X plates in parallel but in anti-phase, or to be more exact, the waveform input to one plate is inverted relative to that applied to the other.

Classification and Use of Paraphase Amplifiers

There are a variety of ways of producing the balanced waveforms required, the majority of which depend upon the use of paraphase amplifiers, but they may be classified under two main headings :—

(*a*) The case where a sawtooth wave of voltage is generated at the required voltage level. In this instance part of the output may be applied direct to one of the X plates, the other part being inverted in a single stage unity gain amplifier before being applied to the other X plate.

(*b*) In other cases the sweep voltage may be generated at low level—perhaps 10 or 15 volts and amplified to the required value in a separate paraphase amplifier. This amplifier has two separate outputs, each identical in waveform and peak values, but the waveform of one is inverted relative to the other.

Fig. 184 shows a circuit to suit the conditions outlined in (*b*).

Suppose that the output from a sweep generator is, say, 15 volts, and that balanced deflection requires a voltage of 165 volts positive and negative going to the X plates respectively. Also let the gain of each amplifier be 11, then the output from V_1 will be 165 volts. (This forms one output.)

If the resistances of R_1 and R_2 are in the ratio of 10 : 1, the voltage applied to the grid of V_2 will be $\frac{1}{11}$th, *i.e.*, 15 volts. The output from V_2 will also be 165 volts, but the input waveform will have been inverted.

This may not be entirely satisfactory, however. Output No. 2 has been achieved by two amplifying operations. In the last case the input capacity C_2 of V_2 is in shunt with R_2 and may cause some distortion of the sawtooth waveform. This can be eliminated by shunting R_1 also with a capacity C_1 which must in this case be $\frac{1}{10}$th capacity of C_2 This does not alter the voltage ratio as long as $C_1R_1 = C_2R_2$.

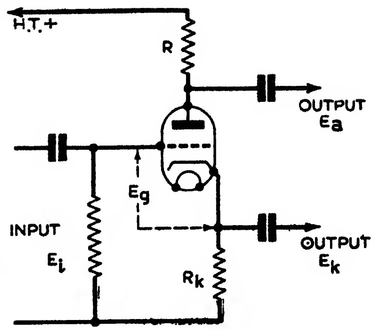


Fig. 185.—BALANCED LOAD PARA-PHASE AMPLIFIER.

Split-load Amplifier

The split-load amplifier (Fig. 185) is a convenient method of supplying from a single input two outputs which are inverted and balanced.

The resistances R and R_k are equal in value, consequently, in the absence of grid current, anode current causes changes of voltage across them that are equal.

Since the values of these resistances must be fairly large, the valve will be biased to cut off. If a smaller bias is required it can be obtained by by-passing part of the cathode resistance.

It is clear that the stage gain for either output must be less than 1, since the gain with feedback is $\frac{A}{1 - AB}$, consequently the overall gain for both inputs is less than 2.

In the equivalent circuit, Fig. 186, μE_g is the voltage across either resistance, E_i is voltage input, and E_o is voltage output.

$$\therefore \mu E_g = I_a (R_a + 2R)$$
$$E_g = E_i - E_o.$$
$$I_a = \frac{E_o}{R}$$
$$\therefore \mu (V_i - V_o) = \frac{E_o}{R}(R_a + 2R).$$

$$\text{Gain due to either load} = \frac{E_o}{E_i}$$
$$= \frac{\mu}{\frac{R_a + 2R}{R} - \mu}$$
$$\frac{\mu R}{R_a + 2R - \mu}$$
$$\frac{1 - \mu R}{R_a + 2R}$$

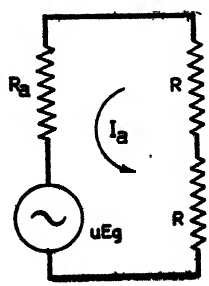


Fig. 186.—EQUIVALENT CIRCUIT FOR ANALYSIS.

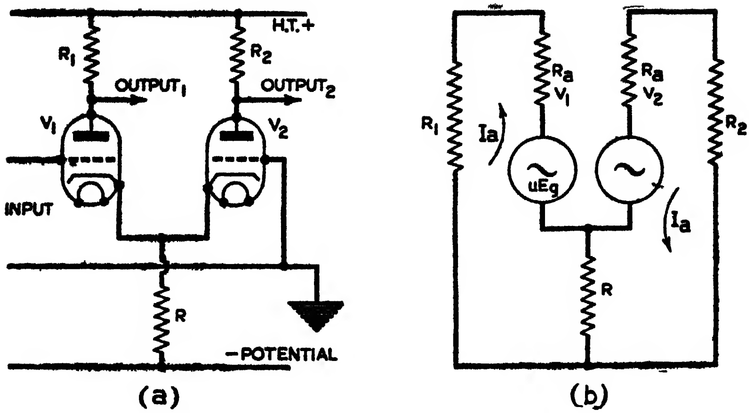


Fig. 187.—DEVELOPMENT OF PARAPHASE OUTPUTS FROM CATHODE-COUPLED AMPLIFIER.

(a) Cathode-coupled paraphase amplifier.
(b) Equivalent circuit.

$$\text{Putting } M = \frac{\mu R}{R_a + 2R}$$

$$= \tfrac{1}{2}\left(\frac{\mu 2R}{R_a + 2R}\right)$$

$= \frac{1}{2}$ the gain of an amplifier with an anode load of 2R.

This amplifier, like the cathode follower handles positive inputs better than negative. If the input is a positive-going sawtooth wave, the negative feedback cuts off the valve, leaving the stray capacities to complete discharge slowly *viâ* R_k, thus delaying the flyback. In cases where the time base working for comparatively long periods, *i.e.*, when the sweep time is short compared with resting time, this does not matter.

Cathode-coupled Paraphase Amplifier

Fig. 187 is the circuit of a cathode-coupled paraphase amplifier. The valves are coupled by a common cathode resistance R which is large enough to develop the e.m.f. necessary to operate, independently, either valve as a cathode follower.

In the static condition both valves are biased to cut off. Equivalent circuit, Fig. 187 (b).

In order to restore the working point to a position in which a small percentage change in the current through R will alter the grid-cathode bias between zero and cut off, R is taken to a negative potential E, which may be -200 volts or so. Thus R and E are interdependent and not affected by the values of other components in the circuit.

The anode loads R_1 and R_2 are usually made small and, not as in a normal amplifier, where they are large as compared with R_a. This minimises amplitude distortion.

The input is applied between the two grids, which are held at a fixed potential—earth in this case.

In the static condition both valves are equally biased by an amount $E - IR$ where $I =$ the combined anode currents. The bias adjusts itself automatically to permit the flow of these currents. Thus :—

$$\text{Grid-cathode bias } E_{gb} = (I_{a1} + I_{a2})\, R$$
$$\text{and } I_{a1} + I_{a2} = \frac{\text{Grid to cathode bias}}{R}.$$

When the input is applied the grid cannot change by an amount greater than the grid bias—say 10 to 15 volts, but since it is small compared with E, it follows that the sum of the combined anode currents is nearly constant.

This means that the two anode currents are complementary, *i.e.*, if one increases by a certain amount, the other will decrease by approximately the same amount.

If E_g is made positive, I_a increases and the potential of both cathodes rises. This change provides the input to V_2 which now by its complementary action endeavours to restore IR to its original value. As a result, the current settles down to an amount where the cathode rises to half the amount that the grid rises relative to earth. Consequently the change in the grid-cathode voltage of the two valves is equal and opposite. Thus if the grid rises to 6 the cathode rises to 3.

$$\therefore\ E_{g1} = +3 \text{ and } E_{g2} = -3.$$

The increase in current to produce this change (if the cathode resistance $= 15{,}000$ ohms) is $\cdot 2$ mA.

Therefore half the input appears as actual voltage applied to the grid-cathode circuit of each valve.

It follows that the total amplification for the whole circuit is the same as that obtained by using one valve as a normal amplifier with a load equal to R.

If the grid of V_1 goes sufficiently positive, a point is reached where the voltage drop across R exceeds E by an amount equal to the grid of V_2. V_2 is then cut off. An increased positive input results in cathode follower action on the part of V_1, grid current will eventually flow but, on account of the protection afforded by the cathode resistance, the results are not as serious as they would be in a normal amplifier.

For a negative-going input, the current through V_1 decreases, whilst that through V_2 increases up to the point at which V_1 is completely cut off. Further input in the negative direction will cause no change in the current of either valve.

Use with a Time Base Amplifier

If this circuit is used with a time base amplifier, the two outputs are taken directly to the X plates. In this case, the X shift is obtained by taking the output of V_2 to the slider of a potentiometer. Thus, changes in D.C. volts on the grid will result in the amplification of D.C. changes at each anode but in opposite directions. If a variable resistance is included with R in the cathode lead, variations will provide changes of D.C. voltage in the same direction at both anodes, and this may be used as an astigmatism correction.

The advantages of this circuit are :—

(*a*) No distortion.
(*b*) Small grid current when overloaded.
(*c*) Freedom from tendency to self oscillation.
(*d*) Freedom from hum.

Another application of this amplifier is that with separate inputs to the grid, the output at either anode is the sum of the two inputs. The voltage across R is equal to the difference of the two inputs.

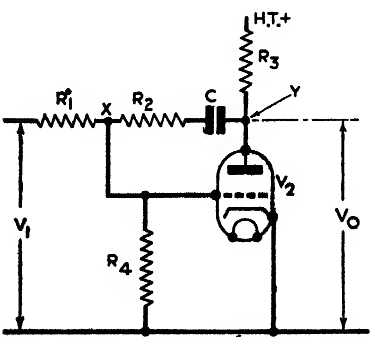


Fig. 188.—FLOATING PARAPHASE CIRCUIT.

R_1R_2 about equal.
R_4 high compared with R_1R_2.
R_3 low compared with R_1R_2.
C is a large condenser.

Floating Paraphase Circuit

In the floating paraphase circuit (Fig. 188), the anodes of the two valves are connected through R_1 and R_2 from the junction of which R_4 goes to the cathodes. The grid of V_2 is connected to X, the common input of the resistance. This is referred to as the "floating point."

Biasing arrangements are not shown, consequently V_2 looks as though it has a

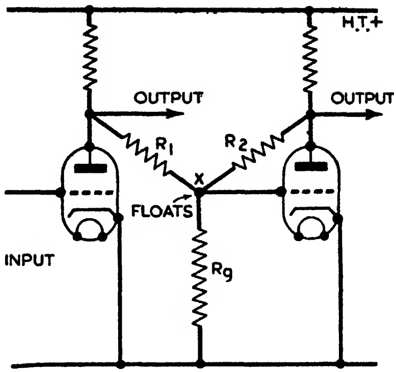


Fig. 189.—COMPLETE DIAGRAM OF FLOATING PARAPHASE AMPLIFIER CIRCUIT.

positive bias. A condenser may be used to separate the junction of R_1, R_2 from R_3 and the grid, also auto bias can be provided for both valves.

The first stage is normal, excepting that the gain is modified by the resistance network in parallel with the load resistance.

Variations at the point X provide the input to V_2 and these variations depend on the *changes* at the anodes of both valves which are in anti-phase.

If $R_1 = R_2$, which is usually the case, equal and opposite changes at the anodes produce zero variation in potential at X.

Consequently, output 2 is less than output 1 by just the amount necessary to produce variations at X which are to establish output 2.

This is an example of voltage-controlled negative feedback since the output is coupled back to the input *via* R_2.

The fraction of output which is fed back is approximately $\dfrac{R_4}{R_2 + R_4}$.

All voltages mentioned are voltage changes relative to the normal voltages :—

In Fig. 189 let a positive voltage V_1 be suddenly applied. The point X will change in voltage for two reasons :

(*a*) The application of V_1 which would make X rise to V_a, *i.e.*, $\left(\dfrac{R_2}{R_1 + R_2}\right)$, if the valve were out of action.

(*b*) A change in anode voltage (the output voltage) V_o which would make X rise to V_b, *i.e.*, $V_o\left(\dfrac{R_1}{R_1 + R_2}\right)$ if V_1 were zero and the input short circuited.

The total change V_x with these two factors operating at the same time is :—

$$V_1R_2/(R_1 + R_2)V_oR_1/(R_1 + R_2)$$

or $$V_1R_2 + V_oR_1/(R_1 + R_2) = V_x \quad . \; . \; . \; . \; . \; . \; . \quad (1)$$

But V_x on the grid must produce a change $A \times V_x$ at the anode. A being the stage gain of the valve under normal conditions.

$$\therefore\ V_o = -A \times V_x \text{ or } V_x = \frac{-V_o}{A}.$$

Putting this value of V_x in the equation (1) above,

$$(V_1R_2 + V_oR_1)/(R_1 + R_2) = \frac{-V_o}{A}.$$

Multiplying through by A $(R_1 + R_2)$

$$A(V_1R_2 + V_oR_1) = -V_o(R_1 + R_2)$$

collecting terms in V_o

$$V_o(AR_1 + R_1 + R_2) = -V_1 AR_2$$
$$\text{or } V_o/V_1 = AR_2/(AR_1 + R_1 + R_2).$$

If $R_1 = R_2$ and A is large this is nearly equal to -1, and so the circuit acts as an inverter.

More exactly, if $AR_2 = AR_1 + R_1 + R_2$, V_o will be exactly equal to $-V_o$.

This is the case if $R_2(A - 1) = R_1(A + 1)$ or $\frac{R_2}{R_1} = \frac{A + 1}{A - 1}$.

The Output Impedance

The impedance R can be most easily deduced by supposing that a sudden voltage E is inserted in a break in the anode lead at say Y. If the change of current in the anode is I, then $R = E/I$.

Suppose $V_1 = 0$ and input short circuited, the voltage E applied at Y will (of itself) produce a small change in anode current. It will, however, produce a change $\frac{R_1}{R_1 + R_2} \times E$ at X, which will produce a change in the anode current (approx.) $I = g\frac{R_1}{R_1 + R_2}E$ where g is the mutual conductance of the valve.

$$\text{Thus, } R = E/I = Z/(g \times \frac{R_1}{R_1 + R_2} \times g)$$
$$= \frac{R_1 + R_2}{gR_1}$$

If R_1 is about equal to R_2, $R = \frac{2}{g}$, *i.e.*, the output impedance is *twice the reciprocal of the mutual conductance.*

If $g = 6$ mA/v $= 0{\cdot}006$ A per volt.

$$\frac{1}{g} = \frac{1{,}000}{6} = 160 \text{ ohms}$$

$$R = \frac{2}{g} \qquad = 320 \text{ ohms.}$$

This junction of R_1R_2 remains at a nearly constant potential.

The amplifier can be regarded as the counterpart of the cathode follower and is sometimes known as an anode follower.

It shares with the cathode follower the advantage of low output impedance.

The Miller Time Base

In general, the Miller time base circuit is used for producing a time base voltage which is far more linear than any other time base circuit using the same voltage of H.T. supply.

This means that equal changes of voltages take place in equal time.

Miller demonstrated that the effect on the grid circuit of a valve in which the grid-anode capacity is C and the stage gain is A, is the same as could be produced by connecting a condenser C (A + 1) from grid to earth. Thus the same effect occurs when this capacity is added to by connecting an actual condenser from grid to anode external to the valve.

In the above case if the grid potential is raised by 1 volt the anode potential falls by A volts. The voltage across the grid-anode condenser C is decreased by A + 1 volts.

In order to make the voltage across a condenser C fall (A + 1) volts, the condenser must be charged with a charge (A + 1) times the charge needed to alter its voltage by 1 volt. But this larger charge would alter the voltage of a condenser of capacity of C (A + 1) by 1 volt. Thus the grid rises 1 volt but current flows into the condenser sufficient to charge a capacity C (A + 1) by 1 volt.

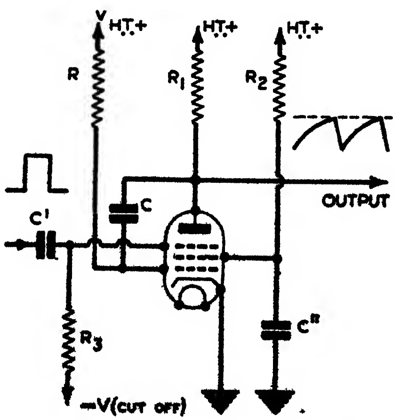


Fig. 190.—Miller time base.

Consequently, the effect on the grid circuit is the same as would be produced by increasing the grid-cathode capacity by C (A + 1).

The basic Miller time base is shown in Fig. 190. The grid leak R is taken to H.T. and the suppressor leak R_3 is taken to a negative voltage sufficient to cut the anode current off.

To start the action a sufficiently great positive-going square pulse is applied through a condenser C^1 to the suppressor and anode current flows.

The anode voltage drops and the grid voltage drops by an equal amount, since grid and anode are tied together by C, which acts as a short circuit for sudden changes.

This drop in grid volts would tend to make the anode volts rise, but due to the condenser charging current the fall in anode volts and grid volts must continue progressively until a state of equilibrium is reached.

When this condition is reached the anode current is only just sufficient to cause a drop in R_1 equal to the amount that the grid has gone negative from its normal potential.

This fall occurs as suddenly as the start of the pulse which triggered the circuit into action.

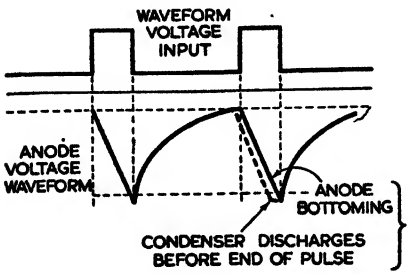


Fig. 191.—Input and output waveforms for a Miller time base generator fired by a square pulse to a suppressor.

The grid voltage is now negative and no grid current can flow. *The development of the time base proper now commences.*

Normally a time base voltage is taken from the voltage developed by the charging current of a condenser during the initial stages of charging, when the rate of growth is at its highest, rising nearly linearly.

These desirable conditions occur only at the very start of the charge, deteriorating rapidly as the voltage across the condenser rises towards the charging voltage.*

If the maximum value of the charging voltage is increased, the whole charging curve is expanded in length and consequently the useable linear portion becomes proportionately longer, thus making a larger and suitable maximum time base voltage available.

* This is obvious since the charging curve is exponential.

Thus in order to secure these conditions for the development of large time base voltages it becomes necessary to use very large charging voltages with all their attendant disadvantages. Alternatively, the Miller time base circuit develops, at comparatively low voltage, a charging current which is made to grow more linearly at the start and continues the linear part of its rise for a longer time, thus providing all the advantages to be obtained from a very high charging voltage. It operates in the following manner :—

When the grid voltage is negative and no grid current flows, the side of C connected to the grid is negative but wants to rise.

The grid voltage (which in an ordinary C.R. circuit, normally rises at V/CR volts per second at the start) will now rise at the slower rate of V/C(A + 1)R volts per second at the start.

However, the anode voltage changes at A times the rate of change of the grid voltage and, therefore, falls at V.A/C(A + 1) R volts per second. This can be written as V/CR × A/(A + 1). If A (the amplification factor) is large enough, A/(A + 1) is nearly = 1, therefore the rate of rise is nearly 1/CR volts per second which is independent of the valve characteristics.

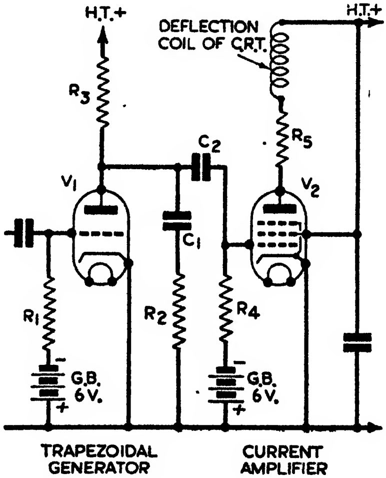


Fig. 192.—CIRCUIT FOR GENERATING A SAWTOOTH CURRENT WAVE FOR ELECTROMAGNETIC DEFLECTION.

Thus at the output- an amplified version of a suitable portion of the impressed charging current curve, developed in the grid circuit, is available as a linear time base ; at a voltage A times that of the grid circuit. In other words, a time base is produced A times as linear as an ordinary C.R. circuit using the same H.T. voltage.

If the square pulse which started the action lasts long enough, the time base action will continue until the anode is just at cathode potential and cannot drop any further. This condition is termed "bottoming."

On the other hand, the linear part of the time base may still be going on when the square pulse ends.

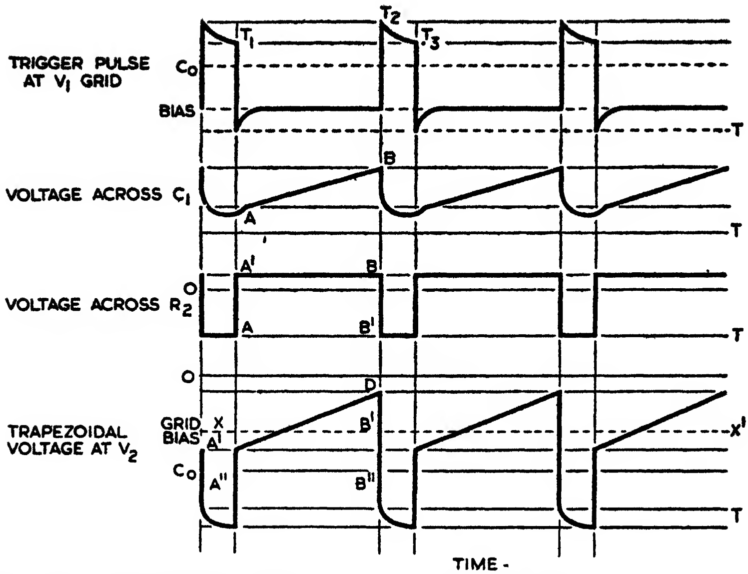


Fig 193.—WAVEFORM IN CIRCUIT OF FIG. 192 FOR GENERATING A SAWTOOTH CURRENT WAVE.

In either case, when the square pulse ends the suppressor cuts off the anode current and the condenser charges at a rate dependent upon C and not upon C(A + 1) and on R_1, since the Miller effect is now non-existent until a state is reached similar to that before the start of the square pulse.

Generating a Sawtooth Time Base for Electromagnetic Deflection

The circuit shown in Fig. 192 is suitable for generating a sawtooth time base for electromagnetic deflection.

The operating characteristics of V_1 are such that C_1 charges on the linear part of the curve, therefore the current flowing in R_3, C_1 and R_2 during the charge must be constant. In these circumstances the voltage across C_1 will rise linearly (AB, Fig. 193). A constant current flowing through R_2 develops a constant voltage across it (A′B, Fig. 193). When V_1 is made to conduct at Time 2 the current through C_1 and R_2 reverses and C_1 starts to discharge through the valve.

The discharge current causes the voltage across R_2 to swing negative very quickly (B—B′, Fig. 193) as the condenser discharges from Time 2 towards Time 3.

The voltage that appears at the anode of V_1 is the sum of the voltages across R_2 and C_1, and so produces the waveform shown.

The axis of the trapezoid is along the line X—X′ after it passes through coupling condenser C_2. The trapezoid of voltage A″, A′, D, B′, B″ produces a sawtooth wave of current in the deflecting coil.

At time B″ the current in V_2 must be reduced to zero, since the valve is cut off by the negative voltage on its grid. If this occurs the current in the deflecting coil situated on the neck of the C.R.T. must be reduced to zero at the same instant of time.

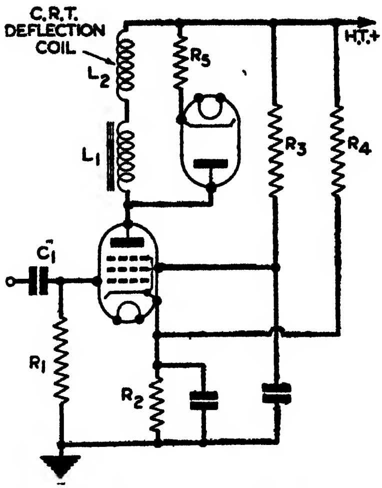


Fig. 194.—SAWTOOTH TIME BASE CURRENT GENERATOR.

In order to provide some means for dissipating the energy of the electromagnetic field so that the current can fall to zero, a resistance is often connected in parallel with the deflecting coil. This resistance also serves to reduce the "Q" of the coil and so prevent the coil from being shocked into oscillation. Unless this is done, oscillations may persist into the next sweep of the time base and so cause distortion. The parallel resistance is sometimes replaced by a diode.

When the current through V_1 is increasing to provide each forward sweep of the time base current, the voltage at the anode of V_1 is less than the supply voltage by the amount of the drop through the deflecting coil; therefore, when a damping diode is used, the cathode is more positive than its anode and the diode will not conduct. When the current falls to zero through V_1 at the end of each sweep of the time base, the voltage of the anode rises above the supply voltage, due to the persistence of the deflecting coil. The energy in the coil is then dissipated by the resistance R_5 and any oscillations that attempt to raise the anode voltage above the supply voltage are quickly damped out.

Sawtooth Time Base Current Generator

Fig. 194 shows an alternative circuit to Fig. 192. In this

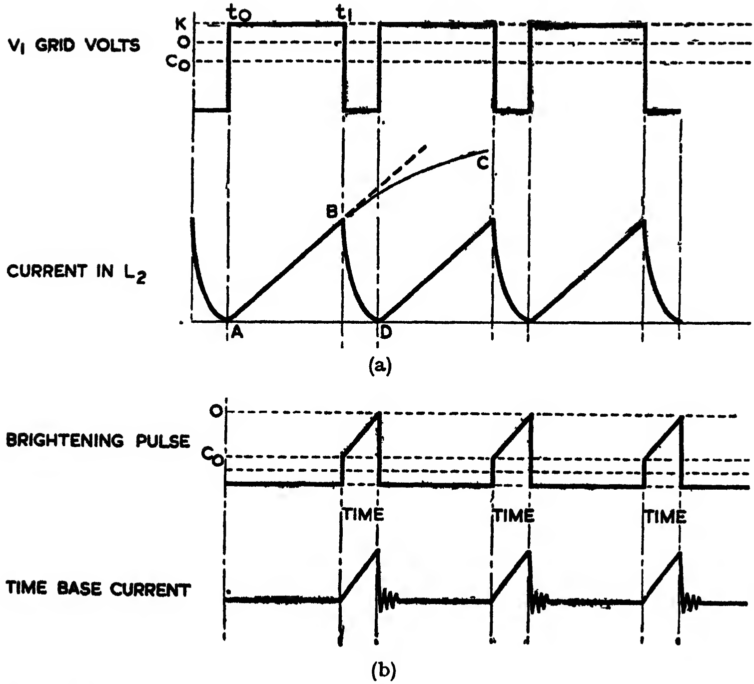


Fig. 195.—WAVEFORMS OF SAWTOOTH CURRENT DERIVED FROM TIME BASE GENERATOR, FIG. 194, ALSO VOLTAGE WAVEFORM OF THE BRIGHTENING PULSE TO THE C.R.T.

(a) Sawtooth current wave derived from square input pulse.
(b) Waveform of brightening pulse voltage to C.R.T. and resultant time base current waveform as exemplified by linear trace on C.R.T. screen. Note (b) is not drawn to the same time scale as (a).

case, however, a linear rise of *current* is generated in an inductance instead of a linear rise of *voltage* across a condenser.

V_1 is normally cut off by bias provided by the potentiometer chain R_4R_2.

At time $t = 0$, the square wave input to the grid is positive and permits a current to flow *viâ* R_2, V_1, L_1 and L_2 (the deflecting coil). This current builds up towards some steady value along an exponential curve such as ABC, Fig. 195. The first part of this curve approaches linearity, so that if the current is stopped at some time t_1 before the curve departs from linearity appreciably, a reasonably linear current has risen in the deflecting coil. At time t_1 the current is stopped by fall of the square pulse to the grid of V_1 and the energy of the magnetic field is dissipated in R_5 and the diode.

Since it is generally desirable to make the time base only a fraction of the time period of the repetition rate, a voltage of the form shown in Fig. 195 (b) must also be applied to the cathode ray tube (positive to the control grid or negative to the cathode) to intensify the beam and brighten the screen for the period of each time base sweep only.

This device has the advantage that it is unnecessary to take precautions to damp out oscillations, since they take place in darkness and are therefore not seen on the screen.

The P.P.I. Display

A form of display often employed is the P.P.I., or prepared plan indicator. This differs from the simple "A" display which has been previously discussed.

In the P.P.I. display the transmitter reference is located at the centre instead of at the extreme left of the screen. The trace of the time base now starts from the centre and moves outwards radially to the circumference, each forward stroke of the time base being the radius of the circular plan which is developed and displayed on the screen of the C.R.T. circle.

For this display an electromagnetic tube with long afterglow is generally employed, and the deflecting coil is made to rotate round the neck of the tube in synchronism with the rotation of a directional aerial. Thus the beam, and hence the trace has two components of motion, viz. a radial movement and an angular movement.

The radial movement differs very little from the diagonal movement discussed for the "A" display, the important difference is that the beam is now positioned initially at the centre of the screen and the trace, starting from the centre, is deflected outwards towards the circumference. The total deflecting force is therefore only half that required for an "A" display on a tube of similar size. The flyback takes place in darkness at the instant that each sweep of the time base is accomplished, *i.e.*, when it reaches the circumference of the display circle.

The aerial is rotated continuously by a motor at speeds which may vary for different installations from 6 or 8 revolutions per minute to 20 or more. The deflecting coil around the neck of the cathode ray tube is rotated by a Selsyn motor controlled by the aerial and is, therefore, maintained in step with the aerial.

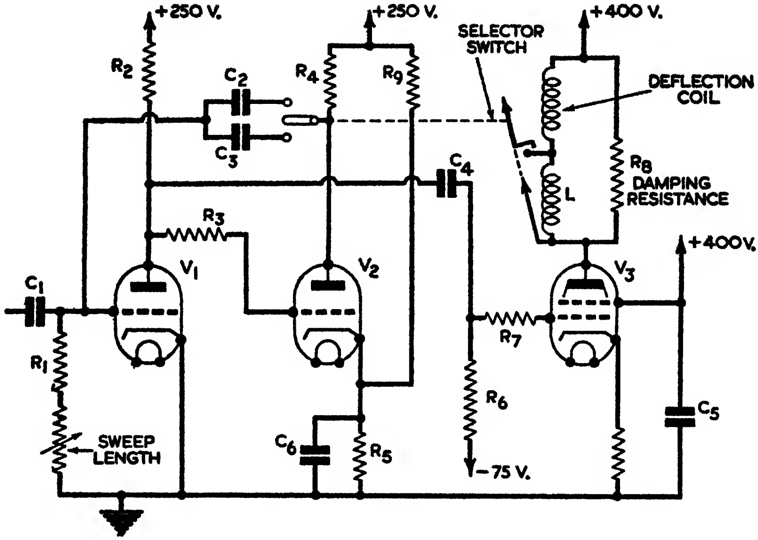


Fig. 196.—P.P.I. TYPE TIME BASE MULTIVIBRATOR AND GENERATOR.

Under the combined influence of the magnetic fields produced by the time base current in the deflecting coil and the angular velocity imparted to the deflecting coil, the beam is made to scan the whole screen area enclosed within the circumference of the display circle once per revolution of the aerial. Thus the circular display is described by a series of forward strokes of the radial scan from centre to circumference, each stroke being successively directed to a point further round on the circumference in synchronism with the rotation of the aerial. Consequently, if a screen with long afterglow is employed and the input from the receiver is applied as a brightening signal to the grid or cathode of the tube, a polar map of the area, of which the transmitting aerial is the centre is produced, and echoes from targets and reflecting surfaces, as seen by the aerial, appear as bright arcs or areas on it in their relative positions.

Thus the range and bearing of all targets, such as surface craft, high land, landmarks and buoys, can be ascertained within a circular area the radius of which is equal to the maximum range of the equipment.

P.P.I. Time Base Generator Employing a Multivibrator Combination

A suitable time base sweep may be generated by a multivibrator combination of which an example is shown in Fig. 196. In this particular case the system is assumed to have a radius of 100,000 yds. and to be used for long-distance warning purposes. Let it also be assumed that provision is made for a fast scan to give a maximum scale reading of 20,000 yds. instead of 100,000, thus enabling readings of increased accuracy to be made when the target approaches within the 20,000 yds. range.

A range of 100,000 yds. at 164 yds. per microsecond is proportional to 610 microseconds. A range of 20,000 yds. is proportional to 122 microseconds, therefore the speed of the time base must be such that deflection of the beam from the centre to the circumference takes place in 610 microseconds for the long range (slow scan) and 122 microseconds for the shorter range (or fast scan).

If the repetition frequency is, say, 800 cycles per second, there is an interval of 1,250 microseconds between pulses. In order to permit the current through the deflecting coils to return to zero, or positioning value, it is desirable that the flyback for the slow-moving time base sweep should occur immediately after the end of the 610 microseconds.

A multivibrator may be used to control the duration of each sweep, since it can generate a pulse of exactly the right duration at intervals which can be regulated by a synchronising pulse.

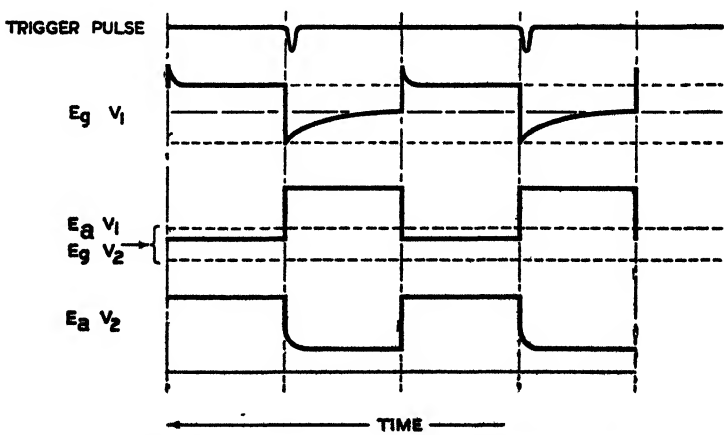


Fig. 197.—WAVEFORM IN SWEEP MULTIVIBRATOR.

The operation of this type of multivibrator has been fully described in Chapter XIII.

The type selected in Fig. 196 is triggered by a negative pulse, and the output is a positive-going square wave.

V_1 is normally conducting, and operating as it does without bias, the anode current is normally about 8 mA., causing a drop of about 180 volts across R_2. The effective peak voltage of V_1 is therefore about + 90 volts. V_1 swings between 70 + and 250 +.

R_3 limits grid current for V_2 to a safe value in event of a positive signal with respect to cathode.

E_k, V_2 is controlled by the voltage divider R_5, R_9, which places the cathode at about 125 volts above earth.

The effective cathode voltage of V_2 when V_1 conducts is — 35 volts, which cuts off V_2.

The multivibrator remains inactive until a Synch. pulse is applied. The operating sequence is described in detail in Chapter XIII, and the waveforms are shown in Fig. 197.

The switch selects C_2 for the 20,000 yds. range, and C_3 for the 100,000 yds. range, C_3 being five times larger than C_2.

The pulse from the anode of V_1 is rectangular. The multivibrator is started by a Synch. pulse and the duration of the positive-going pulse is accurately set by either of two values :—

(*a*) The selector switch.

(*b*) The variable resistance in the grid of V_1 marked "sweep length," which may be regarded as a fine adjustment.

The positive-going rectangular pulse from the anode of V_1 is now applied to the time base generator proper, V_3.

Since two ranges are required, the inductance L is used to limit the rise of current to that which is just sufficient to deflect the beam linearly from the centre of the screen to the outer edge in 610 microseconds.

The current for the short range must build up to the same maximum value as for the longer range, but this build up must also take place at a rate five times faster than for the long range. This is achieved by shorting L.

V_3 is normally cut off by the — 75 volts bias applied to the grid. Its action is similar to that of a switch. It is turned on by the leading edge of the positive wave from V_1 and turned off again by the falling edge.

The anode load must remain constant during the time the valve is conducting, in order that the time base may be linear, therefore the grid potential must not change during this time. In order to accomplish this the time constant of the input circuit C_4R_6 is made long relative to the duration of the positive pulse. Also R_7 is added in order to limit grid current. V_3 is a beam power valve because a large current is required for the deflecting coil.

The valve cannot be cut off by cathode or auto bias, but the resistance between cathode and earth should be less than 1,000 ohms, so that the build up of current in the deflecting coil may be just as linear as possible during the conducting period.

The cut-off bias is therefore a negative voltage to the grid. The T.B. current in the deflecting coil must be reduced to zero at the end of each sweep of the time base, so that the electron beam of the C.R.T. may be repositioned at the centre of the screen. A small circle appears around the centre in the form shown in Fig. 198 (a) when the positioning voltage is correctly adjusted. If the circle appears as in Fig. 198 (b), necessity for readjustment of the positioning or standing current in the deflecting coils is indicated.

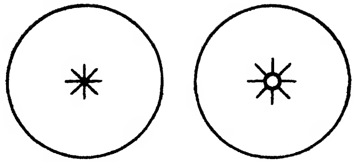


Fig. 198.—CORRECT AND INCORRECT ADJUSTMENT OF THE POSITIONING VOLTAGE FOR CENTERING THE RADIAL TIME BASE P.P.I. DISPLAY.

(a) Indicates correct adjustment of positioning voltage for the time base of the C.R.T.

(b) Incorrect positioning; each radial sweep must be lengthened towards the centre.

When V_3 is cut off the energy in the magnetic field is dissipated in the damping resistance R_8.

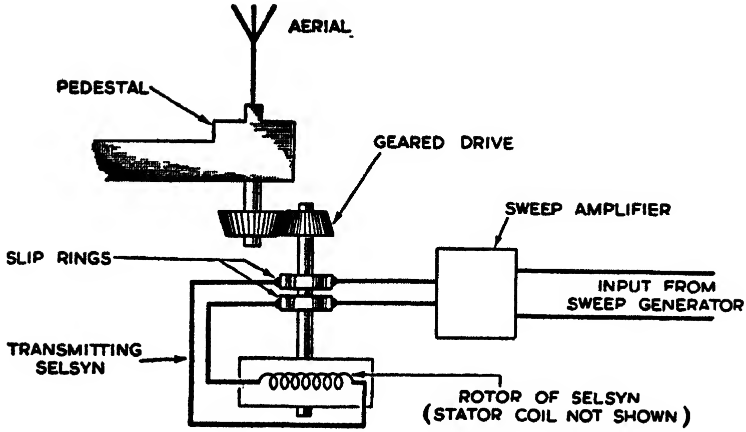
The oscillations produced by shock do not affect the appearance of the screen because the intensity of the beam is reduced at the end of each sweep of the time base.

Employment of a Selsyn Arrangement

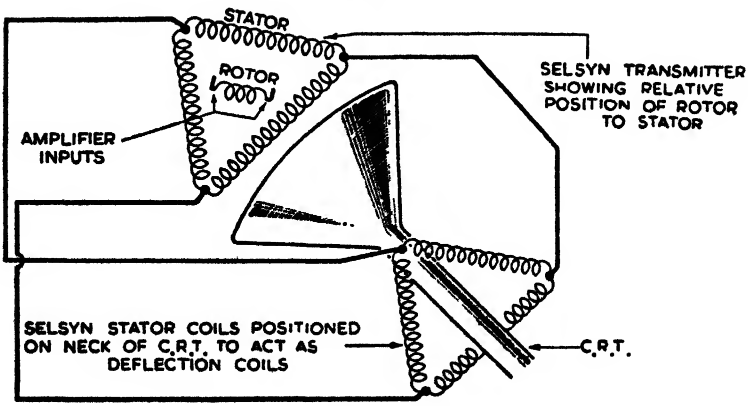
In an alternative arrangement, Fig. 199, using a somewhat similar sweep generator, but followed by a large power amplifier, the time base current is passed to the deflecting element by a Selsyn. The Selsyn is normally used to transmit, electrically, to the operating position, the movements of the aerial relative to the ship's head, or as a position control device.

Briefly, the aerial is mechanically coupled to the rotor of a Selsyn in such a manner that any movement of the aerial causes a redistribution of voltages in the stator of the Selsyn.

These electrical changes which take place in the stator of the



No. 1



No. 2

Fig. 199.—Schematic diagram of application of sweep current to selsyn (transmitter) and transmission of sweep current for deflecting beam radially with an angular motion in step with aerial rotation.

Selsyn at the aerial end of the system, as a result of any movement of the aerial relative to the ship's head, are transmitted by cables to another stator mounted on the neck of the cathode ray tube, this latter then performs the functions of the deflecting element. It is not part of the Servo System.

In this arrangement the redistribution of voltage induced in the stator of the Selsyn at the aerial end for each movement of the aerial is repeated in the stator of the Selsyn mounted on the neck of the cathode ray tube.

It follows, therefore, that if the time base current is passed through the rotor of the Selsyn at the aerial, angular displacement of the combined aerial positioning and radial time base voltages induced in the stator of the Selsyn at the aerial end will be determined by the position of the aerial at any instant of time, wherefore the distribution of current in the stator of the Selsyn mounted on the neck of the cathode ray tube will also be directly controlled by the position of the aerial.

As a result, if the aerial is rotated continuously, the time base currents in the stator mounted on the cathode ray tube will be in step with the aerial, and a rotating, radial time base will be produced. The effect will be similar to that produced when a normal deflecting coil carrying a time base current is rotated by a motor.

This arrangement has many advantages, inasmuch as it avoids the mechanical difficulties encountered in mounting a rotating deflecting coil on the neck of the C.R.T. and it also permits rotation at higher speeds. On the other hand, a large-power amplifier is essential to amplify the output of the time base generator in order to provide the required power to operate the device.

The capacity of the Selsyn rotor windings introduces an undesirable distortion to the early part of the generated time base. In order to avoid this, the distorted portion of the time base generator output is not used. In other words, the brightening pulse is delayed until after the start of the time base generator by a time just sufficient to start the visible time base trace after the unseen distorted portion. It follows, therefore, that when this system is employed, in order that all the other units may be kept in step, that they must be synchronised, by a pulse taken from the time base generator. Also a delay must be introduced equal to the time elapsing between

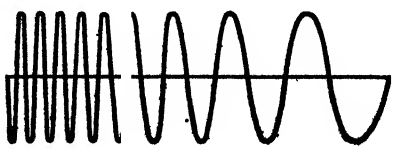


Fig. 200.—EXPONENTIAL TIME BASE PRODUCED BY INPUT DUE TO CURRENT WAVEFORMS WHEN CHARGING A CONDENSER TO ITS MAXIMUM VOLTAGE.

the start of the time base generator and the instant at which the usable portion of its output is brightened.

Non-Linear Time Bases

Non-linear time bases are sometimes used for special purposes. For example, if a time base is not limited to that part of the exponential charging curve which is reasonably linear, a non-linear sawtooth wave is obtained. This produces an exponential time base (see Fig. 200). A sine wave input to the "Y" plates when an exponential time base is applied to the "X" plates produces a trace as shown in Fig. 200.

A sinusoidal time base is obtained by applying a sine wave to the "X" plates. The trace is now swept out alternately (left-right) and (right-left). The speed of this time base varies in a sinusoidal manner across the screen, *i.e.*, fast at the centre and slow at each end. The return trace is made at the same speed as the outward sweep. In this way reversed waveforms can be received.

If the amplitude of the sine wave is increased very considerably, only a small fraction of the circumference of a circle appears on the screen and this is approximately linear. Such an arrangement is sometimes used for showing pulse phenomena that take place over one cycle.

By suitable arrangements of the inputs to the "X" and "Y" plates, circular, elliptical and spiral traces may also be obtained.

Circular Trace

A circular trace is produced from two sine waves of the same frequency and amplitude applied to the "X" and "Y" plates respectively, but 90° out of phase.

If the voltages in the above case are not of equal amplitude, an ellipse will appear. A similar effect is obtained when the phase difference between the two sine waves is not 90°.

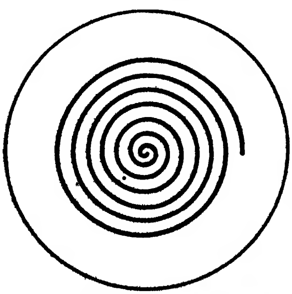


Fig. 201.—SPIRAL TIME BASE.

Spiral Trace

A spiral trace is produced by an arrangement similar to that for a circular trace, but the amplitude of one of the waves must be varied continuously to obtain this effect. A spiral time base can be used when a very long trace is required, *i.e.*, where a comparatively long period of time is to be measured by each sweep of the time base (see Fig. 201).

These time bases are not generally used in radar sets excepting for special purposes. An elliptical trace is very economical to produce, and when a high degree of accuracy is not required it may be sufficiently satisfactory. When an elliptical time base is used for an "A" display, for example, some portion of the reasonably flat part of a large ellipse, between the two foci, may be employed.

In general it may be said that the degree of accuracy with which target distance can be read off a given display depends upon the accurate timing of the sweep for each range, *i.e.* the velocity with which the electron beam moves across the screen and its linearity.

The importance, therefore, of maintaining a constant check, by means of the calibrator, against any departure from each standard range speed through deterioration of circuit components or other causes, cannot be too strongly emphasized.

This operation should be carried out religiously at the commencement of each watch, and in special cases, where circumstances permit, a check should also be made before any reading of major importance is to be taken.

CHAPTER XV

CALIBRATION AND CALIBRATOR CIRCUITS

RANGE MARKS AND RANGE MARK GENERATORS

CALIBRATION of the *screen* of a cathode ray tube is performed in the first place by the production of a trace under the control of a time base circuit as described in Chapter XIV " Radar Time Bases and Time Base Generator Circuits."

If trace time is understood to mean the time elapsing between the commencement of the trace and its completion, and delay time means time elapsing between the transmission and the instant when the echo returns,

$$\text{Trace time} = \frac{\text{Trace length (cms.)}}{\text{velocity of beam deflection (cms. per sec.)}}.$$

For a given cathode ray tube the maximum diameter trace length is a constant, therefore when maximum total trace length is employed, as is nearly always the case, trace time is inversely proportional to the velocity of beam deflection and inversely proportional to the rate of growth of the deflecting voltage.

Thus a trace of a given length may be made proportional to any desired time interval by suitable adjustment of the beam deflection velocity or scanning speed, also by application of the echo delay time equivalent of 164 yds. per microsecond, it may be directly calibrated for range in yards.

In order to translate any fraction of the trace directly into range, the range scale for the entire trace length must first be known.

This is determined by the speed of the scan, *i.e.*, the rate of growth of sweep voltage for which the time base has been adjusted.

Two or more alternative speeds are usually provided, selection being made by a switch.

Hence, the setting of the time base selector switch determines the maximum range to which the trace is made proportional.

At first sight it would appear that it is only necessary to place over the screen a transparent scale, dividing the trace into some convenient number of equal parts, in order to read off directly, or by interpolation, the range of any echo appearing along the trace.

There are two factors that render this simple scheme inadequate, unless supplemented by the provision of some calibrated standard which can be used for checking purposes from time to time :—

(a) Small variations in voltage, or circuit constants, ageing of valves and components, etc., all make it necessary that the speed of the trace as selected by the range switch should be checked, and if necessary adjusted, at frequent intervals.

(*b*) The rate of growth of the trace itself is not ideally linear (although when the Miller time base is employed it is very nearly so), consequently an error must creep in between the geometrically exact divisions of the transparency scale and the corresponding values on the trace with which they are supposed to coincide.

These two factors alone introduce certain complexities, the extent of which depend upon the degree of accuracy required.

Calibrating the Trace

It is therefore, in connection with the problems of initial calibration and routine checking of the scanning speeds and subsequent application to the measurement of range, that this section is concerned.

In the simplest calibrating system a generator (range mark generator) started by a Synch. pulse at the same instant as the time base, produces at equal intervals of time, short sharp signals or pips (at intervals corresponding perhaps to 1,000 yds. for example), and these are imposed upon the trace just as though they were incoming signals.

When this system is employed the calibrator is not normally in use at the same time as echoes are being received. When a check is to be carried out, generally at the commencement of each watch, the calibrator is brought into operation by a switch, and the position of the calibration pips along the trace is compared with the position of their corresponding values on the transparent scale.

Should these not coincide exactly, the general procedure is to line up the direct signal and a calibration pip with the zero mark on the fixed scale, by using the X shift control * and

* The X shift control is shown in diagrams of cathode ray tube connections. In effect it adjusts the starting point of each successive trace on the screen by varying the positioning voltage.

then to adjust the time base or scanning speed until subsequent pips coincide with appropriate ranges on the scale.

Frequent Checking

Frequent checking is necessary to avoid error due to voltage variations or change in circuit time base constants. For example, let the range mark generator be an oscillator, which is started and stopped by a square pulse applied under the control of the master timing control or synchroniser. Its duration is made the same as that of the trace time.

The output from the range mark generator during each sweep period is a train of oscillations at some frequency n and a time period of 1/n seconds. Since such oscillation eventually becomes a marker signal imposed on the trace at equal intervals along its length, to represent some fixed fraction of it, the frequency and time period of the generator output must be determined by the fraction of length which each marker is to represent.

For example, suppose it is required to place a marker on the trace at intervals of 1,000 yds.

Equivalent time interval of the pips which is in other words the time period of the oscillation is

$$1{,}000 \text{ yds.} = \frac{1{,}000 \text{ yds.}}{164 \text{ yds. per microsecond}}.$$
$$= 6{\cdot}1 \text{ microseconds (approx.)}$$

This means that if the time period of the oscillation is made 6·1 microseconds, each oscillation covers 1,000 yds. of the trace; consequently, if the first oscillation in each train commences at the beginning of the trace, successive oscillations must appear along the trace at intervals of 1,000 yds.

Thus if the time base is adjusted to make the trace length proportional to, say, 20,000 yds., 20 oscillations generated with a time period of 6·1 microseconds will exactly cover the trace at equal intervals of 1,000 yds.

A time period of 6·1 microseconds is equal to a frequency of $\frac{1}{6{\cdot}1 \text{ microseconds}} = 163{\cdot}9$ kc/s. (approx.). The frequency to which the output of the range mark generator must be adjusted is, therefore, determined by the required interval between successive range marks.

The time for which the range mark generator must operate is determined by the maximum range to which the trace is made proportional, *i.e.*, $6{\cdot}1 \times 20$ microseconds $= 20{,}000/164$ or nearly, since 6·1 is not quite exact.

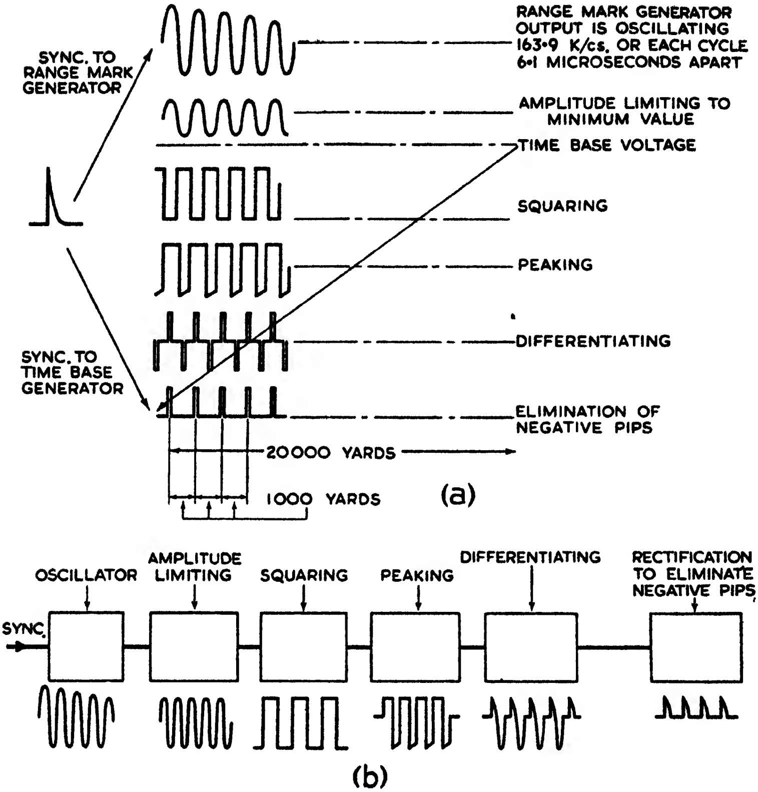


Fig. 202.—TYPICAL SERIES OF OPERATIONS FOR THE PRODUCTION OF RANGE MARKS.

(a) Development of range mark waveform from an oscillation train of the required frequency.
(b) Sequence of operations shown in block form and the output waveforms resulting.

Range Markers

The range marker generator in the above example must, therefore, generate 20 oscillations at a frequency of 163·9 kc/s. approx. (163·86 kc/s. to be exact), each time it is started by a timing pulse.

Fig. 202 shows a series of operations by which range marks can be produced.

The output, consisting of a series of pips spaced at regular intervals, can be applied for calibration purposes to the signal plates of the cathode ray tube, when no received signals are present, in which case they appear as pips along the trace, dividing it into 1,000-yd. lengths. In the case of a P.P.I. they can be applied to the grid or cathode of the cathode ray tube as brightening pulses, in which case bright concentric rings * will appear on the screen, spaced at equal radii.

From the foregoing it should be at once evident that the factors affecting the accuracy of measurement of the C.R.T. display are :—

(*a*) The linearity and correct adjustment of the time base circuits.

(*b*) The synchronisation of the output of the calibration generator with the start of the time base generator.

(*c*) Accurate setting of the oscillator frequency so that the trace is divided into equal parts of known length.

(*d*) The oscillator must have a high degree of frequency stability or freedom from frequency drift.

Determining Exact Range

A number of oscillators of different types are available for use as range mark generators, selection being determined by the degree of accuracy required (frequency stability of the oscillator) and the method employed for determining exact range. This latter point requires some explanation.

When the comparatively small diameter of the screen of a C.R.T. is required to represent several thousand yards there is obviously a limit to which it is usefully possible to calibrate the trace by small subdivisions on a corresponding scale. On the other hand, range requirements call for accurate readings within the limits of $\pm$ 200 yds. for long range warning sets down to a few yards for short range precision sets. Clearly, therefore, the method of direct reading from a scale to the nearest subdivision corresponding to the position of the echo on the trace, and estimating or interpolating for any fraction, is not good enough when extreme accuracy is required.

In consequence of this, devices have been developed such, that by turning a handle, a mark (generally a strobe) † can be made to correspond with the leading edge of the echo to be

* The pips applied to the control grid become rings on the C.R.T. screen because of the angular velocity imparted to the radial time base when the deflecting magnetic field is revolved.

† A strobe is a bright spot or movable range marker, generated by a strobe generator under the control of some variable delay pulse. This is fully dealt with in the chapter on indicators.

measured. In this case the range is read off directly from the dial of a counting mechanism, actuated by the rotating mechanism which causes the mark or strobe to move from zero reference to coincidence with the leading edge of the desired echo.

These systems are not dealt with here, but are more fully described in the chapter on indicators.

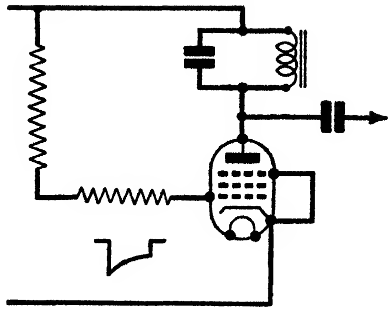


Fig. 203.—CIRCUIT AND INPUT WAVEFORM FOR RINGING OSCILLATOR AS APPLIED TO GENERATOR OF OSCILLATIONS FOR RANGE MARKS.

The immediate point at issue is, that when a high degree of accuracy is required, it is essential that the frequency of the fixed range mark generator should be stabilised by crystal control.

Note. The fixed range mark generator referred to is the generator for calibrating the scale. The movable range mark or strobe generator is quite independent of the range mark generator.

Before passing to crystal-controlled oscillation, it must be pointed out that range marks or markers for calibration do not always appear on the same trace as the echoes. It is sometimes convenient, particularly when continuous calibration is required, to produce a separate trace * which, although independent of the signal trace, is identical in every respect and appears on the screen simultaneously with it. The manner in which this double trace is achieved is dealt with in the detailed chapter on indicators (see footnote). In this case the echoes may appear on the top trace, being measured for range by the range marks on the lower trace with which they are seen to coincide.

Crystal-controlled Oscillators

When crystal-controlled oscillators are used as range mark generators a new problem is introduced. It is no longer

* The double trace is produced on the screen of a C.R.T. by a separation pulse which is applied to the vertical or Y plates to shift the trace periodically to a position higher or lower than the normal. If this is done at some rate above the persistence of vision, the illusion is created of both traces appearing on the screen simultaneously. Echoes appear on one trace and range mark pips on the other.

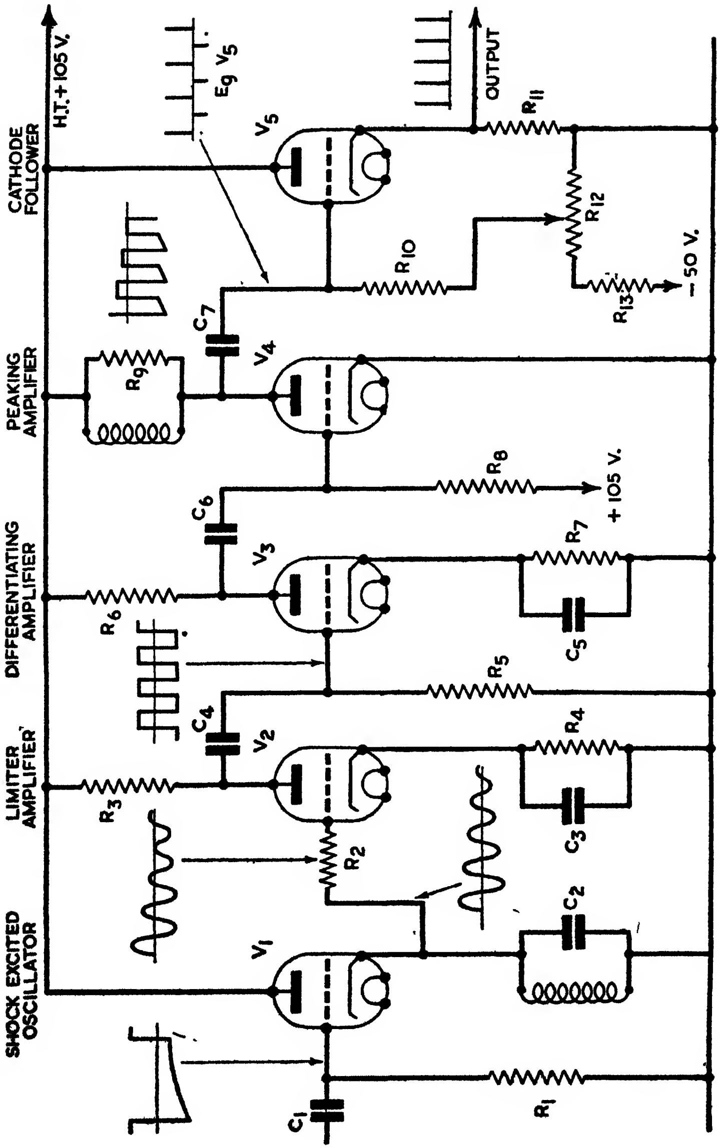


Fig. 204.—SHOCK-EXCITED OSCILLATOR INPUT PRODUCES RANGE MARKERS.

possible to pulse the generator in order to synchronise the start and finish of each train of oscillations with the start and finish of consecutive sweeps of the time base.

This is due to the fact that the crystal-controlled oscillator *does not start up instantaneously when pulsed.*

If a range mark input generated under these conditions were applied to the C.R.T., individual range marks would not always appear in the same place on the trace and the stroboscopic illusion of fixed position would be lost. The range marks would move along the trace—in radar terminology "jitter." The difficulty is overcome by allowing the oscillator to run continuously, but arranging for its output to be gated.*

A simple means of producing range marks by the use of a ringing oscillator (not crystal controlled) is shown in Fig. 203.

The frequency stability of this type of oscillator (although uncontrolled) is very good, because the oscillating circuit is very lightly loaded and the frequency in no way depends upon the action of the valve.

Analysis of Circuits of Fig. 204

A negative-going square pulse drives the grid of V_1 below cut off at the instant the sweep starts, thus setting up oscillations in the L.C. circuit.

The damped oscillations resulting from shock excitation of the L.C. circuit of V_1† are limited and amplified by V_2 so that the output from the anode of V_2 is a series of approximately rectangular pulses all of the *same amplitude*. R_2 limits grid current and thus reduces damping in the L.C. circuit of V_1. The cathode bias across R_4C_3 prevents the grid going extremely positive and also helps to keep grid current down. A low anode voltage is used for the same purpose.

Note. Grid limiting, cathode bias and low anode voltage are all precautions to avoid damping of the ringing circuit oscillations.

V_3 acts as an overdriven amplifier and squares the output from V_2.

V_4 has an inductance in its anode load. This inductance resonates with its self capacity at about 2 megacycles per second.

* A pulse is said to be "gated" when it is allowed to function only under the control of some other pulse, usually a Synch. pulse.

† Note the L.C. circuit of V_1 in Fig. 204 is in the cathode circuit instead of the anode circuit as shown in Fig. 203. The effect is similar with exception that V_1 can be coupled to V_2 as a cathode follower and so avoid damping of the L.C. circuit by the grid-cathode load of V_2 (grid current).

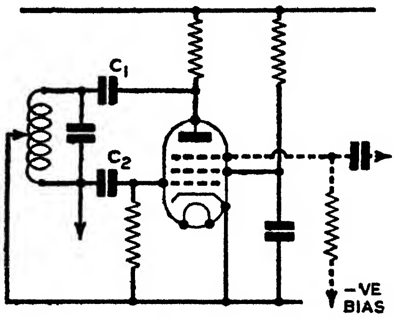


Fig. 205 —HARTLEY OSCILLATOR (DOTTED PORTION SHOWS METHOD OF PULSING).

When V_4 is cut off by the output of V_3 the inductance coil is shocked into oscillation, but the resistance R_9 damps the oscillation almost completely out before one cycle is completed. Thus a positive pulse of about 0·25 microseconds' duration is produced. When the grid swings positive a negative pulse of larger amplitude appears at the anode.

The grid of V_4 is returned to a positive potential rather than earth in order to obtain high conductance in the valve just before E_g goes negative. Grid current is limited to a small value by R_8.

The output of V_4 is fed to V_5 biased beyond cut off, so that the negative pulse is lost. The amplitude of the output pulses can be adjusted by the bias.

If the oscillations in V_1 are damped too much to be useful over the entire sweep, a regenerative circuit can be used to sustain their amplitude.

Other types of oscillators with particular forms of pulsing arrangements may be used, such, for example, as the circuits shown in Fig. 205.

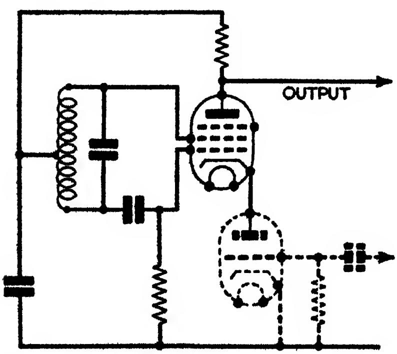


Fig. 206.—ELECTRON-COUPLED OSCILLATOR. (PULSING CIRCUIT DOTTED IN.)

The Hartley Oscillator

The circuit shown in Fig. 205 is that of a normal Hartley oscillator.

Frequency is determined by the L.C. constants of the oscillating circuit, small adjustment being made by varying the position of the core of the inductance. The frequency is affected to some extent by the voltage on the electrodes, the Q of the oscillating circuit and the output loading. It is therefore important to make the Q of the tank circuit as high as possible and to select carefully the values for C_1 and C_2.

The adverse effect of the load in the tuned circuit due to the output is avoided in the circuit of Fig. 206 by the use of electron coupling, tank to output.

The circuit shown in Fig. 206 is a series-fed Hartley circuit. The screen grid of the pentode acts as the virtual anode of a triode oscillator. The oscillating component of the screen current maintains oscillations and the resulting oscillating screen voltage modulates the main electron stream to the anode.

Provided the minimum anode voltage is not too low, the anode current of a pentode is independent of changes in anode voltage. Therefore, the output load has little effect on the oscillator section of the circuit, resulting in improved frequency stability.

In Fig. 205 the suppressor is sufficiently negative to prevent oscillations occurring until the arrival of a positive-going pulse.

Suppressor switching cannot be used in Fig. 206, and cathode switching by means of a low-impedance triode is employed instead.

The triode acts as a cathode resister to the pentode, and prevents oscillations occurring.

A positive control pulse from possibly a multivibrator or some other source is fed to the grid of the triode and simultaneously to the control grid of the pentode. The triode now passes a large current, which means that R_a has been reduced sufficiently to allow the pentode to operate.

All self-oscillators require some time in which to build up maximum oscillations and the time required appears to depend largely upon the magnitude of the initial shock applied to start the oscillation. Consequently, the first few oscillations of each train

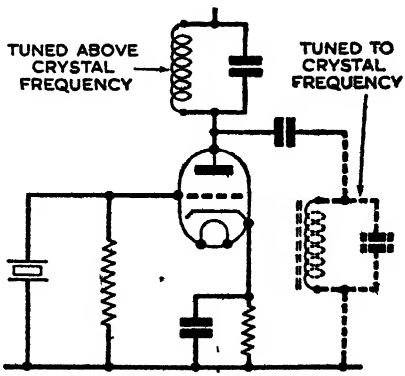


Fig. 207.—CRYSTAL PROVIDES GRID EXCITATION TO THE ABOVE CRYSTAL-CONTROLLED OSCILLATOR AND THE OUTPUT IS NOT A PURE SINE WAVE.

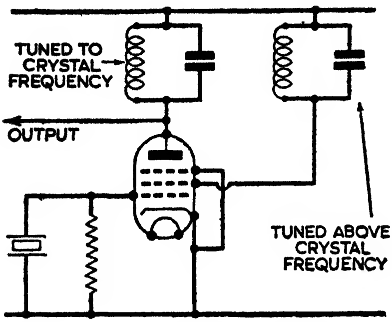


Fig. 208. — CRYSTAL - CONTROLLED OSCILLATOR HAVING A PURE SINE WAVE OUTPUT.

after passing through, limiting and differentiating circuits are not likely to be quite evenly spaced and may be discarded if necessary by introducing a suitable delay.

When a Crystal Oscillator is Used

When greater frequency stability is required, a crystal oscillator is used. This introduces a new difficulty. A much longer time is required for oscillations to build up and die down in such a circuit on account of its high Q, in consequence of which, crystal-controlled oscillators must be run continuously and gated out.

The circuit of a crystal oscillator of the tuned grid, tuned anode type, in which the crystal provides grid excitation, is given in Fig. 207. The following points should be noted :—

(*a*) The anode circuit must be tuned above the crystal frequency.

(*b*) A small control of frequency can be obtained by shunting the crystal with a condenser.

(*c*) The output is taken from the anode and is not a pure sine wave.

In certain circumstances connected with the use of accurate range-measuring circuits (see chapter on indicators and range measurement with phase-shifting transformer), it is desirable that the output waveform should approach as nearly as possible to a pure sine wave. In this case the circuit of Fig. 208 may be employed. The cathode, control grid and screen act as a triode oscillator. The output is taken across the oscillation circuit in the anode lead which is electron coupled.

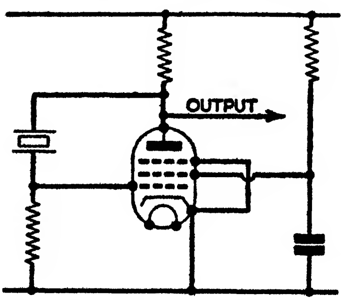


Fig. 209.—CRYSTAL-CONTROLLED OSCILLATOR WITH CRYSTAL INSERTED IN ANODE-GRID CIRCUIT.

A further alternative is shown in Fig. 209. The crystal is connected between grid and anode. This circuit gives a larger output than the previous one and with resistance in the anode circuit the output is rich in harmonics.

Under operating conditions, regardless of whether calibration is carried out continuously or whether single or double traces are used, it is essential that whilst the calibrator is in use, range marks should appear to be perfectly steady as though painted on the screen.

When the range mark generator is pulsed at the exact instant that each sweep starts, the instant at which each marker pip is generated bears a constant relation to the instantaneous voltage or current value of the sweep, thus in the case of an E.S. tube, for example, the instantaneous value of the sweep voltage and the appropriate range marker can be locked together and the required visual illusion is produced.

But if the range mark generator *oscillates continuously* and its *output is gated out* to the C.R.T. at the start of each sweep, the range mark pips (being generated independently and at a much higher frequency) are not locked in any way to individual instantaneous values of the sweep. They will not, therefore, appear on the screen to be fixed, unless the range mark generator oscillates at an exact sub-multiple of the repetition frequency. On the other hand, if the oscillation frequency of the range mark generator is made a sub-multiple of the repetition frequency, this ties down the magnitude of the sub-divisions produced on the trace to probably some odd number, which would be inconvenient for range calculation.

" Gateing "

In order to get over the above difficulty, a gate * is arranged so that two separate pulses are required to open it, *i.e.*, a repetition frequency pulse *prepares the gate* and makes it ready to open, but the *opening* or gateing is actually done by a pip from the range mark generator. Thus a pulse which fixes the repetition, or gate opening frequency performs a sort of " stand by " function and the pip from the fixed range mark generator (running continuously) immediately following receipt of the " stand by," performs the function of opening the gate.

* When the output of a free-running generator, for example, is admitted at given intervals under the control of some other device or circuit, to a third device or circuit, the output of the free-running generator is said to be gated. The intermediate circuit or gateing circuit, in fact, acts like a gate inasmuch as it admits the output of a free-running generator to a third circuit only when ordered to do so by a controlling signal, usually termed the " gateing out " signal.

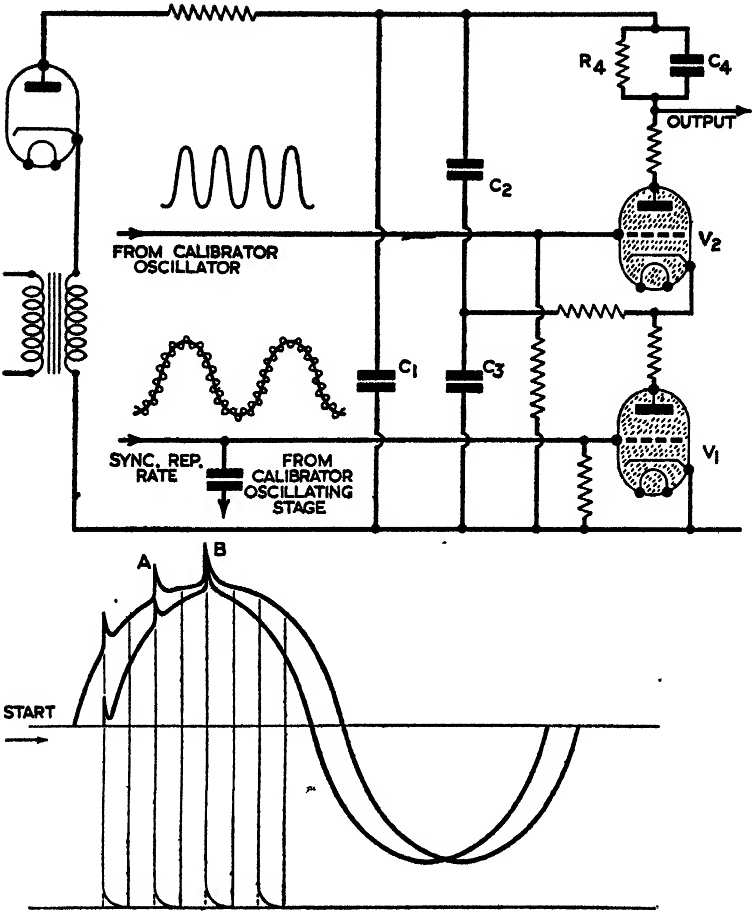


Fig. 210.—Circuit for "gateing out" the output from a free-running crystal-controlled range mark generator, also preparatory and "gateing out" signal waveforms when A.C. mains are used to indicate the desired P.R.F. for the system.

However, when the output of a range mark generator is gated in the above manner, it is obvious that it can no longer be locked to the p.r.f. of the radar system, as fixed by one of the other units of the system, and it follows therefore, that it is also impossible to lock the range mark pip to a definite instantaneous value of the time base sweep, consequently

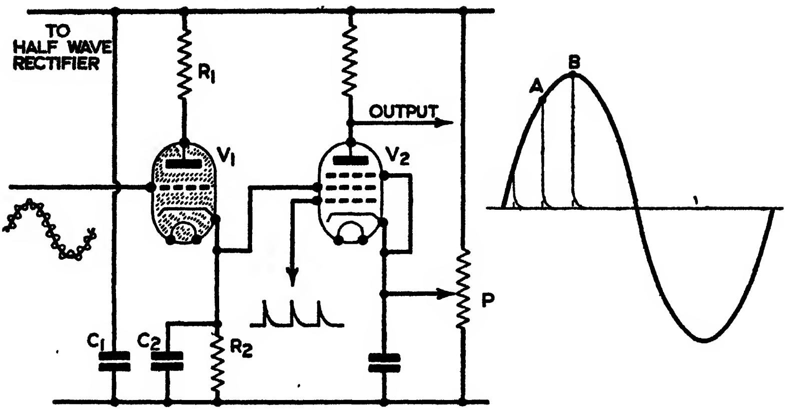


Fig. 211.—" GATEING OUT " CIRCUIT EMPLOYING ONE THYRATRON AND A PENTODE.

the *pips which open the gate* are now made the *synchronising or master control for the entire radar system.*

In effect, the pulse at repetition frequency now acts as a *guide* to synchronisation inasmuch as it selects the particular pip which is to open the gate and simultaneously fires the sweep generator, the transmitter and any other units which require to be synchronised to it.

In this manner the start of the sweep is always locked to the pip which opens the gate, and therefore all the other pips are also locked to instantaneous sweep values.

In these circumstances the calibrations on the trace are quite steady and free from " jitter."

Method of Synchronising the Radar System from a " Gateing Out " Pip

Circuits in which a master synchronising pulse is developed from a calibration pip at the instigation of some wave or pulse which sets the repetition frequency are shown in Figs. 210 and 211. It should be noted in both cases that the A.C. main supply (say 500 ~) is employed to guide the range mark generator and its associated circuits in fixing the frequency of the " gateing out " signal as nearly as possible to give a p.r.f. for the system of approx. 500 ~. It does not follow, however, that this is always the case. The required repetition frequency may be generated by an independent oscillator at some more convenient frequency in which case its output would be used to guide the range mark circuits in fixing the p.r.f.

In Fig. 210 two thyratrons are employed.

The input to V_1 is an A.C. wave from the mains with the differentiated output of the range mark generator superimposed upon it.

The input to V_2 is the output from the range mark generator before differentiation.

H.T. is applied by a half-wave rectifier which charges C_1 during the negative half cycle of the supply. C_2 and C_3 are potential-dividing condensers applying the rectified voltage in correct proportion to V_1 and V_2.

V_2 is biased back by an amount equal to the voltage across C_3. The voltage of the grid input, though large, is insufficient to fire V_2.

The input to the grid of V_1 is approximately in phase with the H.T. supply. E_g V_1 is negative during the negative half-cycle and therefore non-conducting. C_1 is charged to the peak mains voltage until the beginning of the positive half pulse.

Consideration of Fig. 210 shows that the output from the range marker circuit, when superimposed upon the A.C. wave, will occur at different points in each successive A.C. cycle, because the time between pips is much less than the A.C. time period and generally it is not likely to be an exact submultiple. In consequence of this, the instantaneous voltage compounded from the differentiated range mark output and the A.C. voltage will not have the same value in each cycle, therefore the thyratron will not fire at the same point in each consecutive cycle.

In the first cycle it will fire at A and the next cycle it will fire at B. In fact, over a succession of cycles it will fire at points ranging between A and B.

Any Synch. pulse developed in these circumstances must cause the range markers to jitter on the trace—oscillating between values A and B.

However, when V_1 fires (in the circuit shown in Fig. 210) some time at the commencement of the positive pulse it does not fire at the same point in each cycle. Nor does it directly fire V_2. It *prepares* V_2 so that it can be fired by the *next* positive-going pulse from the range mark generator output. It is the pulse generated in the anode of V_2 by the gateing out signal which fires the sweep generator, the transmitter, etc.

In this way the oscillation that fires V_2 locks a range mark pip to the commencement of each sweep of the time base, and since the pips are always spaced the same distance apart, they must appear on the screen always in the same relative position to given instantaneous values of the time base sweep.

Sequence of Operation

The sequence of operation of V_1 and V_2 circuit is as follows :—

V_1 fires and C_3 discharges through it, thus removing the negative bias from V_2.

V_2 is now *prepared* to be fired by the direct output from the range mark generator. This voltage is large and firing takes place at a perfectly definite point in the calibrator cycle.

The first result of the firing of V_1 and V_2 is the production of a large negative pulse at the anode of V_2, due to the discharge of C_1 into C_4.

The back edge of the pulse is formed by the discharge of C_4 through R_4. The time constant C_4R_4 which is of the order of $\frac{1}{2}$ microsecond determines the pulse width.

Alternative Method Using a Pentode

Fig. 211 is an alternative method using a pentode instead of the second thyratron.

As long as V_1 is non-conducting, the screen of V_2 is at earth potential and V_2 does not conduct. When V_1 conducts, its cathode, and therefore the screen of V_2, rises to about +250 volts and V_2 is fired by the next range marker oscillation. Time period C_2 and R_2 is made sufficiently long to prevent the gate opening *too quickly*, so that the *calibrator pulse (on the A.C. pedestal) responsible for firing V_1 does not itself get through to fire V_2.*

C_2 discharges through R_2 and the anode-cathode path of V_2 when it conducts.

A potentiometer adjusts the negative bias on V_2 and is set so that even after the screen of V_2 has reached its working voltage nothing passes through the valve until it is fired by a direct calibrator pip.

The output is smaller than for Fig. 210 but is adequate for firing a multivibrator.

It is possible to use a single pentode with negative bias and a steep-fronted switching pulse applied to its suppressor. In this case, the control grid is also negatively biased and sinusoidal calibrator oscillations are applied to it of such an amplitude that only the tips of the positive half cycles can operate the valve.

As soon as the switching pulse reaches the suppressor, the next calibration pulse to arrive passes through the valve and produces a narrow negative spike at the anode.

CHAPTER XVI

VALVES AND VALVE CIRCUITS FOR VERY HIGH AND ULTRA-HIGH FREQUENCIES

In previous chapters it has been shown that radar sets for accurate ranging and capable of good target separation must employ a narrow beam. This means that the aerial must have highly directional properties.

All other things being equal, directivity varies with carrier frequency. In other words, the higher the carrier frequency the more directional can an aerial of given bulk be made. Thus, in order to obtain highly directive qualities from an aerial of reasonable dimensions, it is necessary to use a carrier wave of some suitable value in the ultra-high frequency range.

It can be said, therefore, that the carrier frequency is determined mainly by the service which a particular radar set is required to perform.

Relation of Carrier Frequency to Range and Accuracy

For example, the main requirements of a radar set designed for long-range searching are high power output for the transmitter and coupled with a sensitive receiving system. Alternatively, sets designed to furnish detailed and accurate data at comparatively short ranges must scan the target area in greater detail, whilst the display must lend itself to more accurate measurement. For this purpose narrow beams are essential. Unless the radar set is to be fitted at a fixed site ashore, consideration of weight, space available and windage limit the maximum dimensions of the aerial. Consequently, in order to obtain the necessary degree of directivity with an aerial of limited dimensions an ultra-high-frequency carrier must be employed.

The carrier frequency employed by long-range searching or warning sets may be as low as 100 megacycles, whereas the carrier frequencies used for short-range highly directional sets are generally of the order of 10,000 megacycles or more.

Since the significance of these ultra-high frequencies is apt to be lost because of their magnitudes, it is more impressive as well as convenient to designate them by their corresponding wavelengths. Consequently frequencies of 3,000 megacycles or more are generally referred to by their wavelength equivalent, *i.e.*, 10 cm., etc.

The transition in technique from metre waves to centimetre waves is gradual until the region of 10 cm. is reached, when an abrupt change takes place, because it is in this region that the use of wave-guides becomes a practical proposition.

In the broadest sense, however, standard radio technique, as applied to transmitters and receivers, approaches a first limit in the region of 100/200 megacycles.

Valve Limitations

As the frequency rises, small values of inductance become important because X_L attains comparatively large values. For the same reason X_C, even for very small values of C, approaches a very low value, such that small capacities become almost a short circuit.

When these effects of increased frequency become important, a limit is imposed on the use of standard valves by reason of inefficiencies due to inter-electrode capacity, inductance of internal leads, growing circuit losses and transit time.

Limitations to the use of standard valves for receiver circuits occur in the region of about 200 megacycles.

Acorn triodes and other special types must be used exclusively for receiving purposes for frequencies from about 200 megacycles to 400 megacycles or more.

Limit for R.F. Amplification

The limit for R.F. amplification in receivers is determined by another factor. Noise generated by the valve itself increases with frequency, consequently when the ratio of the level of noise/signal reaches unity the useful limit of R.F. amplification is reached. For frequencies above this ratio, R.F. amplification is then omitted altogether from receiver circuits.

Special Valves for U.H.F.

Special valves, such as the lighthouse and klystron types, are used exclusively in local oscillators for frequencies above 600 megacycles.

A crystal mounted in a suitable resonant cavity, but without any R.F. amplification, is usually employed as a mixer at frequencies in the region of 2,000/3,000 megacycles.

Standard valves in specially arranged circuits may be used for transmitters up to about 200 megacycles, but whilst they may be used above this frequency, their efficiency drops away rather rapidly. Lighthouse valves may be used for small portable sets, but in general the magnetron is used exclusively for R.F. generators above 400 megacycles.

Components for U.H.F. Circuits

The physical sizes of external capacities and inductances used in transmitter tank circuits tend to become very small at frequencies in the region of 200 megacycles, and it is therefore convenient and efficient to substitute shorted quarter-wave lines * in the tank and oscillating circuits of transmitting generators, in place of the conventional capacity and inductance elements used at lower frequencies.

As frequencies are increased towards 3,000 megacycles (10 cm.), the efficiency of shorted quarter-wave lines in R.F. generators decreases.

This is due to increased losses in skin effect, eddy currents, dielectric losses, radiation, etc., but, fortunately, around this region, the dimensions of wave-guides and cavity resonators reduce to a practical size, enabling them to be substituted for quarter-wave sections of line, with the attendant advantages of high Q, low skin effect, eddy currents, dielectric and radiation losses.

From the foregoing it is clear that when frequencies above 100–200 megacycles are employed, important new considerations arise in respect of :—

(*a*) Valves.
(*b*) Circuit design.

The effect on aerials and aerial systems is discussed in Chapter XXII, whilst the theory of transmission lines, wave-guides and cavity resonators is dealt with in Appendices I, II and III respectively.

The ensuing portion of this chapter is, therefore, confined to a discussion of the limiting factors for valves and valve circuits when frequency increases and the changes that must then be made.

* See Appendix I, " Transmission Lines."

Frequency Limitations for Valves

The main factors limiting the output and efficiency of valves (Fig. 212) with increase of frequency are:—

(*a*) Inductance and capacity associated with the valve electrodes.

(*b*) Increased radio frequency losses.

(*c*) Electron transit time.

At ultra-high frequencies the combined effect of the electrode capacities on the generated frequency is equivalent to an increase in capacity of the tank condenser C by an amount C_i.

$$C_i = C_{ga} + \frac{C_{ak} \times C_{gk}}{C_{ak} + C_{gk}}.$$

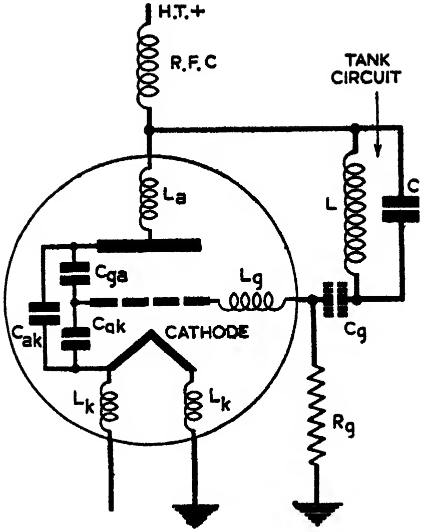


Fig. 212.—EMPHASISED INTER-ELECTRODE CAPACITY AND INTERNAL LEAD INDUCTANCE IN DIAGRAMMATIC FORM.

At lower frequencies when C is large compared to C_i, this effect is negligible, but as the frequency is raised the capacity of C must be made smaller and therefore the proportion which C_i bears to C grows with increase of frequency.

The lead inductances are in parallel with the inductance L of the tank circuit, and this lowers the value of the upper frequency limit for the valve. For example, a wire 0·040 in. diameter and 4 in. long has inductance of approximately 0·1 microhenry. At a frequency of 1 megacycle $X_L = \cdot 63$ ohms, but at 100 megacycles it becomes 63 ohms. Also the inductance of the cathode lead is common to grid and anode circuits and so provides negative feedback, which reduces the amplitude of oscillations.

Because of the increased importance of the reactance of internal C and L elements of the valve at the higher frequencies, the value of L and C of the tank circuit must be decreased until a point is reached where further increases in frequency necessitate reduction of the external capacity of the tank circuit to nearly zero and the inductance shrinks to a straight conductor shorting the anode and grid terminals.

The frequency corresponding to this condition is the resonant

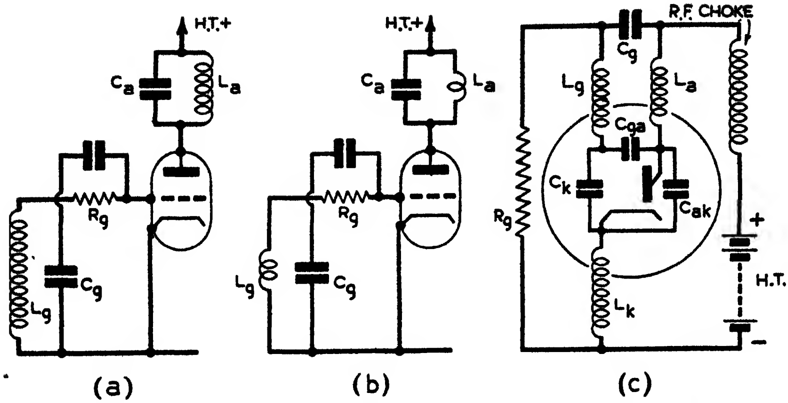


Fig. 213.—INCREASE OF IMPORTANCE OF INTERNAL L AND C ELEMENTS FOR THE HIGHER FREQUENCIES.

(a) Basic tuned anode tuned grid oscillator.
(b) Reduced inductance to permit higher frequency of oscillation.
(c) Lead inductance and internal electrode capacity of valve form a tuned circuit and determine resonant frequency of the valves. Consequently, to make the valve operate at its highest limit all capacity and inductance associated with H.T., R.F. choke C_g, R_g and leads must vanish.

frequency of the valve, and is its upper frequency limit. Fig. 213 (a) and (b) illustrate shrinkage in value of the external L and C elements for frequency increases.

In Fig. 213 (c) C_g has negligible reactance at the frequency of oscillation, and therefore L_g of the grid lead is in series with L_a (anode lead).

The resonant circuit shown in Fig. 213 (c) is determined by the inductance of anode and grid leads in parallel with a combination of the internal capacities of the valve.

Reducing the size of the electrodes and spacing them further apart, would reduce inter-electrode capacity, but this would necessitate the use of very high voltages, also the transit time of the electrons from cathode to anode would be increased.

It is known that if all the dimensions of a valve are reduced by a factor " n," the characteristics of the valve, the operating voltages and currents are not changed, but inter-electrode capacity, lead inductance and transit time, however, all become n times smaller. This reduction in size has one undesirable feature, the power-handling capacity of the valve is decreased.

In addition to reducing the physical size of the valve, lead inductance may be decreased by making the leads of large diameter.

In order to shorten the length as much as possible and avoid the capacity effects of the base, leads are usually brought out directly through the envelope in ultra-high frequency valves, in which case the normal form of base is eliminated.

Acorn types of valves have all these characteristics and can oscillate at 750 megacycles under suitable conditions. They are primarily used, however, in amplifiers and power oscillators, in which case their upper frequency limit is much lower—about 400 megacycles.

Special triodes are obtainable that function at much higher frequencies. A limit of 1,700 megacycles is claimed for one such type.

Losses

Radio-frequency valve and circuit losses increase as the frequency is raised ; these are due to :—

(*a*) Increasing skin effect.

(*b*) Greater condenser charging currents, therefore increased I^2R losses.

(*c*) Eddy currents due to adjacent conductors.

(*d*) Dielectric loss chiefly in the glass parts of the valve.

(*e*) Loss by direct radiation from the circuit.

All these losses cause the loading in the tuned circuit to be increased. This is reflected as a decrease in the Q of the circuit. Unfortunately, the losses increase at the expense of the permissible useful load, causing poor circuit efficiency and less useful output.

Skin effect forces the current to flow in a thin layer on the surface of the conductor. The depth is given for copper by :—

$$d = \frac{2{\cdot}63 \times 10^{-3}}{\sqrt{f}}.$$

d = depth of penetration in inches.

f = frequency in megacycles per second.

Thus the higher the frequency the thinner will be the layer in which the current flows. This is equivalent to reducing the cross-sectional area of the conductor. Therefore I^2R losses increase with frequency.

Skin effect is reduced by using conductors of large diameter. Large-diameter leads for valves and circuits reduces resistance and inductance. A further reduction in resistance can be obtained by silver plating.

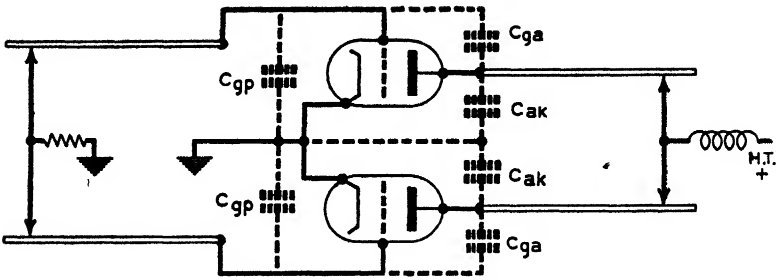


Fig 214.—TRANSMISSION LINE OSCILLATOR.

Demonstrating the use of shorted $\frac{\lambda}{4}$ lengths of transmission line to obtain a Q which is much higher than that which could be obtained with coils and condensers. Note from their phantom arrangement, reduction of total internal capacity brought about by push-pull arrangement.

Since X_c becomes small for inter-electrode capacities at ultra-high frequencies, the magnitude of the charging currents becomes large ; a current of 60 amps. is not unusual for large valves. These large currents contribute nothing to the output but are sources of serious I^2R losses. Another effect of large surface currents is that of excessive heating at the glass seals of the valve envelope, resulting in failures.

It is clearly desirable to keep all values of capacity as low as possible, and for this reason ultra-high frequency circuits are usually designed to have the maximum possible amount of inductance for any particular frequency. Thus, in the case of U.H.F. oscillators, the minimum amount of capacity is needed in addition to the inter-electrode capacity to resonate with the inductance.

Excessive charging currents can be reduced by :—

(*a*) The use of special low inter-electrode capacity valves.

(*b*) The use of push-pull circuits.

Since the inter-electrode capacities are in series in a push-pull oscillator, the effective capacity shorting the resonant circuit is half that of a single valve. The charging currents are correspondingly smaller (see Fig. 214).

Losses from radiation and eddy currents result from incomplete cancellation of electromagnetic fields associated with conductors.

Close spacing reduces these losses, but its application is limited by the R.F. voltage to be employed. Also spacing below certain minima causes an *increase* in R.F. resistance.

Coaxial lines instead of open wires are therefore preferred. Radiation and eddy current losses are almost entirely eliminated by the shielding effect of the inner conductor by the outer conductor at the frequency limit for triodes.

Dielectric losses are reduced by eliminating the valve base and by making connections directly to internal leads by bringing the conductors through the glass at points where the voltage is a maximum.

Use of $\frac{\lambda}{4}$ Lengths Shorted Line in place of Coils and Condensers

The Q of all unloaded circuits must be as high as possible, and for this reason sections of short-circuited $\frac{\lambda}{4}$ transmission lines are preferred to the use of coils and condensers.*

Having regard to skin effect losses, the Q of a large-diameter transmission line section is much higher than that of an inductance coil. A further advantage is obtained by the use of shorted sections of $\frac{\lambda}{4}$ lines, because the valve leads may be regarded as a part of the transmission line sections, and may therefore usefully form part of the oscillatory circuit. From an examination of Fig. 214 it will be seen that as a direct result of using shorted sections of $\frac{\lambda}{4}$ lines, in place of the normal condenser and inductance combination, the inter-electrode capacities as well as the lead inductance of the valves are incorporated in the oscillating circuits.

Replacement of $\frac{\lambda}{4}$ Sections of Line by Wave-Guides

The Q of a wave-guide is again much higher than that of a similar length of transmission line, and losses due to skin effect may be still further reduced by the employment of wave-guides.

It is shown in Appendix II that the minimum dimensions are determined by the applied frequency, therefore it is only at frequencies in the 3,000-megacycle region and above that the

* In Transmission Line Theory (Appendix I) it is shown that a pair of lines $\frac{\lambda}{4}$ in length, shorted at one end, behave like a circuit containing lumped values of inductance and capacity at resonant frequency, *i.e.*, for the wave-length corresponding to four times the length of one of the pair of lines.

minimum dimensions of wave-guides become small enough to render their employment a practical proposition. Consequently, for the transmission of energy at frequencies above that corresponding to about 10 cm. wave-guides are preferred to shorted sections of $\frac{\lambda}{4}$ transmission line.

In both cases effects similar to those exhibited by combinations of normal coils and condensers can be produced and controlled.

Frequency limitation of valves due to transit time has already been dealt with in Chapter III.

Summary.

To summarise, therefore :—

Valves. Inter-electrode capacity, lead inductance, resistance, dielectric losses and transit time may all be reduced by suitable design sufficiently to permit the valve to be used as an amplifier, mixer or oscillator in suitable circuits for ultra-high frequency work. Standard triodes reach the upper frequency limit for R.F. receiving amplifiers at about 200 megacycles, but other special forms of triode may be used for amplification at much higher frequencies. The upper frequency limit for R.F. amplification is reached when signal/noise ratio = unity. Above this point on the curve of ascending frequencies special triodes are available as oscillators, and these supplemented by klystrons, etc., cover the entire range of frequencies in which local oscillators are employed.

Reception. The limit of R.F. amplification is reached when the signal to noise ratio reaches unity. This matter is dealt with fully in Chapter XIX, " Receivers."

Transmission. Triodes arranged in suitable circuits may be used effectively up to about 200 megacycles or more. A limit is reached at about 400 megacycles, in which region efficiency is very low. Lighthouse and similar types of valves may be used for low-power portable sets at much higher frequencies. In general, R.F. for transmitters at frequencies above 400 megacycles is generated almost entirely by the resonant magnetron.

Special Valves for Receiving and Transmitting

Descriptions of special valves for receiving and transmitting now follow, together with typical circuits in which they may be employed.

More detailed circuit information will be found in Chapters XVIII and XIX on Transmitters and Receivers respectively. Appendices I, II and III should also be referred to for the theory of transmission lines, wave-guides and resonant cavities.

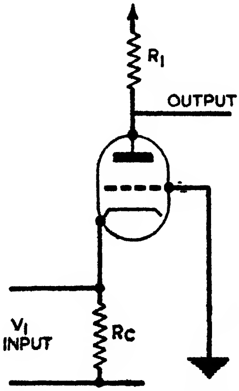


Fig. 215.—BASIC CIRCUIT OF GROUNDED GRID AMPLIFIER.

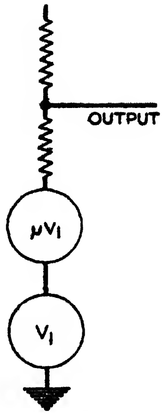


Fig. 216.—EQUIVALENT CIRCUIT OF GROUNDED GRID AMPLIFIER.

The Cathode Input Amplifier or Grounded Grid Amplifier

The input to the amplifier is connected across a resistance in the cathode circuit, the grid is tied down to earth and the output is taken in the normal way from across the anode load.

The most important advantage is its low impedance input, which makes it possible to connect it to the end of a long cable in such a way that the cable is terminated by a resistance equal to its own surge or characteristic impedance.

Another advantage is the low input-output capacity.

The circuit is shown in Fig. 215.

R_1 = anode load, R_c = cathode resistance, R_a = anode impedance and μ = amplification factor.

Stage Amplification

If V_1 volts are applied between cathode and earth this is the same as applying $-V_1$ volts to the grid.

The equivalent circuit is shown in Fig. 216.

The output voltage is $\dfrac{R_1}{R_1 + R_a}(\mu V_1 + V_1)$ and the stage gain is therefore $\dfrac{R_1}{R_1 + R_a}(\mu + 1)$, slightly higher than a normal amplifying stage.

If a battery V_1 is inserted between the cathode and earth, Fig. 216 shows that the change in current is $V_1 \dfrac{\mu + 1}{R_1 + R_a}$.

Since impedance is = to the ratio $\dfrac{E}{I}$

then $$\frac{V_1}{V_1 \dfrac{\mu + 1}{R_1 + R_a}} = \frac{R_1 + R_a}{\mu + 1}.$$

If R_1 is small compared with R_a and μ is large, this is nearly equal to $\frac{R_a}{\mu}$ or $\frac{1}{g}$ where g is the mutual conductance of the valve.

R_c has been omitted, but the combined effect of R_c and the valve is that the input impedance is equal to that of R_c and $\frac{R_1 + R_a}{\mu + 1}$ in parallel.

In the case of a high-impedance pentode with g equal to 6 mA. per volt and $R_c = 200$ ohms, the input impedance is

$$\frac{R_c \times \dfrac{1}{g}}{R_c + \dfrac{1}{g}} = \frac{R_c}{1 + R_{cg}} = \frac{200}{1 + 200 \times \dfrac{6}{1000}} = 91 \text{ ohms.}$$

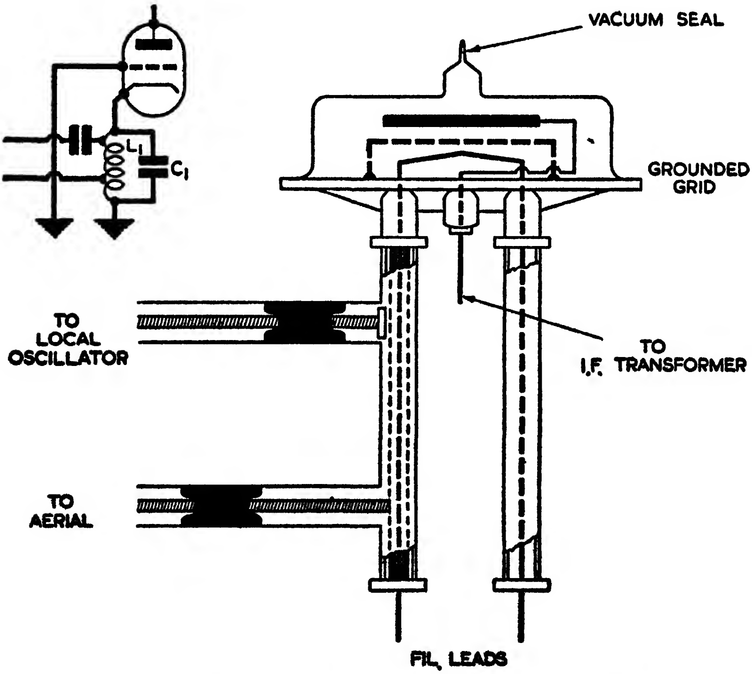


Fig. 217.—DETAILS AND EQUIVALENT CIRCUIT OF GROUNDED GRID TRIODE EMPLOYED AS A MIXER. INPUTS TO CATHODE *via* COAXIAL CABLE.

Low Cathode-Anode Capacity

Since the grid is earthed it acts as a screen between the input and output circuits. The anode capacity may be as low as ·05 picofarad. Fig. 217 shows a form of triode suitable for operation at very high frequencies in conjunction with a suitable oscillator as a mixer-amplifier. The equivalent circuit shows that it is functioning in a grounded grid circuit.

The input signal, *i.e.*, R.F. from the aerial, is coupled into the cathode line (a coaxial cable) at the point where an exact match with the cable is obtained. The output of the local oscillator is, however, deliberately mismatched to avoid heavy loading on the local oscillator and to prevent the R.F. signal from being swamped by the oscillator voltage; also the mismatch prevents the R.F. signal entering the line to the oscillator. This line is usually adjustable in length so a match can be made at the best point for an optimum resultant.

These frequencies, as well as the beat frequency, appear on the inner conductor of the line in the cathode circuit. Thus cathode volts vary accordingly with respect to the grid. This voltage variation in the cathode (which is directly heated) produces the same effect as varying the grid voltage, assuming that the cathode voltage is now held constant.

Although the filament cathode has no D.C. connection between inner and outer conductors of the coaxial line, it is capacitatively coupled at the end, since only a small distance separates inner and outer conductors.

By grounding the grid of the mixer valve, the anode to cathode capacity is considerably decreased. Grounded grid triodes are frequently used as mixers, when a reflex klystron is used as the independent oscillator for heterodyning purposes. The klystron and reflex klystron are described in later paragraphs.

The "Lighthouse" Valve

A sketch of the so-called "Lighthouse" valve is shown in Fig. 218, together with its equivalent in terms of normal condensers and inductances.

This type of valve is often used in grounded grid circuits or in circuits using tuned concentric lines. The valve base and glass are so designed that the valve can be screwed into cylinders in a manner similar to that in which an incandescent lamp screws into its socket. The connections thus made complete the resonant circuits. Tuning is then effected by adjusting the shorting plungers between cylinders.

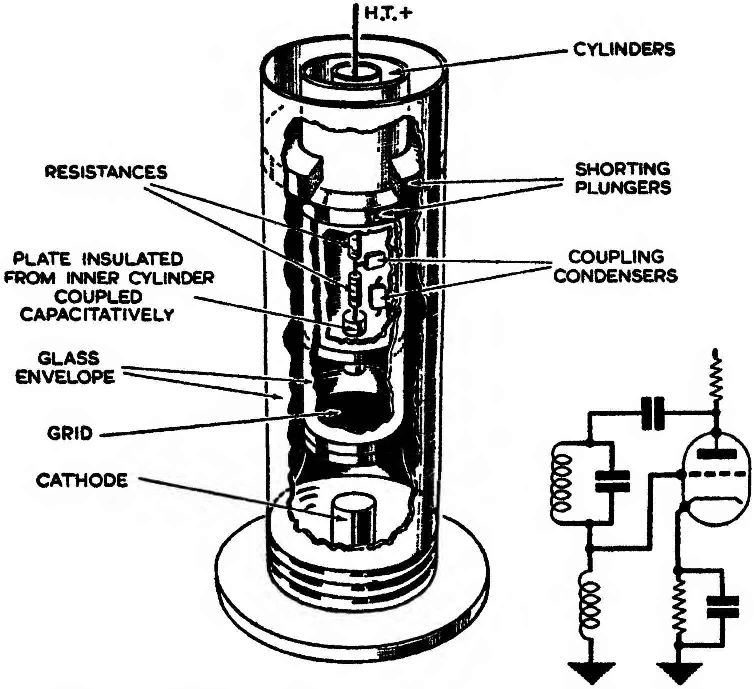


Fig. 218.—GENERAL CONSTRUCTION AND EQUIVALENT ARRANGEMENT OF THE "LIGHTHOUSE" VALVE.

Heater and cathode connections are made to a pin on a standard metal base. The anode connection is made to the cap and the grid connection is made to a ring, encircling the valve about half-way between the cap and the shell. This construction reduces inter-electrode capacity without loss of mechanical strength. Both cylinders are metal coaxial line sections.

R.F. Amplifier with Lighthouse Valve

A circuit using this valve as an R.F. amplifier with concentric lines is shown in Fig. 219 (a). The coaxial cable carrying the signal is connected directly to the cathode shell. The signal voltage across C_1, L_1 (tuned cathode circuit) is coupled to the cathode by capacity between cathode and shell. Since the grid is grounded, application of a signal to the cathode controls the anode current of the valve.

Bias is developed across R_1. The anode coil L_2 is tuned to

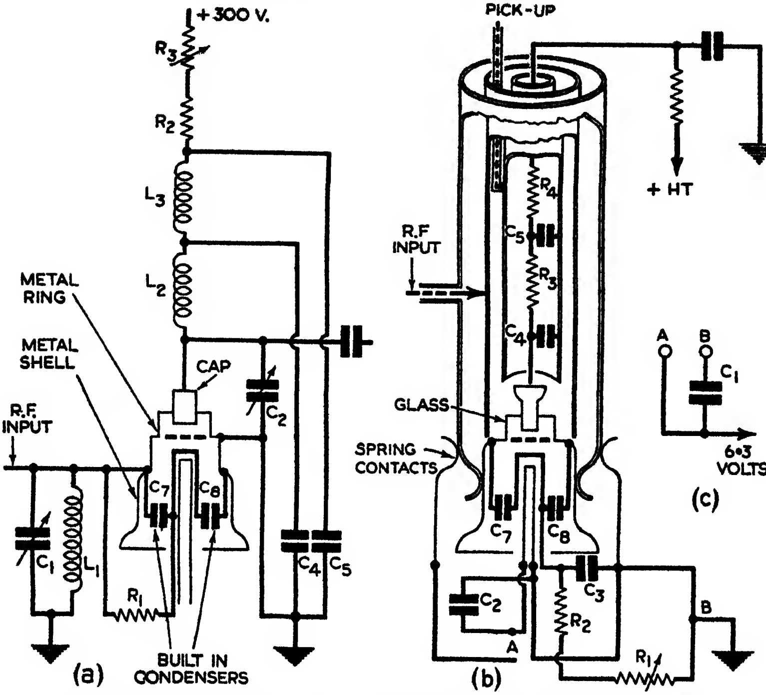


Fig. 219.—"LIGHTHOUSE" VALVE USED AS GROUNDED GRID AMPLIFIER WITHOUT CONCENTRIC LINES AND ALTERNATIVELY WITH CONCENTRIC LINES AND INPUT TO GRID LINE.

(a) Lighthouse valve used as a grounded grid R.F. amplifier.
(b) Lighthouse valve used as an R.F. amplifier. Input to grid line.
(c) Points A and B correspond to A and B (b).

resonance by C_2. L_3, C_4, C_5 (a filter) by-passes R.F. in the anode lead to earth, R_2, R_3 permit adjustment of the operating voltage to the proper value of 200–250 volts.

An alternative arrangement employing concentric lines is shown in Fig. 219 (b). R.F. input is applied directly to the grid line. The cathode is coupled to the outer conductor of the grid-cathode line by C_7 C_8, which are built in and by-pass radar frequency from the metal shell to the cathode. Grid bias is furnished by R_1, which is set to give correct anode current. Anode voltage is supplied *viâ* the filter C_4, C_5, R_4 and R_3. Coupling to the next stage is made by means of a pick-up between grid and anode lines. The grid-anode line may cause some degree of negative feedback. Oscillation is prevented by proper loading and by the location of the grid input tap.

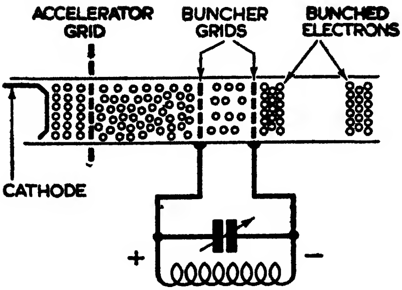


Fig. 220.—MODULATION OF THE VELOCITY OF AN ELECTRON STREAM BY THE ALTERNATING ELECTRIC FIELD BETWEEN A PAIR OF GRIDS CONNECTED TO A RESONANT CIRCUIT.

Velocity Modulation

The klystron and the reflex klystron depend for their action primarily upon velocity modulation of the electron stream.

The action of the reflex klystron will be developed from the following general considerations :—

Consider Fig. 220, electrons are emitted from the cathode and accelerated by the accelerating electrode at a uniform rate.

Upon leaving the accelerator grid the electron stream, moving at uniform speed, is passed through two closely spaced grids termed "buncher grids," each of which is connected to one side of a tuned circuit.

The tuned circuits and the grids are at the same potential as the accelerator grid.

The A.C. voltage which exists across the external tuned tank circuit causes the velocity of the electrons leaving the buncher grids to differ. This difference in velocity depends upon the time in the A.C. cycle at which they pass the grids.

An electron passing the centre of the buncher at the instant the A.C. (R.F.) is zero leaves the buncher at the same velocity as that with which it entered the buncher.

The position electrons occupy are plotted against time in Fig. 221 at A, E and A′. The slope of these lines represents the velocity of the electrons.

Electrons which pass the centre of the buncher a few electrical degrees earlier than the point of zero voltage as at C and D leave with reduced velocity, since the decreased voltage of the buncher slows them up.

Electrons that pass a few electrical degrees after the instant of zero voltage as at F and G leave with increased velocity since the voltage of the buncher is now higher than that of the accelerator grid.

If the space beyond the buncher is field free, the faster electrons F and G will catch up with electron E that left with unchanged velocity and the slower electrons C and D lagging behind will draw near to E.

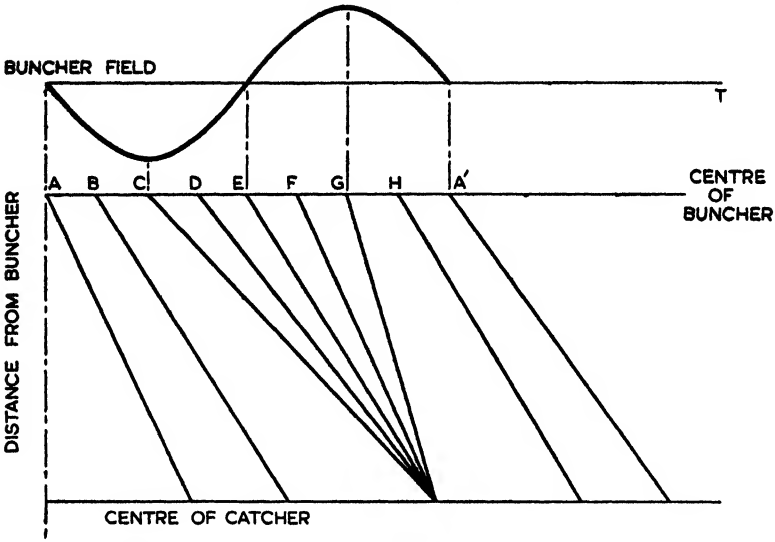


Fig. 221.—ELECTRON BUNCHING (APPLEGATE DIAGRAM).

At some point beyond the buncher grids electrons C, D, E, F and G will be close together in a group.

Another electron as at A which leaves the buncher half a cycle later than E and its neighbours draws away.

As a consequence of this action the electron stream down the tube consists of bunches of electrons separated by regions in which there are few electrons.

These bunches are allowed to pass through a similar set of grids called catcher grids coupled to another oscillating circuit (Fig. 222 (a)).

When each bunch of electrons reaches the first grid of the catcher, the field is such that it slows down the electrons and thus absorbs energy from them.

Note. Since the energy of the moving electron is $\frac{1}{2}mv^2$ and its mass is constant, a decrease in velocity means that its energy is less and, therefore, some of its energy must have been transferred or transformed. Similarly, when an electron is accelerated by an electric field its energy $\frac{1}{2}mv^2$ (m = mass, v = velocity) is increased. In this case it has absorbed energy from the electric field.

By the time the retarded electrons—retarded, say, by the first grid of the catcher—reach the second grid of the catcher, the relative grid potentials have reversed. This is because

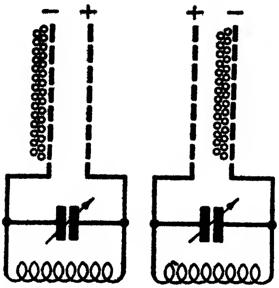


Fig. 222 (a).—CHANGE OF CATCHER GRID POLARITY WITH OSCILLATOR.

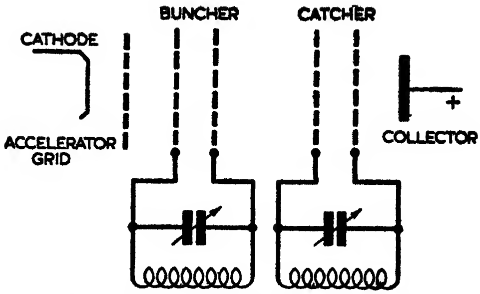


Fig 222 (b) —KLYSTRON SCHEMATIC DIAGRAM.

the spacing is such that it takes the group of electrons approximately half a cycle (in time) to pass from one grid to the other. Therefore the second catcher grid also slows down the electrons, thus absorbing more energy.

Hence, in delivering energy to the tuned circuits of the catcher grids the speed of the electrons is greatly reduced. After passing this second set of grids the spent electrons may be removed from the circuit by a positive collector plate.

Action of the Klystron and Reflex Klystron

The reflex klystron is a development of the klystron. The conclusions reached in the foregoing general analysis can be applied to the klystron proper. A description of the action of the reflex klystron will follow.

Fig. 222 (b) illustrates the general arrangement of elements in a klystron, and this conforms substantially with the general principles of velocity modulation previously described.

If the output from the catcher is fed back into the buncher in the right phase and relationship, the valve acts as an oscillator. The overall requirement for amplification is, that the energy required for bunching must be less than the energy delivered to the catcher. This amplifying action is made possible by the fact that electrons pass through the buncher in a continuous stream but through the catcher in definite bunches. The fundamental principle of operation is, therefore, based upon modulation of the velocity of the electron stream.

Since a continuous stream of electrons enters the buncher, and experience the effects of the alternating fields of its grids, the number of electrons accelerated in one half cycle is exactly equal to the number decelerated by the other half cycle. Thus

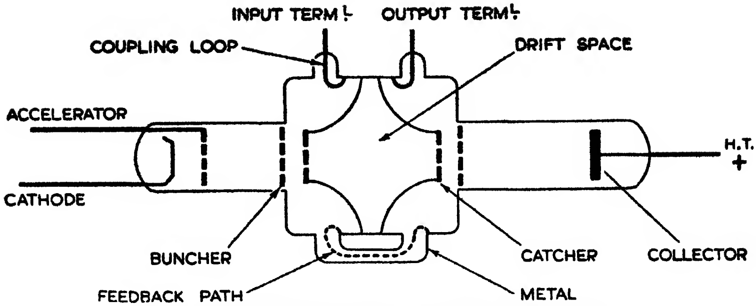


Fig. 223.—Klystron with cavity resonators.

the nett amount of energy exchanged between electrons and the buncher grids must be zero, if there are no losses in the oscillating circuit of the buncher. Moreover, an output is to be taken from the buncher oscillating circuit, consequently the energy fed back by the catcher to the buncher must exactly make good circuit and load losses if the valve is to oscillate.

The conditions at the catcher are such that the continuous electron stream has been broken up into groups. These groups are spaced apart exactly one half cycle, also the phase of the A.C. voltage at the catcher must be such that groups or bunches of electrons enter the catcher grids only during the decelerating half cycle. In this case, energy is absorbed by the catcher from each bunch. The nett result is that enough total energy is absorbed by the catcher to enable it to feed back in phase to the buncher a sufficient amount of energy to make good :—

(*a*) Circuit losses in both buncher and catcher.

(*b*) The load taken from the buncher.

In this way continuous oscillations are started and sustained at the frequency of the tank circuit of the buncher.

Practical Application

The device may be used as an amplifier, oscillator or mixer. When used at high frequencies the resonant circuits are usually resonant cavities, the grids being connected to each side of the cavity.

Resonant cavities are small and highly efficient (see Appendix III). The cavities are sometimes constructed internally, but in other cases they are external.

Tuning of the cavity is performed by screwing plugs into the periphery. Adjustment is critical.

The energy is coupled into or out of the valve by means of single-turn loops as shown in Fig. 223.

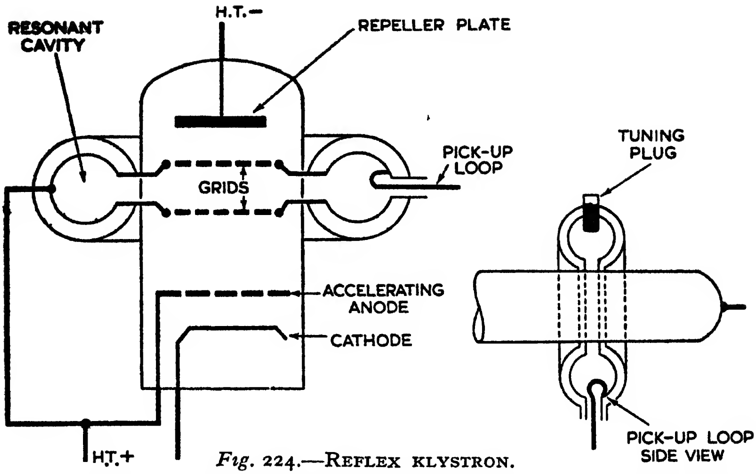


Fig. 224.—REFLEX KLYSTRON.

This is a vertical section. The resonant cavities are formed by a copper tube which encircles the envelope.

A simplified form known as a reflex klystron, is generally used for oscillators (Fig. 224), particularly in U.H.F. reception.

In this case one set of grids together with their associated cavities are used to perform the double function of bunching and catching, and the positive collector is replaced by a highly negative repeller electrode.

By proper adjustment of the negative voltage (which is critical) on the repeller, electrons, which have passed through the buncher field may be made to return again through the resonator grids in the right phase to deliver energy to the circuit, thus performing the feedback function.

Spent electrons are removed by the accelerator or by the grids of the resonator.

Energy can be coupled out of the resonator circuit in the manner shown in Fig. 224.

The operating frequency may be varied over small limits by the repeller voltage. This potential determines the transit time of the electrons between their first and second passages through the grids. Output is, however, affected to a greater extent than frequency by changes in repeller potential, since altering the transit time also changes the phase relations between returning electrons and the grid potentials.

Control of frequency is mainly exercised therefore by adjusting the volume of the cavity by means of plugs.

CHAPTER XVII

THE RESONANT MAGNETRON

The magnetron is now generally thought of as a device for generating R.F. power at frequencies corresponding to centimetric wavelengths. Indisputably this is, at the present time, its main function. It is difficult, indeed, to see how, without it transmission in the centimetre range of wavelengths would be either practicable or efficient.

Nevertheless, it must be remembered that the first magnetron appeared over twenty years ago. In its original form it was a poor competitor with the triode, as regards output, and it did not arouse much interest, outside the laboratory, until its potentialities as a generator of V.H.F. power were realised.

The upper frequency limit, for the original magnetron, is determined by the shortest time taken for an electron to accomplish one complete journey in its orbit in the cathode-anode space, *i.e.*, by its transit time. Its power output and efficiency are both very low indeed, since the R.F. power is derived from electrostatic induction of energy into the anode, due to the motion of electrons in the space between anode and cathode.

Development of the Magnetron

The original magnetron was followed by the split anode magnetron, which functions on the dynatron principle, by virtue of a static negative resistance existing between the two segments of the anode. This is brought about by internal forces, which direct electrons in transit to that segment of the anode which is at a lower potential with respect to the other at any given instant of time.

Whilst the split anode magnetron can be used for wavelengths in the centimetre range, its useful upper frequency limit is generally understood to be in the region of about 600 megacycles, where its power output compares favourably with that of a triode when operated at the same frequency.

The multi-resonator magnetron was developed in 1940 by Drs. Randall and Boot of the Admiralty valve group, working under Professor Oliphant at Birmingham University. This device uses a condition of velocity resonance of the electrons in the anode-cathode space, to excite and build up oscillations in resonant cavities having a high Q. The power to maintain these oscillations and to supply the load is derived from the source of H.T. by means of electrons which escape to the anode.

The resonant frequency is determined primarily by the geometry of the valve and the dimensions of the resonator elements. Since these are of the cavity type, frequencies corresponding to centimetric wavelengths can be generated with large power outputs and at high efficiencies.

Thus there are three steps in the development of the resonant magnetron. In chronological order they are: (1) the original cylindrical anode magnetron, (2) the split anode magnetron, and (3) the multi-resonator magnetron. However, in the light of present knowledge, it is thought by some authorities that the theoretical considerations of the original magnetron and of the split anode magnetron are really special cases of the resonant magnetron. In any case, it is convenient to consider them in chronological order and to develop the theory of the resonant magnetron from that of its forerunners.

THE EARLY CYLINDRICAL ANODE MAGNETRON

The original magnetron consisted of a filament mounted coaxially within a cylindrical anode, the assembly being enclosed in an evacuated glass envelope. The diode thus formed is so positioned between the poles of a suitable magnet, that the magnetic lines of force lie along the axis of the anode, parallel to the cathode (see Fig. 225).

The magnetic field may be provided by either a permanent magnet or electromagnet. The important points are that the magnetic field should be uniform over the entire anode-cathode space and the field density should not vary in strength from time to time.

When the magnetic field strength (H) is suitably adjusted and the cathode made very negative to the anode by the H.T. supply, oscillations can be generated in an orthodox oscillatory circuit, connected externally between anode and cathode (see Fig. 225 (d)). The frequency of the oscillations thus generated is determined by the time required for an electron to go from the cathode out into the cathode-anode space and back again to the cathode.

Since a magnetron of the above type, in the absence of an

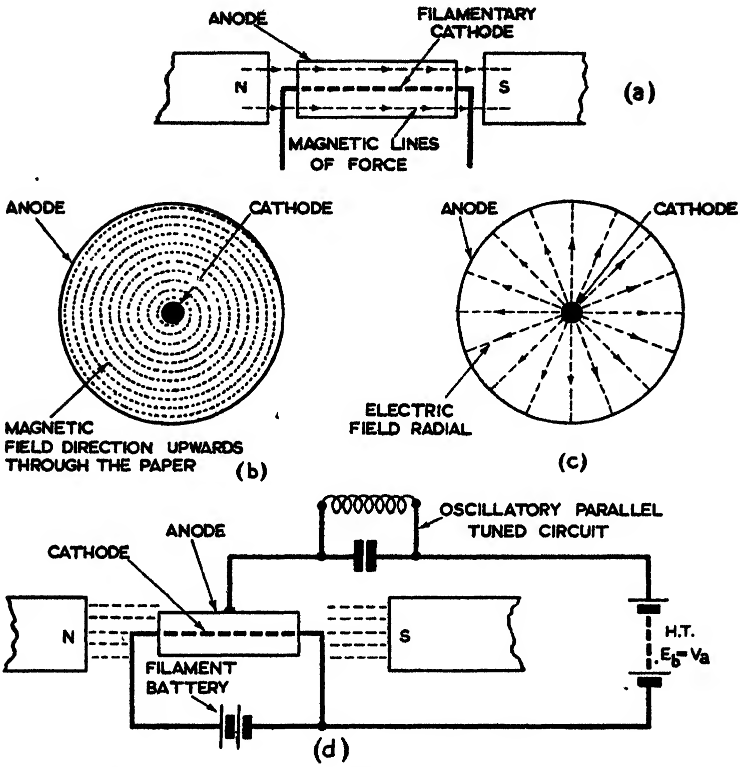


Fig. 225.—DETAILS OF ARRANGEMENT OF ORIGINAL CYLINDRICAL ANODE MAGNETRON, ITS MAGNETIC AND ELECTRIC FIELDS AND CIRCUIT ARRANGEMENT AS AN OSCILLATION GENERATOR.

(a) Disposition of cathode, anode and magnetic systems.
(b) and (c) End view of anode-cathode assembly looking along the axis, showing direction of magnetic lines of force, and direction of electrostatic field respectively.
(d) Circuit arrangement of cylindrical anode magnetron.

axial magnetic field, is basically a diode, electrons emitted from the cathode (when heated), are accelerated along radii towards the anode by the electric field set up by the anode voltage — V_a.

The kinetic energy $\frac{1}{2}mv^2$ (m = mass of the electron and v = its velocity) gained by an electron in its flight from cathode to anode is equal to the work done upon it by the electric field. This is eV_a where e is the numerical value of the charge of the electron and V_a the anode voltage. This quantity eV_a also represents the amount of energy given up as heat to the anode by any electrons which reach it.

When an axial magnetic field of uniform density is present, electrons in flight from cathode towards the anode must traverse it, and since an electron in motion is analogous to an electric current, an electron traversing the magnetic field must be subjected to a mechanical force acting upon it, just as though it were a current-carrying conductor moving through the magnetic field under similar conditions of relative direction for field and motion.

This means that electrons traversing a uniform magnetic field, and impelled by an electric field acting at right angles to the magnetic field are, as can be seen by applying Fleming's rule, subjected to a force acting at right angles to their direction of motion, and also at right angles to the direction of the magnetic field. It can be shown that in a uniform field the curvature is constant so that the motion of the electron becomes circular.

It can also be shown that the radius r of the circular orbit of an electron in a magnetic field is :—

(1) $r \propto \dfrac{\sqrt{Va}}{H}$ . . where H is the field strength.*

* $eV_a = \frac{1}{2}mv^2$ (1)
work done = energy gained

$Hev = \dfrac{mv^2}{r}$ (2)

e = charge on the electron
m = mass of the electron
V_a = anode-cathode potential
v = velocity of electron

$\dfrac{mv^2}{r}$ is the centripetal force, the centrifugal force being a pull of equal magnitude but in the opposite or outward direction.

$$\text{From (2) } r = \frac{mv}{He}$$

$$\text{Put } v = \sqrt{\frac{2eV_a}{m}} \text{ (from (1))}$$

$$r = \frac{1}{H}\frac{m}{e}\sqrt{\frac{2eV_a}{m}}$$

$$= \frac{1}{H}\sqrt{\frac{2mV_a}{e}}$$

$$= \sqrt{\frac{2m}{e}}\,\frac{\sqrt{V_a}}{H}$$

$$\therefore r \propto \frac{\sqrt{V_a}}{H}.$$

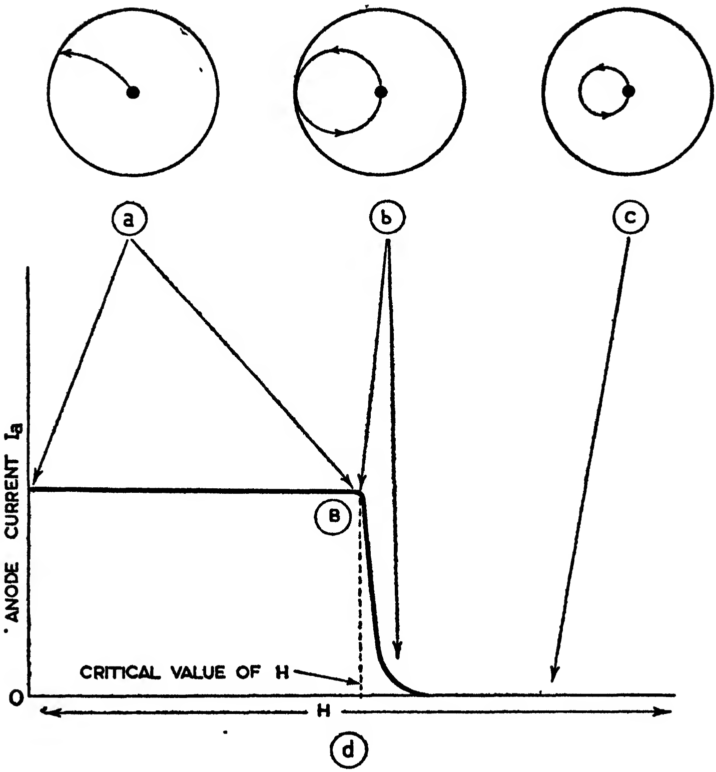


Fig. 226.—I_a–H CHARACTERISTICS OF CYLINDRICAL MAGNETRON WHEN E_a IS HELD AT SOME CONSTANT VALUE AND H VARIED.

(a) Very low value of H, I_a, E_a characteristic similar to diode without magnetic field.
(b) Effect of increased value of H shown as sudden drop of I_a (see (d).
(c) Effect on radius of electron orbit of increasing the value of H.
(d) I_a–H characteristic plotted from results obtained in (a), (b) and (c).

For some given values of V_a and H let the electron path be as shown in Fig. 226 (a). If V_a is held at some fixed value and the value of H gradually increased, the path of the electron shown in Fig. 226 (a) may be modified to that shown in Fig. 226 (b). In this case the radius of the electron orbit has been decreased, so that the electron just misses the anode or grazes it in returning to the cathode to complete its circular orbit.

The value of H relative to the fixed value of V_a which produces these conditions is :—

(2) $H = \frac{\sqrt{180V_a}}{d}$ . . . where H is in gauss
V_a is anode volts
d = anode diameter (cms.).

This *critical value* of H relative to V_a is the condition in which the original circular anode magnetron is operated.

In Fig. 226 (c) H has been increased beyond the critical value, so that electrons now travel outwards from the cathode and back to it again in a circular orbit of much smaller radius.

From a consideration of Fig. 226 (d) it can be seen that when for some fixed value of V_a, anode current I_a is plotted against values of H increasing from point O, I_a remains constant for increasing values of H until the *critical* value is reached at B. At this value of H the electron path just misses the anode and the anode current I_a falls abruptly towards zero. This is the " cut off " value of H for a particular value of V_a. The same effect is produced by holding H constant and varying V_a.

When the V_a/H ratio is adjusted for the critical value of H or I_a " cut off," the original circular anode magnetron generates oscillations in the external circuit joined between anode and cathode (Fig. 226 (d)).

This is made possible by the electric energy, induced electrostatically in the anode, by movements in the cathode-anode space, of electrons relative to the anode.

Briefly, the mechanism by which energy is transferred to the oscillatory circuit is as follows :—

(*a*) Electrons leaving the cathode are accelerated towards the anode by the electric field and *receive* energy from it, the H.T. being the primary source of supply for this energy.

(*b*) Due to the force exerted by the axial magnetic field on the electrons moving towards the anode, they are forced to follow a curved path, which takes them in a circular orbit away from the anode, past it and *back* to the cathode. Since the direction of motion during the *backward* journey to the cathode is in a direction *opposite* to that in which the electric field is urging them, work must be done by the electron upon the electric field, *i.e.*, some of the kinetic energy acquired by the electron during its outward journey is transformed and transferred *viâ* the field to the anode where, in effect, an alternating voltage is superimposed upon V_a by electron excursions.

The frequency of the R.F. energy induced in the anode in this manner is determined by the transit time of the electron, so that if the external parallel-tuned circuit is given the same time period as the transit time, R.F. energy can be built up in the tuned circuit and transferred to a matched load.

The power that can be delivered by this device is comparatively small and the upper limit of frequency which can be generated is limited by the transit time of the electron.

It can be shown that this time period is :—

$$T = \frac{2\pi m}{He} \quad . . . \text{ where m is the mass of the electron.*}$$

THE SPLIT ANODE MAGNETRON

The split anode magnetron followed the cylindrical anode magnetron. This device operates at higher frequencies and is capable of larger power outputs. Under suitable conditions it is said that magnetrons of this type can be designed to cover, in a series of ranges, the frequency band 100 kc/s. to 30,000 Mc/s. But efficiency and power output at frequencies above 600 megacycles fall off very rapidly.

The major differences between the split anode magnetron and the original or continuous anode magnetron are, that in the former, the anode is divided into two or more equal segments, separated by air gaps, also the external oscillatory circuit is now connected between the two anode segments, the centre point of the oscillating system being joined to the cathode as shown in Fig. 227.

As a direct result of splitting the anode, the R.F. fields due to oscillations excited in the parallel-tuned oscillatory circuit now appear across the gaps between segments and exercise

* Kinetic energy gained = work done by the electric field.

(1) $\therefore eV_a$ (work done by electric field) $= \frac{1}{2}mv^2$ . . . m = mass electron
v = velocity

$$(2) \quad \therefore \frac{mv^2}{r} = Hev.$$

$$(3) \quad r = \frac{mv}{He} \text{ or } \frac{1}{H}\frac{\sqrt{2mV_a}}{e}$$

$$(4) \quad r = \frac{\sqrt{V_a}}{H}.$$

If period for 1 cycle is T,

$$2\pi r = vT$$

$$T = \frac{2\pi r}{v} = \frac{2\pi m}{He}.$$

Also from (1),

$$\frac{1}{2}mv^2 = eV_a$$

$$\text{Velocity} = \frac{\sqrt{2eV_a}}{m}.$$

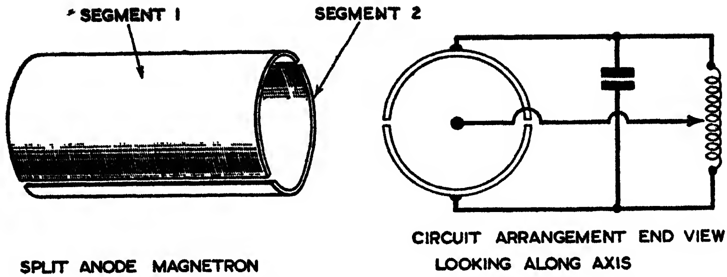


Fig. 227.—CONSTRUCTION OF SPLIT ANODE. ARRANGEMENT OF SPLIT ANODE MAGNETRON AS AN R.F. GENERATOR.

a profound effect in modifying the electric field between cathode and anode.

As a result of this modification, the forces acting on the electrons in the space between cathode and anode are changed and the path of the electron is modified accordingly.

The forces which now act upon the electron in the cathode-anode space are :—

(*a*) The radial electric field due to V_a.
(*b*) The axial magnetic field H.
(*c*) The R.F. field between segments, due to oscillations excited in the external oscillatory circuit. This latter force may be said to modify the electron path by virtue of the distortion which it produces in the radial field set up by V_a.

Note. In the following analysis the effect on a single electron only is considered in order to simplify the initial explanation. This does not take account of the space charge which exists in the region between cathode and anode. The effect of the space charge is fully dealt with in the resonant magnetron where the operation of this device cannot be satisfactorily explained without taking the space charge into account.

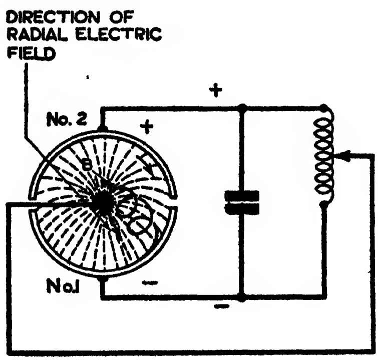


Fig. 228.—CIRCUIT ARRANGEMENT OF SPLIT ANODE MAGNETRON AS R.F. GENERATOR SHOWING THE POINT OF CHANGE OF CURVATURE OF ELECTRON PATH DUE TO MODIFICATION OF THE NORMAL RADIAL FIELD E_b BY E_{rf} AT TIME $\frac{t}{4}$. (DETAILS SHOWN IN FIG. 229.)

Assume that the values of V_a and H are adjusted so that an electron leaving the cathode and accelerated in the direction of the anode, is at some suitable radius of curvature, deflected, so that it tends to move round the cathode. Let it also be assumed that an electron in its first convolution passes by a gap between anode segments at $T = \frac{t}{4}$ (Fig. 228) when (due to oscillations in the external oscillatory circuit), segment No. 2 of the anode is at a higher potential than segment No. 1. In this case V_a at B (Fig. 229 (c)) is weakened so that r proportional to $\frac{\sqrt{V_a}}{H}$ is also less and consequently the radius of curvature of the electron will be shortened.

Since the electron is now turning away from segment 2, see point B, Fig. 228, against the force exerted by the electric field set up by V_a, work is being done on the electric field equal to the kinetic energy lost by the electron. In this condition the electron is retarded. Note that this retardation takes place on crossing the gap, the distortion of the radial field at this point being in effect a retarding field (see Figs. 228 and 229).

For a given aggregate of electrons the current strength

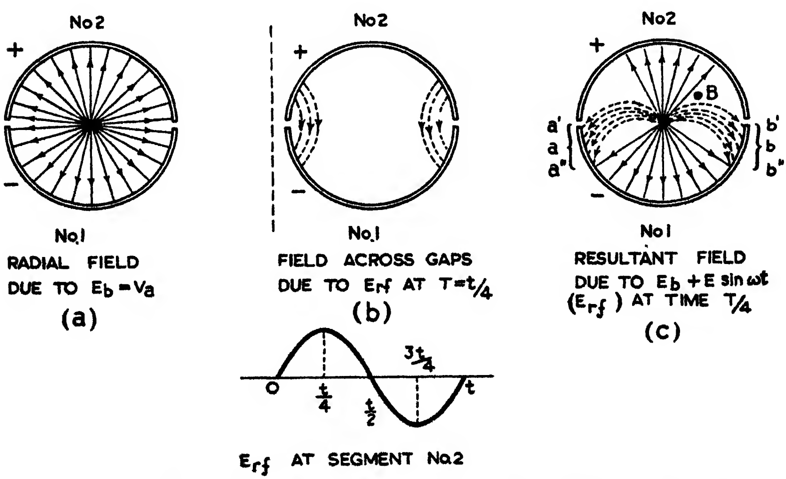


Fig. 229.—DETAILS SHOWING MODIFICATION OF RADIAL FIELD E_b BY THE FIELD OF E_{rf} AT TIME $\frac{t}{4}$ AND RESULTANT FORCE URGING ELECTRONS IN THIS PART OF THE FIELD AT THE MOMENT TO SHORTEN THE CURVATURE OF THEIR PATH. RESULTANT FIELD (C) EXAGGERATED.

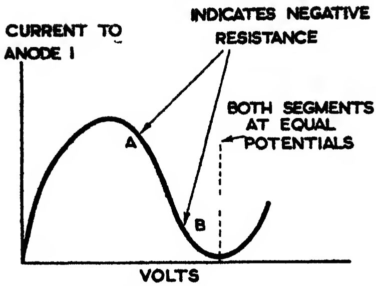


Fig. 230.—ANODE VOLTAGE—ANODE CURRENT CHARACTERISTICS FOR SEGMENT NO. 1 WHEN SEGMENT NO. 2 HELD AT CONSTANT POTENTIAL. VOLTAGE AT SEGMENT NO. 1 IS APPLIED AND VARIED FROM SOME SUITABLE EXTERNAL E.M.F.

depends upon velocity, and since the force acting upon a current-carrying conductor in a magnetic field is proportional to current strength, it follows that when the electron is retarded, due to the kinetic energy which it surrenders to the electric field, the force exerted upon it by H is diminished. Owing to this differential action, the electron must arrive at a point in its backward journey in the direction of the cathode where it will momentarily come to rest, immediately after which it will be acted upon by the electric field, to be accelerated once more towards the cathode. When in the next convolution the gap is again crossed the above sequence will be repeated, causing the electron to approach the anode along a spiral path, finally to land on the anode segment of *lowest potential*, where it gives up its balance of kinetic energy in the form of heat.

Thus when electrons enter the R.F. field in the vicinity of a gap between segments they are either accelerated and receive energy from the R.F. field or else they are decelerated, in which case they give up energy to the R.F. field. The determining factors are the instantaneous potentials of the anode segments and the position, relative to the segments, of the electron in its orbital path.

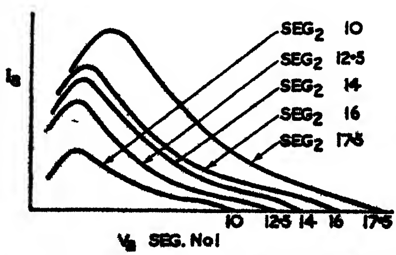


Fig. 231. E_a—I_a CHARACTERISTICS CURVES FOR FIXED E_a VALUES OF SEGMENT NO. 2, SEGMENT NO. 1 BEING VARIED THROUGHOUT ITS RANGE OF E_a FOR EACH FIXED VALUE OF E_a SEGMENT NO. 2.

For example, in Fig. 228, let the values of V_a and H be such that when both anode segments are at equal potential the electron just misses *both* segments. Now let the oscillations in the external resonant circuit be such that at time $\frac{t}{4}$ segment No. 1 is fully negative and segment No. 2 is fully positive. The R.F. field due to this condition will be superimposed on the radial electric field set up by V_a and will modify it by turning it through a small angle (see Fig. 229 (c)). The nett result is that electrons passing the top of their orbits in a clockwise direction are deflected from segment No. 2 towards segment No. 1. Electrons moving *towards* the tops of their orbits are turned back towards the cathode.

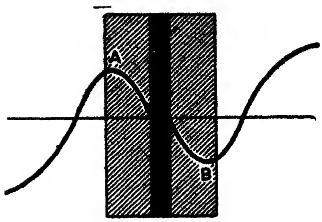


Fig. 232—CURVE OBTAINED BY PLOTTING $I_{a1}—I_{a2}$ AGAINST $V_{a1}—V_{a2}$.

In Fig. 229 (a) the electric field is proportionate to $E_b = V_a$. In Fig. 229 (b) the electric field is proportionate to $E_{rf}t$. In Fig. 229 (c) the resultant field due to superposition is $E_b + E_{rf}t$.

In this way it can be shown that electrons reaching the anode always land on that segment of the anode which is at the lowest potential. The current and voltage relations of segments 1 and 2 are shown in Figs. 230–233.

The usual split-anode magnetron characteristic diagram is shown in Fig. 230. Here segment No. 2 is held at constant potential with reference to the filament, the magnetic field strength H being above the "cut-off" value. The potential, V_a, of segment No. 1 is now varied by an external source and the current I_a is measured and plotted along the Y co-ordinate over the whole range of V_a variations for segment No. 1. The downward slope of the characteristic from A to B shows a falling current for a rising voltage and clearly, therefore, indicates the existence of a negative resistance effect.

The diagram shown in Fig. 231 is an extension of that shown in Fig. 230. In this case a family of curves is plotted for various fixed values of V_a for segment No. 2, when the voltage of segment No. 1 is varied throughout its range

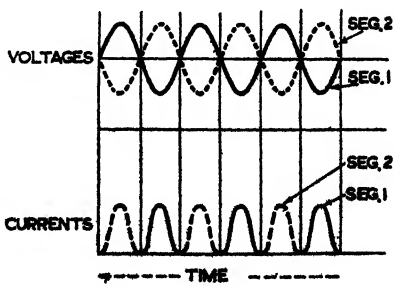


Fig. 233.—INSTANTANEOUS VALUES OF V_a AND I_a FOR SEGMENTS 1 AND 2 PLOTTED ON TIME SCALE.

Note V_{a1} in anti-phase V_{a2}; also I_{a1} and I_{a2} are in anti-phase to their respective voltages, $V_{a1}—V_{a2}$.

Note that as the fixed values for segment No. 2 are increased from 10 to 17·5, the point at which the anode current I_a approaches zero on the X co-ordinate (V_a for segment No. 1) increases.

If values for $(I_{a1}—I_{a2})$ are plotted against values for $(V_{a1}—V_{a2})$, the diagram in Fig. 232 is obtained. In this case the presence of a negative resistance is very marked at A—B. The significance of the heavily shaded part of this curve is dealt with in later paragraphs, in which the instability and threshold voltages for the resonant magnetron are discussed.

Finally, if the instantaneous values of V_a and I_a are plotted on a time scale, Fig. 233, it is seen that I_a for each segment is in anti-phase to V_a in each case.

Conclusions

Detailed examination of Figs. 230, 231, 232 and 233 lead to the following general conclusions :—

(1) As the potential of segment No. 1 falls and the potential of segment No. 2 rises, I_a in segment No. 1 increases, whilst I_a for segment No. 2 falls.

(2) When segment No. 1 and segment No. 2 have the same potential, I_a in both segments is zero.

(3) When for the next half cycle of R.F. segment No. 1 rises, the potential of segment No. 2 falls and I_a for segment No. 1 decreases, but I_a for segment No. 2 increases.

(4) In Figs. 230 and 232 the existence of a static negative resistance is clearly demonstrated and this, taken together with the conclusive results obtained in Fig. 233 make it quite clear that the device can be used as an oscillation generator, since the voltage and current relations of the segments satisfy the required conditions for oscillation.

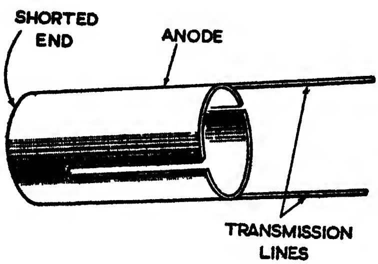


Fig. 234.—SHORT-CIRCUITED $\frac{\lambda}{4}$ TRANSMISSION LINES USED IN PLACE OF INDUCTANCE AND CAPACITY OF THE CONVENTIONAL TUNED CIRCUIT. RESULT HIGHER Q.

(5) The looped trajectory of electrons passing the gaps at favourable times suggests, in Fig. 228, the existence of a state in which the transit time of the electron is related to the frequency of the R.F. excited in the anode resonant circuits.

Since the oscillation frequency of the split anode magnetron is determined by the frequency of the external resonant circuit, shorted $\frac{\lambda}{4}$ sections of transmission line can be used in place of the orthodox parallel-tuned circuit when V.H. or U.H. frequencies are to be generated, see Fig. 234. An improved value for Q is thus obtained. The fact that there is a finite anode current, due to nett emission of electrons from the cathode, permits the frequency range and power output of the split anode magnetron to be increased greatly over those of its predecessor. Considerations of an improved value for Q suggests possibilities for the use of resonant cavities to obtain still greater efficiency at the higher frequencies.

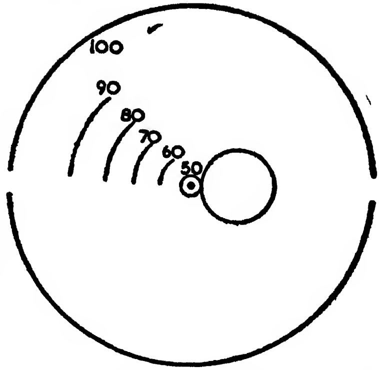


Fig. 235.—EQUIPOTENTIAL SURFACES OF THE CATHODE-ANODE SPACE.

($V_a = 100$)

Note. The large circle tangent to the 50 equipotential surface is assumed to be the orbit of the electron for the given conditions of E_a and H when no R.F. is present.

The power input is limited by voltage and heat dissipation restrictions.

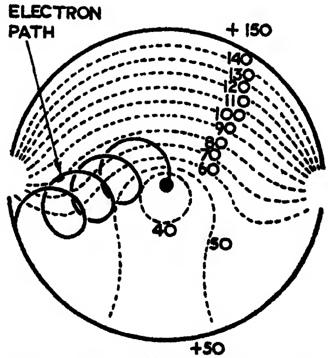


Fig. 236.—MODIFICATION TO DISPOSITION OF EQUIPOTENTIAL SURFACES BY AN R.F. FIELD DUE TO 50 VOLTS (MAX.). R.F. GENERATED IN EXTERNAL CIRCUIT BETWEEN SEGMENTS.

When E_{rf} is max. the upper segment is $E_a + E_{rf}$ (max.), *i.e.* (100 + 50) + 150 volts. Segment 2 = $E_a - E_{rf}$ max., *i.e.*, + 50 volts.

Equipotential Surfaces in Cathode-Anode Space. (No R.F. Present)

Before considering the resonant magnetron it is useful to examine in further detail the orbit of a single electron in the cathode-anode space. In doing so, the assumption is made that the oscillation period of the R.F. is long, compared with the transit time of the electron, so that the electron can be considered as being in a steady R.F. field during its passage from cathode to anode.

Consider the equipotential surfaces between anode and cathode when the anode voltage is, say, 100 volts positive and when the external oscillatory circuit is disconnected. The distribution is shown in Fig. 235.*

Modification of Equipotential Surfaces by R.F. Field

It will be noted that the electron path in the cathode-anode space follows the equipotential surfaces.

Now, superimposed upon the steady potential of 100 volts, let the potentials of the anode segments increase and decrease, sinusoidally—say 50 units. The equipotential configuration and the path of the electron will now be modified as shown in Fig. 236.

Note. There is a general drift of electrons following the equipotential surfaces, and since this drift is towards the more negative region, all electrons must land on the anode segment of lowest potential.

Power Output. Split Anode Magnetron

The split anode magnetron is capable of developing power outputs of the order 100–200 watts at frequencies in the region of 50–80 cm. with efficiencies of 40–70 per cent.

THE MULTI-RESONATOR MAGNETRON

It will be convenient to consider the operation of the resonant magnetron in the first place by considering the path of a single electron in the cathode-anode space and the principle by which it may transfer energy to the resonant elements of the magnetron anode. It is then proposed to modify this simple explanation to meet the conditions imposed by the presence of the space charge. Here again, for the sake of simplicity, operation in the simplest mode only is considered in detail. Later other possible modes of operation are mentioned together with the factors by which they are governed.

It should, however, be said at the outset, in order to prevent confusion at a later stage, that it is generally accepted that the conditions in the cathode-anode space are such, that in the absence of disturbing or distorting factors, *i.e.*, R.F. in the anode circuit, electrons emitted from the cathode proceed along their orbital paths in streams or orderly processions. When at switching on, the boundaries of the space charge have

* In Fig. 235 the electron receives potential energy when it is close to the cathode. On its journey to the anode this potential energy is converted to kinetic energy.

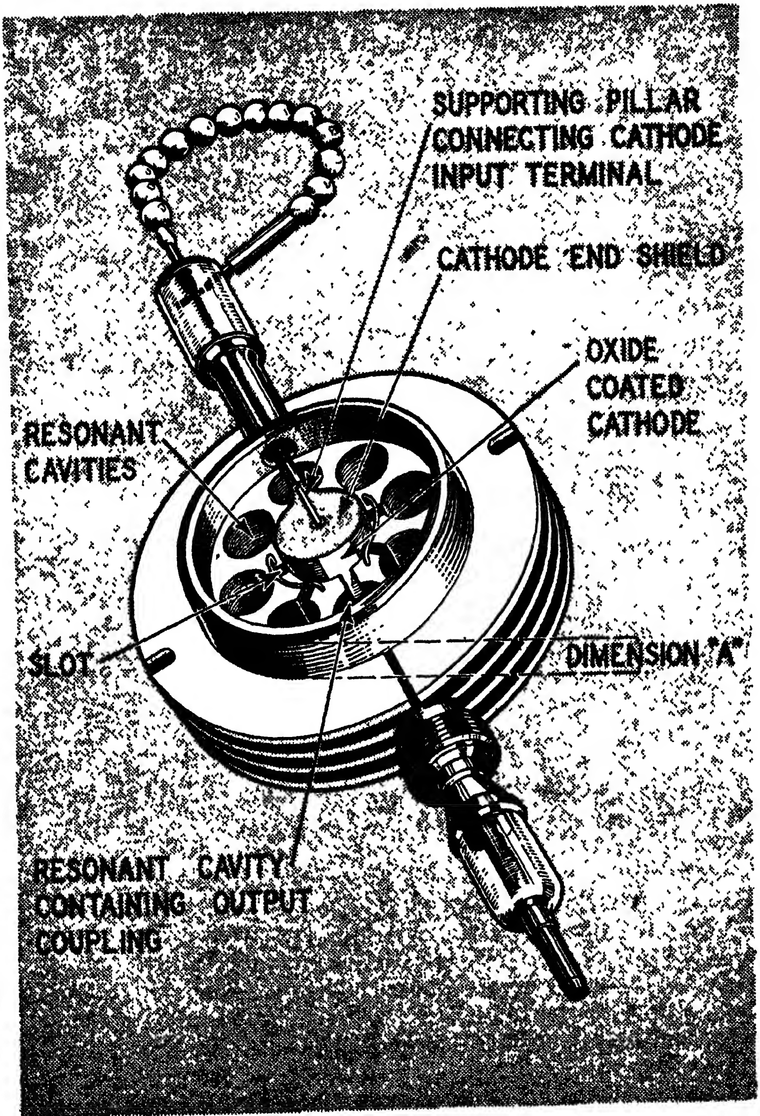


Fig. 237.—ANODE BLOCK AND CATHODE ASSEMBLY OF A MODERN MAGNETRON (END BLOCKS REMOVED).

[*Reproduced by courtesy of the Admiralty.*

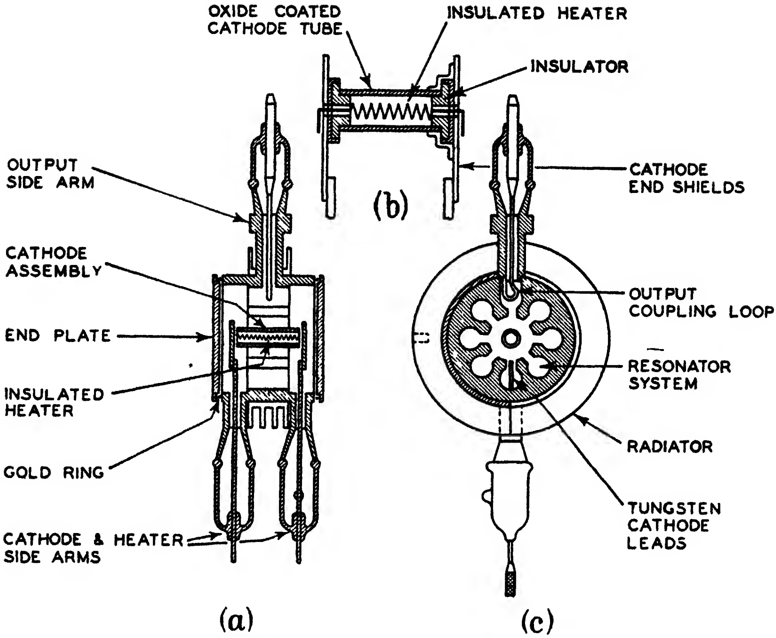


Fig. 238.—DETAILS OF A MULTI-RESONANT MAGNETRON.

(a) Side elevation of magnetron (cross-section).
(b) Details of the cathode, heater and cathode suspension (cross-section).
(c) Plan of magnetron looking along the axis of the cathode (end block removed).

been established there is no hurly-burly in the electron cloud, the to-and-fro traffic being controlled and confined generally to shells concentric to the cathode.

General Arrangement

Apart from constructional details, the obvious differences between the split anode magnetron and the resonant magnetron are that the anode is split into more segments (usually 8 to 16), also the external resonant circuit is replaced by a cavity resonator connected between each pair of segments. These resonating elements are incorporated as an integral part of the anode block.

The principal electrical difference is that a resonance velocity effect, perceived and noted by Megaw and others for certain operating conditions of the split anode magnetron, is exploited in the resonant magnetron, so that this device operates by means

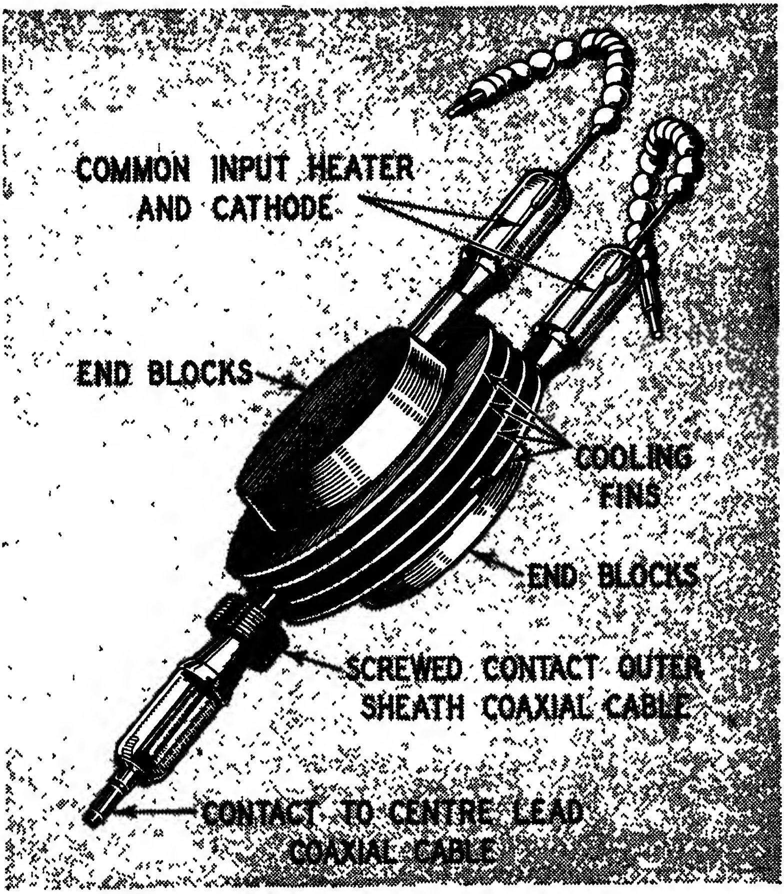


Fig. 239.—MAGNETRON WITH END BLOCKS FITTED OVER THE ANODE BLOCK.
[*Reproduced by courtesy of the Admiralty.*

of this velocity resonance effect, and not by the dynatron or negative resistance effect employed by the split anode magnetron.

The anode consists of a machined copper block, in which are cut the hole and slot combinations that form the resonant elements and which divide the anode into some even number of equal segments. The outer circumference of the anode is fitted with radiating fins or some alternative cooling provision (Fig. 237).

The cathode, with a diameter usually half that of the anode bore, is mounted co-axially within the anode itself, but insulated from it entirely. End shields are provided in order to prevent electrons emitted by the cathode reaching the anode by paths parallel to the lines of magnetic force of the axial magnetic field.

The cathode is usually constructed from a nickel mesh, to which the oxide coating is keyed. In this form, the cathode is mechanically strong enough to withstand the intense back bombardment of electrons and this assures a reasonably long life for the valve.

The cathode-heater assembly is suitably mounted and insulated as shown in Fig. 238. The heater leads being brought out through the copper glass seals, which are also shown (Fig. 239).

It is convenient to operate most magnetrons with their anodes at earth potential, consequently, under operating conditions, the cathode has to be made very negative to the anode and must therefore receive special attention with regard to insulation.

The output is obtained by magnetic coupling between the pick-up loop and the resonant cavity in which it is situated.* The output lead is also taken out through a special copper-to-glass seal, special arrangements, which will be referred to later, being made in this connection.

The Vacuum Seal

The vacuum enclosure is completed by copper end blocks which are clamped against the ends of the anode block, Fig. 239.

In order to effect a suitable seal a pure gold ring is clamped between each end plate and the anode block. During the baking process, on the pumps, the gold diffuses into the copper, and a permanent vacuum-tight joint is formed (Fig. 238).

Distance between End Plates and Anode

Since the anode block is turned with a raised rim, when the end plates are fitted into position, there is a space between these end plates and the recessed portion of the ends of the anode block, marked " Dimension A " in Fig. 237. The distance across this space between the end plates and the

* The cavities are linked by their fields and may be regarded as a common source of energy. Note space between anode and end blocks through which cavity coupling may take place.

anode block is an important dimension in the design of the magnetron, so also is the length of the anode. These matters will be referred to in later paragraphs.

The number of segments into which the anode is divided must be some even number, generally between 8 and 16 inclusive.

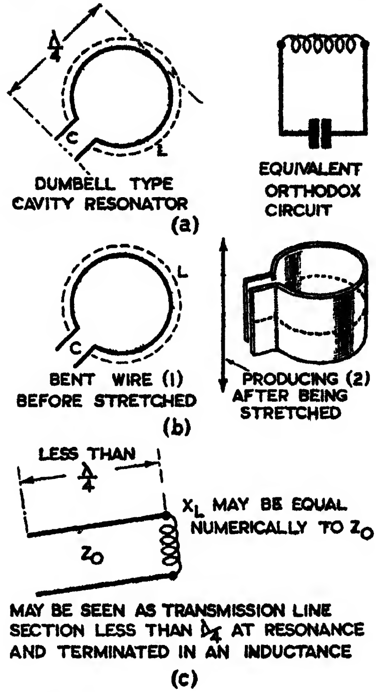


Fig. 240.—DETAILS OF HALF A DUMB-BELL RESONANT CAVITY SUCH AS MAY BE EMPLOYED IN SOME TYPES OF RESONANT MAGNETRONS, ALSO DETAILS OF ITS DEVELOPMENT FROM AN EQUIVALENT ORTHODOX CAPACITY INDUCTANCE COMBINATION.

The Resonant Elements

The hole-and-slot combination form half a "dumb-bell"-type cavity resonator. The dimension of this resonating element, measured from the inner diameter of the anode to the outer diameter of the hole, is a quarter wavelength (see Fig. 240 (a)), at the resonant frequency for which the magnetron is designed.

In the hole-and-slot combination, the region in the circumference of the hole provides the inductance component and the faces of the slot provide the capacity, see Fig. 240 (b). This combination may be thought of as a piece of wire bent into the shape of the hole and slot and then expanded axially, in the direction of the ends of the anode. The result is a thin skin of metal, shaped like the hole-and-slot combination, having a comparatively large surface dimension in the direction of the axis of the anode and therefore offering a considerable surface area to R.F. currents (Fig. 240 (b)).

Looked at from another point of view, the slots may be seen as some section of a transmission line, physically less than a quarter wavelength (at resonant frequency), terminated by an inductance which may have reactance equal in value to the characteristic impedance of the lines (necessarily fairly small on account of the axial dimension * (Fig. 240 (c)) This combination will then resonate at some frequency that makes the transmission line section about one-eighth the wavelength. For these dimensions the terminating inductance is then reflected to the open end as a capacity (see Appendix I, Theory of Transmission Lines).

Comparison with Split Anode Magnetron

Since the hole-and-slot combination is electrically equivalent to a parallel-tuned circuit at resonant frequency it is seen that the principal physical differences between the resonant magnetron and the split anode magnetron are :—

(*a*) The number of segments are usually not less than four pairs—always, an even number of sections of 2π.

(*b*) Each pair of segments is connected together by a resonant cavity. The resonant element is incorporated as an integral part of the anode itself and replaces the external oscillatory circuit of the split anode magnetron. The result is a greatly improved value of " Q."

(*c*) In place of the filamentary cathode normally employed by the continuous anode magnetron, a much more substantial cathode of larger radius is provided. The oxide coating is commonly capable of emitting at the rate of 40–50 amps. per sq. cm. The mechanical construction is such that early disintegration is prevented under normal working conditions.

Simple Analogy to Describe Action of the Resonant Magnetron

Before proceeding to any general discussion on the theory of the resonant magnetron, it may be helpful if an analogy is given to demonstrate the general principle involved in the transformation of D.C. energy into R.F. energy by this device. In doing so, however, two points must be borne in mind: (1) Nearly all analogies are only partially true and must not be carried too far; (2) in the comparison that follows, the movement of a single electron only is considered, the effect of the space charge (which is very important) being neglected in preliminary explanations for the sake of simplicity.

Let it be assumed that a number of empty jars, similar in

* When the wire is expanded into a cavity resonator as shown in Fig. 240 (b), the depth of the hole and slot is determined by the length of the axis of the anode. Obviously it would be beneficial to make this dimension infinitely great. It is limited, however, by the optimum air gap across which a magnet of reasonable proportions can drive the high flux density required.

size, form and material are arranged in a circle with all their mouths pointing inwards, and that an electrically driven fan is arranged to blow a current of air over the mouths of these jars. The air currents passing over the mouths of the jars will, under favourable circumstances, cause air vibrations within the jars, *i.e.*, produce a resonance within them.

A large amount of energy, represented by these vibrations, is now available *in the jars*; consequently, if this energy can be transferred efficiently to some suitable device or load so that a state of matching exists throughout, the energy set up and maintained in the jars by the air currents and the resonant properties of the jars, can be utilised and made to do a proportionately large amount of work.

Since a very small amount of energy is taken from the air currents to establish and maintain the state of resonance in the jars, provided that the losses in the jars are low, the fan is only required to provide energy in the form of D.C. air current, to cover the alternating air current load, plus the small amount necessary to cover losses in transformation. The efficiency of the device may therefore be very high and large inputs can be handled.

The large amount of energy represented by the alternating air current output is of course supplied by the fan in the analogy, just as the energy for the R.F. output of the resonant magnetron is supplied by the H.T. source. If, therefore, the jars, the D.C. air current and the fan of the analogy are replaced in the resonant magnetron by the resonant cavities, the electron motion in the cathode-anode space and the source of H.T. voltage respectively, it can be seen that the function of the electron processions which take place in the cathode-anode space of the resonant magnetron is to establish and maintain electrical oscillations in the resonant cavities of the anode system. In connection with this, a discrepancy must be pointed out. In the analogy, the air current set in motion by the fan acts directly upon the jars. In the resonant magnetron, the energy imparted to the electrons by the H.T. *viâ* the electric field is transferred to the resonant cavities in the first place *viâ* the electric field, contribution to the finite anode current is only made as and when electrons finally land each on its appointed segment of the anode.

Factors Affecting Power-Handling Capacity

With resonant magnetrons of suitable design, when the

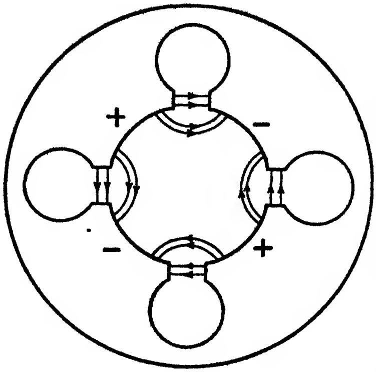


Fig. 241.—ELECTROSTATIC FIELDS AT ANODE FACE OF A FOUR-SEGMENT MAGNETRON WHEN OSCILLATING AT RADIO FREQUENCY IN THE π MODE AT THE INSTANT WHEN THE R.F. VOLTAGE IS E MAX. THE SIGNS OF THE SEGMENTS CHANGE OVER AT E MIN.

magnetic field strength is correctly adjusted,* pulse voltages of the order of 20 kV. or more can be employed at the input, and large R.F. powers can thus be obtained at the output. These large powers can be realised because the pulse length is usually very short (1 microsecond or less), thus permitting the use of an oxide-coated cathode which provides large primary emission.

Also there are no insulators to cause excessive losses in the R.F. circuit, wherefore voltages of the order of 20 kV. can be maintained at frequencies of 3,000 megacycles per second or more. All these factors also contribute to the large power-handling capacity of the valve.

Frequency Range

Valves covering the range of 8·5 to 10·9 cm. have been made for input power levels of 150 kW., 1,000 kW. and 5,000 kW. For wavelengths in the region of 3 cm., peak power inputs of 500 kW. represent the maximum generally available. Magnetrons, in general, used for this wavelength are for 150 kW. peak power input.

Initial Considerations

For reasons which will be given later, and as a prelude to any examination of the mechanism by means of which, in the resonant magnetron, energy supplied by the H.T. source is transferred to the resonant cavities by electronic action, it is necessary to assume that some initial disturbance has produced a state of instability between each pair of anode segments, so that oscillations of radio frequency are taking place.

* It can readily be seen that if high-voltage pulses are applied before the magnetic field is applied (where the field is supplied by an electromagnet), the magnetron will behave as a simple diode. Very large anode currents will therefore tend to flow and in all probability the valve will be destroyed. Hence high-voltage pulses must not be applied unless the magnetic field strength is normal.

It is also necessary to assume that the magnetron is oscillating in the simplest mode,* so that alternate segments are positive and negative at any given instant, each segment changing from maximum positive potential to maximum negative potential and back again once per cycle. In these circumstances the electrostatic fields due to R.F. segment voltages is shown in Fig. 241, where a four-segment valve is shown for convenience.

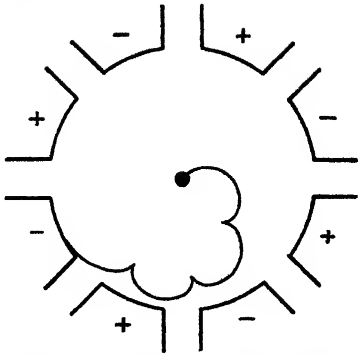


Fig. 242.—POSSIBLE TRAJECTORY OF AN ELECTRON FROM CATHODE TO ANODE OF AN EIGHT-SEGMENT MAGNETRON, OPERATING IN THE π MODE WHEN THE R.F. VOLTAGE AT THE ANODE IS, SAY, E MAX.

Means by which Energy is Transferred to Resonant Cavities

In the above conditions it is now proposed to examine the means by which electrons, under the influence of a D.C. potential between anode and cathode (anode positive to cathode), and an axial magnetic field, can transfer energy to the resonant system.

In the absence of a magnetic field, electrons, as already shown in the discussion on the continuous anode and split anode magnetron, emitted by the cathode are accelerated towards the anode by the radial electrostatic field, and the kinetic energy which they thus acquire is given up as heat on reaching the anode.

In the presence of an axial magnetic field, electrons in transit to the anode experience a force at right angles to the direction of acceleration and, therefore, follow a curved path or trajectory (Fig. 226 (a)). In the cathode-anode space the direction of motion of the electrons may change progressively, so that they may actually be turned back towards the cathode (Fig. 226 (b) and (c)), in which case the electrostatic field exerts a force on the electron which tends to retard its return journey to the cathode, so as to bring it momentarily to rest. In this condition, since the motion of the electron is zero, the deflecting force exerted on it by the magnetic field is also zero in conse-

* The simplest mode is the π mode, which causes a voltage distribution on the face of the anode as described above (Fig. 241). It is shown in later paragraphs that a magnetron may oscillate in other modes in which case the voltage distribution on the anode face is not the same as for the π mode.

quence of which the electric field tends to accelerate it in the opposite direction away from the cathode and towards the anode again.

It can now be seen that the electron path is determined by a sequence comprising a backward curving journey dictated by H, retarded to zero and converted to a forward journey by the action of the combined radial and R.F. fields, and that this sequence is periodic. It is also apparent that there is a relationship between the periodic evolutions of the electron in its orbital path to the anode and the R.F. fields by which the radial field is modulated (*i.e.*, a velocity resonance).

Hence, in a multi-segment resonant magnetron the trajectory shown in Fig. 242 may occur for certain values of high voltages and strong magnetic fields. Also for a given geometrical shape of anode and cathode, and for a given value of H, just as in the split anode magnetron, there is a critical anode voltage below which electrons cannot escape to the anode *in the absence of an R.F. voltage on the anode.* This is known as the Hull cut-off voltage. It is also to be noted that a condition exists somewhere below the Hull cut-off value, where electrons cannot escape to the anode even when *small* R.F. voltages do exist at the anode.

If at a particular instant in the high-frequency cycle when the valve is oscillating in the simplest mode, the segments are positive and negative as in Fig. 242, then some electrons emitted from the cathode will describe trajectories as in Fig. 242 because they enter regions where the fields due to R.F. in the anode are large enough to modify their motion.

But, in general, electrons which enter these regions may be either accelerated, and therefore acquire energy from the R.F. field, or they may be retarded, and so give up energy to the R.F. field, according to the directions of the electrostatic fields and the orbital position of the electrons, *i.e.*, relative directions of E.S.field and electron motion.

In the latter case, R.F. energy is transferred to the resonant cavities and may be built up. Assuming that (*a*) conditions favour a build up, and (*b*) that electrons can escape to the anode at a rate sufficient to maintain these oscillations under load conditions, then, by a suitable matching system, the energy in the resonant cavities can be extracted and put to work on a suitable load.

In order that the device may have a positive efficiency, more energy must be given up by the electrons to the R.F.

field than is imparted to them by it and the means by which this is accomplished must now be examined.

In making the complete journey from cathode to anode an electron is acted upon by three forces :—

(*a*) A force due to the R.F. electrostatic field.

(*b*) A force in an outward radial direction due to the D.C. electrostatic field produced by the H.T. between cathode and anode.

(*c*) A force at right angles to its motion due to H which tends to make the electron follow a curved path.

Thus, for any fixed values of V_a and H, any factor that tends to weaken the effect of the force due to the action of the magnetic field makes it easier for the electron to be accelerated towards the anode and so to contribute to a finite anode current.

But there are two distinct and separate issues, viz. :—

(*a*) The excitation and building up of oscillations in the cavities. This requires a nett gain of energy by the R.F. field from electrons in the cathode-anode space dealt with above.

(*b*) Energy to maintain the oscillation build up (condition (*a*)) and to supply the load.

Requirement (*a*) is fulfilled by successive retardations of electrons in the cathode-anode space under conditions of velocity resonance. Requirement (*b*) can be fulfilled only by nett cathode emission, *i.e.*, by electrons reaching the anode, but in view of condition (*a*) they must not absorb energy from the R.F. field to do this. It will be shown, in later paragraphs, that the force which imparts the necessary component of radial velocity to enable retarded electrons to reach the anode, may be derived from the electromechanical condition of the space charge itself.

In (*a*) we are concerned with the means by which more energy is imparted to the R.F. field by electrons than they absorb. Since the force exerted on the electron by the magnetic field is proportional to the *velocity* of an aggregate of electrons (velocity of the electron being analogous to current), it follows that electrons which are *retarded* by the R.F. field, *i.e.*, give up energy to it, are *less affected by the magnetic field than if there had been no such interaction with the R.F. field.* In other words, electrons that have been retarded and so lost radial velocity are deflected less by the magnetic field. Conversely, electrons which receive energy from the R.F. field, and are thereby accelerated, *experience a greater measure of force due to H and are in consequence turned away from the anode towards the cathode.*

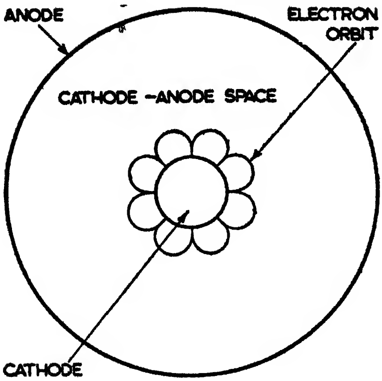


Fig. 243.—CONDITIONS FOR LOW VALUES OF V_a (ELECTRON ORBITS).

Condition of Space Charge when no R.F. Present on Anode

It can be shown that for the Hull cut-off conditions, when *no R.F. field is present*, electrons emitted from the cathode revolve round it in circular orbits at uniform speeds, but when an R.F. field is present of sufficient strength to affect them, some electrons follow the path shown in Fig. 242. *Abrupt changes of trajectory* imply retardation, and therefore energy delivered to the R.F. field. Fig. 242 shows that these changes take place periodically and that they occur at the instant when the electron is in the vicinity of a gap.

Conditions for Resonant Excitation

It is seen, therefore, that resonant conditions are obtained when an electron crosses a gap between segments in an R.F. retarding field and reaches the next gap at the same time as the R.F. field changes from accelerating to retarding, and so on for each subsequent gap. This can lead to a condition where some electrons are retarded by the R.F. fields of all the gaps throughout an entire orbit, whereby they give up the greater part of their energy to the R.F. field.

Electrons which are initially accelerated by the R.F. field do not survive to be again accelerated at subsequent gaps, because, in receiving acceleration they also experience an increased force due to H, in consequence of which they are turned back from the anode to the cathode, where they are removed from circulation by collision with the cathode before they have had sufficient time to absorb much energy from the electric field.

The nett result is that only those electrons which by successive retardations have transferred the greater part of the energy acquired from the H.T. source, can be favourably placed to escape to the anode and so contribute to the working anode current. There is also nett gain in cavity excitation.

Operating Conditions

In practice, the operating voltage for all normal magnetrons is *well below the Hull cut-off voltage and is in the region where electrons would never escape to the anode unless they could give up a substantial part of their energy to the R.F. field.*

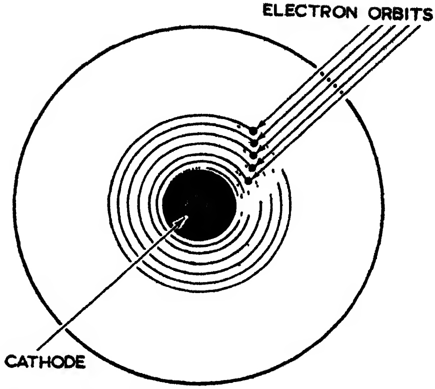


Fig. 244.—CONDITIONS IN SPACE CHARGE WHEN V_a IS INCREASED, IMMEDIATELY STABILITY IS ESTABLISHED AFTER SWITCHING ON

The electrons which complete unfavourable paths and are turned back, collide with the cathode, where they give up their kinetic energy in the form of heat. This factor frequently limits the mean power input for a given cathode size. If the mean power rating is known in the design stage, it is usually possible to reduce the size of the valve until the back bombardment voltage is just sufficient to maintain the cathode temperature without overheating. The valve heater can then be switched off for the time during which the valve is in operation. Since it is also easier to obtain good stability with a small valve than with a large one, this is a desirable feature.

Classification of Regions

From the foregoing it is evident that further detailed examination of the operation of the magnetron must be conducted, in the first place, by an exploration into the conditions existing in three separate and distinct regions :—

(*a*) The cathode and the space immediately surrounding it, *i.e.*, the space charge.

(*b*) The region between the boundaries of the space charge and the face of the anode.

(*c*) The anode, its resonant elements, the load and the matching system.

Analysis of Conditions in Region (*a*). Assume a condition where there is no R.F. set up in the anode and therefore no R.F. field to affect the electrons. In these circumstances electrons are emitted from the cathode and a space charge cloud will be formed in the region immediately surrounding the cathode.

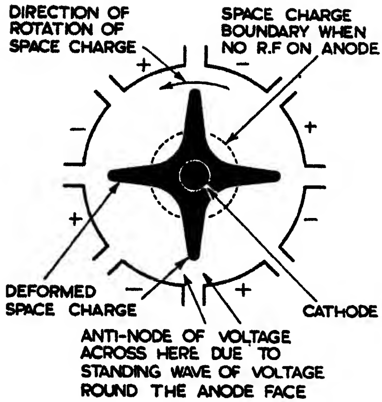


Fig 245.—CHANGE OF SHAPE OF THE ROTATING SPACE CHARGE UNDER THE INFLUENCE OF THE R.F. FIELD.

The fields across the gaps are those corresponding to antinodes of voltage across the gaps, which arise from a stationary voltage wave round the anode face.

For low values of V_a, the value of H being fixed, electrons describe paths as shown in Fig. 243. As the value of V_a is increased, and because of the mutual repulsion that exists between them, electrons do not return to the cathode, but revolve round it at speeds dependent entirely upon the value of H and the radii of their orbits (Fig. 244).

It is thought that the motion of the electrons in the space charge itself takes the form of an orderly procession, at varying radii, between the surface of the cathode and the space charge boundary, forming a series of concentric shells revolving with the cathode as axis as shown in Fig. 244.

The distribution of potential in the anode-cathode space and therefore the radii of the various concentric revolving shells is given by the expression

$$V = V_a \frac{\log \frac{r}{r_c}}{\log \frac{r_a}{r_c}}$$

where V is the potential at any point distant r from the axis

V_a = anode potential
r_a = anode radius
r_c = cathode radius.

Whilst for a given design the orbital radius of the electron path is seen to increase when V_a is increased, it can be shown that the angular velocity of the electrons along their orbits (ω) is independent of the strength of the radial electric field V_a and that it is dependent upon the value of H and the radius

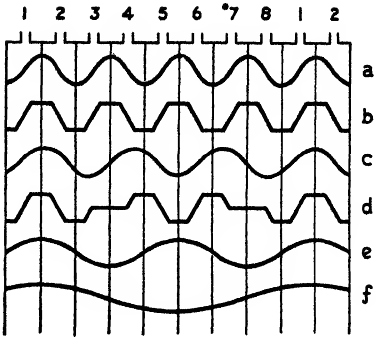


Fig. 246.—PATTERNS OF VOLTAGE DISTRIBUTION OBTAINED AT ANODE FACE WHEN THE INPUT FROM AN EXTERNAL OSCILLATOR IS VARIED IN FREQUENCY.

Note (a) shows the position and sinusoidal amplitude of the voltage antinodes for the π mode whilst (b) indicates the distribution of voltage for this mode at the anode face. (c) and (d) are complementary for the No. 3 mode. (e) and (f) show the position and voltage variation of the antinodes for the No. 2 and No. 1 modes respectively.

of the cathode*. The maximum fall of potential is in the region of the cathode; examination of the equipotential surface diagram, Fig. 235, shows that electrons receive the greater part of their potential energy immediately on leaving the cathode.

It must be realised, however, that from the instant of switching on until a condition of equilibrium has been reached, there is probably some confused motion in the space charge region, but it may be assumed that compared with a time of perhaps 1/50th microsecond, this period of instability is short, consequently it is not necessary to consider the special case when a magnetron is pulsed.

The space charge boundary for fixed values of electric and magnetic quantities is determined by the geometry of the valve.

Effect of R.F. Fields upon Space Charge

When R.F. is set up in the anode, the R.F. electric fields cause deformations of the space charge cloud. The nature of this deformation and its magnitude determine, very largely, the operating conditions of the magnetron. Since it is to be

* If F is the force accelerating electrons towards the centre, *i.e.*, the centripetal force, m = mass of the electron, v = its velocity and r the radius measured from the centre of gravity, then

$$F = \frac{mv^2}{r}$$
$$= m\omega^2 r$$
$$\omega^2 = \frac{F}{mr}$$
$$\omega = \sqrt{\frac{F}{mr}}.$$

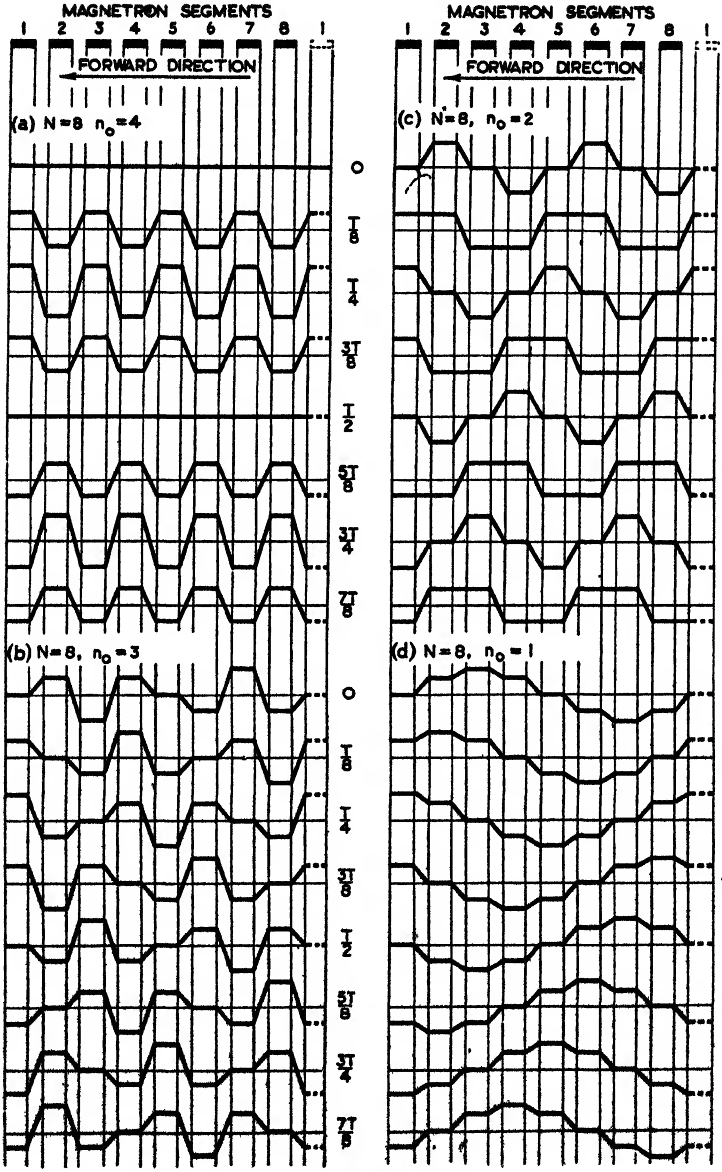


Fig. 247.—AMPLITUDE VARIATIONS OF RECTANGULAR STANDING WAVES. N = number of segments, n_0 = mode number.

shown subsequently that the space charge deformation and the distribution of voltage on the anode face are similar, it is necessary to examine the voltage distribution on the face of the anode segments for various modes of oscillation (see Fig. 245). In this connection it is essential to bear in mind the fundamental fact that the characteristics of the resonant elements (hole and slot) are such that when R.F. oscillations at resonant frequency take place an antinode of voltage appears across the gaps at the anode face, *i.e.*, a standing voltage wave.

R.F. Voltage Distribution at Anode Face

If at the output terminal of the magnetron an input is applied from a local oscillator under conditions where the magnetron is itself not generating, the voltage distribution at the faces of the segments can be examined by means of a probe just protruding from the surface of the dummy cathode and the aid of an oscilloscope. It is found that as the frequency of the oscillation generator is varied, a number of sharp resonances may be noted, at particular frequencies where large segment voltages are built up, with no segment amplitude at intermediate frequencies. These resonant frequencies are those at which the oscillating magnetron can be made to work by choice of suitable operating conditions.

Voltage Patterns on Anode Face

Referring to Fig. 246, as the frequency of the local oscillator is changed, pattern (a) may disappear and pattern (c) may take its place. Here, instead of having four maxima and four minima (for an 8-segment block) it is found that there are three maxima and three minima which cannot coincide exactly with the segment faces. If these conditions are investigated close to the anode face, the more exact condition (b) would be obtained for pattern (a) and (d) for pattern (c),* further comparisons of pattern with input show that patterns similar to (e) and (f) can be obtained. Pattern (a) is obtained at *one wavelength only*, while patterns with one, two and three maxima are each obtained at two wavelengths, usually with a small difference between them. The two patterns with the same

* This is readily understood when it is remembered that the potential between any two points of any one segment must be the same. Since there is no appreciable drop between any two points on the same segment, the whole of the difference of potential between adjacent segments appears across the gaps. Hence in going round the anode a rectangular wave is obtained for each pair of gaps passed over. These waves are standing waves.

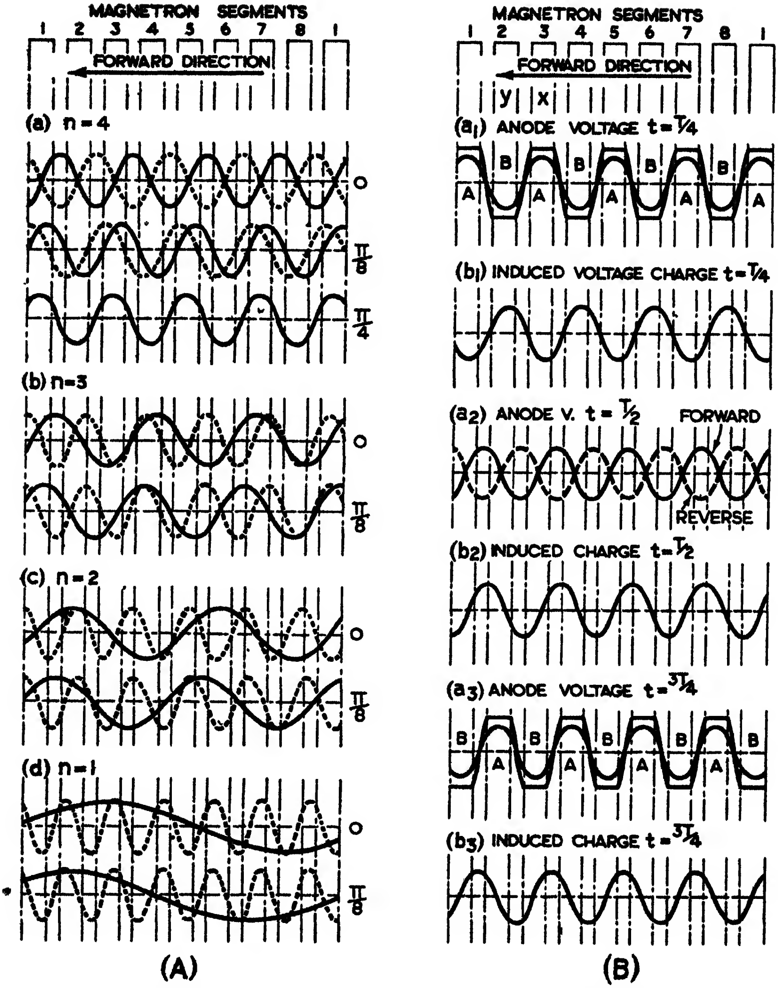


Fig. 248.—(A). Phase relations forward and reverse waves. (B). Phase relations standing wave components and E induced in anode by space charge reaction (modulating function).

number of maxima are found 180° out of phase with one another.

If the block as illustrated in Fig. 246 is now made up as an oscillating valve it is found that the different mode wavelengths obtained are exactly the same as those required to produce the various patterns of Fig. 246. These patterns are described by the number of maxima round the block. Thus (a) is $n = 4$, (c) is $n = 3$, etc. In (a) there is a phase difference of π in the R.F. voltage on successive segments and this pattern is usually called the π-mode pattern.

The curves in Fig. 246 for the voltage distribution patterns a and c, for example, show that the *amplitudes* of the standing waves between segments, which are really square waves (as in b and d), vary sinusoidally.

Thus the rectangular-shaped wave b (Fig. 246), for example, represents the distribution of voltage between segments of an eight-segment magnetron when it is operating in the π mode, corresponding to pattern $n = 4$. Rectangular standing waves of voltage for other modes are also shown in Fig. 247.

Each rectangular wave, which is also a standing wave, may be thought of as composed of a large number of *standing sine waves*, each of which are harmonics of (or strictly speaking in this case secondaries to) a primary sine wave, the latter having the same frequency as the rectangular wave.

Fig. 247 (a) shows the variations in amplitude at the anti-nodes of the standing wave for the π mode or $n_o = 4$ mode, for time increments of $\frac{T}{8}$ throughout a complete cycle.

Fig. 247 (b), (c) and (d) show in manner similar to (a) the position and sinusoidal variation of the anti-nodes of voltage round the anode face for modes $n_o = 3$, $n_o = 2$ and $n_o = 1$ respectively

π Mode

By transmission line theory these standing sinusoidal waves (components of the rectangular standing wave) are each caused by two travelling square waves, *i.e.*, a forward wave and a reflected wave, travelling anticlockwise and clockwise respectively, but of equal amplitude and with the same velocity (see (a), Fig. 248 (A)) This means that standing secondaries plus the standing primary or fundamental sine wave component of these rectangular waves are caused by forward sine waves rotating round the anode in an anticlockwise direction (signified by +), and reverse or reflected waves of equal amplitude and similar velocity, travelling round the anode in a clockwise direction (signified by —). Thus, when oscillating in the π mode the primary standing sine wave of the rectangular wave is produced by two rotating sine waves $n = +4$ and $n = -4$ (see (a), Fig. 248 (A)).

The rectangular standing wave has, however, many other component standing sine waves, besides the primary standing sine wave, each of which is secondary to the primary or fundamental wave. For an eight-segment magnetron, where the primary wave has four repeat maxima round the anode, the secondary waves would have repeats maxima equal to 12, 20, etc.

These secondary *standing* sine waves having repeats of $n = 12$, $n = 20$, etc., are, like the primary wave, caused by *travelling* sine waves of $n = +12$, $n = -12$ and $n = +20$ and $n = -20$, and so on respectively.

The general case then is covered by the formula :—

$$n = n_o + \mu N$$

where n = number of repeats of the original patterns
n_o = any integer between 1 and $(N - 1)$
N = number of segments
μ = any integer + or — including 0.

In Fig. 248 (A), (a) to (d) show the primary forward and reverse components the interference of which produces the primary standing sine component of the standing tiapezoidal wave for modes $n = 4$, $n = 3$, $n = 2$, $n = 1$ at times 0 and $\frac{\pi}{8}$.

Other Modes

For modes other than the π mode it is clear from consideration of Figs. 246 and 247 that the above general formula cannot be applied. It has already been stated that the voltage distribution is constant over the whole of each segment, consequently the system of primary and secondary travelling waves, which provide the primary and secondary standing components of the standing rectangular wave, must all contribute to the potential of each segment in their proper relations.

This imposes the following conditions on the component waves (b) to (d) in Fig. 248 (A) :—

(*a*) Each pair of travelling waves must combine to produce a wave which will rotate with the primary wave.

(*b*) The sum of each pair of components taken at the centre point of each segment must give the value of the potential of the segment, *i.e.*, they must make some contribution to each gap, in turn, and also to each segment.

It can be shown that these conditions are satisfied by pairs of travelling waves having maxima of $n_1 = n + \mu N$ and $n_2 = n - \mu N$.

Secondaries or Harmonics

Thus if an eight-segment magnetron is operating in the π mode the forward waves are $+ 4$, $+ 12$, $+ 20$, etc., and the reverse or reflected waves are $- 4$, $- 12$, $- 20$, etc. ((a), Fig. 248 (A)). If the same magnetron is operated in the $n = 3$ mode, the forward waves are $+ 3$, $+ 11$, $+ 19$, etc., whilst the corresponding reverse or reflected waves are $- 6$, $- 13$, $- 21$, etc. The formula for $n = 3$ mode applies to all the other modes except the π mode.

The values of V_a and H having been suitably adjusted, the primary forward and reflected waves arising therefrom are the predominant factors in exciting any desired mode; consequently, these are the only ones that need be considered in detail.

All the secondaries are, however, present in the output to a greater or less degree, and they have, therefore, been considered at some length, because of the important bearing which they have on the matching and loading conditions of the valve.

It will be shown that the selection of the mode to be established is made by initial adjustment of the values of V_a and H for any given magnetron, *that is always assuming that the design of the magnetron* (its geometry) is such that it is possible for the given mode to exist.

Conditions for Build-up and Mode Excitation

In order that oscillations can be excited and built up in the resonant elements, it is essential that there should be a nett loss of energy from the space charge cloud. It is also essential, as a supplementary condition, that in order to maintain these oscillations and to supply power to a load, electrons must escape from the space charge cloud to the anode, thus enabling power to be supplied from the H.T. source, by reason of a nett loss of electrons from the cathode which are replaced by the H.T.

Fulfilment of these conditions implies that a deformation of the space charge cloud must therefore take place, wherefore it now becomes necessary to consider the nature of this deformation of the space charge cloud and the factors that control it.

When an eight-segment magnetron is oscillating in the π mode, the voltage distribution over the face of the anode may be summarised as follows :—

(*a*) There is a difference of potential between segments which modifies the radial electric field distribution.

(*b*) At any given instant the potential at all points in any one segment is the same.

(*c*) There is a phase difference between adjacent segments.

(*d*) The total phase difference distributed round the anode must be some multiple of 2π.

(*e*) For an 8-segment magnetron the phase difference between segments is some multiple of $\frac{2\pi}{8} = \frac{\pi}{4}$.

(*f*) Corresponding to different multiples of $\frac{\pi}{4}$ there are different patterns (modes) of voltage distribution possible.

(*g*) For anticlockwise rotation in an 8-segment magnetron there is a phase lag to the left and this is given by

$$\frac{n \times 2\pi}{N}$$

where n is the number of maxima corresponding to a particular mode and N is the number of segments.

Thus for the simplest case, viz., the π mode, the phase difference between segments is :—

$$\frac{4 \times 2\pi}{8} = \pi.$$

It can be seen from the foregoing that the π-mode pattern is a simple standing wave and it has been shown that these standing waves are due to a pair of travelling waves revolving in opposite directions round the face of the anode. The standing waves are caused by the interference between the two travelling waves.

It can be shown that the travelling waves rotate round the face of the anode with velocity $\frac{\omega}{n}$,* the forward wave being assumed to rotate round the face of the anode in the same direction as the space charge cloud revolves round the cathode, *i.e.* anticlockwise.†

Note. The expression $\frac{\omega}{n}$ indicates that the travelling waves for different modes, No. 4, No. 3, etc., rotate round the anode

* $\omega = \frac{v}{r}$ where v = velocity and r = radius.

† Let ω of the periodic component be ψ relative to stationary axes, then

face with different velocities, depending upon the value of n for any particular mode.

Space Charge Deformation and Oscillation Build-up

Let it be assumed that a magnetron of 8 segments and suitably adjusted for V_a and H, commences to oscillate in the π mode, due to some disturbance or transient at the anode. The voltage distribution on the face of the anode will then be similar to that which has already been described for this mode. At the instant before the onset of oscillations, the space charge cloud will be symmetrical about the cathode, with orderly processions of electrons rotating in their orbits. The angular velocities of the electrons rotating in these concentric shells or orbits decreases when the radii of their respective

the time required to travel once round the anode is $\frac{2\pi}{\psi}$. Consider fluctuations of potential at particular points on the anode. Since this pattern has n repeats the potential will have carried out a sinusoidal oscillation in that time.

$\therefore$ Time for 1 oscillation is $\frac{2\pi}{\psi n}$ frequency $= \frac{\psi n}{2\pi}$ $\therefore$ $\psi = \frac{2\pi f}{n} = \frac{\omega}{n}$. (Fig. 249.)

The anode wave system may alternatively be thought of as a simple travelling wave having n repeats represented by :—

$$E\psi = \mathcal{E}\psi \cos (\omega t - n_0\psi).$$

where $E\psi$ = instantaneous value of the tangential component of the electric field.

$\mathcal{E}\psi$ = maximum value.

This wave can be made good to cover the correct distribution of field across the gaps, provided they are satisfactorily narrow, but fails over the anode segments where there is zero field.

The actual wave in the magnetron can be regarded as the above wave *modulated* by a suitable function of ω, *i.e.*, one which gives zero over the segments and unity over the gaps.

Representing this function as $f(\omega)$, the complete wave can be given by $\mathcal{E}\psi = f(\omega)E\psi\,(\cos \omega t - n_0\omega)$.

This modulating function can be analysed into a series of sine terms :—

$$f(\omega) = \frac{n_0}{2} + \omega \cos N\omega + 2 \cos 2N\omega + a \cos 3\omega, \text{ etc. . . .}$$

where there are N gaps and ω is measured from a gap centre.

The wave formula can be rewritten :—

$$E\omega = E\psi \cos (\omega t - n_0\psi)$$

which constitutes the actual primary forward wave. The other terms mean that the wave represented by this general term can be regarded as arising from the interference of the simple travelling wave with mode numbers $(n_0 + \mu N)$ and $(n_0 - \mu N)$.

This complete wave can be regarded as the resultant of a number of travelling components whose mode numbers are given by $n = (n_0 + \mu N)$ where n is allowed to be zero or any positive or negative integer.

The π mode is a special case of this in which $N = 2n_0$.

As the slot width increases, the mode numbers of waves set up for any pattern remain the same as in the above analysis but the amplitude will be different.

orbits increase. This follows from equation for time period of 1 cycle $= \frac{2\pi r}{v}$.

At the onset of oscillations a standing wave of voltage appears on the face of the anode, and this, as has been shown, can be resolved into two waves travelling round the anode in opposite directions with a velocity of $\frac{\omega}{n}$ (n being equal to 4 for an 8-segment magnetron oscillating in the π mode).

The space charge cloud is deemed to be rotating in an anti-clockwise direction round the cathode; consequently, the wave that travels round the anode in the same direction—the forward wave—will produce the major reaction on the space charge, the reverse wave, rotating in the opposite direction will have comparatively little effect upon it excepting for unfavourably placed electrons and secondary modes.

The forward wave travelling round the anode face with an angular velocity of $\frac{\omega}{n}$ will directly affect to the greatest extent some particular shell of the space charge in which the angular velocity of the electrons is also approximately $\frac{\omega}{n}$. *Note* that the radius at which the synchronous shell resides depends upon the value of n, *i.e.*, the mode of operation. The larger the value of n, the further from the cathode will the synchronous or resonant shell be found and the smaller will be the angular velocity of its electrons.*

The crests and troughs of the wave travelling round the face of the anode impress hills and dales on the resonant shell so that the anode wave becomes locked to it, and there is a resultant change in distribution of the electric fields.

In order that the space charge cloud may conform in shape to the hills and dales impressed upon its periphery by the crests and troughs of the anode waveform, with which it rotates in step, some of its electrons must be retarded and others accelerated.

Thus the condition for the *build-up* of R.F. oscillations in the resonant cavities of the anode is that the space charge

* From this it follows that the resonant shells corresponding to the lower values of n are to be found nearer the anode than those for the larger values of n. $\frac{v}{2\pi r} = \frac{\omega}{n}$.

cloud must give up more energy to the resonant circuit of the anode, *i.e.*, to the R.F. field, than it receives from it.

This condition of a nett loss of energy by the space charge cloud requires that V_a and H, for a particular magnetron, be adjusted initially to make the radius of the space charge cloud (before oscillations begin), sufficiently large to ensure some nett loss of energy when oscillations commence.

Since the nett energy loss is limited by the initial radius of the space charge cloud, the amplitude to which the build-up can ultimately attain is limited by the same consideration. This is only true however when the resonant cavities are not loaded. When the cavities are loaded, the energy transferred from the space charge to the resonant cavities (by what might be called the electromechanical action of the valve), is insufficient by itself to maintain the oscillations and to supply the load.

Requirements to Sustain Oscillations under Loaded Conditions

Under loaded conditions, oscillations initially excited by the anode wavespace charge reaction must be maintained by an escape of electrons from the cathode to the anode.*

The conditions for this to occur are, of course, dependent, primarily, upon the initial excitation of R.F. energy in the anode. Without this condition electrons could never escape to the anode, since V_a and H are adjusted initially for the valve to operate well below the Hull cut-off condition.

For a magnetron of given design, this means that, for maximum efficiency, the values of V_a and H must be chosen to fulfil the following conditions :—

(*a*) A suitable value for H having been selected and fixed, V_a must then be adjusted so that oscillations are *excited* in the desired mode. This voltage is the *instability voltage*.

(*b*) Oscillations having been excited in the desired mode, they must be maintained under loaded conditions by an escape of electrons from the space charge cloud to the anode, R.F. having been established in the anode circuits, the conditions for escape of electrons to the anode now depends upon a further adjustment of V_a. In the interests of efficiency, it is

* These two requirements correspond to a priming function which excites oscillations in the required mode and a power supply function which maintains them under loaded conditions. The escape value for V_a may be equal to, greater or less than the excitation value for V_a; thus the final adjustment of the common source V_a must be determined by independent consideration of the two factors that determine it.

desirable that the adjustment of V_a, necessary to enable electrons to escape, should coincide as nearly as possible with the value of V_a required under condition (*a*). The escape voltage is known as the *threshold voltage*. The conditions for this are dealt with in a later paragraph.

All component waves affect the space charge to some extent, but as all components of V decrease in amplitude towards the cathode, depending on $\left(\frac{r}{r_a}\right)^n$, the contribution made to velocity by low values of n, will be greater than that for large values of n.

Thus the primary forward component, the lowest value of n_0 for the particular mode, will cause the greatest distortion of the space charge cloud, and will be the major effective component in the initiation of oscillations.

Alternative Theory

By some authorities it is thought that the resonant shell is nearer the anode than the shell with which the anode wave is in step. In this case the conditions for building up oscillations are that V_a and H must be adjusted so that the radius of the space charge cloud is large enough to contain the resonant shell. As in the previous case, however, the value of V_a to produce instability in the required mode for a given value of H must exceed a certain minimum. The instability potential reckoned on this basis is somewhat larger than it would be if calculated on the previously mentioned basis.

In either case, however, the deformed shape of the space charge cloud is much the same, and there is in consequence a nett loss of kinetic energy from it.*

Electrical Conditions in the Synchronous or Resonant Cavities

Before considering the conditions under which electrons can escape to the anode (threshold potential), it is necessary to examine the distribution of voltage and current in and about the resonant cavities.

Consider the effect on the anode when a deformed space charge cloud revolves in step with the primary forward component of the anode wave.

At any given instant the hills on the space charge cloud, which are formed by accumulations of electrons, and are

* It is offered as a suggestion that the deformed shape of the space charge may be caused by a piling up of favourably placed electrons, due to retarda-

therefore negatively charged, induce in the anode, at points immediately opposite them, local increases in the steady positive charge due to H.T. Similarly, at points on the anode corresponding to each dale in the space charge cloud, there will be at the anode a local reduction of the charge due to H.T.

Thus, superimposed on the steady positive charge on the anode, due to H.T., alternating charges are induced which are in phase with the alternating potentials of the anode segments.

Let the electrical conditions of the resonant cavities now be examined. Take, for instance, the cavity marked x—y, Figs. 248 (B), 250 and 251.

In Fig. 248 (B) the letters A and B in the segments indicate positive and negative maxima of the trapezoidal standing wave across the gaps and xy is the particular gap under examination. (a_1), (a_2), and (a_3) show the phases of the forward and reflected component waves, in phase at $\frac{T}{4}$ and $\frac{3T}{4}$ and in antiphase at $\frac{T}{2}$. (b_1), (b_2), and (b_3) show the variations and distribution of charges induced in the anode at the above times by the electro-mechanical action of the space charge and the phase relations of this space modulating wave to the anode primary wave. (Fig. 249.)

The curve in Fig. 249 (c) (1) shows the manner in which the potential of x changes relative to y across the gap xy. Since the distribution of the induced charges is such that they move anticlockwise round the anode, a component of I_a must flow at every point on the anode proportional to the amplitude of the corresponding component of instantaneous charge density at the point. Since graphs b_1, b_2 and b_3 in Fig. 248 (B) may be interpreted as current distribution round the anode instant by instant, and assuming I in the anti-clockwise direction to be regarded as positive, then the current I_a flowing from x to y round the cavity xy can be deduced and plotted as in Fig. 249 (c) (2).

When x is positive with respect to y, the direction of the electric field intensity across the gap E is as shown in Fig. 251, being reckoned as positive in the direction of the arrow.

When I round the cavity is positive the direction of the magnetic field is as shown by the arrow head, *i.e.*, out of the paper, Fig. 251, therefore if H be reckoned positive in this

tion by the major anode wave at different rates for orbital radii. In which case waves may be raised on the surface of the space charge, rotating in step with the primary forward component of the anode wave.

desirable that the adjustment of V_a, necessary to enable electrons to escape, should coincide as nearly as possible with the value of V_a required under condition (*a*). The escape voltage is known as the *threshold voltage.* The conditions for this are dealt with in a later paragraph.

All component waves affect the space charge to some extent, but as all components of V decrease in amplitude towards the cathode, depending on $\left(\frac{r}{r_a}\right)^n$, the contribution made to velocity by low values of n, will be greater than that for large values of n.

Thus the primary forward component, the lowest value of n_0 for the particular mode, will cause the greatest distortion of the space charge cloud, and will be the major effective component in the initiation of oscillations.

Alternative Theory

By some authorities it is thought that the resonant shell is nearer the anode than the shell with which the anode wave is in step. In this case the conditions for building up oscillations are that V_a and H must be adjusted so that the radius of the space charge cloud is large enough to contain the resonant shell. As in the previous case, however, the value of V_a to produce instability in the required mode for a given value of H must exceed a certain minimum. The instability potential reckoned on this basis is somewhat larger than it would be if calculated on the previously mentioned basis.

In either case, however, the deformed shape of the space charge cloud is much the same, and there is in consequence a nett loss of kinetic energy from it.*

Electrical Conditions in the Synchronous or Resonant Cavities

Before considering the conditions under which electrons can escape to the anode (threshold potential), it is necessary to examine the distribution of voltage and current in and about the resonant cavities.

Consider the effect on the anode when a deformed space charge cloud revolves in step with the primary forward component of the anode wave.

At any given instant the hills on the space charge cloud, which are formed by accumulations of electrons, and are

* It is offered as a suggestion that the deformed shape of the space charge may be caused by a piling up of favourably placed electrons, due to retarda-

therefore negatively charged, induce in the anode, at points immediately opposite them, local increases in the steady positive charge due to H.T. Similarly, at points on the anode corresponding to each dale in the space charge cloud, there will be at the anode a local reduction of the charge due to H.T.

Thus, superimposed on the steady positive charge on the anode, due to H.T., alternating charges are induced which are in phase with the alternating potentials of the anode segments.

Let the electrical conditions of the resonant cavities now be examined. Take, for instance, the cavity marked x—y, Figs. 248 (B), 250 and 251.

In Fig. 248 (B) the letters A and B in the segments indicate positive and negative maxima of the trapezoidal standing wave across the gaps and xy is the particular gap under examination. (a_1), (a_2), and (a_3) show the phases of the forward and reflected component waves, in phase at $\frac{T}{4}$ and $\frac{3T}{4}$ and in antiphase at $\frac{T}{2}$. (b_1), (b_2), and (b_3) show the variations and distribution of charges induced in the anode at the above times by the electro-mechanical action of the space charge and the phase relations of this space modulating wave to the anode primary wave. (Fig. 249.)

The curve in Fig. 249 (c) (1) shows the manner in which the potential of x changes relative to y across the gap xy. Since the distribution of the induced charges is such that they move anticlockwise round the anode, a component of I_a must flow at every point on the anode proportional to the amplitude of the corresponding component of instantaneous charge density at the point. Since graphs b_1, b_2 and b_3 in Fig. 248 (B) may be interpreted as current distribution round the anode instant by instant, and assuming I in the anti-clockwise direction to be regarded as positive, then the current I_a flowing from x to y round the cavity xy can be deduced and plotted as in Fig. 249 (c) (2).

When x is positive with respect to y, the direction of the electric field intensity across the gap E is as shown in Fig. 251, being reckoned as positive in the direction of the arrow.

When I round the cavity is positive the direction of the magnetic field is as shown by the arrow head, *i.e.*, out of the paper, Fig. 251, therefore if H be reckoned positive in this

tion by the major anode wave at different rates for orbital radii. In which case waves may be raised on the surface of the space charge, rotating in step with the primary forward component of the anode wave.

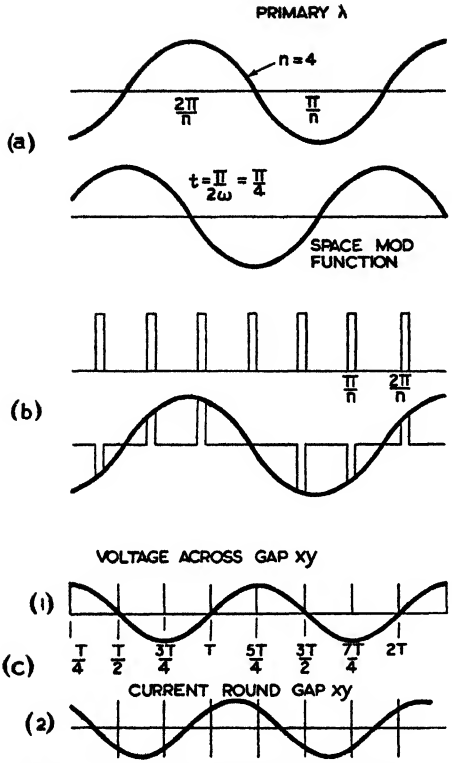


Fig. 249.—(a) ILLUSTRATES THE EFFECT OF THE ELECTROMECHANICAL ACTION OF THE SPACE CHARGE IN MODIFYING THE INTENSITIES OF CHARGES IN THE ANODE. (b) GAP CURRENT FOR FORWARD WAVE.

(c) (1) is voltage across gap xy developed from (a_1), (a_2), (a_3), Fig. 248 (B).

(c) (2) is current round gap xy developed from (b_1), (b_2), (b_3), Fig. 248 (B).

direction the graph shown in Fig. 249 (b) may be taken to represent the magnetic field intensity.

If the cavity is resonating without gain or loss of energy the time phase relations of E and H are in quadrature. Superimposed on the magnetic field due to resonance, however, there is an additional magnetic field due to the currents resulting from charges induced by the rotating space charge cloud. From the graphs in Figs. 249 (c), it can be seen that H lags E by a phase angle less than $\frac{\pi}{2}$ so that H has components which are in quadrature and others which are in phase with E. This means that the electromechanical system of the magnetron involving the rotating space charge) modifies the resonating field in the cavity and hence affects the frequency of oscillation. *This is the reason why the magnetron frequency is so sensitive to changes in V_a* (Fig. 249 (a) and (b)).

Thus due to variations in the values of $\frac{\omega}{n}$ each mode will cause a slightly different frequency to be generated.

The in-phase component indicates a flow of energy and it can be seen that a power flux is directed into the cavity as indicated by the right-hand system for E, P and H, Fig. 251 (b).

Summary

When reading the summary that follows it should be borne in mind that there is still a great deal to be learned about the magnetron. The explanations offered are simplified versions of conclusions reached by eminent authorities, consequently there are, inevitably, discrepancies which do not matter very much from a practical point of view, but which may be rather wide of the mark from a purely academic standpoint.

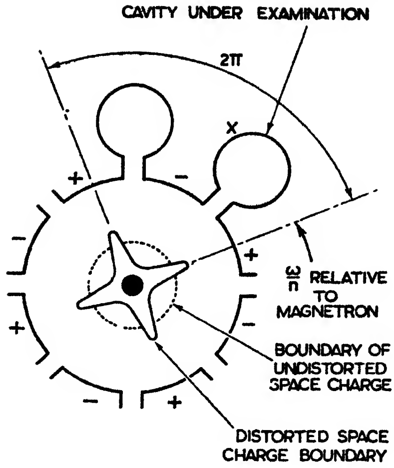


Fig. 250.—DETAILS OF THE GAP AND CAVITY XY UNDER CONSIDERATION.

(1) In normal practice the values of V_a and H are adjusted for any given type of magnetron so that it operates well below the Hull cut-off condition and in this state electrons cannot escape to the anode without the help of an R.F. field.

(2) In the absence of R.F. on the anode, electrons revolve round the cathode at speeds which increase with the radii of their respective orbits or shells.

(3) This state is not very stable, since any small change of anode voltage affects electrons in the space charge and, *if large enough*, causes the shells to lose symmetry, thereby causing electrons to be retarded, pile up and thus create hills and dales in the space charge.

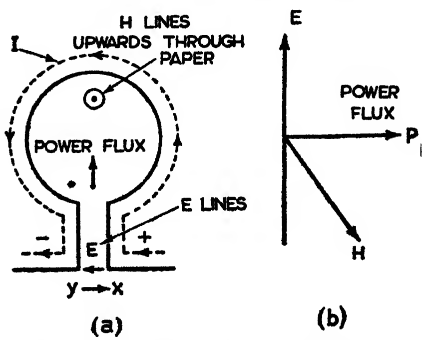


Fig. 251.—PHASE RELATIONS OF E AND H IN CAVITY.

(a) Direction of power flux into cavity xy.

(b) Vector relations of E, P and H governing the direction of power flux into xy.

(4) R.F. at the anode results from any disturbance which shocks the resonant cavities into electrical oscillation. When this occurs, a pattern of voltage is set up on the anode face, with standing

waves of voltage across the gaps. The forward component of the primary sinusoidal standing wave has a major effect in exciting anode oscillations and travels round the anode face in the same direction as the electron shells revolve round the cathode.

(5) If the amplitude of oscillation due to the initial disturbance is *large enough* to cause deformation of the space charge, the anode wave locks to some shell in the space charge which has approximately the same angular velocity and the space charge deformed by the action of the anode wave revolves in step with it.

(6) As a result of operation (5), some of the electrons in the space charge are retarded and others are accelerated, so that if the radius of the space charge boundary (as fixed in (1)) is *large enough*, nett energy is given up by the space charge to the resonant cavities and the initial oscillation is built up.

(7) The mode in which these oscillations are built up for a given magnetron may be thought of as being predetermined by the initial values chosen for V_a and H within the requirements of paragraph (1).

(8) The potential at which the state of instability occurs decreases as n diminishes, since the shells that excite the lower values for n are nearest the cathode.

(9) Thus for any given geometrical design of anode and cathode, operation in the desired mode is predetermined by the initial adjustment of V_a and H, within the requirements of paragraph (1), but in such manner that the conditions for the sequence of operations (4), (5), (6) and (7) are also fulfilled.

Threshold Voltage

Paragraphs (1) to (9), inclusive, are concerned solely with the requirements which must be met in order that oscillations may be excited and built up in any desired mode. In order that oscillations, when established, may be maintained, circuit losses and energy supplied to the load must be replenished from the H.T. source, since the energy transferred to the resonant cavities by the electromechanical action of the valve is insufficient for this purpose.

In order that the H.T. may be enabled to function in the above manner, there must be a nett loss of electrons from the cathode to the anode, so that the H.T. may replenish this loss.

Since the anode current of the valve is initially cut off, see paragraph (1), electrons leaving the cathode to take up their positions in the space charge cloud (prior to the onset of R.F.

oscillations at the anode), receive insufficient potential energy from the radial electric field, to enable them to reach the anode for the given value of H; consequently (after the onset of oscillations has deformed the space charge, in course of which the R.F. field absorbs energy), additional radial velocity must be imparted to the retarded electrons from some source other than the R.F. field to enable them to reach the anode.

It has been shown that the nett effect of R.F. on the anode in deforming the space charge is to absorb kinetic energy from favourably placed electrons, and to weaken the deflecting action of the magnetic field. Consequently electrons, acted upon in the above manner, are enabled to acquire a *component* of radial velocity from their angular velocity $\frac{\omega}{n}$ which, if sufficiently large, enables them to reach the anode.

Comparison of Conditions for Instability and Threshold

Consequently it follows that the requirements for escape of electrons from the deformed space charge to the anode are not identical with those for excitation and build-up of oscillations in a given mode. Wherefore, according to the position of the resonant shell, the threshold or escape voltage required may be greater or less than the instability voltage, as determined by the position of the resonant shell from the cathode (the mode excited). Thus conditions may arise where the R.F. on the anode is sufficient to provoke oscillations in the desired mode, but owing to the situation of the resonant shell, electrons do not escape from the space charge or do not acquire a component of radial velocity sufficiently large to enable them to reach the anode. The reverse condition may also arise.

Clearly, therefore, it is essential for the initial values of V_a and H to be so chosen as to satisfy (*a*) the requirements for excitation and build-up of oscillations in the desired mode, and (*b*) the requirements for that minimum value of threshold voltage, which permits the electrons just to reach the anode. This means that for a given design of magnetron and a given mode of *excitation* there is just one radius for the resonant shell that satisfies conditions for both instability and threshold and therefore the values of V_a and H are fixed for optimum efficiency in any given mode.

It has been suggested that a mental picture can be formed of the manner in which electrons escape to the anode by

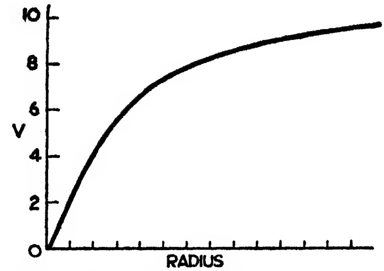


Fig. 252.—DISTRIBUTION OF POTENTIAL IN THE CATHODE-ANODE SPACE.

the assumption that favourably placed electrons, after experiencing a series of retardations during a brief sojourn in the electro-mechanical space charge and having, therefore, surrendered to the R.F. field the greater part of the energy they receive from the radial electric field (H.T.), experience successive weakenings of the deflecting force of the magnetic field. At some point, as the influence of the magnetic field deteriorates, these electrons are literally " flung out " of the space charge by an action similar to that of a centrifugal force, $m\left(\frac{\omega}{n}\right)^2 r$. This force must impart on expulsion a component of radial velocity to the electrons, sufficient in magnitude to enable them to reach the anode along a spiral path. A condition of velocity resonance is thus established and maintained by this electron stream.

When electrons strike the anode, any excess of kinetic energy which they possess in the form of radial velocity is transformed into heat. It is, therefore, desirable, from the point of view of overall efficiency, that the radial velocity component imparted to electrons on leaving the space charge should be only just sufficient to enable them to reach the anode.

The relation of instability potential to threshold potential is shown for several values of n in Fig. 254 for an unstrapped magnetron. (Strapping and its effects are dealt with in subsequent paragraphs.)

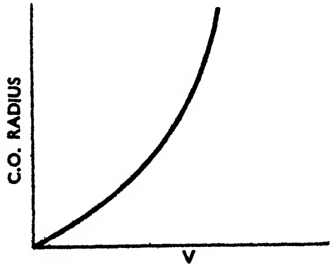


Fig. 253.—INCREASE OF RADIUS OF CUT OFF WITH INCREASE OF VOLTAGE FOR FIXED VALUE OF H.

If this potential is too low, electrons will not be able to escape to the anode at all; on the other hand, if it exceeds the critical value :—

(*a*) Electrons will strike the anode with a surplus of radial velocity, generating an unwanted amount of heat, thereby lowering the overall efficiency ;

(*b*) If the threshold potential

is increased much above the critical value, an unsuitable mode may be excited, in which case the wavelength must alter and the magnetron will operate inefficiently.

Choice of Mode

Consider Fig. 254. Let some value of H be fixed, then in moving upwards from R in the direction of S, increasing values of V_a are encountered.

Proceeding then from R, it is seen that for values of $n = 7$ and $n = 6$ in that order, the instability co-ordinate is reached and crossed in each case, just prior to that for the threshold voltage. Continuing in the upward direction, the conditions for $n = 5$ are reversed, and the threshold voltage co-ordinate is reached and crossed in advance of the instability value.

Advancing still further, it is seen that in the conditions for $n = 4$ the above phenomenon is still more marked; the threshold potential for this mode is reached just above the point P, whilst the instability potential for the mode is not reached until the point T.

For the particular valve and for the conditions given, it is clear that it should be operated, for maximum efficiency, in the $n = 5$ mode. Modes $n = 6$ and $n = 7$ can co-exist, but $n = 5$ being nearest the cathode, is the most effective in initiating persistent oscillations. It can be said therefore that the mode $n = 5$ is the mode in which oscillations are most likely to persist.

Were it desired to operate this valve in the $n = 4$ mode, it can be seen from Fig. 254 that it would be necessary to increase the minimum value of V_a for the threshold potential (just above point P) to point T in order that the instability potential for the $n = 4$ mode for this valve can be attained. Hence, although for mode $n = 4$ conditions as shown by Fig. 254 are suitable for electrons to escape to the anode at point P, oscillations would be excited, built up and maintained in the $n = 5$ mode at reduced overall efficiency, the latter being due to excessive amount of heat generated by electrons with surplus radial velocity.

This means that if the $n = 4$ mode must be used, the value of V_a required for the instability potential is greatly in excess of that required for the threshold potential, in consequence of which electrons would again arrive at the anode with a surplus of radial velocity. The kinetic energy proportional to this excess would be transformed into heat, thereby leading to

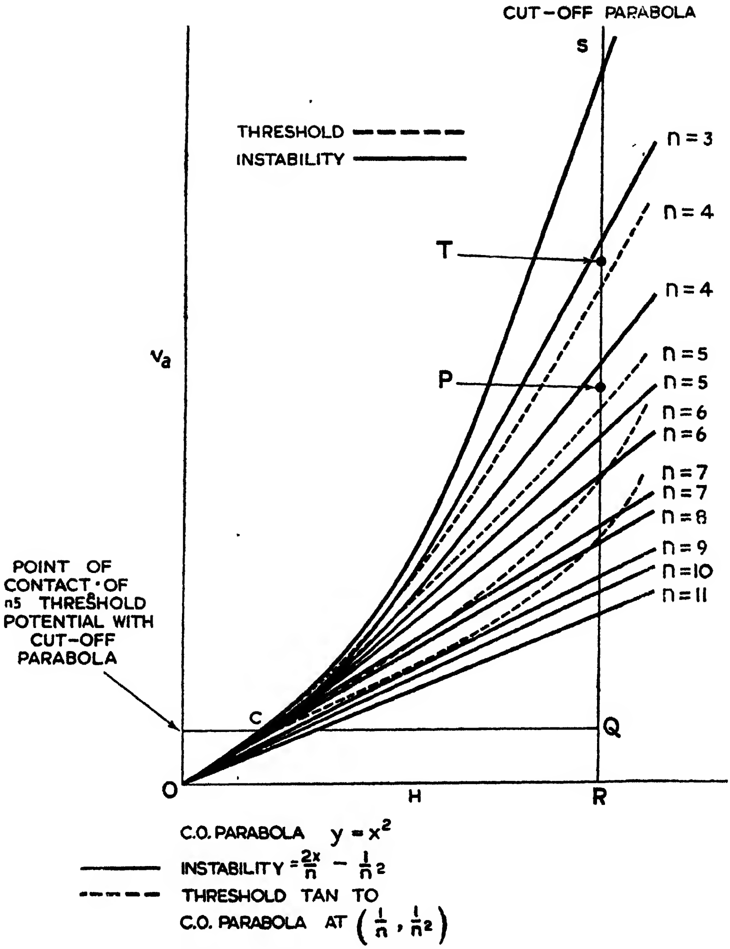


Fig. 254.—THRESHOLD CONDITIONS PLOTTED TO SATISFY $n = 5$, SO THAT OSCILLATIONS (IF INITIATED) CAN BE MAINTAINED. THIS APPLIES ALSO TO $n = 6$, $n = 7$, ETC. BUT DEFORMATION OF SPACE CHARGE IS GREATEST FOR SMALL VALUES OF n, THEREFORE $n = 5$ IS MORE LIKELY TO BE INITIATED THAN $n = 6$, ETC.

Note. If V_a exceeds the threshold value for a given mode, further increases cause operating point to cross threshold and instability lines, therefore mode changes accompanied by small change in frequency. More than one mode can exist at one time in which case an undesirable frequency appears in the output. These curves are for an unstrapped magnetron.

Since the ordinate for the point of contact of the threshold line for $n = 5$ is $\frac{1}{n^2}$ it can be shown that the operating point for maximum efficiency is $\frac{PR-QR}{PR}$.

overall inefficiency of the valve when operated in the unsuitable $n = 4$ mode.

If V_a is increased above the point T it is clear that the valve will operate in yet another mode, *i.e.*, $n = 3$. Still further increasing V_a will lead to conditions where oscillations must cease altogether, since the cathode-anode space will, at some value of V_a, be ultimately filled entirely by the space charge cloud.*

To fix the relationship of the instability and threshold potentials it is convenient, if not strictly accurate, to think of a triode connected as an oscillation generator. Oscillations will not commence until the conditions for feedback are complied with, particularly in regard to phase relationship. This condition is analogous to the instability potential requirements of the magnetron. Similarly, oscillations cannot be sustained unless the feedback is sufficiently large to make good circuit losses. This condition is analogous to the threshold potential.

The Magnetron as a Fixed Frequency Valve

Since the angular velocity of the anode wave decreases with decreasing values of n, and because the angular velocity of the electron shells round the cathode varies inversely with their radii, it follows that for a fixed value of H, the radius at which the resonant shell occurs in a given magnetron is determined by V_a.† Wherefore it again follows that the magnetron, as known at present, is, within very narrow limits, a fixed frequency valve. The mode and frequency at which a given magnetron operates efficiently is determined by its design, *i.e.*, the geometry of anode and cathode, the axial length of the anode, the distance between the anode and end plates together with the measures taken to ensure *electrical asymmetry* of the anode (strapping, etc.). Also it must be adjusted so that for certain definite values of V_a and H, the radius of the resonant shell which locks to the anode waveform, coincides, as nearly as possible, with the radius at which electrons can escape to the anode and arrive there with a minimum of surplus radial velocity.

* It may be seen that owing to oscillation conditions anode current can flow for values of V_a less than the value of V_a (cut off). This affords an explanation of the shaded portion of the curve (Fig. 232). This is only true for unstrapped magnetrons.

† The relations regarding V_a and H are less stringent for strapped magnetrons. The above remarks apply to unstrapped magnetrons.

Scaling

It should now be noted that, like the triode, a good design for a magnetron, at some given frequency or wavelength, may be used as the basis of design for another, having corresponding characteristics, but operating at a different wavelength. This can be effected by changing all dimensions in the ratio of the new to the old wavelength, the only limiting factors being the voltage gradients, emission and power dissipation.

Efficiency

The efficiency for operation under any given condition can be obtained from diagrams similar to Fig. 254, constructed for each magnetron.

The energy supplied by H.T. when an electron travels from cathode to anode is eV_a, electrons on reaching the anode, revolve with the angular velocity of the anode waveform. The kinetic energy due to this is $\frac{1}{2}mv^2\left(\frac{\omega}{n}\right)^2$. The electron also has kinetic energy due to radial velocity, but on arrival at the anode under ideal conditions this approaches zero and can therefore be neglected.

The difference between energy supplied by the source and the energy converted to R.F. is :—

$$\text{Efficiency} = \frac{eV_a - \frac{1}{2}mv^2\left(\frac{\omega}{n}\right)^2}{eV_a}.^*$$

Strapping

It has already been shown that in an unstrapped magnetron, particularly, several modes, all having slightly different wavelengths may co-exist.

If the frequency discriminating properties of the magnetron *load* are insufficient to reduce the unwanted modes to insignificance, unwanted frequencies of some importance will appear in the output.

In the diagram for the particular unstrapped magnetron, shown in Fig. 254, the operating point is fixed as near the $n = 5$ threshold value as possible in order to obtain maximum efficiency. This means that whilst by no means predominant, amongst others, the $n = 6$ and the $n = 7$ modes are both

* But neglects loss at the cathode due to bombardment by unfavourably placed electrons.

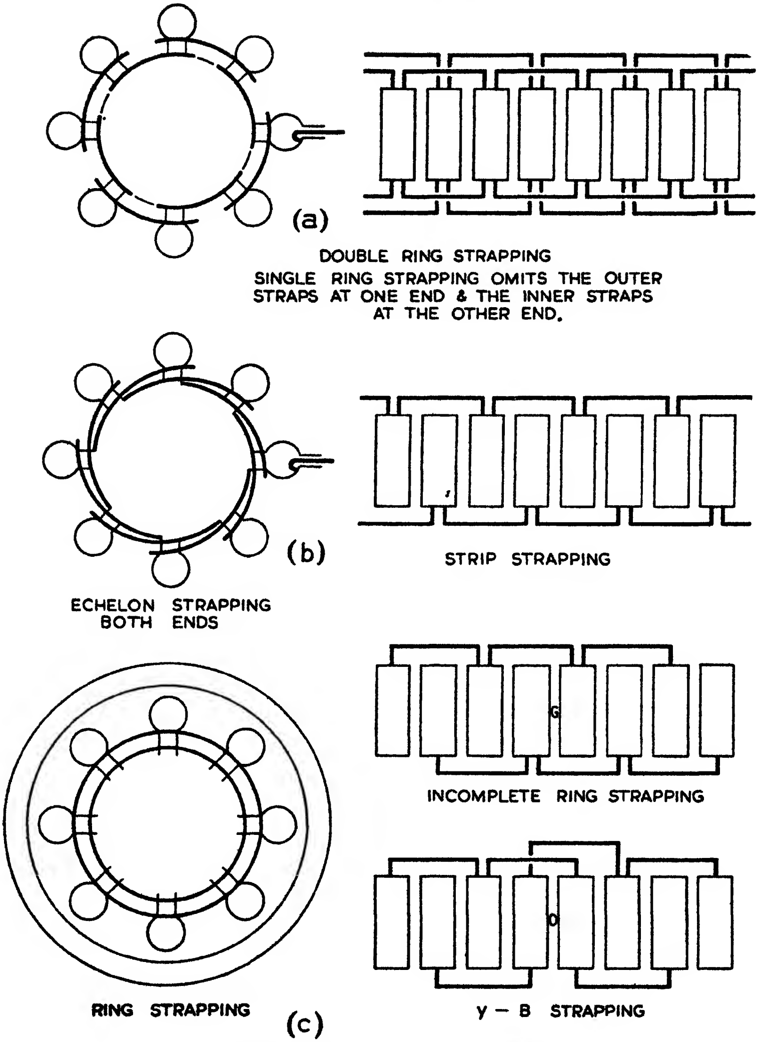


Fig. 255.—STRAPPING SYSTEMS.

present and, therefore, in contributing to the output, make for overall inefficiency.

It is obviously desirable therefore to eliminate as far as possible all modes other than the particular one required. This function may be performed, to a greater or less degree, by one or other of the various arrangements for strapping segments together, the idea being first suggested by Dr. J. Sayers.

Alternate methods of strapping are shown in Fig. 255 (a), (b) and (c).

In the case of an 8-segment magnetron when short conductors, as nearly as possible $\lambda/2$ in length join two segments, separated by a third segment, the π mode is favoured since a phase difference of π between strapped segments is implied. With this arrangement the phase difference for the $n = 5$ mode would be π, consequently this mode cannot exist under the alternate segment system of strapping shown in Fig. 255 (a) and (b).

Strapped valves can be designed to operate over a wide range of inputs for the π mode, and since this mode is the easiest, and in many cases the only one that can be made efficient, most valves are now designed for π-mode operation.* The unstrapped magnetron, diagram for which is shown in Fig. 254, normally operates in the $n = 5$ mode. When strapped, however, it will operate most efficiently in the $n = 4$ mode.

Strapping and Asymmetry

It would appear that in early models all the segments were strapped together in one or another of the various arrangements available, and electrical conditions, although modified by strapping, still remained symmetrical around the anode. Later it was found that the introduction of some measure of asymmetry into the èlectrical conditions of the anode could be employed, to discriminate still further against unwanted modes.

This led to breaks in the straps, particularly in the omission of straps across the resonant element in which the output coupling loop is situated, and immediately adjacent to the input to the cathode. In certain cases other straps are also omitted but, in general, knowledge available at the present stage of development appears to be insufficient to enable the ultimate strapping pattern for any particular type to be forecast accurately. The final arrangement is, therefore, at present, determined by experiment.

Straps cannot be regarded as true short circuits. It can only

* Strapped magnetrons are less frequency sensitive to changes of peak pulse voltage, and in practice appreciable changes in V_a may be made without affecting, seriously, the performance of strapped valves.

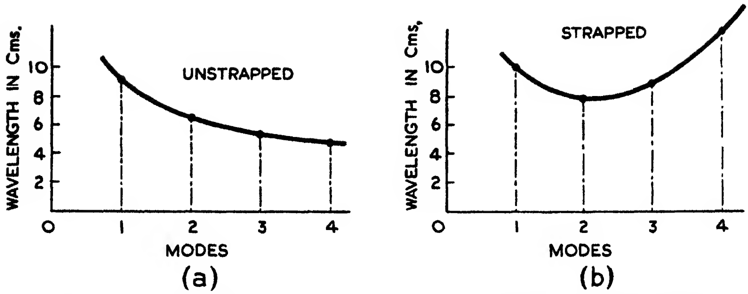


Fig. 256.—COMPARISON OF FREQUENCY SEPARATION BETWEEN UNSTRAPPED AND STRAPPED MAGNETRONS.

Comparing (a) and (b) note that 1 is slightly depressed but 3 and 4 are noticeably accentuated, *i.e.*, the overall separation is improved.

be said of them that they favour the establishment of some modes and discourage others, without necessarily eliminating them entirely.

Strapping and Frequency

Each strap, being a conductor, has inductance, and since it passes over segments between those which it connects together, some capacitance must exist between it and this segment. *Thus a strap may alter the frequency at which a particular mode arises.* It can be shown that λ^2 varies linearly with strap capacity.

In the π mode particularly, if the length of the strap is $\frac{\lambda}{2}$, both ends of the strap are at the same potential and carry no current, excepting for a small capacity component. Strap inductance is therefore negligible. The capacity effect is, however, not negligible, since the frequency is modified by the capacity due to the relative positions of the strap and the anode block.

If, for example, the mode n = 5 arises there is current through the strap and, therefore, the strap has inductance and capacity and the frequency is still further modified.

Mode Separation

The significant effect of strapping is seen by comparison of the exciting wavelengths necessary to produce the various patterns of Fig. 256 (a) and (b) for unstrapped and strapped

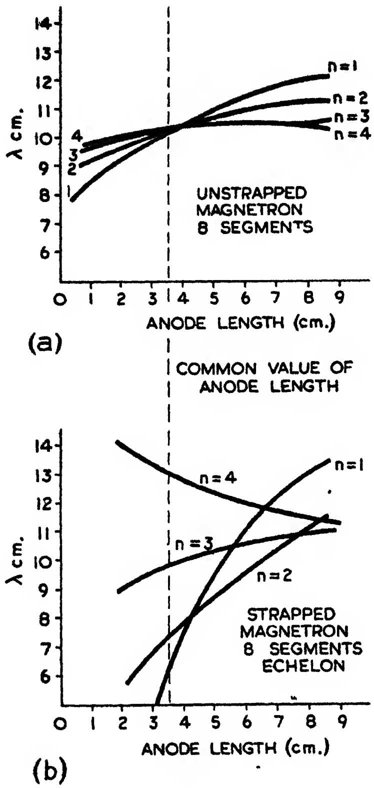


Fig. 257.—(a) AND (b) PLOTTED TO SAME SCALES AND SHOW THE FREQUENCY SEPARATION FOR VARIOUS ANODE LENGTHS.

The similar scales enable frequency separation to be compared for unstrapped and strapped magnetrons and for different lengths of anode.

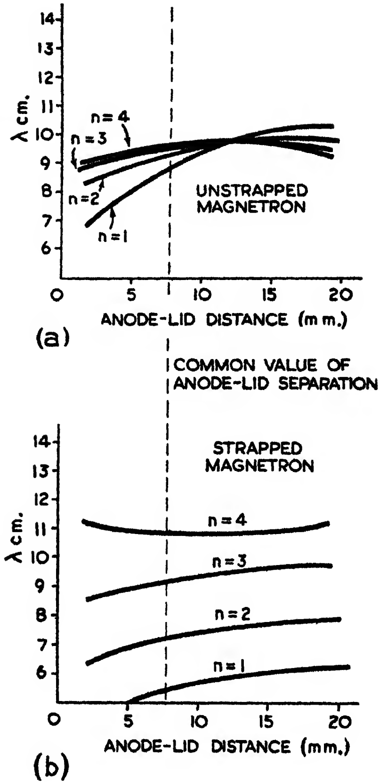


Fig. 258.—ANODE LID DISTANCE PLOTTED FOR SAME SCALE IN (a) AND (b) ENABLING UNSTRAPPED AND STRAPPED MAGNETRONS TO BE COMPARED FOR FREQUENCY SEPARATION FOR VARIOUS VALUES OF ANODE LID DISTANCE.

anodes. It will be seen from Fig. 256 that the effect of strapping is to increase, considerably, the resonant wavelengths of the higher frequency modes whilst tending to reduce those of lower frequency, thus increasing the separation of the modes in frequency.

Conditions for the magnetron as a whole, and in particular the selectivity of the load, should be adjusted so that matching is achieved at the frequency of the wanted mode—normally the lowest (*i.e.*, $n = 4$ for the 8-segment magnetron and $n = 5$ for a 10-segment magnetron).

Note. The reason why the efficient operating point for mode $n = 5$ for the unstrapped magnetron in Fig. 254 does not apply when it is strapped now becomes clear. Changes in wavelength which take place due to strapping alter the threshold potential diagram, so that $n = 4$ now becomes the efficient mode for operation.

Separation of modes in frequency is found to be accentuated by asymmetry in the magnetron anode structure, and this, as already stated, accounts for the omission of certain straps and other devices to produce similar results.

Other Important Factors of Design

Other important factors that influence the separation of modes are : (*a*) anode length, and (*b*) distance of the end plate from the anode block. Figs. 257 (a) and (b) and 258 (a) and (b) show the effect of these factors upon mode separation.

Static Characteristic

Fig. 259 shows magnetron characteristics drawn on cartesian co-ordinates, peak voltage being plotted against peak current, superposed are contours of constant applied values of H and contours of constant R.F. peak power output. Efficiency contours may also be included.

Consideration of the curves shown in Fig. 259 shows that the H contours are fairly straight and that in order to obtain high efficiency it is necessary to employ high peak voltages and large values of H. Under these conditions the impedance presented to the source of anode volts is high and the frequency instability is great.

Also, for constant values of H, and increasing values of anode voltage, the efficiency increases to a maximum and then falls.

Input Conditions and Requirements

From Figs. 238 and 239 it can be seen that the connections to the heater filament are brought out through two glass seals and extended by flexible copper leads connected to the pillars which carry the filament. The cathode is connected internally to the filament legs.

The magnetron current is measured by means of a current transformer. The primary of this transformer is connected in the anode lead to earth, whilst the secondary is connected

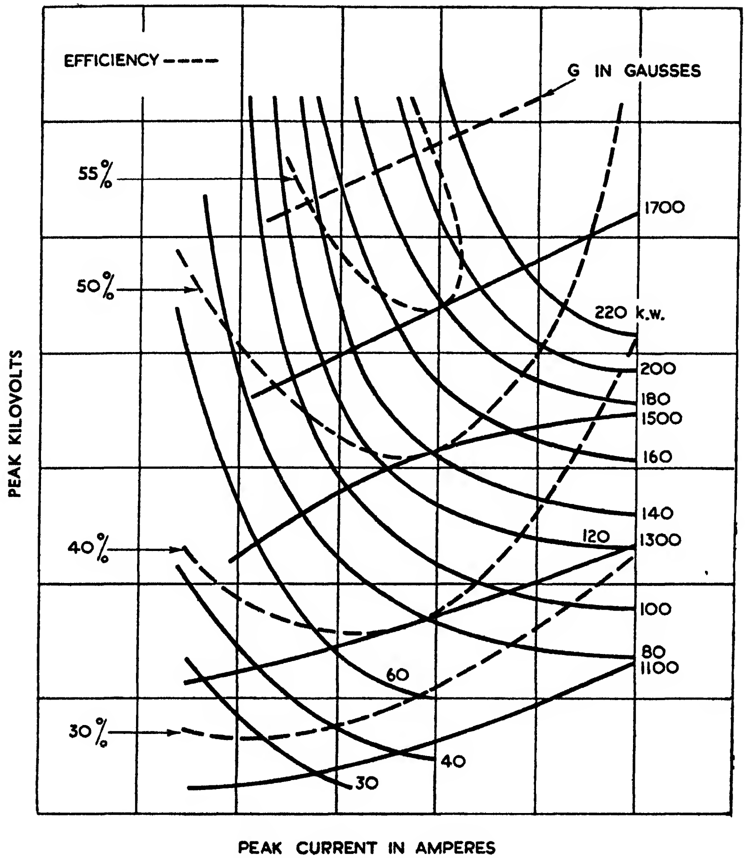


Fig. 259.—MAGNETRON STATIC CHARACTERISTIC.

to a monitoring point, across a suitable resistance network. The measuring instrument, usually a cathode ray tube, is calibrated to give a deflection proportionate to a given number of volts across the secondary for each ampere of current in the primary.

The filament supply may be about $1\frac{1}{2}$ amps. and the insulation between primary and secondary of the filament transformer has to withstand the maximum peak pulse voltage.

(1) It has been shown that magnetron frequency is very sensitive to changes in anode voltage ; consequently, in order to avoid changes in frequency, efficiency and output power, it is

essential that pulses applied to the input should have a constant value of voltage for the duration of the pulse.

(2) The magnetron shows a variable impedance for different values of pulse voltage across it. In its oscillating state, with the normal working voltage across it, the impedance of a certain type may be about 1,100 ohms, the ratio of the pulse transformer must therefore be chosen to match this impedance to the characteristic impedance of the coaxial cable which carries the pulse from the modulator to the primary of the pulse transformer.

(3) Also, from the point of view of the input characteristic, it is desirable to have as many segments as possible.

In order to meet the conditions outlined under paragraph (1), the input pulse of voltage from the modulator must have a steep leading edge, a flat top and a trailing edge with a rapid fall.

A steep rise and fall is essential in order to avoid energy waste and overall inefficiency due to oscillations in unwanted modes, which are set up when a change takes place in the anode volts, such as might take place at the commencement of a pulse the rise of which is insufficiently steep.

V_a must be maintained as steady as possible by each applied voltage pulse for its duration, since a change in value seriously affects the conditions of oscillation. This, all other things being equal, will depend upon the regulation of the source from which the H.T. pulse is drawn.

An artificial line is generally used to store the energy for each pulse and to control the rate at which this energy is given up to each pulse. The suitability of the pulse shape depends almost entirely upon the goodness of the artificial line, or how nearly its characteristics approach to those of the real line which it simulates.

From the impedance-current curves in Fig. 260 it can be seen that at the beginning and end of each pulse the magnetron presents a very high impedance and, therefore, completely mismatches the pulse line and the pulse-forming device (generally an artificial line).

Unless correcting circuits are applied there are, therefore, reflections of the initial wave front up and down the pulse cable, and the pulse cable being mismatched looks like a capacity to earth.

This effect lets the initial current pulse rise very sharply, shocking the oscillatory circuits formed by the inductance and

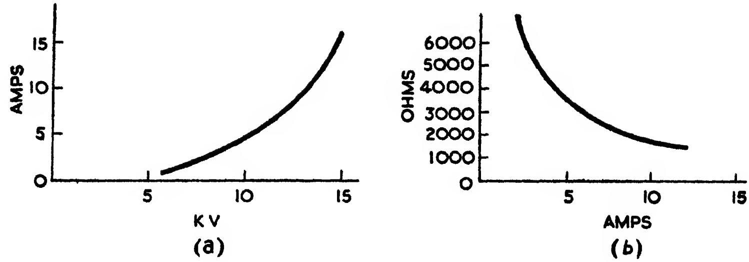


Fig. 260.—MAGNETRON IMPEDANCE AT BEGINNING AND END OF EACH PULSE.

(a) Magnetron current plotted against peak pulse voltage (V_a).
(b) Impedance plotted against magnetron current.

distributed capacity of the pulse transformer into oscillation. An oscillation of frequency of about 2 megacycles appears on the flat top of the voltage and current pulse. The amplitude of these oscillations increases when the length of the pulse cable is decreased, since a long length of cable has a damping effect due to resistance.

A compensating circuit may be incorporated in the modulator, consisting of a suitable resistance and capacity across the transformer. Before the pulse the capacity is uncharged, and during the rise of the pulse across the transformer the capacity charges through the resistance, the value of which is chosen to match the artificial line. By the time that the condenser has charged, the magnetron is conducting and the pulse cable is matched at both ends.

Pulse Transformer

Oscillations due to mismatch at the end of the pulse are not serious provided that the voltage on the magnetron cathode does not go negative again, causing it to conduct. To avoid this, the inductance of the pulse transformer is chosen so that the voltage at that point swings sufficiently positive to prevent the appearance of any negative peak.

If it is assumed that all impedances are correctly matched and that there is no energy dissipated by the pulse transformers or by the cable, then the following voltages would be obtained for the magnetron connected as in Fig. 261.

Assume voltage to which the discharge line charges = 10·5 kV.

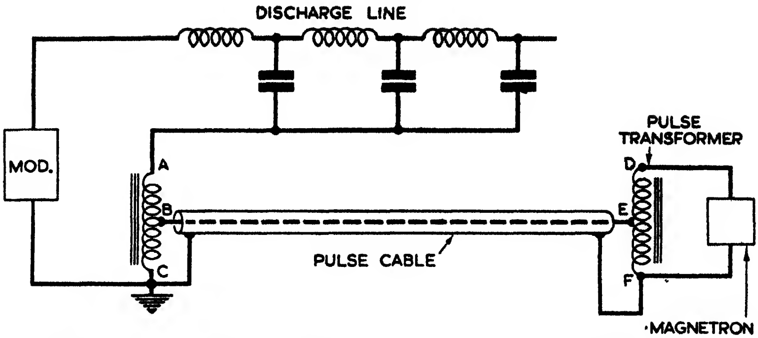


Fig. 261.—DIAGRAM OF ARTIFICIAL LINE, VOLTAGE TRANSFORMER, PULSE CABLE AND MAGNETRON TO ILLUSTRATE THE VOLTAGE DISTRIBUTION ACROSS THE PRIMARY AND SECONDARY SECTIONS OF THE PULSE TRANSFORMER FOR MATCHED CONDITIONS.

Impedance of the discharge line = 150 ohms.
Ratio of voltage transformation = 1 : 1·5
Ratio of windings = 1 : 4

Voltage between A and C $= \frac{V}{2} = -(5{\cdot}25)$ kV.

B and C $= \frac{V}{2 \times 1{\cdot}5} = \frac{V}{3} = (3{\cdot}5$ kV.).

E and F $= \frac{V}{3} = -(3{\cdot}5$ kV.)

D and F $= \frac{4V}{3} = -14{\cdot}0$ kV.

Current Pulling

It has been shown that a variation of the input conditions produces a change in frequency, and since this also results in a change of anode current, the effect is termed "current pulling." The effect on frequency may be of the order of from 0·1 to 1·0 megacycles per second per ampere in the centimetric range.

Variations in anode block temperature also cause a frequency change, occasioned by expansion of the anode block itself. This change is roughly proportional to the change in temperature and the coefficient of linear expansion of copper.

Output Conditions and Requirements

The output from the magnetron is by way of a concentric line system of which the projecting tungsten pin is a continuation of the inner conductor. The outer conductor ends at the *copper to glass seal* and connection cannot, therefore, be made at this point. It is found satisfactory, however, to connect to the external threaded portion, which ends about 1 centimetre before the seal. In certain types the output pin is pushed into the springy end of the inner conductor of a 70-ohm coaxial cable, the outer conductor of which is connected at one end to the threaded portion of the magnetron and at the other end to a flange connected to the middle of the wider side of a rectangular wave-guide.

The inner conductor may cross the guide, and after reaching the other side, continues as the inner conductor of another coaxial line which is short circuited at the far end by a plunger. The function of the plunger has been dealt with in another chapter, but it may be useful at this juncture to recall that it is essentially a tuning adjustment, employed to match the output conditions of the magnetron to the coaxial line or wave-guide.

When a feeder system (either coaxial cable or wave-guide) is coupled to a magnetron, there are two resonances to be considered, *i.e.*, the internal resonant circuits of the magnetron anode and the output line, the two being coupled together by the loop or magnetic pick-up. When matching * has been achieved, the magnitude and phase of the reflected portion of the forward wave are as small as the nature of the load permits. The magnitude depends upon the value of the reactance component of the load and the phase is also determined by this quantity and whether such reactance is inductive or capacitative in character (see Appendix IV).

Since in the matched condition the line is deemed to be terminated in a load matched to Z_0, but which may have either inductive or capacitative reactance, any change in the impedance of the load or of its composition must cause a change in the reflected portion of the forward wave. This appears as a change in the magnitude and phase of the standing wave ratio (S.W.R.). Since a change of phase in the line or guide is seen by the coupling loop as a change of frequency in the line or guide, the effect is to pull the frequency of the magnetron slightly off its natural resonant frequency, bringing about thereby a reduction in efficiency of operation and consequent loss of power output.

* Matching is usually understood to mean that the load is resistive and equal to the characteristic impedance of the line. A radar load is nearly always reactive.

Causes of Frequency Pulling

In addition to the loss of power involved, a change due to frequency pulling may affect, seriously, receiving conditions, more particularly is this the case if the receiver is not designed for automatic frequency control, or if the A.F.C. provided is not capable of handling a frequency swing because of its magnitude.

It is therefore desirable that a high degree of frequency stability should be achieved in order to avoid loss of power and adverse effects on receiving conditions; consequently the factors must now be examined which cause the impedance of a radar load to change.

As aerials are rotated for scanning purposes, a varying amount of energy is reflected from the windows of the aerial house, from parts of a ship, or from large objects in the immediate vicinity of the aerial. These varying amounts of energy are reflected back through the wave-guide system to the magnetron, where they produce effects similar to those which would be produced by varying the reactive components of a physical load terminating the wave-guide or coaxial feeder. The effect on the magnetron of changing the reflected portion of the forward wave is to pull its frequency slightly away from the true resonant frequency. This effect is known as "frequency pulling."

Reflected energy resulting from waves of the H mode finding their way into a circular guide, when the latter is used in conjunction with a rectangular guide to form a rotating joint, and similar causes, may alter the value of the reflected portion of the forward wave and the phase angle which it makes with the forward wave.

In order to minimise this effect as much as possible, suitable filters are inserted in the circular guide to exclude the H mode. These filters may take the form of irises (see Appendix II).

Need for Care in Tuning

In matching the feeder system to the magnetron impedance, unless the proper procedure is followed, a state of resonance or matching may result in which the power output is far from the optimum.

If in the process of tuning the feeder system to the magnetron frequency, changes made to the impedance of the feeder system pull the magnetron frequency away from the true resonant frequency, the final result may be a state of resonance, it is true, but not at the natural resonant frequency of the magnetron.

In this state, overall efficiency and power output are both below the optimum.

For this reason pre-plumbing is favoured, since the feeder system with its T.R. switching devices, etc., can then be carefully calculated, designed and manufactured for a frequency to match the resonant frequency of the magnetron with which it is to be employed.

Frequency Stability and Automatic Frequency Control

The relationship between the frequency stability of the magnetron and the receiver design (with particular reference to A.F.C.) can be seen from consideration of the following case :—

Suppose that an aerial system, in scanning, constitutes a variable termination to the line or guide, thereby causing the standing wave ratio to vary between 1 and 1·5. This will cause variation in the magnetron frequency. Suppose also that the design of the receiver A.F.C. is such that it cannot handle a frequency change of more than $\pm$ 5 megacycles at the rate of scan of the aerial system, *then it is necessary that the magnetron be operated under such conditions that any change in standing wave ratio, due to varying load, produces a frequency change of not more than* 10 *Mc/s. in the transmission carrier frequency.* This consideration leads to the conception of a *pulling figure.**

Pulling Figure

The pulling figure is defined as the difference between the highest and the lowest frequency obtained as the load is varied in any possible manner, to produce a standing wave ratio between maximum permissible value (as determined by the maximum frequency change that the A.F.C. circuit of the receiver can handle), and the reciprocal of the maximum permissible value. The expression relating the standing wave ratio to that part of the forward wave which is reflected is

$$\text{S.W.R.}\dagger = \frac{1 - r}{1 + r}.$$

r = that portion of the forward wave that is reflected, *i.e.*, the reflection coefficient (see Appendix IV).

Normal values for pulling figures are from 10 to 15 Mc/s.

* The pulling figure may be thought of as a tolerance imposed upon the frequency stability of the transmitter by the receiver associated with it, the magnitude of this tolerance being determined by the upper and lower limits of frequency variation which the receiver can handle.

† In the United States, standing wave ratio is taken $\frac{\text{voltage min.}}{\text{voltage max.}}$, in this country S.W.R. $= \frac{\text{voltage max.}}{\text{voltage min.}}$

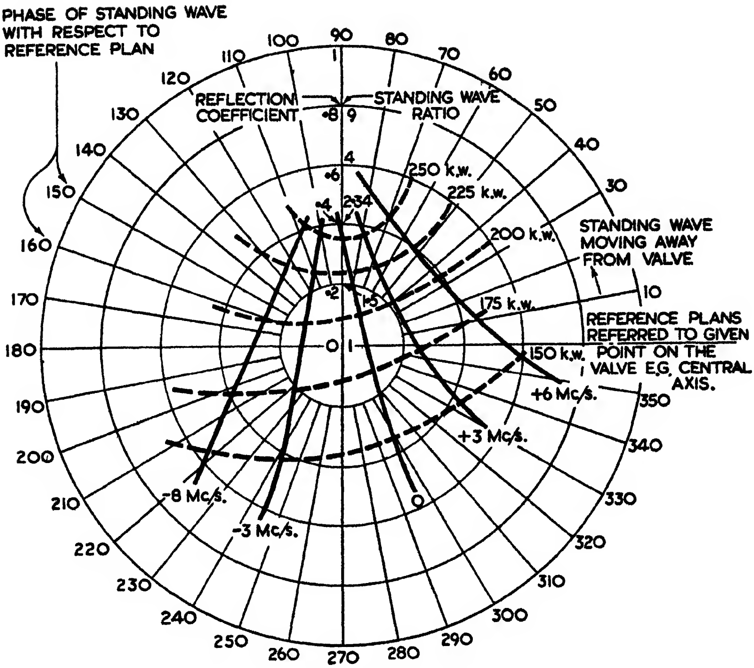


Fig. 262.—TYPICAL LOADING DIAGRAM.

Loading Diagrams

Loading diagrams show how changes in loading effect changes in frequency for various outputs and efficiencies. Thus the operating condition of a magnetron and the efficiency at which it can be operated depend upon the characteristics of the magnetron and its load.

One form of loading diagram is shown in Fig. 262. In this method of representation contours of *constant power* or *efficiency* and *constant frequency* are drawn against a polar co-ordinate system. The " r " co-ordinate represents the fraction of power reflected from a discontinuity placed in the matched output line, and the angle varies with the standing wave set up by the discontinuity. These conditions repeat themselves when a discontinuity is moved half a wavelength along the line and a variation of 360° corresponds to this movement.

The performance of a load system under operational conditions may always be specified in terms of r and θ (see following note), and with this loading diagram for the valve the variables in wavelength and power, when the valve is coupled to the load, is readily predicted.

Note. It has been pointed out in the Appendices and elsewhere in the book that orthodox methods of determining, in terms of voltage and current, the distribution of electrical energy at ultra-high frequencies in lines and wave-guides is impracticable generally. Therefore, investigation is usually conducted in terms of the electric and magnetic fields. Whilst this does not alter the fundamental conception of impedance as being the ratio of current to voltage, it is not so easily applied. In consequence of this, therefore, matching problems at ultra-high frequency require different treatment.

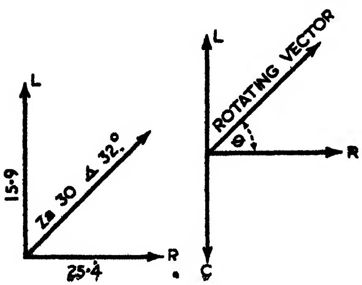
In ultra-high frequency measurement, the ratio between the electric and magnetic fields is usually determined by an examination of the standing wave which is always present, when a line or guide is not terminated by a purely resistive load matched to the characteristic impedance of the line or guide.

By the use of a slotted line in the case of a coaxial cable or some similar device in the case of a wave-guide, the standing wave ratio can be measured. With the help of charts or tables and a knowledge of the characteristic impedance of the line or guide, it is possible to determine the ratio of load impedance to line impedance. Since the line impedance is known, the load impedance can be determined when the above-mentioned ratio has been found. The portion of the forward wave which is reflected and its phase relation to the forward wave can also be determined, thus the load impedance may be expressed in terms of resistive and reactive components and a phase angle, alternatively, in terms of the fraction of the forward wave that is reflected and a phase angle. Thus :—

Let $Z_0 = 40$ and let the ratio of load impedance/line impedance $= Z_a/Z_0$ (as determined after measurement of the standing wave) = say ·75,

$$i.e., \frac{\text{load impedance } Z_a}{\text{line impedance } Z_0} = 0{\cdot}75.$$

This may be expressed in resistive and reactive components,



$$i.e., Z_a = 30 \cos 32° + J30 \sin 32°$$
$$= 25{\cdot}4 + J15{\cdot}9.$$

In the case of a steady reactive load, the reflected portion of the forward wave has fixed magnitude and phase, the reactance being either inductive or capacitative. This means that the first maximum of standing wave voltage will appear, from the load end of the line or guide, somewhere between the load end of the line or guide and half a wavelength along it, towards the generator end of the line or guide. If the load is inductive the first maximum will appear between the load end of the line or guide and a quarter of a wavelength along the line or guide. If it is capacitative, the first maximum will appear at some point between a quarter of a wavelength from the load end of the line or guide and half a wavelength. Any change in the reactance of the load will affect these conditions and will therefore be reflected at the generator end of the line or guide.

Since the phase of r (that portion of the forward wave which is reflected) relative to the forward wave, is determined by the reactive condition of the load, changes in the reactive components of a load may be indicated by a fixed vector of fixed magnitude, revolved through 360° relative to three fixed vectors representing resistance, X_L and X_C respectively (see diagram, p. 378). It will be noted that the conditions from 180° to 360° merely repeat the conditions found between 0° and 180°.

In this way the instantaneous condition of a load may be expressed in terms of r and θ, θ being the phase angle between the forward wave and the reflected fraction.

Valves can be designed to give wide variation in the loading diagram and the degree of loading. Briefly expressed in terms of frequency when r is maintained at 0·2 and varied from 0° to 360°, this is the frequency-pulling figure. Fig. 262 shows that low frequency pulling can only be obtained at some expense of efficiency.

For known values of r and known limits between which frequency can vary, θ can have certain values shown by that part of the diagram parcelled off by limit lines, but as the angle θ varies with changes in the impedance of the load, within the above-mentioned boundaries, the power output will vary. Hence the limits of power output, *i.e.*, the operating conditions of the valve, can be determined within which the swing of frequency, due to frequency pulling, can be contained.

The magnetron characteristics shown in Fig. 259 indicate performance under varying conditions of V_a and H, the load being always adjusted for maximum power output.

It has been shown, however, that maximum possible power output (as distinct from maximum power output for good operating conditions) must generally be sacrificed, in order that frequency pulling, due to a varying load impedance, may be contained within the limits of the handling capacity of the A.F.C. circuit of the receiver.

Rieke Diagrams

Whereas the magnetron characteristic shown in Fig. 259 indicates magnetron performance under varying conditions of V_a and H, the load being always adjusted for maximum power output, Rieke diagrams show performance for fixed values of I_a or V_a and fixed values of H, for variations in the load impedance.

Consider H and the pulse amplitude V_a to be constant, then variation of the load impedance over all possible values of resistance and reactance (inductive and capacitative) cause changes in the value of the peak current, peak power output, efficiency, frequency of operation and frequency stability.

Corresponding changes occur if the peak current is fixed and V_a varied. The advantage of the Smith circle diagram is that the loaded output line or wave-guide can be taken into account on the same diagram, when performance is being deduced in a particular case.

In a Smith circle diagram, complete circles represent constant resistance, and the arcs such, for example, as that marked "*ZZ*" on the diagram shown in Fig 263, represent reactances.

An important property of the circle diagram is that when a line or guide is terminated in a load represented by the point P, the magnitude of the reflected coefficient r of the load is represented by OP and the phase change on reflection is given by ϕ.

Circles centred at O are circles of constant r, these are also circles of constant standing wave ratios and inter-related by the expression :—

$$\text{S.W.R.} = \frac{1 - r}{1 + r}.$$

The vertical axis is scaled in terms of standing wave ratios.

The following important points arise from a consideration of Fig. 263 :—

(1) In the example taken the region of highest power lies at A. On following the constant r circle from A to the vertical axis, it is seen that the corresponding standing wave ratio is not unity. This means that the line or wave-guide is mismatched, and this is due to the fact that the optimum load for the magnetron at the output loop is different from the characteristic resistance of the line or wave-guide.

To match the magnetron output impedance to the line or wave-guide, there must be a built-in matching device at the output loop.

(2) In the region of A the frequency stability is poor, because

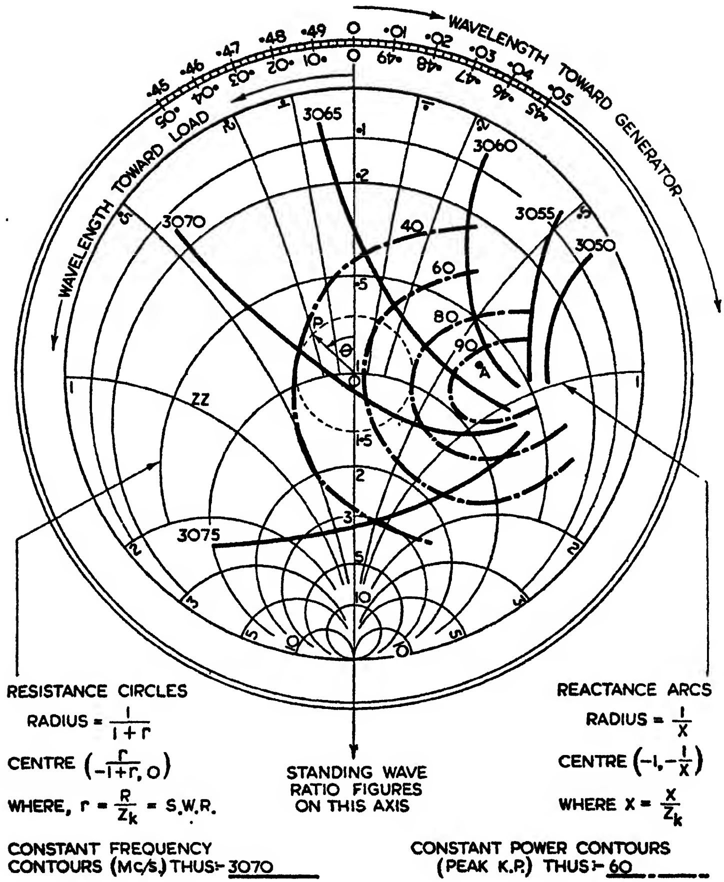


Fig. 263.—RIEKE DIAGRAM.

Wave impedance approach must be considered in three distinct regions :—

A. All circuits beyond the anode radius.

B. Interaction space between space charge and anode radius.

C. Space charge cloud.

The general approach involves solution of Maxwell's equations. If the resulting equations are linear, however, impedance concepts may replace field concepts so that the consistency conditions reduce to impedance matching.

1. Calculate Z looking in to region A from its internal boundary.

2. Z is transformed *via* region B looking out from the inner boundary of B, Z = f (frequency).

3. Solve for C and evaluate Z looking in from B to C boundary.

4. Equate Z_c to Z_b by adjustment of parameters (including n).

The expression for the parameters gives the operating conditions.

the frequency contours converge, so that small changes in loading cause comparatively large changes in frequency. Hence working the magnetron at maximum efficiency is incompatible with good frequency stability.

(3) The operating condition of the magnetron and the maximum efficiency at which it can be operated depends upon the characteristics of the magnetron design and its load.

(4) The particular features of the radar system that are important are the input voltage and its pulse shape, type of receiver employed, with particular reference to A.F.C. and the maximum standing wave that may be set up in the feeder system, due to joints, spacers and variations in loading imposed by the aerial system.

(5) The pulling figure, interpreted in terms of the Rieke diagram, is given by the frequency difference of those constant frequency contours which just embrace the r circle, corresponding to the maximum permissible standing wave ratio. This is commonly the value taken in the following example :—

If the circumference of the constant r circle, having the value $r = 0{\cdot}2$ is followed, the corresponding frequencies may be read off and plotted against the angle which the radius makes with the vertical axis. From this it is apparent that the pulling figure is about ·5 megacycles. Fig. 264 also makes it clear, that for reasons of frequency stability, a short line or guide is preferable to a long one.

Length of Line or Guide

For suppose that the line or guide is of physical length l and has a load such that the standing wave ratio $= 1{\cdot}5$, the input impedance depends not only upon the load, but also upon the electrical length of the line or guide and hence upon frequency.

Suppose that a line or guide be energised from a source, the frequency of which can be varied, then the point on the Rieke diagram corresponding to the input impedance moves round a constant r circle. For any given position of the point in the circle, the corresponding angle θ is in fact the phase angle ϕ between reflected and transmitted waves in the line or guide at the input terminals and hence there is the phase change on reflection at the load.

This phase angle is $-\frac{4\pi lf}{\mu}$ where μ is the velocity of propagation on the line (or phase velocity in the guide).

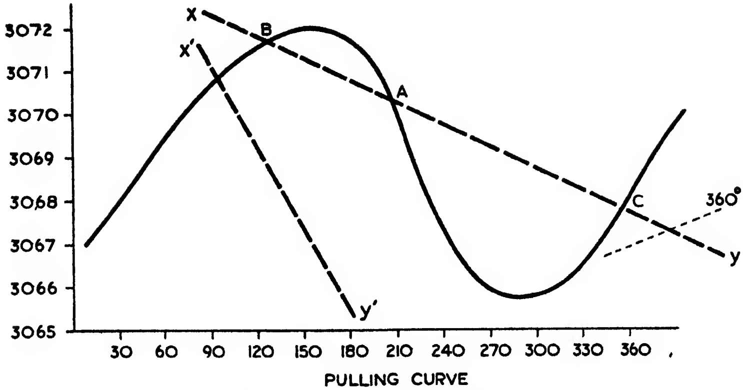


Fig. 264.—PULLING CURVE.

In Fig. 264 frequency is plotted against ϕ, giving the line xy. Suppose the line or guide is now energised by a magnetron, then there is the fundamental relation between F and ϕ plotted as above and operating conditions are given by the intersection of the two frequency — ϕ curves.

Frequency Splitting

In the example taken, there are three intersections, only two are possible because, although three points of intersection give three conditions of equilibrium, the one at A is unstable; for if some small change in the system were to occur, resulting in a small increase of frequency, then ϕ is decreased which, from the frequency-pulling curve, would result in a further increase in frequency.

But at B and C an increase in frequency causes a decrease in ϕ which tends to pull the frequency back. Since these two frequencies are equally possible, the magnetron may jump from one to the other. This is termed "frequency splitting."

If this is to be avoided, the line xy must be sufficiently steep for the curves to intersect at one value of frequency only, as indicated by x′y′. The slope of the line xy is $\frac{\mu}{4\pi l}$ so that to make xy steep enough to avoid frequency splitting l must be sufficiently small.

Broad Spectrum Operation

For a certain critical length of line or guide the slope of xy will be equal to the maximum slope of the frequency-pulling curve, and it is then possible, with a suitable value of ϕr, for xy to be tangential to the frequency-pulling curve at its steepest part. This leads to very considerable frequency instability and is termed " broad spectrum operation."

R.F. Spectrum

The shape and duration of the magnetron pulse, as has been shown in the chapter on receivers, has a very important bearing on receiver design. Harmonic analysis of a magnetron pulse reveals that a certain number of components must be handled by the magnetron output circuits if a good pulse shape is to be produced. The reasons for the somewhat detailed analysis of the components of the anode wave for various modes now becomes obvious.

If the band width involves a frequency spread for a 1-micro-second pulse of at least 2 Mc/s., this would be diminished if the pulse duration were increased

Although the shape of the modulator pulse has an important bearing on frequency spread, poor rotating joints may also materially affect the output pulse shape.

Accordingly it is now general practice to examine the R.F. envelope by spectrum analysis. This is accomplished by the use of a superheterodyne receiver, the local oscillation frequency of which is swept in synchronism with a cathode ray oscillo-scope time base. The detected signals are fed into the Y plates to produce an amplitude-frequency spectrum of the components in the magnetron pulse.

CHAPTER XVIII

TRANSMITTERS, INCLUDING R.F. GENERATORS

THE transmitter considered as a unit may be deemed to incorporate :—

(*a*) Its power supply.

(*b*) The R.F. generator and the circuits which must be associated, in order to pulse it at the selected frequency and with a pulse of the necessary voltage amplitude and duration.

The circuits of the power pack from which H.T. and L.T. are derived are conventional and similar to those employed in radio communication transmitters.

Since 400 megacycles and in certain cases 600 megacycles marks the highest limits of frequency at which triodes can be successfully and efficiently operated (generally however the limit is about 400 megacycles), choice of the R.F. generator to be employed is governed directly by the choice of carrier frequency.

At frequencies above the upper limit for triodes, the magnetron stands alone as an R.F. generator for ultra-high frequencies, with a power output capacity of a megawatt or more.

Selection of Carrier Frequency

The main factors which determine selection of the carrier frequency are :—

(*a*) Maximum and minimum ranges at which targets are to be detected.

(*b*) The required degree of accuracy for range and bearing readings.

(*c*) Discrimination between lumped targets (definition).

(*d*) Permissible weight.

(*e*) Accommodation space available for the equipment.

(*f*) Restrictions on size and weight of the aerial.

(*g*) In some cases the power supply available.

Unless (*d*), (*e*), (*f*) or (*g*) become the overriding factors, as may be the case in airborne equipment, (*a*), (*b*) and (*c*) will generally be the determining factors.

In the case of long-range warning sets, the maximum range at which targets must be detected is the most important consideration, and provided restrictions are not imposed by (*e*), (*f*) and (*g*), a comparatively low frequency is generally chosen in order to obtain an adequate echo from targets at extreme range with minimum peak power output from the transmitter.

In these circumstances the R.F. generator is usually a circuit arrangement of triodes similar to one of those shown in Figs. 265–267.

If the power output required and the type of valve selected necessitates an R.F. generator circuit containing more than two valves, the valves must be arranged as for a ring oscillator and as shown in Fig. 269. Additional valves now required for increased power output must be added to the basic ring circuit in pairs, for reasons explained in later paragraphs.

When restrictions are imposed by size, weight and wind resistance of the aerial, plain-surface and even perforated-surface reflectors may have to be abandoned and simple rod-type reflectors substituted, with or without directors, and generally with a loss of directional effect.

In general, when the performance specification lays down requirements for minimum range, accuracy and definition and conditions impose such restrictions as may be classified under (*d*), (*e*) or (*g*) ((*f*) already dealt with above), choice of carrier frequency and possibly of R.F. generator must, inevitably, result in a compromise.

When triodes are used for the R.F. generator any system of pulsing may be employed, with reservations as to suitability in regard to any restrictions imposed on stability of the repetition rate.

If no restrictions are imposed in respect of repetition rate stability, and the carrier frequency is to be in the triode range, a self-pulsing oscillator arrangement can be used for the R.F. generator. Any of the oscillators shown in Figs. 265–269 may be modified for pulsing by an independently synchronised modulator by adjusting the values of a resistance and one condenser.

Apart from such advantages possessed by the self-pulsing oscillator as minimum space requirements, minimum weight and cost, simplicity of operation and maintenance, its tendency towards a somewhat irregular repetition rate is not always as undesirable as might appear. It may, in fact, be a distinct advantage when accuracy of range and bearing readings is not of primary importance. This is due to the fact that signals transmitted at an irregular repetition rate cannot be easily jammed by enemy action at the receiving end of the system.

R.F. Generators using Triodes

When triodes are used in radar transmitters they are generally arranged in circuits which take the form shown in Figs. 265–269.

In all these circuits short-circuited quarter-wave and half-wave sections of transmission line are employed in place of the conventional condensers and coils normally used in R.F. generators at lower frequencies.

The limits imposed upon the use of coils and condensers in high-frequency circuits have been discussed in Chapter XVI, where it has been shown that when frequency is increased the physical dimensions of condensers and inductances must shrink towards zero and a straight piece of short wire respectively. Since the inter-electrode capacities and inductance of the internal elements assume reactive values which must ultimately determine the resonant frequency of the valve, all external capacity and inductance in this condition become practically zero.

In any case, the above condition is a limiting factor for the use of triodes, even when internal inductance and inter-electrode capacity has been reduced to the absolute minimum, by special valve design, as outlined in Chapter XVI.

Within the limits imposed by frequency on the use of triodes of special design or otherwise, it is possible to design circuits having a reasonably high Q for the oscillating circuits, and to secure output efficiencies comparable with those obtained in R.F. generators at lower frequencies. This can be achieved by utilising the resonance effects of short-circuited lengths of quarter-wave transmission lines, and by arranging them in circuit with due regard to the conditions essential for generating and maintaining oscillations, as laid down in Chapter V.

Circuit Arrangements of R.F. Generators

It is shown in Appendix I that short-circuited lengths of quarter-wave transmission line *behave* at *resonant frequency* in a manner similar to that of a parallel resonant circuit at its resonant frequency.

Advantage is taken of this phenomenon in the circuits shown in Figs. 265–269.

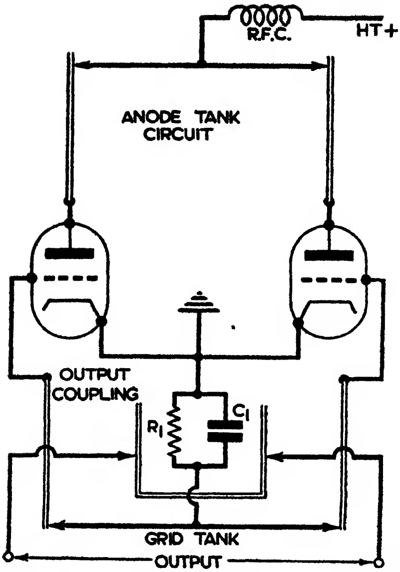


Fig 265.—Tuned anode tuned grid push-pull oscillator. Load inductively coupled to grid.

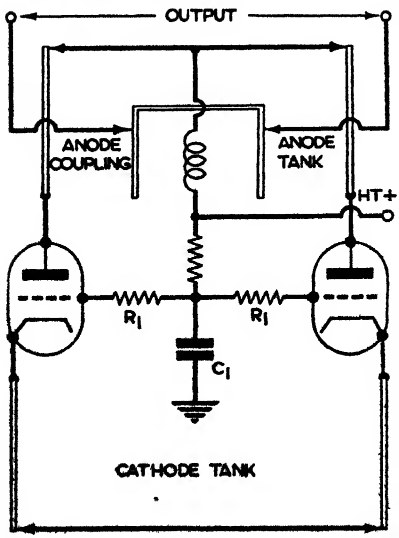


Fig. 266.—Tuned anode tuned cathode push-pull oscillator. Load inductively coupled to anode.

The push-pull circuits shown in Fig. 271 differ from the push-pull circuits shown in Figs. 265-267 by reason of the fact that concentric cable sections are used in place of two-wire transmission line sections. The principle of operation is the same in both cases, the only difference being that the Q obtained with the concentric cable is higher than with two-wire transmission line owing to reduced R.F. losses.

The expedient of changing from short-circuited quarter-wave lengths of transmission line to similar lengths of concentric cable is adopted in order to prolong the upper frequency limit of triodes, because in the region of 300 megacycles the losses incurred by transmission line sections are considerable, and can be greatly reduced by the employment of coaxial cables for reasons which are inherent to their construction (see Appendix I).

Briefly summarising then, the Q of the tuned circuits of an R.F. generator for ultra-high frequencies can be improved by the use of sections of quarter-wave shorted transmission line in place of coils and condensers. These may be silver plated if necessary further to improve the Q of this arrangement.

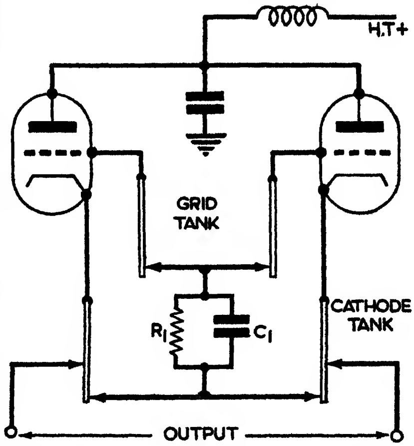


Fig. 267.—TUNED GRID TUNED CATHODE PUSH-PULL OSCILLATOR. LOAD COUPLED *directly* TO THE CATHODE.

The improved Q is obtained because the inductance and capacity of the system is distributed and the skin effect which causes heavy losses in normal conductors is minimised by using large-diameter silver-plated rods for the quarter-wave lines.

It will be noticed that the length of the quarter-wave line in the grid circuit is physically less than the quarter-wave line in the anode circuit. This arises because the inter-electrode capacity from grid to cathode is larger than the inter-electrode capacity from anode to cathode. This larger capacity causes more of the tuned line to be effectively inside the valve envelope and is a contribution to electrical length by the elements themselves. Physically, therefore, the grid line is shorter, although the grid and anode lines have the same *electrical length.*

Voltage Distribution

The voltage distribution along the anode and grid lines is such that energy fed back from the anode to the grid *viâ* the inter-electrode capacity is in phase with the grid voltage, and therefore oscillations occur and are maintained at the frequency which causes the lines to become exact quarter lengths of the wave corresponding to the oscillation frequency.

When lines * are employed for the tuned circuit of an oscillator the high voltage supply to the anode can be connected to the centre of the shorting bar. Theoretically the shorting bar has zero potential to earth, but because of its physical length, and therefore the presence of some fraction of the standing wave of voltage across it, the true zero point is determined by the capacity to earth of the anode of each valve,

* The term "lines" is used in the general sense of short-circuited quarter-wave or open-circuited half-wave sections of either two-wire transmission line or coaxial cable. They are often spoken of by the original laboratory name as Leicher wires. See any standard text-book on radio communication for further information regarding their use in this sphere.

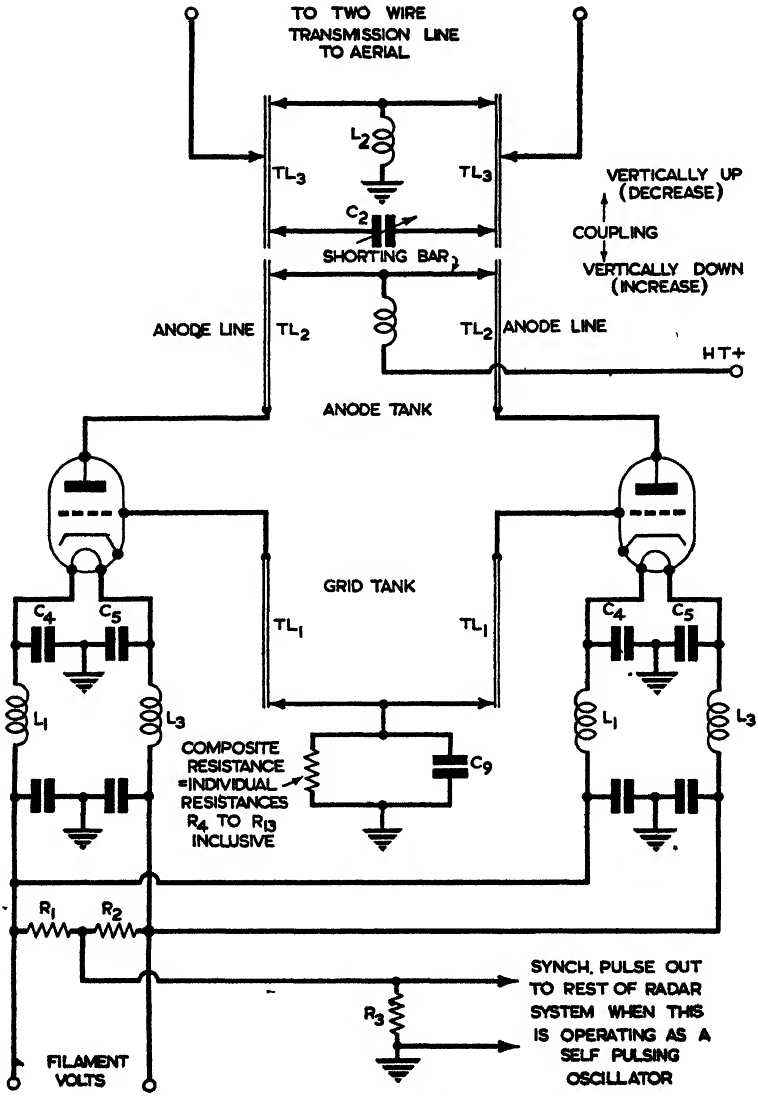


Fig. 268.—PUSH-PULL OSCILLATOR INDUCTIVELY COUPLED OUT OF THE ANODE TANK CIRCUIT TO AERIAL.

R_4–R_{13} is in sections which can be strapped across to vary the total resistance. Conversion from self pulsing to modulated is effected by adjusting the values of R_4 and C_9.

which changes slightly for different valves. In practice, the connection is made to the centre of the shorting bar and an R.F. choke takes care of any feedback to the supply due to any off-zero position of the tap.

When these circuits are connected as self-pulsing oscillators, as in Fig. 268, R_1 and R_2 form an impedance in the cathode circuit and conditions for negative feedback effects are present. This is to be avoided, as it reduces the positive feedback, and therefore the efficiency and power output of the device as an R.F. generator. This is prevented by the addition of by-pass condensers at the valve base to provide low impedance paths to earth for the R.F. currents, thus avoiding unwanted feedback.

The filament leads of the particular type of valve here employed are about 2 in. long, and therefore the actual R.F. filament voltage cannot be reduced to zero.

C_4 and C_5, together with L_1 and L_3, form a filter and so prevent R.F. voltage from being fed back to the power supply.

Alternative Methods of Coupling

In Fig. 268 energy is coupled out of the anode tank circuit by the tuned line TL_3. The shorting bar of TL_3 being placed directly above the shorting bar of the anode tuned line.

The coupling may be adjusted in the usual way by raising or lowering the secondary from the primary. In order that adjustment may be made for maximum energy transfer, C_2 can be varied to make the coupling lines resonant at the transmitter frequency (Appendix I).

The two-wire transmission line is coupled to the coupling line at the point at which their respective impedances match. L_2 keeps the transmission line, the coupling line and the aerial at earth potential and prevents accumulation of a static charge.

The R.F. generator shown in Fig. 268 is assumed to be adjusted for self-pulsing operation and is modulated by the self-pulsing action (described in Chapter V).

The duration of the pulse is determined by the time required to charge C_9. This time also affects the magnitude of the grid voltage, therefore the time required to charge C_9 is affected by any of the factors which determine the magnitude of the grid voltage. These are the tuning of TL_2 relative to TL_1, the magnitude of the applied anode voltage, the tuning of C_2 to the coupling line, the position of the transmission line taps and the coupling between primary and secondary circuits.

The tuned circuits may be connected to the oscillator in any of the ways shown in Figs. 265–268. Other connections may be used when it is necessary to take particular precautions to avoid negative feedback to the cathode circuit or feedback to the power supply line.

All these oscillators may be arranged for self-pulsing or for Synch-pulsing by any of the common forms of modulator. The change from a self-pulsing state being effected by adjusting the value of the grid leak or the capacity of the grid condenser (Chapter V).

Inductive coupling to the grid as in Fig. 265 is not frequently employed, since it reduces the Q of the grid circuit so much that it ceases to be reliable as the frequency-controlling factor of the oscillator.

In Fig. 266, because of the high voltage existing between the coupling line and tank circuit, the degree of coupling is limited so that in some cases the optimum coupling cannot be achieved. This difficulty is avoided in the direct coupling arrangement shown in Fig. 267. There is no H.T. on the cathode, therefore the transmission line may be coupled to the cathode tank to any desired degree by adjusting the output tap.

Pulsing Control

When the transmitter is tuned for maximum power output the duration of the transmitted pulse is controlled *only* by the capacity of C_9.

The pulse repetition rate of the oscillator can be varied by adjusting the values which control pulse length.

Since variation in any of the adjustments affecting pulse length must cause the condenser to charge to a different voltage, it follows that the *time* taken by the condenser to discharge to the voltage value at which the valve can again conduct must also vary, wherefore the repetition rate is changed accordingly. The usual method of effecting this adjustment is to short circuit one or more of the resistances which form the chain R_4—R_{13}.

Power Output

The power output of a two-valve circuit in push-pull is limited by the anode current and the anode dissipation. To increase the power output it is necessary to increase either the current-carrying capacity of the valves or the number of valves employed. Because of the limiting factors of inter-electrode capacity, transit time and lead inductance it is

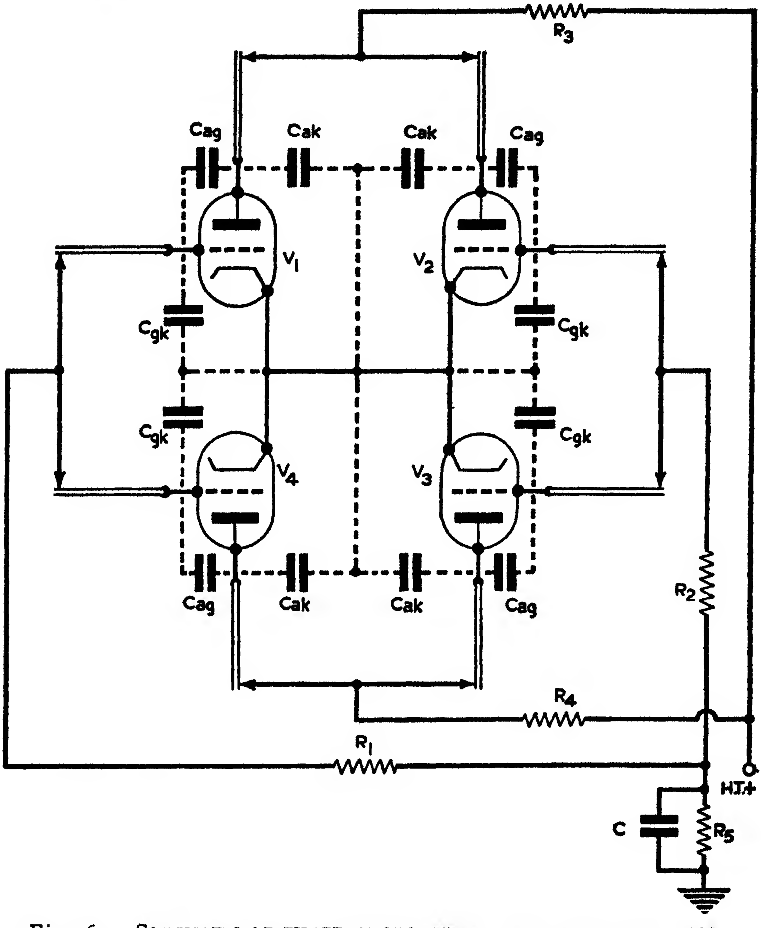


Fig. 269.—SCHEMATIC OF TUNED ANODE TUNED GRID RING OSCILLATOR. Note that the essential feature of the arrangement is such that pairs of inter-electrode capacities are in series as shown, thus halving the total inter-electrode capacities C_{ag}, C_{ak} and C_{gk}.

undesirable to increase the physical size of the valve itself. Increasing filament temperature helps, but paralleling increases the inter-electrode capacity and is, therefore, to be avoided.

The Ring Oscillator

When it is necessary to increase the power of a triode oscillator, valves are added in series and in pairs to form a ring circuit, see Fig. 269.

The ring oscillator circuit shown is a modification of that shown in Fig. 268, but the inter-electrode capacity is half that of the same valves when connected in push-pull. The ring circuit permits a given valve to be used at a higher frequency *i.e.*, it permits the use of a valve which is physically large.

The ring oscillator functions because energy is fed back from anode to grid through the inter-electrode capacity, and because of the voltage distribution on the anode and grid lines.

Proper adjustment of the tank circuit causes a voltage to be fed back from the tank circuit *viâ* the inter-electrode capacity, at the same time standing waves of voltage are established on the sections of transmission line, thus causing voltage at one end of the line to be in anti-phase with the voltage at the other end. The general requirements for the production of oscillations are thus complied with.

Let the grid of V_1 (Fig. 269) be positive at a given instant, the anode of V_1 will be at a minimum or negative maximum. The anode of V_2, approximately half a wavelength away, is positive because of the standing wave of voltage on the tuned transmission line. A signal on the grid of V_2 is produced (by inter-electrode capacity) as a negative maximum, in order to reinforce the anode current. Inter-valve anode relations are maintained by always adding valves in pairs if a larger number is required in the ring to obtain large power outputs.

Fig. 270 shows one method of obtaining energy from the tank circuit. The anode lines are supported vertically between the valves, with the shorting bars at the lower end. A transmission line with a loop in the end is lowered between the anode lines so as to place the loop near the shorting bars to provide inductive coupling.

The line and loop can be made resonant by using a shorting bar a half wavelength from the loop to ensure maximum transfer of energy. The feed line is tapped across the resonant section in this case. Any of the oscillators shown in Figs. 265-267 may be connected in a ring circuit to produce high peak power output.

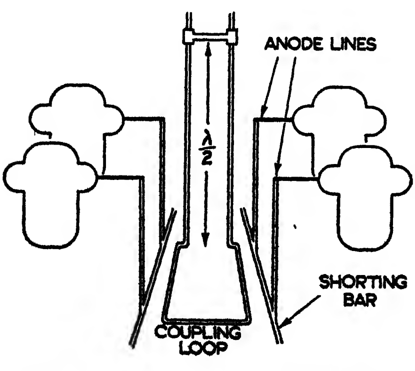


Fig. 270.—METHOD OF COUPLING THE AERIAL TO A RING OSCILLATOR.

Since the ring connection reduces the effect of inter-electrode capacities, valves of fairly large physical size may be used to produce high peak power output at frequencies up to 400 megacycles.

The ring oscillator has the disadvantage of having many initial tuning adjustments in its circuits. If these adjustments are not made properly, inefficiency must result. The ring oscillator, however, permits mechanical symmetry in construction, and therefore the tuning adjustments may be ganged to a few controls, and the oscillator is then relatively simple to handle.

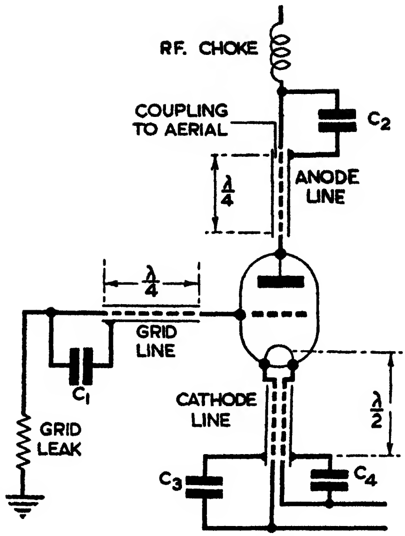


Fig. 271.—ULTRA HIGH FREQUENCY OSCILLATOR USING $\frac{\lambda}{4}$ CONCENTRIC LINES.

Aerial is coupled by hairpin loop, the position of which is determined by matching considerations.

Oscillator Employing Sections of $\frac{\lambda}{4}$ Concentric Lines

Fig. 271 shows a tuned grid tuned anode oscillator in which the grid and anode circuits are tuned by a quarter-wave short-circuited concentric line. A small trimming capacity across the open end can be adjusted for the resonant frequency. A concentric line is well adapted for the frequency control of ultra-high frequency oscillators, because it is completely shielded and offers the required high Q. The Q is highest for a quarter-wave section.

This follows from consideration of the following facts :—

(*a*) The resonant frequency of a line is proportional to the product of L and C, and this is true for *all* resonant lengths of line.

(*b*) At a specified frequency, however, Q is proportional to $\frac{\omega L}{R}$.

(*c*) L is the same for *all* resonant lengths.

(*d*) Resistance varies *with* the length of line and is different for various lengths.

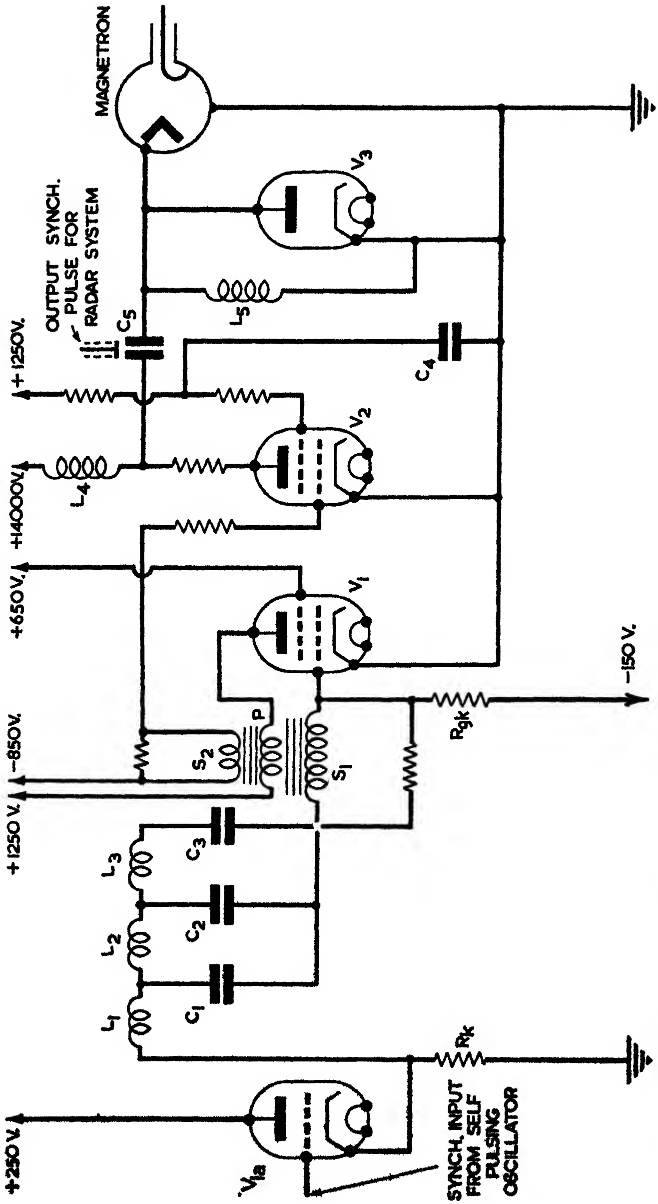


Fig. 272.—SCHEMATIC DIAGRAM OF A TRANSMITTER IN WHICH A POWER TIME CONTROL PULSE FOR THE MAGNETRON IS DERIVED FROM THE DISCHARGE OF C_5 THROUGH V_3.

An indication of Synch. is suggested by a self-pulsing oscillator (not shown) but Synch. for the radar system is taken from C_5. The artificial line in this case checks any tendency to frequency instability of the p.r.f. and ensures uniform pulse width.

(*e*) $\frac{\omega L}{R}$ or Q is largest when R is smallest.

(*f*) The shortest resonant length is a quarter wavelength.

The power-handling capacity of a concentric line depends upon the magnitude of the breakdown voltage of the tuned circuit rather than the energy stored in the tank circuit, consequently the use of a tuned line longer than a quarter wavelength does not permit the oscillator to handle more power.

In Fig. 268, which shows a push-pull oscillator, sliding bars are used for tuning. The grid bars may be varied for maximum stability, whilst the anode bars are varied for maximum output.

In the circuit shown in Fig. 271 coupling is accomplished by a hairpin loop about one-quarter of the length of the line. C_1, at the end of the grid line, serves to short the line effectively at this point, thus making the line equivalent to a tank circuit. C_2 performs the same function on the anode line.

At ultra-high frequencies the filament cannot be earthed directly by a by-pass condenser because of the high filament lead reactance at these higher frequencies. The effect is eliminated by making the filament lead inductance part of the tuned transmission line. The line is a half length from the cathode proper to the earth, and is shorted by C_3 and C_4. Since a half-wave line shorted at the far end appears to be also shorted at the near end (see Transmission Line Theory, Appendix I), C_3 and C_4 have the same effect as though they were connected directly from the cathode to earth, inside the valve, and at the cathode itself.

Application of Self-pulsing Oscillator to Provide a Trigger Pulse

Fig. 272 shows the output of a self-pulsing oscillator used with a discharge line. In this case the line is used to control the duration of the pulse, since it is assumed that the oscillator frequency stability is not within the limits required for some particular application where precise measurement is required.

V_1 is normally cut off and a positive pulse of approximately 40 volts is applied to V_{1a} in order to start each pulse at approximately the required repetition rate.

This positive pulse, which is coupled across the discharge line, to the grid of V_1, suffers some attenuation in the delay line network, but is large enough, initially, to fire V_1.

The voltage induced in the secondary of the transformer by the anode current flowing in the primary is in such a direction that it reinforces the trigger pulse and drives the grid of V_1 very positive. The transformer ratio is of the order of 2 : 1. Thus, when 1,000 volts are developed across the primary, 500 volts are applied across S_1.

The sum of the grid-cathode resistance of V_1 and the output resistance of the cathode follower is made equal to Z_0 of the discharge line, *i.e.*, $R_k + R_{gk} = Z_0$.

Thus the current that flows in the circuit R_{gk} R_k produces a difference of potential of approximately 175 volts across the line and remains constant at this value as long as current flows. The terminal voltage of 175 volts is then large enough to maintain V_1 in conduction for the duration of the pulse.

The effective $E = S_1 + 500$ (induced $-$ ($-$ 150)), *i.e.*, $\frac{350}{2} = 175$ volts. *Note.* Since $R_k + R_{gk} = Z_0$, the effective voltage divides equally (see Chapter XII, Artificial Lines).

The additional voltage of 175 volts impressed on the line is of the same polarity as the original charge of 150 volts, therefore a wave of voltage travels down the line and is reflected without change of sign (since the line is terminated in Z_0),

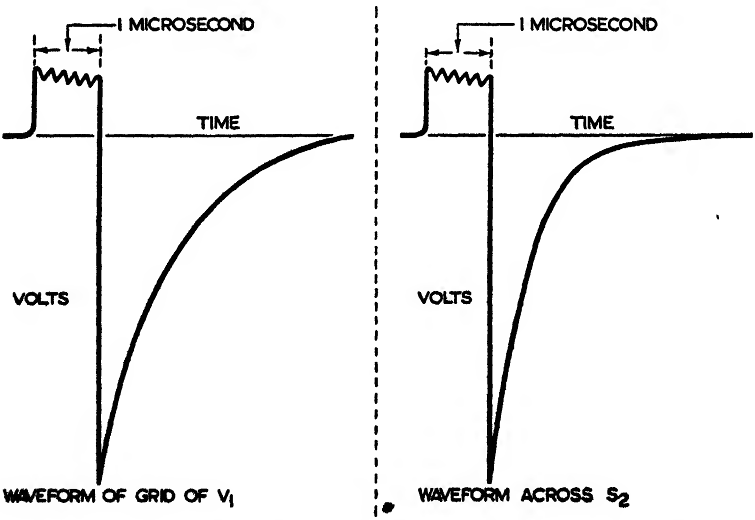


Fig. 273.—COMPARISON OF WAVEFORMS AT THE GRID V_1 AND ACROSS S_2. **Note the accelerated rise of the S_2 waveform compared with the V_1 waveform.**

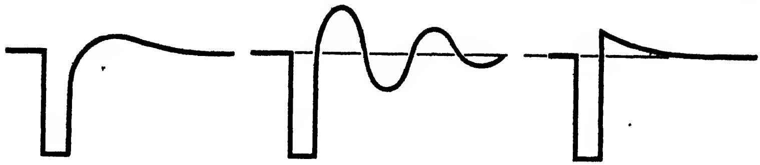


Fig. 274.—ILLUSTRATING THE EFFECT OF THE DELIBERATE INTRODUCTION OF OSCILLATIONS AND THE DIODE COMBINATION IN STEEPENING THE TRAILING EDGE OF THE INPUT PULSE TO THE MAGNETRON.

raising the voltage of the line from 325 volts to (325 + 175 volts) 500 volts.

If the voltage induced in the secondary of the transformer S_1 is 500 volts, then two equal voltages are in opposition and the result is cancellation or zero.

In this way the duration of conduction of V_1, and hence the duration of the pulse, is determined by the artificial line. For example, if the line is designed for a delay of ½ microsecond (travelling in one direction), the voltage wave raising it to parity with the voltage induced in the secondary of the transformer takes 1 microsecond to go and return, consequently the duration of conduction of V_1, and therefore of the output pulse, is 1 microsecond. As the anode current falls off the voltage induced in S_1 falls to zero and the grid of V_1 is sharply cut off.

The cut off of V_1 is accelerated by the action of the 150 volts impressed on the negative bias and by the sharp negative surge set up in S_1. These factors tend to drive the grid sharply to a very negative value.

The waveform shown in Fig. 273 is irregular at the top, this is due to the fact that the performance of the artificial line only approximates to the performance of a real line. It is also an effect of the sharp negative surge set up in S_1.

The relatively slow recovery of the grid of V_1 to normal is caused by the action of the 150 volts and the capacity of the artificial line, but the waveform applied to the grid of V_2 is modified by the action of S_2 as shown in Fig. 273.

For simultaneously with the foregoing a pulse is induced in S_2. The voltage ratio is 1 : 2, so that if the voltage across the primary is, say, 500 volts, a pulse of 1,000 volts is applied across S_2 to the grid of V_2. If at this instant V_2 is cut off and

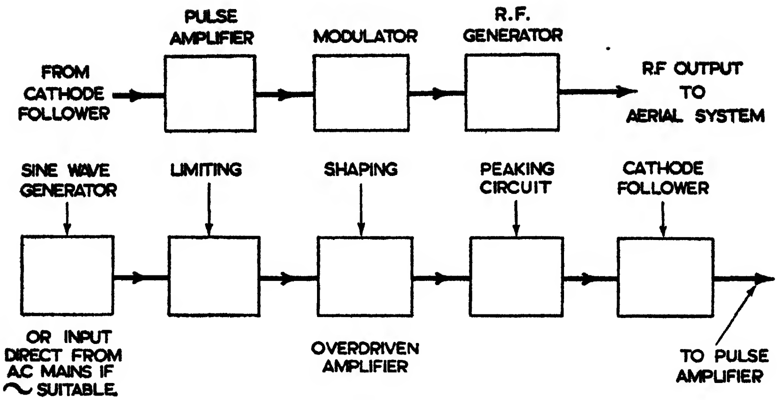


Fig 275—BLOCK DIAGRAM SHOWING SEQUENCE OF OPERATIONS WHEN THE POWER PULSE FOR THE R F. GENERATOR IS DEVELOPED FROM A SINE WAVE

The advantage is a stable p r f which is essential to good definition. Details of the circuits involved are given in Chapter XII

C_5 is charged to 14,000 volts, then application of 1,000 volts to the grid of V_2 causes V_2 to conduct and C_5 discharges through the magnetron and V_2.

When the pulse to the grid of V_2 ceases, V_2 is sharply cut off, oscillations are set up in L_5 and the distributed capacity between the magnetron filament and earth by the sudden change. The inductance is included in the circuit in order to cause these oscillations, thereby ensuring that the pulse to the magnetron filament shall be given a sharp falling edge. Damping of the oscillations thus set up may be provided by a diode which acts as a very low resistance across the resonant circuit during the positive alternation of the oscillation, effectively preventing any negative alternation of the oscillation which, if of sufficiently large amplitude would cause unwanted operation of the magnetron and more than one pulse.

Since there is always a slight delay in starting the pulse in a circuit of this description it is not desirable to use any part of the output of the initial stages as a synchronous pulse for timing the remainder of the radar system. To avoid this difficulty, a timing pulse is coupled out of the transmitter from the case of C_5. Thus it is ensured that whenever the magnetron is pulsed a negative synchronising pulse is coupled out of the transmitter to the other units of the system at exactly the right instant.

Fig. 275 is a block diagram showing the development of a high-voltage pulse from a sine wave.

Relations between Carrier Frequency and Accuracy of Measurement

Precise range and bearing measurements necessitate the use of narrow beams. The beam must be made narrower as the requirements for precise measurement become increasingly exacting. However, the minimum dimensions for any aerial unit are fixed by the carrier frequency selected, and since the overall dimensions of the aerial must be expanded by the addition of multiple units in order to reduce the beam width for any given carrier frequency, then it is clear that in order to produce a narrow beam and at the same time to keep the overall aerial dimensions within reasonable limits a very high carrier frequency must be employed.

For example, at 100 megacycles, in order to obtain a beam of approximately 20° in the horizontal plane, the aerial array must be two dipoles wide, *i.e.*, about 6 metres, irrespective of spacing between the dipoles themselves. If a sheet of metal is to be used as a reflector it must be evident that the overall width of the aerial system will be about 9 metres. In addition, if the vertical beam width is to be about equal to the horizontal width, the overall dimensions of the aerial would seem to be in the region of 25 to 27 ft. square. This assembly has to be mounted as a very stable structure capable of withstanding wind pressure, vibration, etc., and at the same time to be capable of rotation in azimuth and elevation; consequently its weight, including pedestal, is likely to be prohibitive, excepting in the case of very large vessels or for installation on shore.

Compare with this the "cheese" shown in Fig. 326(c), which is capable of radiating a beam of approximately 2° in the horizontal plane at a frequency lying between 9,000 and 10,000 megacycles per second.

Microwaves, however, suffer from the disadvantage that the service range for the direct ray for surface targets is generally limited by the line of sight distance to the horizon (as seen from the height of the aerial above earth) multiplied by a rectifying factor of 4/3.* Microwaves are therefore unsuited for use with long-range warning sets for the detection of

* The ratio 4/3 is a figure usually taken to allow for the curvature of the earth's surface.

surface targets on account of limitations which they impose on range.

Magnetron Requirements

Pulses for application to a magnetron must be negative going, since they are applied to the cathode, the anode being held at earth potential. This arrangement is adopted in order to avoid the necessity for high insulation of the heater transformer, and the necessity for special insulation design and particular care in manufacture.

The H.T. or power pulses should, in general, have steep vertical sides and a flat top. This is particularly the case when very short ranges are to be measured. To fulfil this requirement the pulse time must be short, and the falling edge must be as steep as possible, in order that the end of transmission may cease abruptly, and so avoid masking echoes returning from nearby targets.

The magnitude of the voltage peak should remain reasonably constant for the duration of the pulsing period. This means that the source from which the pulse is obtained should be such that the load imposed by the magnetron does not cause any very considerable voltage drop. The impedance of the source from which the pulse is drawn must therefore be low in order to supply power to the load with good voltage regulation.

This requirement is brought about by the need to avoid the possibility of any change taking place in the normal operating mode of the magnetron due to H.T. voltage drop on load.

Any change in the normal operating mode, or the introduction of some other mode, causes frequencies to be generated by the magnetron other than the required frequency. This must result in a loss of power and a decrease in overall efficiency (broad spectrum operation, see p. 383).

The various methods by which high voltage rectangular pulses may be developed for use with a magnetron have been discussed in Chapter XII. Selection of the sequence of circuits to be employed for any given case depends largely upon the type of modulator which is to be employed, and this in turn is governed by the application and the conditions under which the system is to be worked.

Discharging Devices for High-power Pulse-forming Devices

When a thyratron or other similar device (such, for example, as a trigatron) * is used to modulate the pulse for a magnetron the repetition frequency is controlled by a Synch. pulse, thus the stability of the repetition rate for the system can be ensured by employing a suitable source or generator for the Synch. pulse. This system has another advantage over the spark gap modulator. If the shape of the pulse developed in the pulse-forming circuit is not satisfactory it can be modified before application to the magnetron *viâ* the modulator. This is not possible when a spark gap is used as a modulator.

On the other hand, the spark gap modulator has the advantages of simplicity and high power-handling capacity.

The power-handling capacity of thyratrons and similar modulating devices is limited by the maximum voltage they can withstand, without arcing over taking place between the elements ; consequently, when very high power outputs are required from the R.F. generator a spark gap discharging device is nearly always employed. In general, then, the method for obtaining a high-power output is fixed but for applications requiring a very square rectangular pulse and medium power output there are several alternatives offering more control of the pulse shape.

Apart from the type of modulator chosen, the pulse-forming device may utilise the discharge of a condenser of sufficiently large capacity, or an artificial discharge line may be employed as an alternative.

Comparison between Uses of a Condenser and Artificial Line

When a large condenser is employed, the width of the pulse (duration of discharge) is controlled by a modulator. The modulator is in turn controlled by the duration of the pulse supplied to fire it by the sub-modulator and associated circuits. The repetition rate is also controlled by this pulse, which is itself under the control of the Synch. generator.

Thus, when a condenser of large capacity is employed the pulse formed by its discharge is under the direct control of the Synch. generator for repetition rate, but its duration is governed by the sub-modulator and its associated shaping circuits.

On the other hand, when an artificial line is employed the line itself constitutes the pulse-forming circuit by virtue of its

* A trigatron is a discharger enclosed in a glass envelope such that each discharge can take place under constant pressure.

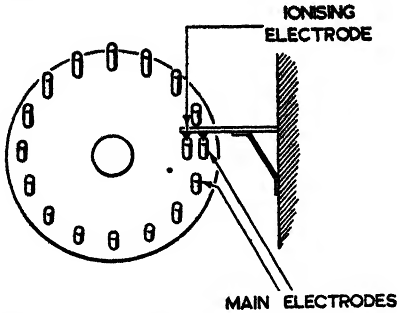


Fig 276.—FORM OF ROTARY SPARK GAP WITH IONISING ELECTRODE.

This may be driven independently or synchronised. (For details see standard works on radio transmission.)

characteristics. The length of the artificial line times accurately the duration of the pulse, whilst its general design and construction determine the squareness or quality of the pulse.

If the discharge of the artificial line or condenser is performed by a thyratron or similar type of modulator, the repetition rate is under the direct control of the Synch. generator. If this function is performed by a spark gap monitoring device the repetition rate is determined by the time period and regularity of operation of the discharger itself and there is only limited control by any Synch. source that may be employed.

When the discharge of a condenser or artificial line by means of a thyratron or similar device is employed to provide the H.T. pulse for the magnetron, the condenser must be charged from a D.C. source, and a charging diode is included in the circuit arrangements.

When a spark gap or similar device is used to discharge a condenser or artificial line the charging source can be either a suitable D.C. supply, or *viâ* a transformer operating on an A.C. supply, the p.r.f. then approximates to the A.C. frequency.

When a spark gap modulator is employed and the charging source is D.C., a large iron core inductance is usually included in the charging circuits to cause, by shock, a transient oscillation to take place at the natural frequency of the charging circuit and at the commencement of each successive recharge. The result of this is to cause the energy in the circuit to oscillate between its inductance and capacity, so that at one moment the energy is all associated with the inductance and $= \frac{1}{2}LI^2$ and a quarter cycle later it transfers to the condenser where it appears with a value of $\frac{1}{2}CV^2$.

If the spark gap is caused to break down when the whole of the energy has been transferred to the condenser, the voltage across it due to oscillatory action will be $2E$ where E is the applied voltage.

When charging is performed by a transformer operating from an A.C. supply the inductance and the first diode are not required. The frequency of the circuit can be adjusted by transformer design as nearly as possible to the frequency of the supply, so that the voltage across the line capacity rises to 2E during each half cycle of the supply. The spark gap is adjusted to break down at this voltage and the discharge rate, which in effect becomes the repetition rate for the system, is approximately the same as the frequency of the A.C. supply.

Spark gaps may be of the fixed or rotary type, the great advantage of the latter is that its life is longer than that of the fixed type. Production of ozone and nitrous oxide at each discharge tends to corrode the electrodes of the fixed gap and reduces its effective life to about 200 hours. The life of a rotary spark gap may be in the region of 1,000 hours or more.

The Fixed Spark Gap

The fixed spark gap consists of two spherical electrodes. Inserted in one of the spheres (which is hollow) and insulated from it, is a third electrode. The main gap is made large so that the voltage will not break it down unaided, the breakdown is controlled by a trigger pulse to the third electrode. This initiates by ionisation a spark at the main electrode and the gap breaks down to discharge the condenser or artificial line.

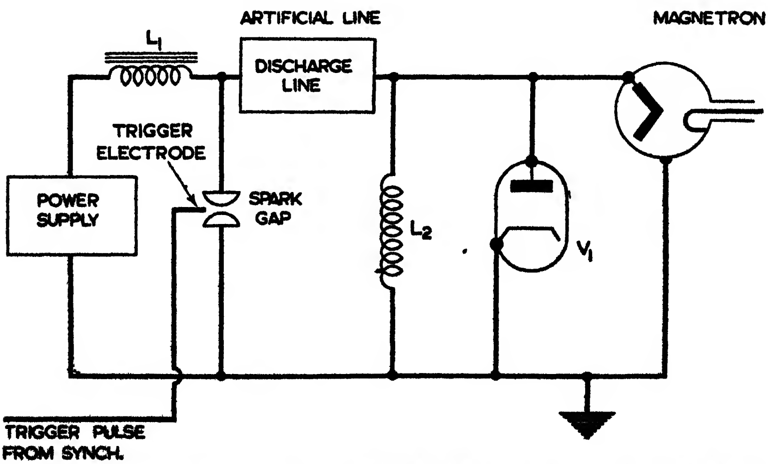


Fig. 277.—BASIC CIRCUIT USING A FIXED SPARK GAP, MODULATOR, ARTIFICIAL LINE AND D.C. OR A.C. CHARGING SOURCE.

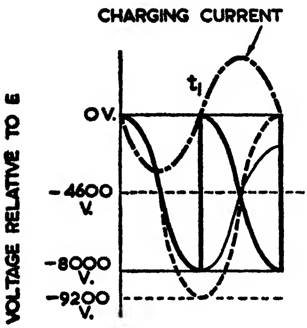


Fig. 278.—CAPACITY CHARGING AND DISCHARGING CURVES.

The voltage at discharge is approximately twice that of the charging voltage.

The Rotary Spark Gap

The rotary spark gap is as shown in Fig. 276. The small pointed electrode mounted on the same bracket as the fixed electrode, ionises the main gap at the critical instant by a corona effect, and ensures it breaking down as the moving electrode passes the fixed electrode.

The life of this form of spark gap is longer than that of the fixed gap, as corrosion is distributed over several electrodes and cooling is promoted.

In order to prevent interference with other parts of the radar system, the spark gap must be enclosed in a metal screening box, also chokes and filters are inserted in all leads to and from the gap to prevent transmission of R.F. energy to other units of the system.

Since the pressure of the air or gas in the spark gap controls the breakdown voltage for any given length of gap, it is important that this should be kept as constant as possible. This is particularly important in the case of aircraft, where the breakdown voltage decreases at high altitudes.

The basic circuit of the spark gap modulator (fixed gap variety) is shown in Fig. 277. The pulse-forming line is charged through L_1 and V_1. The inductance L_1 and the capacity of the line form a series resonant circuit so that the voltage across the line tends to oscillate.

At the peak of the positive swing the line is charged to a voltage considerably higher than that of the charging voltage. At this moment the spark gap is caused to break down.

The characteristic impedance of the artificial line is made equal to the impedance of the magnetron, so that one half of the voltage across the line is impressed on the magnetron for the duration of the pulse.

A fixed type of spark gap is shown in the basic circuit, and the application of a trigger pulse from Synch. to the auxiliary or third electrode promotes ionisation and causes the gap to break down at intervals controlled fairly effectively by the trigger pulse and therefore to the same degree by the Synch. source.

When an artificial line is charged, the capacity of the line

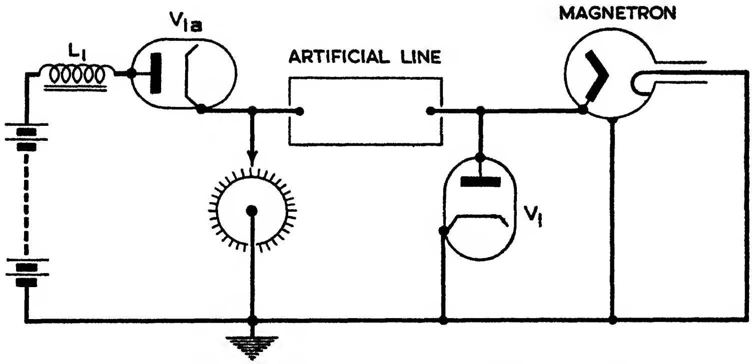


Fig. 279.—CHARGING FROM A D.C. SOURCE AND EMPLOYING AN ADDITIONAL DIODE V_{1a}.

offers a capacitative reaction and in consequence an inductance is employed to establish a state of L.F. resonance.

The behaviour of spark gaps * in general, and the analysis of the charge and discharge of capacity in a spark gap circuit, are dealt with fully in standard works dealing with W/T spark transmitters. It is sufficient to state, therefore, that if the electrical constants of the circuit and the revolutions of the spark gap are adjusted so that the discharge takes place at the moment when the energy of each charge resides wholly within the condenser, the voltage across the capacity at that instant is twice that of the charging source (see Fig. 278). Thus, in the case where an artificial line is employed as a discharge line, and providing the load in the discharge circuit is made equal to Z_0, the voltage applied across the load at each discharge will be approximately equal to that of the charging source, *i.e.*, to half the *resonant voltage* across the line at the instant of discharge.

In Fig. 279, L_1 and V_{1a} are employed to produce these conditions, the charging source being D.C.

L_1 resonates with the line capacity so that the charging energy oscillates as a transient.

The main function of the diode V_{1a} in the charging circuit is to prolong, by its unilateral conductivity, the favourable

* The use of a spark gap to produce damped oscillations by the discharge of a condenser in early W/T systems must not be confused with the case where it is employed to discharge an artificial line. Pulses are generated in both cases, but the condenser has no control over the rate of release of the stored energy, whereas the artificial line does control the rate of release of its stored energy according to its length.

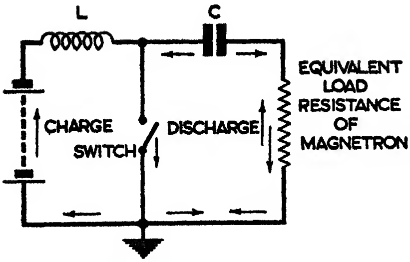


Fig. 280.—EQUIVALENT CIRCUITS FOR CHARGE AND DISCHARGE OF ARTIFICIAL LINE OR CONDENSER (LEAVING OUT DIODES).

The switch is introduced to simulate the action of the spark gap.

period for discharge in case the spark gap fails to fire at the critical instant when the peak voltage occurs.

When the charging source is A.C., the necessity for V_{1a} disappears. Also in this case it is no longer necessary to oscillate the charging circuit, since the phases of the A.C. charging current and voltage can be adjusted by suitable design of the transformer and by a fine adjustment in the primary or secondary circuits to give a power factor approaching to unity.

When a rotary spark gap is employed it is not essential to synchronise it, as must be done in a W/T transmitter, since this function is performed by the ionising gap.*

Pulse Transformer

Instead of connecting the magnetron directly into the discharge circuit, a pulse transformer of special design may be used with considerable advantage (Figs. 282 and 283).

(*a*) The output impedance can be adjusted for high- or low-

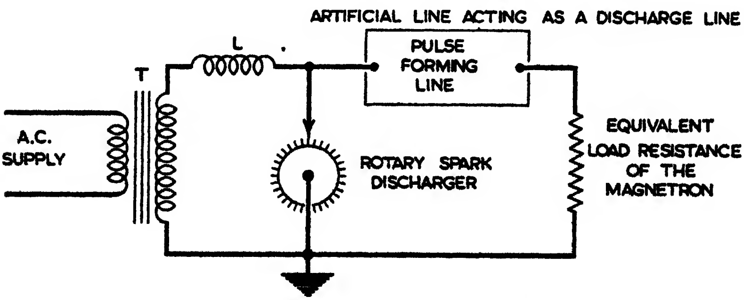


Fig. 281.—ARTIFICIAL LINE CHARGED *viâ* TRANSFORMER AND CHARGING CIRCUIT RESONATED BY L. DIODE V_{1a} IS NOT REQUIRED.

* In W/T, if the number of electrodes on the disc × revolutions per min. = $\frac{\text{A.C. frequency}}{2}$, a position can be found for the fixed electrode where the moving electrodes pass under it at the instant when the A.C. supply passes through zero and E across the condenser is a maximum value.

level operation according to the design of the transformer.

(*b*) The transformer may be used to reverse polarity of the pulse to the magnetron if required, *i.e.*, a positive-going pulse can be generated and fed to the magnetron as a negative-going pulse.

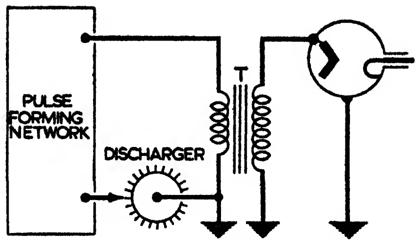


Fig. 282.—SCHEMATIC INTRODUCING PULSE TRANSFORMER.

(*c*) It may be used to step up the voltage of the generated pulse before it is applied to the magnetron.

(*d*) The artificial line is not required to stand the high voltage of the pulse reaching the magnetron. It can also have a lower characteristic impedance, since the impedance of the magnetron as seen by the line is stepped down.

(*e*) The voltage of the source for charging the line may be made less.

(*f*) It enables the magnetron to be located near the aerial and remote from the pulse generating circuit, thereby reducing the length of wave-guide between the magnetron output and the aerial itself (see p. 382).

Since the output from the magnetron must be fed to the aerial and because the aerial is generally situated a considerable distance from the transmitter, a transmission line or wave-guide of considerable length would usually be required. Loss

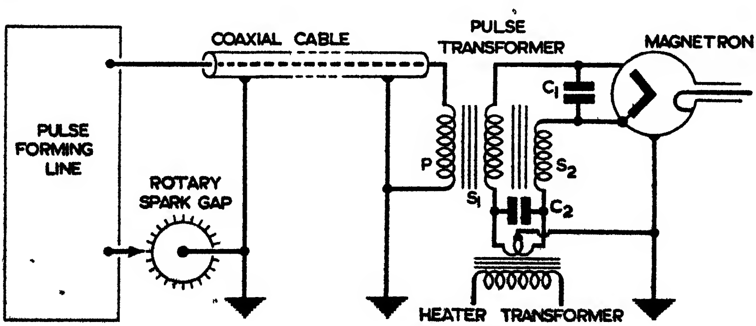


Fig. 283.—SCHEMATIC SHOWING HIGH-VOLTAGE HIGH-POWER PULSE GENERATOR SEPARATED FROM THE MAGNETRON POSITION BY COAXIAL CABLE AND PULSE TRANSFORMER.

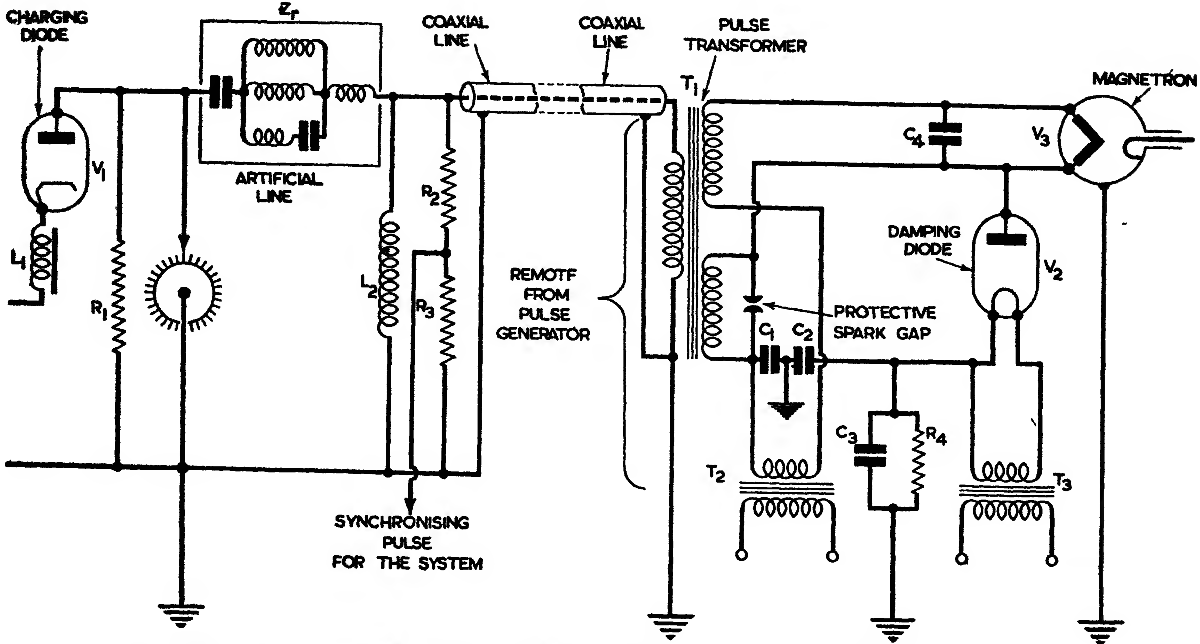


Fig. 284.—CIRCUIT DIAGRAM OF ROTARY SPARK GAP MODULATOR AND R.F. TRANSMITTER (REMOTE).

Note.—The charging source is D.C. so V_1 is included and L_1 resonates the charging circuit in conjunction with L_2; also note that the Synch. pulse for the other units is in this case taken from the transmitter. The coaxial cable and pulse transformer saves a long run of wave-guide from the transmitter position to the aerial, since it permits the magnetron to be situated close to the aerial, and only a short wave-guide is now required to feed from the magnetron to the aerial.

and attenuation over such a line would be considerable, also the problem of matching arises.*

Efficiency requires the transmission line to be terminated in its own characteristic impedance. The impedance of the magnetron is higher than that of any practical coaxial line, consequently the transmission line must be terminated in a transformer on account of matching considerations in any case. It is, therefore, desirable to employ a special pulse transformer and to locate it and the magnetron as near to the aerial as possible. In this case the pulses generated in the modulator circuit are fed over a coaxial line to the primary of the pulse transformer.

Two secondary windings may be wound on the pulse transformer, the voltages induced in each secondary by the pulse fed into the primary is negative at the cathode end of the coil.

The two secondary windings are connected in parallel for the pulse voltages by C_1 and C_2 so that both windings aid in driving current through the magnetron (Fig 283).

The lower ends are nearly always at earth potential and the heater transformer need not therefore be insulated for high voltages. These connections permit the use of a standard heater transformer, thereby effecting economies in design and spare parts.

Fig. 284 shows the schematic of a complete rotary spark gap modulator and transmitter.

The artificial line which acts as a discharge line is charged to 8,000 volts by a voltage doubler through the resonant circuit, L_1 V_1 Z_r L_2, and the primary winding of the transformer. Most of the discharge current flows through the primary of the pulse transformer, but some flows through R_2 R_3 and L_2.

The output is a positive-going 4,000-volt pulse, and a pulse of approximately 125 volts is developed across R_2 to be used as a synchronising pulse for the rest of the system.

A negative-going pulse of 18,000 volts is induced in the secondary windings of the pulse transformer. To prevent damage to the transformer, and in case the load is accidentally disconnected, a protective spark gap is adjusted for 25,000 volts. The firing of this gap would load the transformer heavily and prevent any further rise of voltage due to bad regulation.

* In Chapter XVII it was shown that the length of the guide or line from the magnetron output to the aerial should be as short as possible.

The pulse voltage in the two secondaries is equalised by C_4.

C_1 and C_2 are earthed in order to keep the lower ends of the secondary windings at earth potential.

The magnetron anode current passes through the magnetron to earth *viâ* R_4 and C_3 in parallel and the pulse transformer secondaries.

The voltage developed across R_4 is proportional to the current flowing in the magnetron and the damping diode, because C_3 filters out the surges. A meter may, therefore, be connected across R_4, roughly to indicate the anode current.

If no meter is required the by-pass condensers can be dispensed with and the centre tap of the heater transformer may be earthed to provide a closed circuit for the anode current.

The arrangement shown in Fig. 284 is a variation of the arrangements shown in Fig. 261, Chapter XVII and elsewhere. Nevertheless, the principle is the same, the main difference involving questions of insulation, particularly as applied to the heater transformer and the connections of the artificial line to the primary of the pulse transformer. The latter may be of the auto-transformer type as shown in Fig. 261, or it may have separate primary and secondary windings. The former is favoured by British engineers, whilst the latter more or less conforms to standard American practice.

The connections between the artificial line, the pulse transformer and magnetron must always be such that a negative-going pulse is applied to the cathode of the magnetron.

CHAPTER XIX

RECEIVERS

THE energy of the received signal is necessarily small, in some cases at a micro-volt or less; consequently, all other things being equal, the efficiency of a radar system, as a whole is, to a large extent, dependent upon the efficiency of the receiver. This means that a comparatively small improvement in receiver efficiency may increase the maximum range of a radar set to a greater extent than a comparatively large increase in transmitter output power.

A certain amount of heterogeneous interference is present in all receivers, and this is collectively classified as "noise." The minimum useful received signal can, therefore, be regarded as that signal strength which enables a signal to be just distinguishable from the background noise of the receiver.

Receiver efficiency may, therefore, be thought of as useful signal output/signal input, and the effective range of a radar set is consequently limited by the ability of the receiver effectively to utilise weak signals.

Theoretically, it is possible by using many stages of amplification, to build up any signal, no matter how weak, to any desired amplitude. This does not help very much if noise, due either to pick-up or inherent to high amplification, is amplified at the same rate as the signal.

It is abundantly clear, therefore, that if the signal amplitude at the input is not at least as large as the noise input voltage, it cannot be recognised at the receiver output, and is therefore useless. This holds true whether the noise is due either to generating action or to pick-up or to both causes.

The criterion of receiver efficiency in the above respect is, therefore, the signal-to-noise ratio.* All other things being equal, the efficiency of a radar set depends upon this value.

Generally, the sensitivity of a receiver must be such that it will accept signals of the order of a micro-volt or less, and amplify them to a useful value, consequently the general noise level must be well below this figure if reception is to be efficient.

* When signals are strong compared with noise the ratio is *high*. This is the desired condition for efficient operation, *i.e.*, a *high* ratio.

Noise Voltages

Noise voltages which are generated in an amplifier stage include three types :—

(a) Thermal agitation.
(b) Shot effect.
(c) Induced effects.

All these include frequency components throughout the entire frequency spectrum, and the amount of noise is therefore affected by the choice of band width for the receiver.

In general, a reduction of the band width reduces the noise level, but does so at the expense of pulse shape.

Thus the amount of pulse shape distortion that can be tolerated is the limiting factor in improving signal-to-noise ratio by reduction in band width.

Wide Band Response

Fig. 285 shows the output of a receiver with wide band response. The pulse shape is faithfully reproduced, because all the principal harmonics are present in the wide band, but the noise is equal to the signal in amplitude.

On the other hand, for strong signals, the echo signal will have a sharp leading edge which is conducive to accurate range measurement.

In applications where accuracy is more important than range, the band width of the receiver may be kept as large as possible, consistent with the maximum range required.

Narrow Band Response

Fig. 287 shows the output obtained when the same signal is fed to a receiver with narrow band response.

In this case all the important harmonics are not present, consequently the pulse is considerably distorted and has

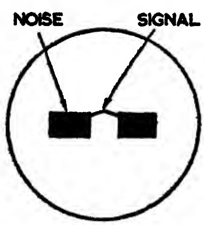


Fig. 285.—WIDE BAND RESPONSE.

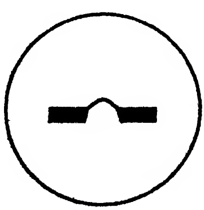


Fig. 286.—MEDIUM BAND RESPONSE.

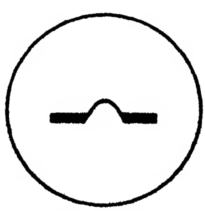


Fig. 287.—NARROW BAND RESPONSE.

smaller amplitude, but the noise is reduced well below signal level. The leading edge is ill-defined.

Hence a much weaker signal can be recognised in the noise. The receiver with narrow band-pass is, therefore, effective for extremely long-range warning sets, where the accuracy of range measurement required may be as low as 2 to 3 per cent. of the total range.

Medium Band Response

Fig. 286 shows a compromise selection which produces a reasonable pulse shape and a tolerable signal-to-noise ratio in the receiver.

In order to receive and reproduce satisfactorily, rectangular pulses sufficiently greater in amplitude than the noise voltage, the receiver band width should be approximately $\frac{2}{\text{pulse width in seconds}}$ for optimum conditions. One megacycle per second is usually added to the band width thus obtained to compensate for frequency drifts in either the magnetron or the local frequency oscillator.

For example, let the pulse width be, say, 2 microseconds, then the required receiver band width is :—

$$\text{Band width } \frac{2}{2 \times 10^6} \text{ plus } 1, \textit{ i.e.}, 2 \text{ megacycles.}$$

Since the control of signal-to-noise ratio by design of band width is limited by the factors mentioned in the foregoing discussion, it is now necessary to examine and analyse the factors that produce thermal agitation, shot effect and induced effects, in order to ascertain the extent to which they can be minimised by careful design.

Thermal Agitation

At any instant, due to the random motion of electrons in a conductor, there are likely to be more free electrons moving in one direction than in the other. This causes a voltage to be developed across the conductor. If the temperature of the conductor is now raised, the agitation of electrons in both directions increases and the instantaneous currents are therefore greater. In consequence the IR drop (thermal agitation noise voltage) increases with temperature and with resistance.

Shot Effect

Shot effect is caused by irregular emission from the cathode. The electron flow in the anode circuit varies slightly in regard to the number of electrons reaching the anode from one instant to another, and also in the velocities of individual electrons. This produces a small voltage variation across the load impedance, which produces noise when these irregularities are entirely random.

When a positive grid is placed in the electron path to divide the electron flow to the anode, the shot effect is magnified, because the division of electrons is also irregular. For this reason, therefore, pentodes and multi-grid valves are more noisy than triodes.

High mutual conductance triodes are therefore used to minimise shot effects, since signal control causes no division of the electron stream and, therefore, excludes a source of irregularity which must tend further to aggravate the unwanted shot effect.

An increase in the space charge by using higher filament temperatures may help in smoothing out emission variations.

Induced Effects

Strong electromagnetic and electrostatic fields may induce voltages and currents in resistances, leads, and even in the valves themselves. Irregularities in the flow of the electron stream may also induce current flow in the *grid circuit.*

Electrons moving past the grid induce charges relative to their positions. If the flow is constant in velocity the nett result will be constant, but random variations occur, resulting in a to-and-fro movement of electrons on the grid, which is grid current. This action takes place due to transit of the electrons from cathode to anode and must be distinguished from, and is not due to, electrons transported directly from the cathode to the grid.

All these voltages produced in the grid circuit are amplified in due course by the valve. The magnitude of the noise voltage on the grid, due to the above causes, increases with frequency and therefore noise voltages from this source depend upon the band width and the frequency.

Stray fields external to the valve may produce noise effects by induction in addition to which voltages resulting from insufficient filtering of the anode voltage supply may also prove troublesome. Long feed lines are particularly susceptible to pick-up from stray fields.

In general, therefore, induced noise voltage is reduced by the use of low impedance circuits, shielding, filtering and short leads.

R.F. Amplification

The difficulty in obtaining useful amplification becomes more difficult as the frequency increases. Screen-grid valves are less useful at the higher frequencies :—

(*a*) For reasons given in previous paragraphs they produce larger noise voltages than triodes.

(*b*) The inductive reactance X_L of the cathode leads at ultra-high frequencies causes negative feedback.

Signals, and part of the noise voltage, are affected equally by negative feedback. The noise voltage resulting from random division of current between anode and screen is however, not affected, because both currents flow in the cathode. The result is generally an overall reduction of signal-to-noise ratio from this source.* Other factors controlling the useful limit for amplification are discussed in Chapter XVI. Already it has been pointed out that irregular increases in the flow of electrons from the cathode to the anode induces grid currents. The effect of this is to introduce into the grid-cathode circuit an apparent leading impedance. The higher the carrier frequency the lower becomes the grid input impedance. Furthermore, since the electron stream is varied by the applied signal, this tends further to lower the input impedance.

Thus the physical grid circuit constructed from inductance and capacity elements is paralleled by the apparent input resistance and capacity of the valve. For these and similar reasons discussed in Chapter XVI, R.F. amplifiers are not used extensively in the micro-wave region.

Paralysis

Paralysis occurs when, notwithstanding protection provided by the T.R. and T.B. devices,† strong signals enter the receiver from the transmitter. These signals may overdrive the valves and paralyse the circuits of the receiver by virtue of residual charges, and so render it insensitive to signals which arrive shortly after the transmitted pulse has ended.

This effect may seriously impair the *minimum* range effi-

* A reduction of signal-to-noise ratio means that the condition is unsatisfactory. A high signal-to-noise ratio means increased efficiency.

† May be thought of as means for automatically switching the aerial from transmitter to receiver for the period of time that each unit functions.

ciency of the radar system and affects adversely the ability to detect signals received from targets at short ranges.

Temporary paralysis of the receiver results generally from excessive bias developed on the grid of one or more of the receiver valves by the transmitter pulse (it usually occurs in the video stage, which is resistance coupled).

This unwanted bias may be generated by grid current or by excessive current through the cathode biasing resistance if it is not by-passed by a condenser.

The receiver stages preceding the second detector are not usually subject to this trouble, because the signal has not then been amplified sufficiently to overdrive the valve. Inductances, rather than resistances, are therefore used for grid leaks and cathode-biasing resistances are by-passed.

Receiver paralysis can be minimised by the application of a square pulse of suitable width to one or more of the stages to bias the valve or valves beyond cut off, or to remove entirely the H.T. supply voltage for the duration of the transmitted pulse. This latter procedure is adopted when it is desired to receive signals at the earliest possible moment after the end of the transmitted pulse (minimum range detection).

If the second detector produces negative pulses, the first video stage may be used as a limiter. If the bias is supplied from a power supply through a voltage divider, which itself draws a comparatively large current, conduction by the valve produces only a negligible voltage change in the biasing resistance.

When paralysis does occur its duration is determined by the time constants involved. The effect can therefore be reduced by making the time constants and therefore the recovery time as short as practicable. By leaving the cathode resistor unby-passed the time constant is made practically zero, so that recovery time is reduced, but unless H.T. is suppressed for the duration of the transmitter pulse, paralysis is promoted by overdriving into grid current.

Circuit Arrangements for U.H.F. Reception

It is assumed that the reader is familiar with the general principles of the superheterodyne receiver for W/T and R/T applications. In case of difficulty, reference should be made to some standard work on radio communication.

Superheterodyne receivers for radar applications differ from those employed for W/T and R/T applications in several respects :—

(*a*) Receivers for the centimetre wave range have, in general, no R.F. amplifying stage, since the signal-to-noise ratio is so low, that considered in conjunction with the low gain obtainable at these frequencies, R.F. amplification is not worth while. At U.H. frequencies a klystron is usually employed as local oscillator, and the mixing function is performed by a crystal mounted in a resonant cavity.

(*b*) Another point of difference is that the local oscillator, the crystal mixer and the first two or three stages of the I.F. amplifier are usually located as close to the receiving branch of the wave-guide system as possible, *i.e.*, near the aerial. This is done in order that a signal of reasonable magnitude can be passed along a coaxial cable or transmission line, from the aerial position to the receiver position, the intervening distance in some cases being 50 to 60 ft. or more. If this is not done, very weak signals received in the aerial would be so greatly attenuated in passing over the connecting cable that they would be useless on arrival at the receiver position. Thus the first stages of the I.F. amplifier situated at the aerial position may be regarded as a remote grid-frequency pre-amplifier.

(*c*) Some form of frequency control, preferably automatic, must be applied to the local oscillator, in order to correct any change in the intermediate frequency caused by variations in frequency of the magnetron, and therefore, of the transmitted signal or variations in the frequency of the local oscillator itself.

(*d*) In radar, fading is important only when a particular echo is under observation and being tracked. Automatic gain control or A.V.C. is not therefore applied to the output signals of the receiver indiscriminately, as in the case of W/T and R/T. Provision is made, in radar,. for automatic gain to be applied *only* to the *particular signal which is being tracked*. In other words, a signal-selecting operation must first take place before automatic gain is applied.

(*d*) In common with receivers for television, the output is at video frequency, *i.e.*, ranging between a few cycles per second and perhaps 4 megacycles per second. Such band widths are made necessary on account of the pulse form and consequently the harmonics that must be faithfully reproduced in order to preserve pulse shape.

Local Oscillator

The theory of operation of the klystron has been described in Chapter XVI. This type of valve must be used as the local

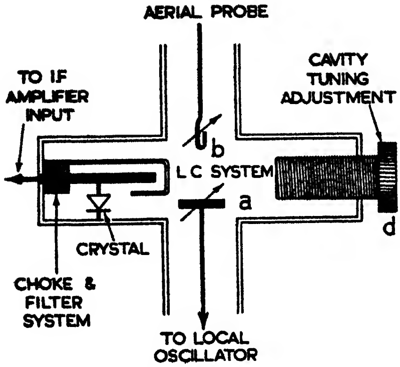


Fig 288.—SHOWING ONE METHOD OF MOUNTING A CRYSTAL MIXER IN A CAVITY RESONATOR.

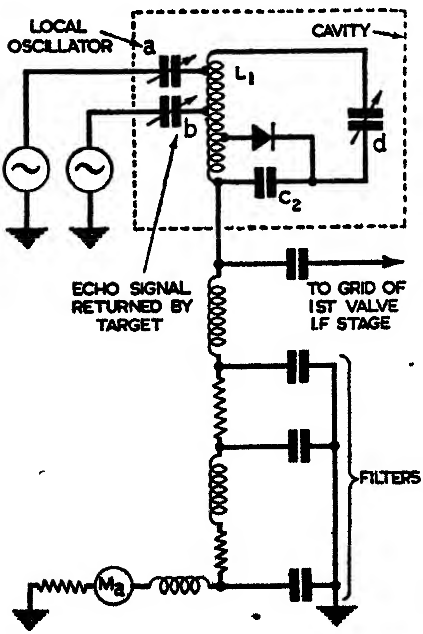


Fig. 289.—EQUIVALENT CIRCUIT OF FIG. 288.

C_2 and L_1 are derived from the L.C. system of the resonant cavity.

oscillator for reception of waves with frequencies corresponding to 10 cm. or less, in order to generate oscillations to beat with the incoming signals. Mixing is performed by a crystal mounted in a resonant cavity or a resonant section of the receiving branch of the wave-guide system (Fig. 288).

The cavity can be used, as explained in Appendix III, as a resonant circuit in place of coils and condensers, the equivalent circuit being shown in Fig. 289.

Since no R.F. amplifier is employed, the mixer, as the input stage, is the chief source of noise at the input, consequently the crystal to be used is selected with due regard to minimum noise contribution.

The gain of the mixer stage is generally of the order of plus or minus 1.

The output of the local oscillator may differ from the received frequency by a frequency in the region of 30 megacycles or so. This output and the received signal beat together in the mixer to produce a current with several frequency components. These components of current contain frequencies of the received and local oscillator signals, their higher harmonics and their sum and difference. The difference frequency is selected as the intermediate frequency, and is fed to the input of the I.F. amplifier.

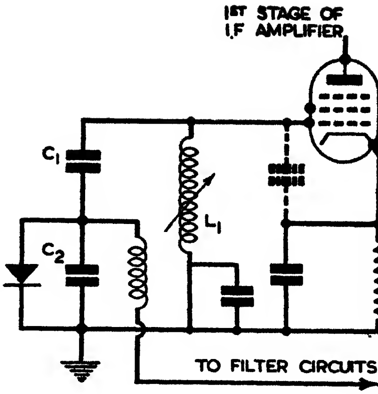


Fig. 290.—ONE FORM OF INPUT TO INTERMEDIATE CIRCUIT STAGE.

The signal from the local oscillator is fed into the mixer cavity by a probe and the mixer cavity is adjusted until the crystal current as read on the meter is of the required magnitude. This value in some cases may be in the region of plus or minus 0·3 milliamperes. A final adjustment can be made by the probe which couples the aerial input into the cavity. The high-frequency components are by-passed by the filter shown in the equivalent circuit diagram, an additional filter circuit formed by the *elements* indicated in the equivalent circuit diagram filters the selected intermediate frequency out of the crystal current (Figs. 288 and 289).

It will be noted that all the capacities, inductances and resistances shown in the equivalent circuit diagram do not consist of normal coils, condensers and resistances. Some are structural elements, equivalent to capacity, inductance or resistance, by virtue of the similar effects which they produce at microwave frequencies.*

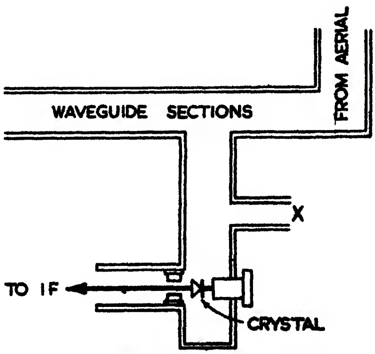


Fig. 291.—ALTERNATIVE FORM OF MOUNTING CRYSTAL IN RECEIVER BRANCH OF THE WAVE-GUIDE SYSTEM.

T.R. and T.B. switches are not shown. Position of T.R. is roughly indicated by X.

A form of input to the I.F. amplifier is shown in Fig. 290. L_1 might be tuned to resonance at 30 megacycles with its dis-

* In general the frequency range to which this technique applies is from about 3,000 megacycles and upwards (see Appendices II and III).

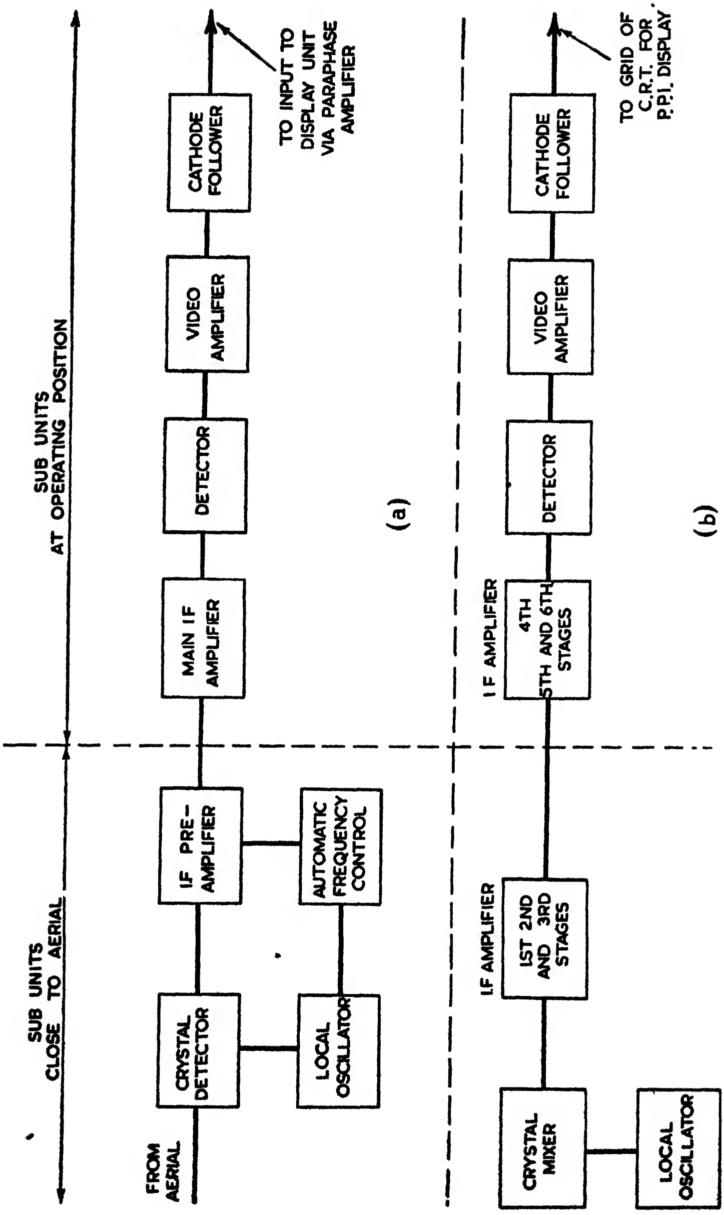


Fig. 292.—TYPICAL ARRANGEMENTS FOR UNITS AND SUB-UNITS OF A RECEIVER ASSEMBLY.
(a) With A.F.C. and (b) without A.F.C. Generally A.F.C. is regarded as indispensable in modern technique.

tributed capacity, input valve capacity plus C_1 and C_2 in series.

The crystal mixer acts as a resistance shunted across C_2. Its reflected resistance * loading the tuned circuit is somewhat greater because of the step-up ratio of C_1 to C_2. The voltage applied to the grid of the valve in the first stage of the I.F. amplifier is, therefore, the crystal voltage stepped up several times. Fig. 291 shows a somewhat similar alternative arrangement.

The I.F. amplifier circuit is conventional.

A block diagram of the circuits of two receivers is shown in Fig. 292 (a) and Fig. 292 (b) for comparative purposes.

Automatic Frequency Control

Any frequency drift of the carrier or local oscillator frequencies causes the intermediate frequency to change by the same amount. To compensate for this drift, the I.F. band width may be increased in the design of the receiver. Any increase in band width will, however, raise the noise level, but if this increase is not made the pulse will be distorted, owing to loss of high-frequency components.

The performance of the receiver can be improved under these conditions by automatic frequency control, which avoids both difficulties. If for any reason the intermediate frequency changes, the automatic control brings it back to its proper value by readjusting the frequency of the local oscillator.

In Fig. 293 the input signal is at an intermediate frequency of 60 megacycles. A drift of the I.F. produces a D.C. voltage change which, after amplification, is impressed on the repeller grid of the klystron oscillator to return it to the correct frequency.†

Fig. 293 contains an I.F. amplifier with input from the normal channel. The secondary of T_1 is tuned to resonance at intermediate frequency. The coupling of primary and secondary, and tuning of the primary, are adjusted to give a voltage across the secondary that differs by 90° in time phase with the

* The term "reflected resistance" is used in this instance because of the voltage transformer action which is brought about by the input circuit arrangements and the voltage across L is inversely proportional to the magnitude of $\frac{C_1C_2}{C_1 + C_2}$ as compared with C_2. C_1C_2 are in series with L as far as the input is concerned.

† See description of the klystron, Chapter XVI, in which it is pointed out that the frequency of this device can be slightly varied by a change of voltage on the repeller grid.

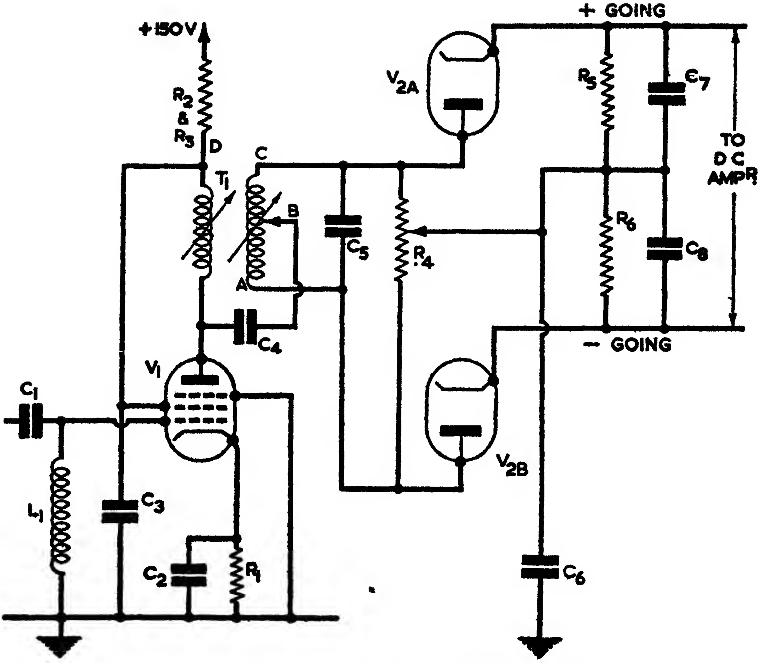


Fig. 293.—FREQUENCY DISCRIMINATOR FOR AUTOMATIC FREQUENCY CONTROL.

primary voltage. The primary is connected to the outer tip of the secondary by means of C_4. Fig. 294 (a) shows the equivalent circuit for the discriminator circuit which follows the I.F. stage and the various voltage relations for the frequencies above, below and at resonance.

As shown in the equivalent circuit, Fig. 294 (a) and (b), the voltage applied between the upper diode anode and cathode is the drop across the resistance between A and D (Fig. 293). This is the vector sum of the voltages across the upper half of the secondary A to B and the voltage of the primary B to D. Similarly, the voltage applied to the lower diode anode is that across the lower half of the secondary C to B (Fig. 294) plus the primary voltage B to D (Fig. 294), or B to D of T_1 in Fig. 293.

When the I.F. signal is at its proper frequency the voltage across the secondary A to C is 90° out of phase with the voltage across the primary B to D. Thus E_{CB} leads E_{BD} by 90°, while E_{AB} lags E_{BD} by 90°. Since the secondary is centre-

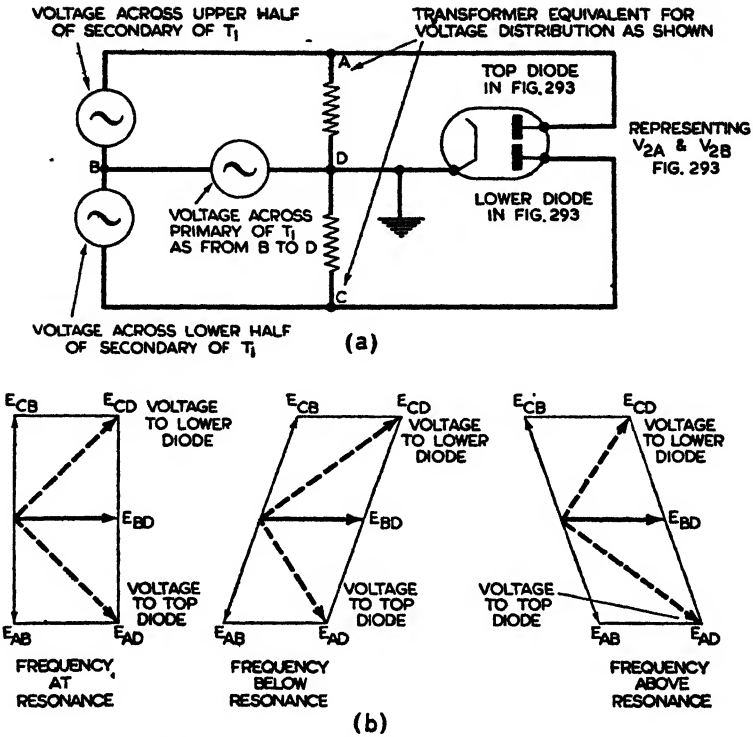


Fig. 294.—AUTOMATIC FREQUENCY CONTROL OBTAINED BY APPLICATION OF AMPLIFIED OUTPUT FROM THE DISCRIMINATOR TO THE REPELLER OF THE KLYSTRON LOCAL OSCILLATOR.

(a) Equivalent R.F. circuit of discriminator.
(b) Vector diagram of equivalent circuit of discriminator.

tapped to make E_{AB} equal to E_{CB}, the vector sums are equal in magnitude.

Equal signals in the two diode anodes produce equal currents in the cathodes, which in turn produce D.C. voltage drops across R_5 and R_6 (Fig. 293), which are equal, but of opposite polarity. The output to the D.C. amplifier which follows the discriminator of Fig. 293 is therefore zero.

If the I.F. changes in frequency, the secondary is no longer tuned to resonance voltage, and voltages from A to C (Fig. 294) no longer differ by 90° each lag relative to B D. If the frequency decreases, E_{AB} lags more than 90°; if the frequency increases, a lag of less than 90° occurs

In the former case the voltage applied to the lower diode anode is greater, and the output to the D.C. amplifier is negative going. In the latter case conditions are reversed and a positive-going output is produced (see Fig. 293).

After being amplified by a Loften-White circuit, the D.C. voltage is used to correct the klystron or local amplifier.

Automatic Gain Control

The amplitude of the input to a receiver, of the echo from a distant target may vary, because of fading and changing in position of the target at a more or less rapid rate. If the receiver has a constant gain there is a corresponding variation in the amplitude of the output. Automatic gain control may be used to give greater gain when the signal is weak than when it is strong. Thus the amplitude of the output is maintained relatively constant.

Frequently A.G.C. is obtained by using the signal voltage to control the bias on one or more valves.

Alternatively, the voltage developed from the received signal may be used to control the gain by letting it regulate the supplies for the anodes and screens. When this method is used, several stages of D.C. amplification are required to furnish the A.G.C. voltage, the last stage being a power amplifier.

In radar sets where A.G.C. is used, *the controlling voltage is developed from the signal of the particular target under observation,* because the signals of all targets displayed at the same time vary independently of one another.

Thus it is essential initially to adjust a signal selector circuit in order to select the interval of time which includes the particular echo under observation. Only the signal received during this selected time interval influences the receiver gain. Thus echoes from targets at ranges appreciably different from that of the selected target have no effect on gain.

In practice, the operator places on the indicator screen a strobe against the echo of the target to be tracked and continuously adjusts the position of the strobe to coincide with the movements of the target.

Gain Control of a Selected Signal

Automatic gain control of a selected signal can be accomplished in the following manner :—

A Synch. pulse starts up a square-wave generator (strobe generator control) at the same time as the time base generator

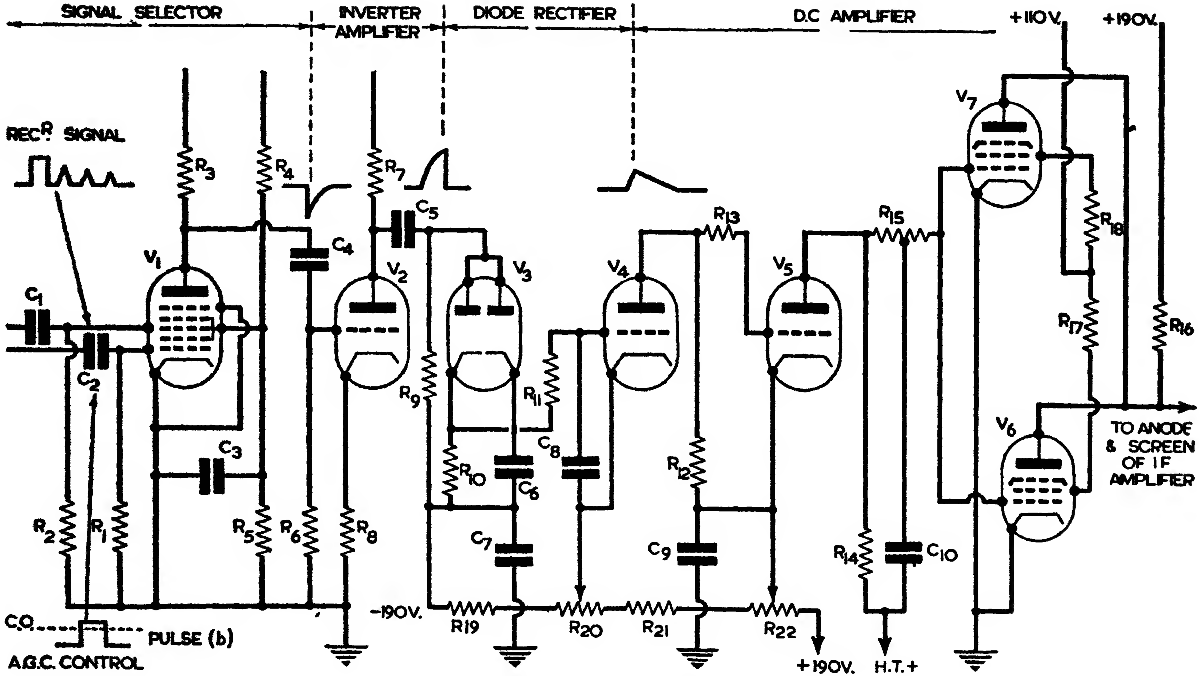


Fig. 295.—AUTOMATIC GAIN CONTROL.

The typical receiver output shows the ground wave and echoes from three separate targets. A.G.C. may be applied to any of the three echoes by manipulating the manual control so as to place a strobe against the echo to be tracked. A.G.C. can only be applied to one echo at a time. To apply A.G.C. to the most distant echo, A.G.C. pulse must be moved to right.

commences each successive sweep of the time base. The output from the square-wave generator is a pulse, the width of which (delay time), can be varied by the operation of a manual control to make its falling edge coincide with the trailing edge of the selected echo. The falling edge of the square pulse which occurs at some instant of time after the start of the time base, provides a trigger pulse to start up a strobe generator which may have two outputs :—

(*a*) A short, brightening pulse, which causes a bright spot to appear on the indicator screen at some point along the time base selected by the manual control.

(*b*) A pulse of the same duration as (*a*), which actuates the gain control circuit and causes automatic gain control to be applied to any signal received at the instant of time corresponding to the position of the strobe on the time base and for the duration of the actuating pulse (*b*).

Thus when it is desired to apply gain control to any particular signal under observation, the operator continuously operates the manual control to cause the strobe to follow the echo signal as the target range closes or lengthens.

In these circumstances the position of the strobe at any instant of time, with respect to each successive sweep of the time base, coincides with the point on the time base at which the selected signal appears. The pulse from the other output of the generator (*b*) actuates the gain control circuit of the receiver at this same instant, and therefore causes automatic gain control to be applied to the receiver output, *i.e.*, to the selected signal for the duration of pulse (*b*). Pulse (*b*) is made just long enough to cover the duration of the selected echo signal for each sweep of the time base. No other signals are affected by the A.G.C. circuit.

Analysis of Fig. 295

The output of the receiver and the short square pulse (*b*) developed by the strobe generator are applied to V_1 as shown in Fig. 295. The receiver output is not sufficiently large, normally to make V_1 conduct (V_1 is biased back), and the gain control circuit is therefore inoperative until the arrival of pulse (*b*), which is large enough to cause V_1 to conduct, but for its duration only. The signal selected is amplified and inverted by V_2 to produce positive pulses. V_3 is a diode which rectifies the pulses to produce an approximately steady D.C. voltage across the filter R_{10} and C_6 The magnitude of this signal is

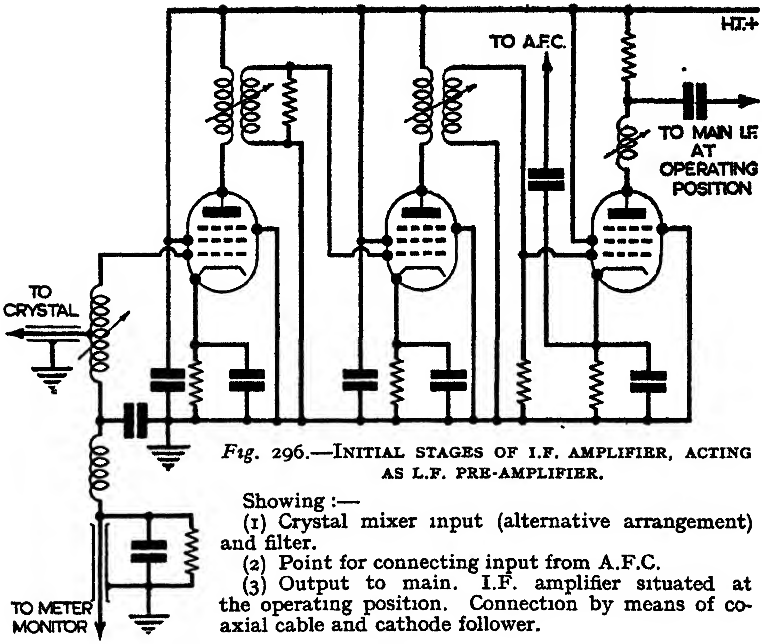


Fig. 296.—INITIAL STAGES OF I.F. AMPLIFIER, ACTING AS L.F. PRE-AMPLIFIER.

Showing :—
(1) Crystal mixer input (alternative arrangement) and filter.
(2) Point for connecting input from A.F.C.
(3) Output to main. I.F. amplifier situated at the operating position. Connection by means of co-axial cable and cathode follower.

proportional to the strength of the R.F. signal returned by the target which is being tracked.

The output of the diode is applied to the grid of V_4, which is directly coupled to the grid of V_5.

In order to obtain direct coupling to these valves and to the parallel amplifier V_6, V_7, the required voltage is obtained from the voltage divider R_{19}, R_{20}, R_{21} and R_{22} ; C_8, C_9 and C_{10} form a filter to remove ripple.

The magnitude of the grid voltages of V_6 and V_7 depends upon the echo signal strength received from the selected target. If the signal is large, these valves draw a heavy anode current and the anode voltage is low. If the signal is weak, these conditions are reversed.

The anodes and screens of the I.F. amplifiers are connected to the anodes of V_6 and V_7 so that the gain of the I.F. amplifier is controlled directly by signal strength.

The speed of response of the A.G.C. circuit is fixed by the time constant of the ripple filters; these are set so that successive pulses can be received with little change to the D.C. output so long as their amplitudes are equal. If the average value changes, a change is produced in the output, which will correct the gain of the R.F. amplifier and return the amplitude of the selected signal to its proper value.

Fig. 296 shows the circuit of the initial stages of an I.F. amplifier when used as an L.F. pre-amplifier. An alternative method of connecting the crystal mixer is shown, together with the junction for A.F.C. and the output to the main I.F. amplifier at the operating position.

Some Important Notes on Receivers

The importance of the efficiency of the receiver in regard to maximum range cannot be too strongly stressed.

(*a*) A small increase in sensitivity of the receiver has a greater effect on increased range than a comparatively large increase in transmitter output. From this it follows that every care must be taken to maintain the receiver at the highest point of efficiency at all times. Apart then from the maintenance of this unit in a high state of electrical efficiency, since the ratio signal/noise determines the minimum echo signal that can be distinguished from *noise*, care must be taken to check, in the ordinary course of maintenance, that this ratio has not deteriorated.

(*b*) It is quite possible to observe plenty of " grass " on the screen of the C.R.T. and yet have very poor echo response from targets. (In radar terminology, " grass " is the fringing of the time base resulting from " noise." When a green display is used there is considerable resemblance.) Therefore the presence of plenty of " grass " *does not necessarily indicate that the receiver is functioning satisfactorily*.

In cases, therefore, where direct echoes are not available, the receiver should be tested within the limits of the test equipment provided. See notes on Test Equipment, Appendix IV, " Echo Box," which is a help.

CHAPTER XX

DISPLAY UNITS OR INDICATORS

In the broadest sense, an indicator is a device in which an electron beam is used to plot any two variable quantities whose values can be made proportional to either voltage or current.

There are two general cases, *i.e.*,

(*a*) When the beam is not modulated.
(*b*) When the beam is modulated.

When the beam is *not* modulated the entire trace is visible. When the beam is modulated the plot is for the most part traced in darkness,* selected points being indicated by bright spots. The co-ordinates of these selected points are instantaneous values of the two variables, taken along the X (horizontal) axis and the Y (vertical) axis respectively.†

The "A" and "B" Displays

As examples of these two cases, take the "A" display and the "B" display.

In the case of the "A" display, time, and range proportional to time, are plotted along the horizontal or X axis by the deflection of the electron beam under the influence of a linear time base voltage. The entire trace is visible, and signal amplitude is plotted against time or range by vertical deflections of the beam (and therefore of the trace), the magnitude of the deflection being proportional to the amplitude of the received signal.

In the case of the "B" display, azimuth is plotted along the X axis and time or range is plotted along the Y axis. To accomplish this, voltages proportional to the angular displacement of the aerial from the reference position (compass north on land, or ship's head at sea) are applied to the X plates, and a linear time base voltage is applied to the Y plates.

Instantaneous values of these two variables are plotted by the beam in comparative darkness, points for illumination being

* "Darkness" is used here in a comparative sense and may actually mean very low degree of illumination as compared with normal fluorescence.

† This condition can be expressed in the j notation as $X_0 = A + jB$, assuming the time base to be absolutely linear.

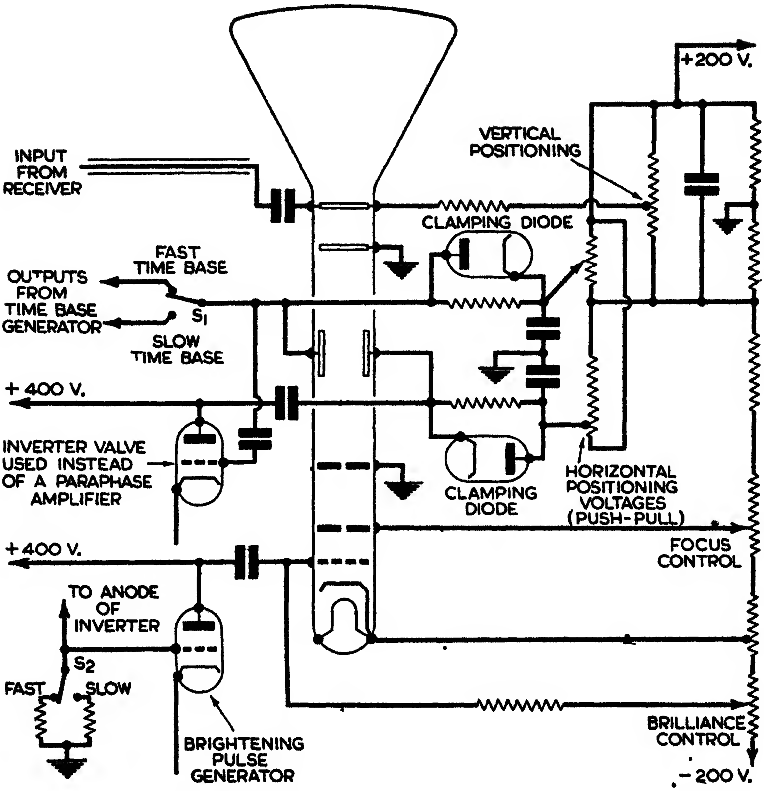


Fig. 297.—CONNECTIONS TO CATHODE RAY TUBE (ELECTROSTATIC TYPE) FOR TYPE "A" DISPLAY.

S_1 and S_2 are ganged. The clamping diodes limit the minimum and maximum swing of the time base voltage and fix the points at which the trace commences and ends.

selected by echo signals returned from targets which are applied to the control grid to intensify the beam at the instant of their arrival at the receiving aerial.

The P.P.I. Display

The other basic indicator, which is the P.P.I., is really a modification of the "B" display. In this case angular displacement of the aerial, which is rotated in azimuth continuously, is plotted against range. This is accomplished by a magnetic field proportional to linear time base current in one set of

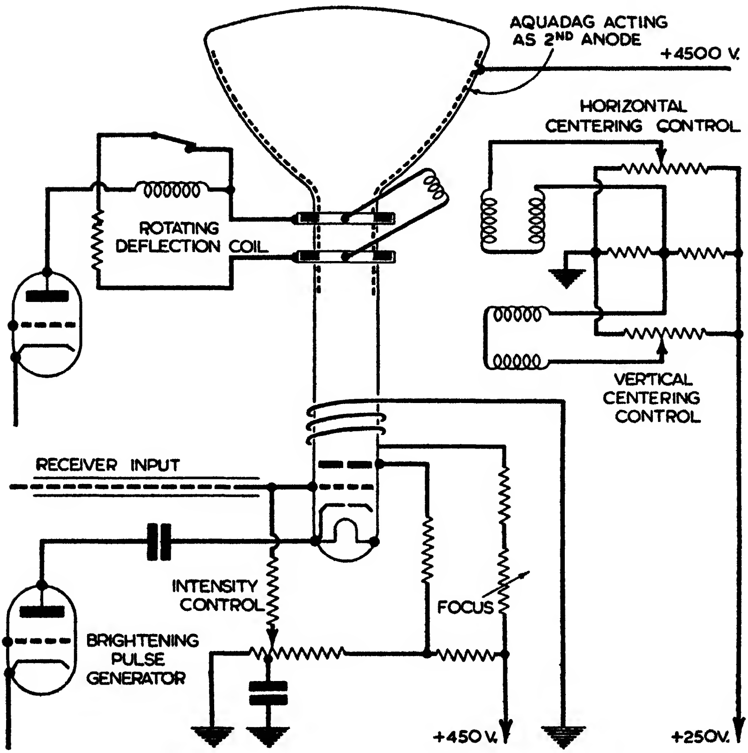


Fig. 298.—ELECTROMAGNETIC TUBE ARRANGED FOR P.P.I. DISPLAY.

Inputs from time base brightening pulse generators are schematic. Coils of vertical centering control are mounted on the neck of the tube.

deflecting coils which is rotated mechanically, in step with the aerial. The signals returned by reflecting objects in the area scanned by the aerial during each revolution are applied to the control grid.

In this manner the screen of the cathode ray tube is scanned by the beam (in semi-darkness), just as the aerial scans the area surrounding its location. Since these two scans are in step, signals returned by reflecting objects scanned by the aerial, cause bright spots to appear at corresponding points on the screen.

The position of these points is determined for range and bearing by the corresponding instantaneous receipt time measured along the time base and the angular displacement of the time base respectively.

The intensity of the bright spot will vary with target range, height above the waterline (not altitude) and nature of the reflecting object.

Thus a P.P.I. display is a form of presentation similar to the outline of a polar map of the area with an indication of comparative contours provided by a brilliance/range ratio, *i.e.*, the echo returned from a tower situated inland may be greater than that returned by the foreshore.

Height-finding Indicators

Indicators for height finding may be modifications of either the " A " or the " B " basic displays.

Height-finding indicators are subject to certain limiting conditions. In Chapter I it was pointed out that there are two cases :—

(*a*) The case where the vertical width of the beam is so narrow and the elevation is such that no part of the beam touches the earth. There is no reflected component. In this case the slant range can be measured by the direct ray.

(*b*) The other case is that in which the vertical pattern of the beam is so wide or the elevation is so low that the beam does touch the earth. In this case the vertical field pattern of the aerial is no longer regular and the lobe representing the field strength distribution is broken up.

The modified pattern of the vertical field can, however, be mapped and shown diagrammatically, as explained in Chapter VII. In case (*b*), this pattern plays an important part in both range and height finding.

In case (*a*) both slant range and angle of elevation can be ascertained from the display. If θ is the angle of elevation and R is the slant range, the height of the reflecting object in space is R sin θ (not allowing for the curvature of the earth).* Height-finding indicators are available for determining height directly from a special display and by means of a special courser manipulation of which gives a direct reading of the altitude of a target above earth. This, together with the slant range and bearing determines fully its position in space.

In case (*b*), no really direct reading is possible, and it is

* Indicators designed for direct reading of height employ a time base which is not quite linear and which compensates for the curvature of the earth. Thus R sin θ gives a correct reading of height.

necessary to consult the vertical polar diagram for the particular aerial system in use.

If the amplitude of the signal returned by an object in space is measured, this information together with the measured range, enables the height of an object to be estimated from comparison with similar standard data for the vertical polar diagram. Special types of indicators are supplied for this purpose.

An approximate estimate of height can be arrived at by referring the data available from an " A " display directly to the vertical polar diagram of the aerial system. This method can, however, produce some very misleading results, particularly when the target consists of several aircraft bunched together.* The foregoing are not necessarily the only methods employed for height finding, but they do serve to indicate some of the difficulties of the problem and the general principles involved in its solution.

Thus the basic indicators are :—

(*a*) The " A " display.
(*b*) The " B " display.

The P.P.I. and height-finding indicators are special cases of the " B " display and the " A " and " B " displays respectively.

Modifications for Special Applications

Each of these basic indicators may be further modified by the addition of refinements and facilities for special applications.

In the simplest form of indicator, range is estimated by interpolation from the time base scale calibrated by range marks. The various methods of producing and checking these range marks are described in Chapter XV, " Calibration."

Precise measurement of range requires the use of a more accurate method than interpolation or estimation. This is dealt with in detail in later paragraphs subsidiary to which one or more of the following facilities may be incorporated for special applications :—

(*a*) Selected range. (This becomes a selected sector for a P.P.I.)
(*b*) Strobing or movable range markers.
(*c*) Time base expansion.
(*d*) Multiple trace displays.

* The energy of an echo returned from a number of aircraft bunched so closely together that definition of individual planes is not possible, must be larger than that returned from a single aircraft at the same range. Consequently a distinguishable echo may be received in this case where normally a useless echo would be returned by a single plane.

For search or early warning sets, it is desirable to scan continuously as large an area as possible, in order to be able to track several targets simultaneously. For this application the P.P.I. display is very useful. It is generally supplemented with an " A " display, so that the character of any particular echo can be more carefully scrutinised for range, identification, etc.*

Selected Range

When short-range accurate sets are in operation, the target may be either on the surface or in the air, and the observer is generally interested in a particular target selected from the main display. For this application, therefore, an indicator of the selected-range type may be included in the system.†

If the set is associated with an anti-aircraft gun, the display must be arranged to follow very rapid changes in three dimensions, and some form of " B " display, with or without selected range facilities and embodying height finding, may be supplied.

Aircraft Displays

Radar sets are installed in aircraft for various purposes, each requiring a different form of presentation or display :—

(*a*) To detect and locate other aircraft.

(*b*) For detection and location of surface craft.

(*c*) For recognition through cloud and darkness permanent features and landmarks on the earth's surface.

In case (*a*), the display must be arranged in such a manner as to facilitate the pointing of one aeroplane at another.

As already explained, in fighter interception applications, aircraft are, in the first place, directed by R/T from the ground to a zone in which the range of the radar set installed in the fighter plane becomes effective. This enables the observer to pick up the target on a type " B " display. If the pilot's position is fitted with a remote indicator of the selected range display type, the target, located and followed by the observer on the " B " display, can be selected for display at the pilot's position (all others being excluded). This enables the pilot to point his plane at the target with the object of bringing it into visual range.

* The word " identification " is used here in a general sense. It may refer to the approximate or comparative dimensions of the target, or to distinguishing " friend " from " foe." For obvious reasons, the latter is not dealt with in this volume.

† A selected range indicator is capable of being adjusted manually, to give a full-scale display of some selected part of the main display, when the movements of a particular target are under close observation.

In cases (*b*) and (*c*) some form of P.P.I. display is employed. In case (*b*), however, this is usually supplemented with an "A" display which is not essential for application (*c*).

Complex Displays

From the foregoing it can be seen that a display assembly for a system may be simple or complex. By the use of a number of displays of various types, very complete information can be made available about the movements of all targets within range. Furthermore, this information can be collated and simultaneously distributed to any number of remote points.

In this respect, therefore, a radar set with its remote indicators resembles a broadcast receiving station with a number of extension loudspeakers, composed of a variety of types, each suited to a particular application.

Accuracy of Range Measurement

In the case of long-range warning sets, accuracy of the order of 2 to 3 per cent. of the total range, plus or minus, may be tolerated, but a much higher degree of accuracy is necessary for fire control. The angular accuracy of bearing is also subject to similar limitations.

Thus for long-range warning sets, estimation by direct reading or interpolation may be sufficiently accurate for extreme ranges. When a higher degree of accuracy is necessary, several alternative methods are available. The underlying principle of accurate range measurement is, however, the same in all cases, and the various methods employed differ only in respect of circuits and components which are used.

The underlying principle of accurate range measurement is that of aligning the echo signal with a fixed reference mark, or alternately moving a strobe or marker into alignment with the echo signal as it appears on the trace ; in either case, range is measured by the time delay that must be introduced in order to effect alignment.

Amongst the various methods by which this may be accomplished there are the following :—

(*a*) When the output from the synchronising source is a sine wave, phase shifting may be applied to move the signal into alignment with a fixed reference mark or to move a range mark from zero position to coincide with the echo signal.

(*b*) With an "A" display, range can be ascertained by measuring the value of the time base voltage at the instant that the echo signal appears. In order to secure accuracy by this method, the time base must be approximately linear (see Fig. 299).

(*c*) As a third alternative, a portion of the whole range may be expanded to occupy the full width of the screen. Thus any portion of the trace may be carefully scrutinised and range can be more accurately determined.

Precision Measurement (Alternative Methods)

Methods of calibrating the time base for range estimation have already been dealt with in Chapter XV.

Precise measurement of range involves the measurement of extremely short time intervals This is accomplished by

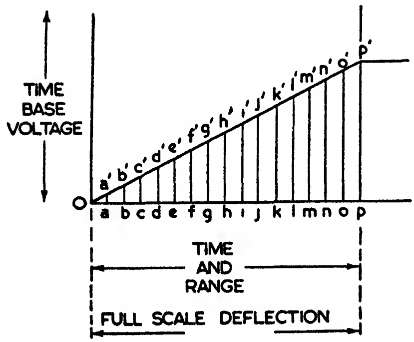


Fig 299—pp^1 IS PROPORTIONAL TO op SINCE THEY ARE SINE AND COSINE RESPECTIVELY OF THE ANGLE pop^1 WHICH IS CONSTANT CONSEQUENTLY aa^1, bb^1, cc^1, ETC, MUST ALL BE PROPORTIONAL TO O_a, O_b, O_c, ETC

The time base voltage measured at any instant is, therefore, proportional to range Thus the range of an echo appearing at, say, f on the trace is proportional to the voltage of f, f^1 at that instant.

introducing an accurately calibrated variable time delay between the transmitted pulse and a particular echo.

When the Synch. pulse for the radar system is developed from a sine wave, accurate range measurement can be accomplished by shifting the phase of the sine wave which times the time base generator, relative to the sine wave that times the transmitter.

This phase shift is used to delay the action of some circuit in the indicator by a measured number of electrical degrees and thereby to measure the time delay in terms of the period of one cycle of the synchronising sine wave. The device usually is geared to a dial, calibrated in suitable units of range.

In case (*b*) the magnitude of the time base voltage is measured at the position at which the echo appears on the time base. This method depends upon comparison of this time base voltage with the output from a ganged and calibrated D.C. potentiometer, comparison being made at the point on the time base

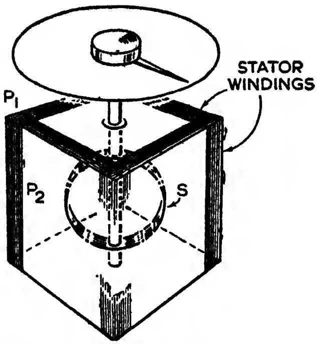


Fig. 300.—ARRANGEMENT OF THE HELMHOLTZ COIL OR PHASE-SHIFTING TRANSFORMER.

P_1 and P_2 are two primary or field coils, and S a secondary or search coil wound on a pivoted former.

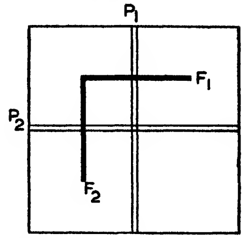


Fig. 301.—DIRECTION ONLY OF FIELDS F_1 AND F_2 OF COILS P_1 AND P_2 RESPECTIVELY (NOT SIMULTANEOUS).

The windings are shown intersecting at 90° at the centre of the cube. This may be taken as the equivalent of the winding shown in Fig. 300.

where an echo occurs. This alternative is not as accurate as the phase-changing method.

Other methods include the use of a delay multivibrator, and calibrated shift of the time base (and therefore the echo), relative to a fixed marker on the face of the indicator.

Measurement by Phase-shifting Method

The voltage obtained from a phase shifter should be of constant amplitude at all phases, in order to produce uniform operation of the circuits with which it is associated.

To make accurate measurements, one degree of rotation of the phase shifter should produce a change in phase of one electrical degree or less.

The Helmholtz coil, or phase-shifting transformer, is a device which enables the user to change or shift the phase of an A.C. voltage by an amount that is accurately known, in a smooth and continuously adjustable manner.* It is sometimes

* It is important to have a very clear understanding of the physical meaning of the term "phase." Actually, it represents the amount of displacement as a fraction of the maximum displacement, either of a particular particle, or at a particular point in a wave train.

Thus, at a distance l from the crest of a wave, $\phi = 2\pi\frac{l}{\lambda}$.

If l is a whole number or some integral multiple of 2π.

If $l = (N + a)\lambda$, where N is an integer and alpha some fraction $\phi = 2\pi(N + a)$.

In the following description the reference point is taken as zero, instead of maximum amplitude, but the same reasoning applies.

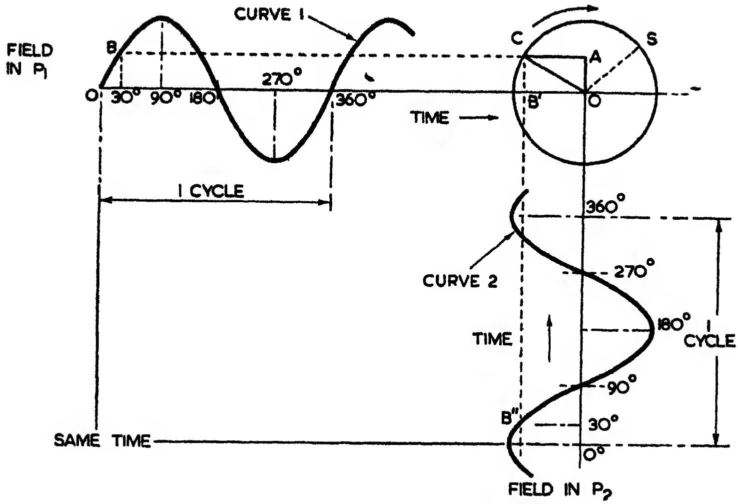


FIG 302.—DIAGRAM EXPLAINING THE PHASE-SHIFTING ACTION OF THE HELMHOLTZ COIL

The phase in the rotating coil shifts 360° for each revolution or 2π radians Hence if the rotor is revolved $7\frac{1}{2}$ times and the movement of C and S plotted horizontally the lag would be $7 \times 360° + 180°$ or $7.5 \times 2\pi$ radians, i.e., $2\pi\frac{l}{\lambda}$ (see footnote, p. 439)

referred to as a goniometer, or gonio., being similar in construction but different in application.

The device consists of two primary or field coils P_1 and P_2 wound on a cube, the direction of the two windings being at right angles (see Figs. 300 and 301).

Inside the cube a secondary or search coil is wound on a pivoted former, to which is fixed a pointer moving over a circular scale and divided into degrees.

If currents flow in P_1 and P_2 they produce fields F_1 and F_2. The field inside the pivoted coil will therefore be due to the combined fields of F_1 and F_2.

Let P_1 and P_2 be fed with two alternating currents differing in phase from each other by 90°. Fig. 302 shows the combined field.

In curve 1, the variation of the field of P_1 is shown on a time scale which is marked in degrees 0–360° for a complete cycle of the A.C. producing the field.

In curve 2 the variations of the field of P_2 are shown. The zero marks on the two curves indicate the same moment.

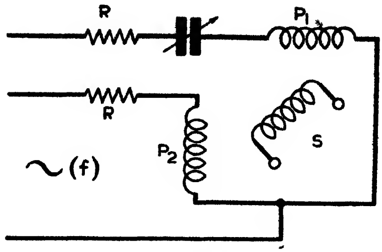


Fig. 303.—ARRANGEMENT FOR PRODUCING A PHASE DIFFERENCE OF 90° BETWEEN THE INPUTS TO P_1 AND P_2.

This is a condition necessary to the operation of the phase-shifting device as described in text.

In order to find the direction of the combined fields at a time corresponding to, say, 30° after zero, a circle is drawn with centre "O," where the two time axes cut one another, and with radius equal to the maximum height of the sine wave.

A line parallel to the time axis is now drawn through point "B" (of curve P_1) to "A." On curve 2 draw a line parallel to its time axis through "B." It can be shown that the point "C" at which these lines cut is on the circle.

OC represents the direction of the combined fields—expressed mathematically, if the line through A cuts the time axis of curve 1 in B, OA and OB′ are equal to the two fields, OB and OB″ respectively and OC, the diagonal of the parallelogram OA—OB′ is equal to their resultant.

It can also be shown that the angle COB′ is 30°, which corresponds to the point taken on curves 1 and 2.

If the same construction is carried out for all corresponding points on curves 1 and 2, a series of points "C_s" will be found and since this succession of points on curves 1 and 2 occur at steadily increasing times, it will be seen that C rotates at a constant speed round the circle—moving 1° for each 1° on curves 1 and 2, and completing the circle in one cycle of A.C.

If the secondary of the transformer is set at a definite angle, the magnetic flux through it will become zero when the axis of the coil is at right angles to the flux. At the moment corresponding to C, 30° in Fig. 302, the flux through the coil will be zero if the axis is along OS.

Thus by taking a number of consecutive positions or settings for C and S, it can be seen that turning the search coil through a definite angle in the direction of the arrow makes the flux become zero at a later time—the change in time expressed as degrees of the cycle of A.C. being equal to the angle (in degrees) through which the search coil has been turned. In other words, the phase change in degrees is equal to the rotation of the search coil in degrees.

It is this property that gives the device its name and makes it so useful in radar, when a time delay corresponding to a phase change can be used to measure a radar range.

The phase shift is continuous, *i.e.*, if the search coil is turned, say, $7\frac{1}{2}$ times, the phase shift would be $7 \times 360° + 180°$, or $2{,}700°$.

Fig. 303 is a circuit for producing two fields 90° apart in phase.

If the inductances of P_1 and P_2 are both L henrys, and C capacity expressed in farads, the frequency is f and the resistances of each are "R."

$$R \times 2\pi f L = \frac{1}{4\pi f C}.$$

The accuracy of the device as a delay measurer depends upon quantitative design, the accuracy to which the circuits are adjusted to comply with these conditions.

Two Methods using Phase-shifting Device

Range measurement by the use of some phase-shifting device may be carried out in two ways :—

(*a*) The entire trace may be shifted relative to a fixed reference mark.

In this case the reference mark is *zero* range, which coincides with the leading edge of the transmitter pulse. The counting mechanism attached to the phase-shifting device indicates, in yards, the corresponding number of electrical degrees through which the synchronising voltage of the time base generator must be delayed, relative to the Synch. pulse for the transmitter, in order to make the leading edge of the measured echo coincide with the leading edge of the transmitted pulse. In other words, it measures the delay in degrees, which must be introduced to make the time base generator (as normally synchronised to the transmitter) start at the same moment as the appearance of the leading edge of the echo undergoing measurement for range (Fig. 304).

(*b*) The time base is initially synchronised with the transmitter in the normal way and a reference mark (generally a strobe) is moved along the trace from the zero range position to the leading edge of the echo desired by the phase shifter; the output from the phase shifter being now applied to produce the time delay to control the Synch. pulse which starts up the range mark or strobe generator. The counting mechanism, operated by the phase shifter, indicates the yard equivalent of

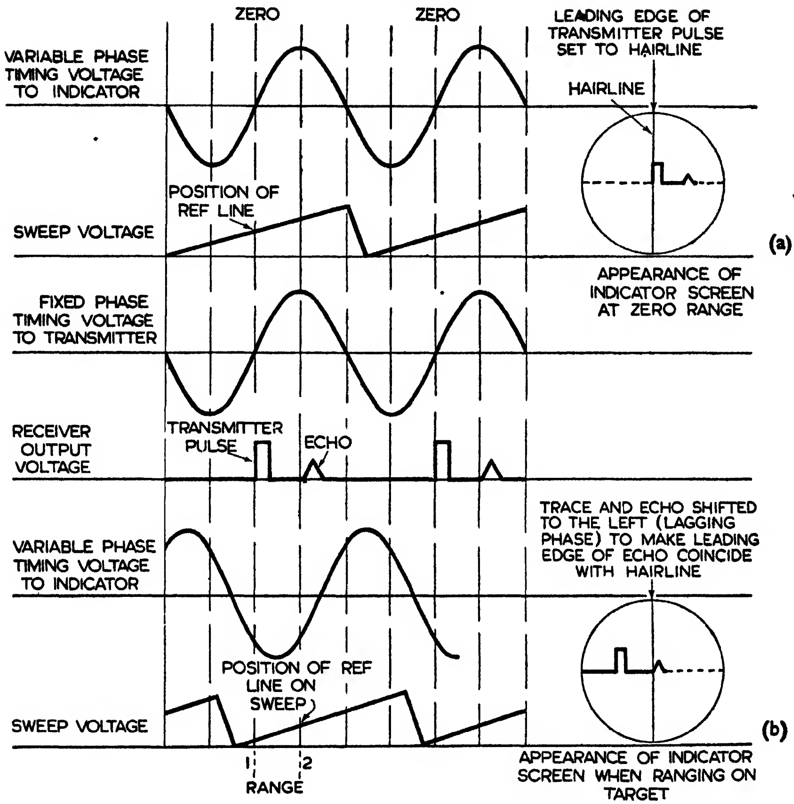


Fig. 304.—THE TRACE AND ECHO SHIFTED RELATIVE TO THE HAIR-LINE (POSITION (a)) TO POSITION (b) BY APPLICATION OF A PHASE SHIFT TO THE A.C. VOLTAGE WHICH SYNCHRONISES THE TIME BASE GENERATOR WITH THE FIRING POINT OF THE TRANSMITTER. POSITION (a) OBTAINED BY MANIPULATING THE HORIZONTAL SHIFT VOLTAGE CONTROL.

The range of the echo measured in yards is indicated on the phase shifter scale which is calibrated in terms of the number of revolutions required to align the leading edge of the echo with the hair-line as in position (b).

the number of electrical degrees through which the strobe or range mark generator must be made to lag the start of the transmitter pulse or time base generator, in order to make the output from the strobe or range mark generator coincide with the leading edge of the echo to be measured for range.

Figs. 304 and 305 indicate the basis of two methods of range measurement by the phase-shifting device. Summarising,

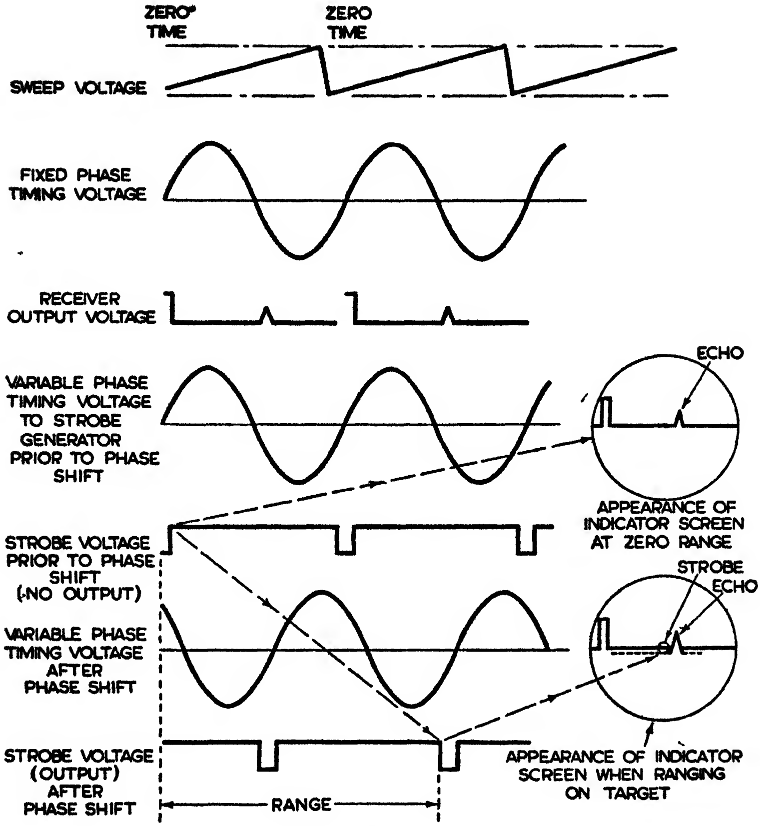


Fig 305—STROBE TIMED BY PHASE SHIFT.

Range determined by phase shift required to move an independently generated strobe from the firing point of the transmitter on the trace out to the leading edge of the echo

in the first method the reference mark—Fig. 304—is fixed and the trace is moved along to make the echo coincide with the fixed reference mark. In the second method the trace is produced under normal conditions and measurement is carried out by moving a strobe or variable range marker, controlled by the variable time delay, up to the leading edge of the echo. The lag of the strobe generator measures the range.

In applications of method (*a*) the reference mark is not *zero* range. In some cases a distinguishing reference mark may be provided by a transparency fitted over the

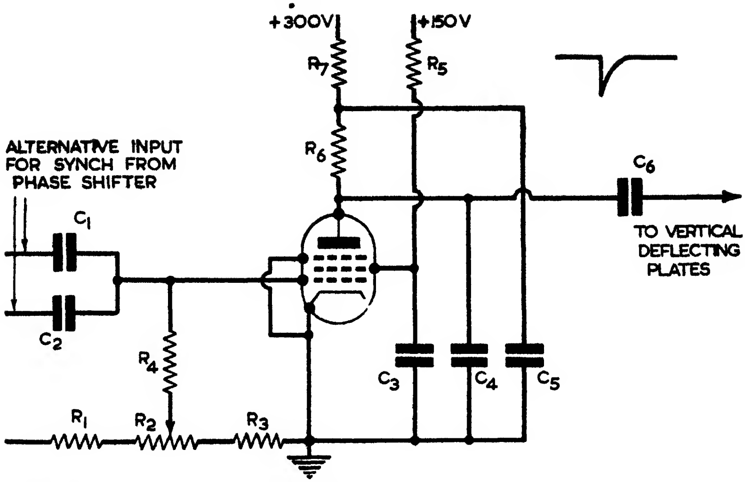


Fig. 306.—CIRCUIT FOR GENERATING MOVABLE "STEP" RANGE MARKER TO BE OPERATED BY PHASE SHIFTER.

The output is applied to the vertical plates to produce a "step" on the trace which can be moved along the trace to act as a strobe or range marker.

screen or by a marker, such, for example, as a hair-line across the screen, or again by generating a fixed step or notch in the trace. Similarly, in case (*b*), the form taken by the strobe or movable range marker may be a bright spot, a dark spot, a step, produced on the trace or a notch. The principles are identical and only the form which the strobe takes may differ.

In radar systems where the synchronising pulse is not a sine wave, phase-shifting devices may still be employed for range measurement (Fig. 307). The Synch. pulse may be used to control a multivibrator which produces a fairly square output wave. During the negative portion of this wave, a timing wave is generated by a shock-excited oscillator. The sine wave thus generated is applied to a phase-shifting device. The shifted output is amplified and differentiated to form a train of positive and negative pulses, which are then superimposed upon a sawtooth wave. The resulting waveform is applied to the input of a pulse-selector circuit, Fig. 307. The sawtooth output of V_1 is synchronised with the time base for start, duration and rate of rise, but V_2 is fired by some pip on the sawtooth and its output is, therefore, delayed by a corresponding length of time. The pip that fires V_2 is determined by the potential of the cathode of V_2 relative to its grid and this, in turn, is

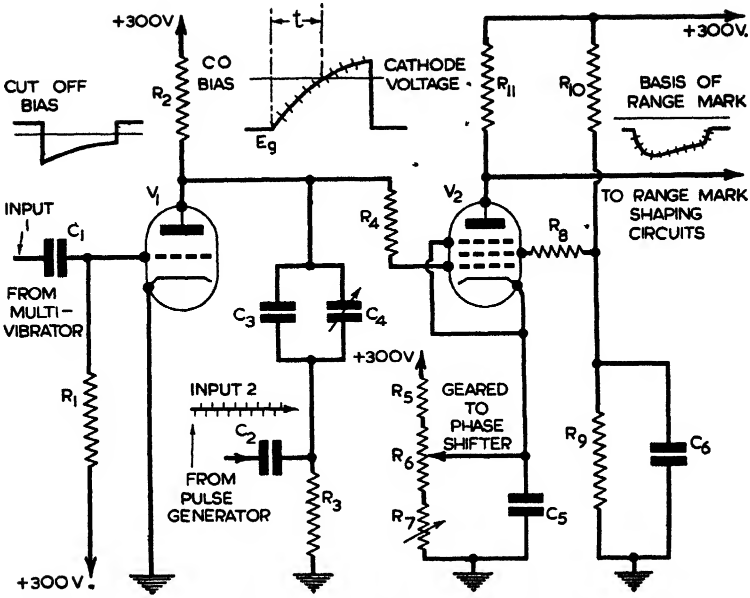


Fig. 307—SCHEMATIC CIRCUIT DIAGRAM OF PULSE SELECTOR CIRCUIT.
Input 1 is the output from the multivibrator Input 2 is the differentiated and phase-shifted output from the shock-excited oscillator or generator.

controlled by a potentiometer geared to the spindle of the phase shifter. Hence V_2 can be adjusted to fire at the leading edge of any echo along the trace by manipulation of the phase shifter. In consequence, the leading edge of the strobe, developed by the range mark shaping circuits, can be made to correspond with the leading edge of the echo on which the operator ranges. When echo and strobe are aligned, the phase shifter scale indicates the range in yards.

Range Measured by Instantaneous Time Base Voltage

Accurate measurement of range can be made by means of a movable range marker developed from the time base voltage and which enables the time base voltage to be compared at any instant, for each successive sweep, with the voltage output from a special potentiometer (supplied from a separate D.C. source), accurately calibrated in yards.

In this case the controlling device which enables the range marker to be applied to the leading edge of the echo, simultaneously causes the time base voltage to be compared with the calibrated potentiometer output and in this way the range of the echo is obtained.

The strobe or marker frequently takes the form of a bright spot which can be manipulated to move along the trace by a manually operated delay generator to the leading edge of any particular echo.

A handwheel operating the delay circuits of the strobe generator is turned to move the strobe or range marker out to the leading edge of the echo selected for measurement, and this operation causes two effects to take place simultaneously :—

(*a*) A contact arm is moved over an accurately calibrated potentiometer (range marker comparator).

(*b*) An "M"-type transmitter geared to the shaft of the marker control potentiometer transmits to remote indicators (if any are in use) a signal which places a strobe against the echo selected at the operating position for particular attention.* A scale is coupled directly to the arm of the moving shaft of the potentiometer from which the range of the echo can be read for any setting of the strobe.

In Fig. 308 an input taken from the sweep generator is applied to the grid of V_1, and the grid potential of V_2 is determined by the position of the slider of the potentiometer. V_2 conducts only when the anode voltage on V_1 passes (in the negative direction) the voltage on the grid of V_2 at some point during each time base sweep. This point is determined by the position of the potentiometer arm when the range marker or strobe is aligned with an echo.

The anode of V_2 contains an inductance so that when V_2 conducts, ringing is produced by the action of the inductance in conjunction with stray capacity. This results in the production of a train of damped oscillations at the anode of V_2.

When V_2 conducts, the anode volts fall and the resulting pulse is negative going on the first half cycle. The grid leak of V_3 is effectively in parallel with the anode inductance of V_2 and causes heavy damping. Accordingly, at the anode of V_2 there appears a single half cycle of oscillation or, in other words, a negative pulse of duration equal to half the time period of one oscillation for the ringing circuit. Thus the position in time of this negative-going pulse is determined (for a particular setting of the range switch) by the position of

* This focuses attention to a particular echo on all indicators, remote or otherwise.

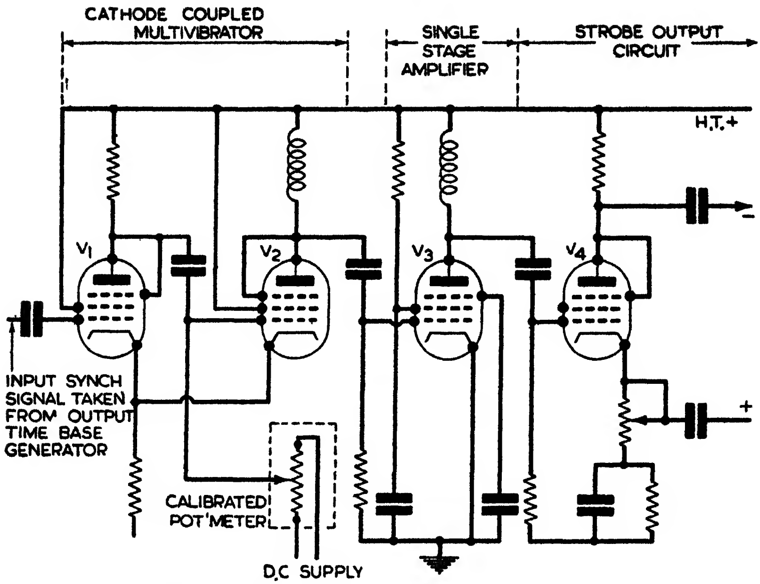


Fig 308.—BRIGHT SPOT RANGE MARK AND STROBE GENERATOR.

The input to V_1 may be taken *viâ* a buffer stage consisting of a cathode follower, from a Miller time base generating circuit in order to ensure optimum linearity The buffer stage avoids overloading and distortion of the time base generator. The suppressors of V_1 and V_2 are strapped to their anodes and the valves function as tetrodes.

the slider on the potentiometer. The delay imposed on V_2 is proportional to the sweep voltage at the leading edge of the echo.

The anode output of V_2 is a short negative pulse, the leading edge of which is made to occur at a given time after the commencement of the Synch. pulse. This delay is regulated by turning the handwheel attached to the spindle of the potentiometer.

The output from V_2 is amplified by V_3, which is loaded inductively. The result of this is to preserve the H.F. harmonics and so to make the Synch. pulse produced as short and as steep as permissible. The output from the anode of V_3 is a positive pulse followed by a small negative-going kink.

This output is fed to V_4, which has two outputs in anti-phase. Since V_4 is biased back to cut off, only a positive pulse appears at the anode, and this is applied to the C.R.T. to produce a bright spot which acts as a strobe or range marker.

A circuit which uses the time base voltage to provide a time

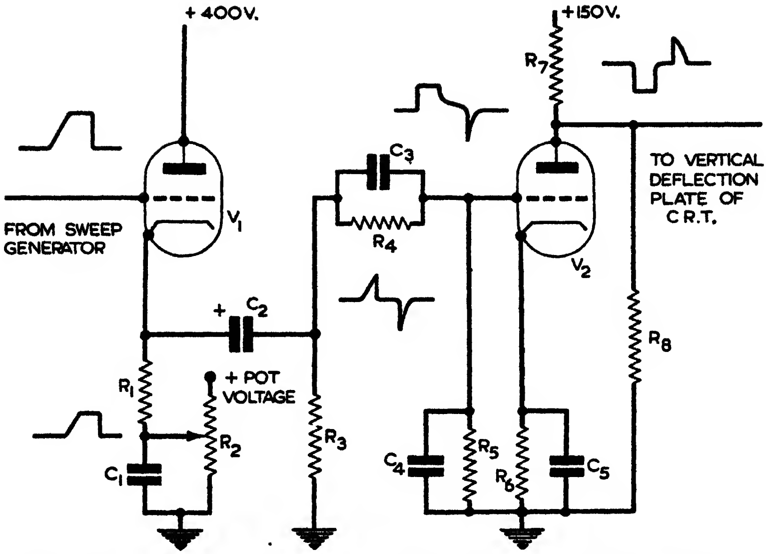


Fig. 309.—CIRCUIT GENERATING MOVABLE "STEP" RANGE MARKER FROM SWEEP VOLTAGE.

This circuit shows an example of repeated differentiation. It will be seen that the output from V_1 is differentiated twice in order to derive the waveform at E_gV_2.

Note.—The parallel circuit C_3R_4 can be replaced by an equivalent series circuit where $R_4 = \frac{1}{\omega^2 C^2 R^2_4}$.

delay pulse to the generator of a "notch" or "step" range marker or strobe is shown in Fig. 309.

The time base voltage is coupled directly from the time base generator to the grid of V_1, which is operated as a biased cathode follower. The bias is adjusted to cause V_1 to commence conducting at any instant from the beginning to the end of each sweep of the time base generator.

The voltage at the cathode of V_1 remains constant at the level selected by delay control R_2 (comparative potentiometer), until the instant that the time base voltage drives the grid more positive than cut off. The cathode voltage then begins to rise at the same rate as the time base voltage. The cathode voltage is differentiated by C_2 and R_3.

The linear rise of voltage causes a sudden rise of voltage in the differentiated output and the waveform at E_gV_2 is shown. This is amplified by V_2, and the positively clipped result is applied to the vertical deflecting plates of the C.R.T. to serve as a range marker in the form of a "step." Range is measured by varying R_2 until the step coincides with the leading edge of the echo, the range of which is required. The shaft of R_2 is geared to the counting mechanism and the range is shown in yards or other suitable units.

Other Indicator Modifications

The following are some of the other modifications that may be incorporated in indicators to meet special requirements :—

Generally a voltage is applied to the cathode of the C.R.T. to intensify the beam for the duration of the forward sweep of the time base. In some cases, however, provision may be made to limit the intensification of the screen to some selected portion of the trace. This means that only part of the trace is illuminated during each successive sweep. In order to accomplish this, the intensifying wave of voltage applied to the cathode or grid is made narrower (in time) than the full period of duration of the time base sweep. It may be delayed and/or suppressed by a pulse actuated by the output from a strobe generator so that any section of the trace may be selected for intensification and unwanted echoes are therefore not illuminated.

In some indicators it is possible to expand a portion of the time base in the vicinity of the range marker (fixed or variable), in order to facilitate accurate ranging.

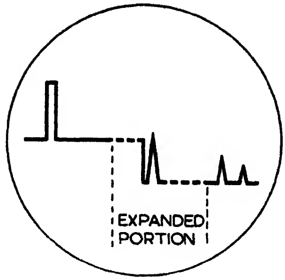
This is accomplished by increasing the speed of the sweep of the electron beam across the screen for a selected portion of each sweep of the time base. In other words, the velocity of the time base sweep is speeded up over a selected portion of each sweep. The section to be operated upon in this manner is selected and marked by manually operated range mark generators and strobes.

The appearance of the screen under these conditions is as shown in Fig. 310.

This result is achieved by causing the pulse which produces the range marker or strobe to fire simultaneously a sweep expansion generator. The duration of the expanded portion of the sweep depends upon the width of the control pulse. The pulse to the range mark generator is however generally required to be as short as possible—much shorter than the pulse controlling the duration of the expanded portion of the time base, *i.e.*, the expansion generator. To meet these

Fig. 310.—SWEEP EXPANDED IN VICINITY OF RANGE MARKER.

The portion of the trace to be expanded can be controlled by a pulse from the strobe generator to start the sweep expansion generator. The strobe employed in this instance takes the form of a "step."



conflicting requirements the pulse to the range mark generator may be formed into a sharp peak and superimposed or mounted on the broader pulse controlling the time base expansion generator. The sharp-pointed peak of this composite pulse provides a short sharp pulse to the range mark generator, and the broader part of the pulse; forming the base of the pedestal, controls the time of operation of the time base expansion generator—one composite pulse firing two different generators and controlling one only.

The expansion generator may be arranged to operate from the pedestal pulse in such a manner that the resulting discharge of a condenser causes the rise of the normal sweep voltage to be accelerated by its discharge. This action is equivalent to decreasing the apparent capacity of the controlling condenser in the time base circuit for the duration of the control pulse.

Another form of expanded time base sweep is when any portion of the whole range occupies the full width of the screen.

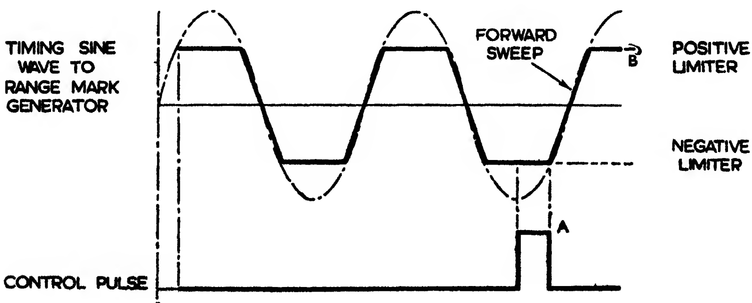


Fig. 311.—PRODUCTION OF EXPANDED SWEEP FROM SINE WAVE GENERATED BY A RINGING OSCILLATOR. A. CONTROL PULSE FROM THE TRANSMITTER. B. WAVEFORM FROM RINGING CIRCUIT AFTER LIMITING OPERATION STARTS FORWARD SWEEP AT END OF TRANSMITTER PULSE.

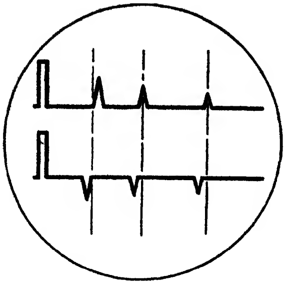


Fig. 312.—INDICATOR SCREEN WITH DOUBLE TRACE.

The signals on the lower trace in this figure have no specific meaning; they only illustrate the principle. A display of this type may be used to correlate identification signals with the echoes to which they refer.

In this case, a sine wave generated by a ringing circuit is greatly amplified, and limited for both amplitudes by a double diode in order to produce the waveform shown in Fig. 311. The portion of the sine wave between the upper and lower limits is nearly linear and is used as the accelerated time base sweep.

The sine wave generating circuit and amplifier are controlled by a phase shifter, so that the echo to be examined marked by a step always appears at the centre of the expanded sweep and occurs, in time, during the rising portion of the limited sine wave.

The acceleration is such that under control of the phase shifter any 2,000-yd. section of the total range may be examined in detail, and the range of the target itself can therefore be determined with great accuracy.

Multiple Trace Displays

For certain applications, indicators may be arranged to display two or more traces simultaneously as in Fig. 312. This form of display is used when it is desired to show recognition signals or to display supplementary data on a trace separate from the main trace.

The effect is obtained by a separation pulse which is applied to the Y or *vertical plates* to shift the second trace above or below the main trace periodically, the main trace being suitably positioned initially.

The separation pulse may be applied to shift the main trace every other sweep of the time base or at some sub-multiple of the repetition rate according to requirements. As long as this shift takes place at a frequency above the persistence of vision, the two traces appear on the screen together. In this manner, data supplementary to individual echoes appearing on the main trace can be displayed in correct relation on the second trace. This avoids congestion of the main trace and facilitates observation.

By employment of various input combinations, in conjunction with intensity modulation, one variable can be plotted against another as already explained in the first part of this chapter. For example :—

(*a*) Elevation plotted against range.
(*b*) Elevation and range plotted against bearing.
(*c*) Elevation error plotted against bearing error.
(*d*) Range plotted against bearing and elevation, etc.

An indicator which plots signal strength against bearing is a type used when beam switching is employed.

A "B" display plots range against bearing in azimuth. Usually a cathode ray tube employing magnetic deflection and with a long persistence screen is used for this display. The tube is intensity modulated by the signal to cause the position of the target to be indicated by a bright spot on the screen.

The trace covers a rectangular area on the screen with the range shown vertically and azimuth shown horizontally. Clamping may be used on the signal voltage so that the instantaneous potential at the grid of the cathode ray tube depends only upon the signal amplitude and not on the waveform.

Bearing Potentiometer

When a display of the "B" type is to be produced by plotting, say, range along the vertical or Y axis against bearing plotted along the horizontal or X axis, a normal time base similar to that provided for an "A" type display is applied to the Y plates, and a deflecting voltage proportional to the bearing of the aerial relative to the ship's head is applied to the X plates. Since scanning in azimuth is performed by the aerial with comparative slowness, the deflecting voltage proportional to aerial bearings can be supplied by an output taken

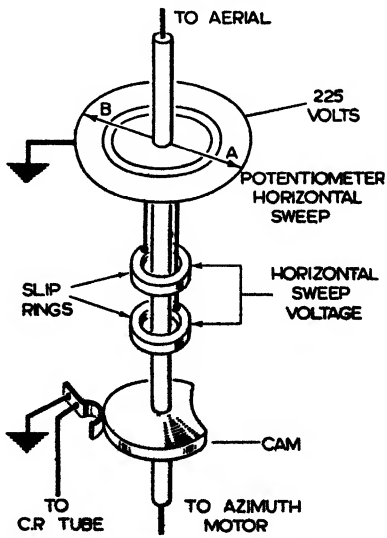


Fig. 313.—ARRANGEMENT OF POTENTIOMETER TO GIVE VOLTAGES PROPORTIONAL TO BEARING OF THE AERIAL RELATIVE TO SHIP'S HEAD.

from the rotating arm of a potentiometer. Maximum deflection to the right centre occurs when A, see Fig. 313, is at the point of application of the positive voltage and B is at earth potential. When the conditions of A and B are reversed, the direction of deflection of the beam is also reversed. A cam-operated switch is provided to prevent signals from the rear appearing on the indicator screen.

Alternative Method for Magnetic Deflection (Fig. 314)

The sweep for the range time base may be a sawtooth wave of voltage coupled to the grids of V_2 and V_3 in parallel. V_1 clamps the starting voltage for each sweep of the time base. Both valves are biased near to cut off so that the anode currents are small for zero signal.

Anode current flows *viâ* L_1 to 450 volts positive (electron currents), simultaneously a current flows from the 450-volt supply through the vertical deflecting coil to the 300 volts positive. The magnetic field of this last current deflects the beam towards the bottom of the indicator screen. R_7 and R_8 are adjusted to cause the time base current to start from the reference line zero.

When the positive sawtooth voltage is applied to the grids of V_2 and V_3, their anode currents tend to rise rapidly. The inductance L_1 is sufficiently large to prevent the 450-volt source from supplying the additional electron current. This additional current must therefore be supplied by the 300-volt source, as a result of which the current through the deflecting coil in the static state must be reversed, *i.e.*, it must fall towards zero and build up again in the opposite direction. Accordingly a current that rises from a negative value to some equal positive value, in proportion to the applied sawtooth voltage, is produced in the vertical deflecting coil.

Display showing Elevation Plotted against Azimuth

Another type of display presents azimuth along the horizontal abscissa and elevation angle along the vertical or ordinate. The beam is intensity-modulated by the signal. This display is used in fighter interception work. The rate of screen area coverage may be twenty lines in azimuth for one sweep in elevation. This is essential in order to get good coverage of the area which is being scanned. The potentiometer from which proportional voltage output is obtained for the azimuth sweep is essentially the same as for the " B " display (Fig. 313).

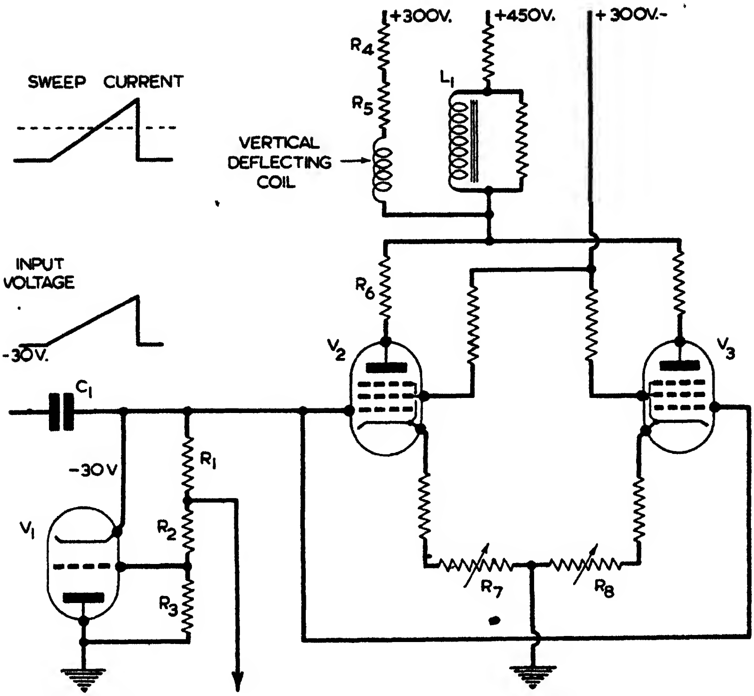


Fig. 314.—RANGE SWEEP GENERATOR FOR MAGNETIC DEFLECTION SUITABLE FOR APPLICATION TO "B" TYPE DISPLAYS.

Examples such as this must, in general, be regarded as indications as to the manner in which certain functions can be performed. They are not always circuits in actual use.

A potentiometer may also be used to provide voltage proportional to the angle of elevation for the elevation deflections. An identical electrical and mechanical arrangement can be used for elevation as for azimuth.

Angular coverage may be 70° for elevation as against 180° for azimuth; consequently, a gear drive can be used that drives the aerial through 2·5° of elevation for each completed rotation in azimuth.

Because the motion of the beam in elevation is slow, the area scanned relative to the pulse-repetition frequency is small; therefore all signals or noise that reach the screen during each period of the p.r.f. tend to pile up at some point, and unless the signal-to-noise ratio of the receiver is very good, the result of this piling up is such that the signal under observation cannot be distinguished.

Most of the noise due to unwanted signals can be eliminated by using range selection when the range has been determined.

This result is obtained by the production of a square pulse with a duration that is a small part of one repetition period and a starting point that is controlled by some range-measuring device as, for example, a phase shifter. The range-measuring device is then adjusted so that the time base sweep starts at some intermediate point between zero range and the particular echo in which the pilot is interested. The falling edge of the short pulse which starts the time base of the pilot's display also determines the moment of cut off, in consequence of which there appears on the pilot's display, the echo in which he is interested and a small portion of the trace of the main display containing this particular echo.

Thus the effect of noises which would be produced before the commencement of the trace (as controlled by the range-measuring device) and after the end of the selected trace (as determined by the falling edge of the time-delay pulse) is reduced.

The P.P.I. Display

In the case of the P.P.I. display, the time base for the sweep for the range time base starts at the centre of the screen and moves outwards along a radius, in a direction corresponding to the direction in azimuth of the radiated beam. Generally a tube of the magnetic deflection type having a long afterglow is used. The sweep for the range time base is made linear, so that range is measured by the distance from the centre of the screen of the bright spot produced by an echo.

The deflecting coils, mounted on a yoke, are made to rotate round the neck of the tube and, by so doing, to rotate the deflecting field and the radial trace. Since the coils are rotated in step with the aerial, the radial trace follows the aerial. The grid is intensity-modulated by the received signal, wherefore a polar map of the surrounding territory is produced on the screen if the radiated beam is directed by the aerial in the plane of the horizon.*

* The output from a range mark generator or strobe generator which normally appears as "pips" or bright spots on the trace of an "A" or "B" display, when applied to the rotating time base of a P.P.I. appear as bright concentric circles. Range is measured along the radius of any given circle, and the mechanism for strobing follows along the same general lines as that already described for "A" and "B" displays. Any desired sector of a P.P.I. display can also be selected for detailed examination.

Most P.P.I. indicators use a pair of coils which may be rotated round the neck of the C.R.T. to produce the rotating time base. If, however, it is desired to avoid the mechanical difficulties involved (especially for speeds above 10 r.p.m.), a rotating field may be obtained, using fixed coils as shown in Fig. 315.

The Selsyn Transformer *

The rotor of a special Selsyn transformer is arranged to be turned mechanically by a shaft as the aerial is rotated in azimuth. Two stator secondary windings are placed at right angles, consequently any position of the rotor (as determined by the aerial) that causes voltage in stator No. 1 to be maximum is 90° away from the position that induces maximum voltage in stator No. 2. Although the alternations of the voltages induced in the two stators are always in phase, variations in amplitude are 90° out of phase, because of the 90° separation of the stator windings.

Since the output of stator 1 controls horizontal deflection and the output of stator 2 controls vertical deflection, a circular motion of the radial trace on the screen of a cathode ray tube is produced which must be in step with the rotating aerial.

The radial displacement of the beam at any instant is controlled by the waveform of the signal applied to the rotor of the Selsyn transformer. To produce a linear rise of current in the deflecting coil, a trapezoidal voltage must be applied to the Selsyn rotor. For a short-range time base the trapezoid is distorted by the addition of an extra peak at its leading edge to prevent fast charging of the distributed capacity associated with the deflecting coil, and so that the time base may start with proper velocity at the time of the transmitted pulse. A sawtooth voltage can be used for long-range time bases, because non-linearity near zero range is of no consequence.†

Conventional Selsyns may be employed in place of the special Selsyn transformer just described. The problems involved are much the same.

* The Selsyn or Magslip is fully dealt with in the chapter on aerial servo systems. The above application is not, however, a Servo function.

† An alternative method of compensating for the distortion introduced by capacity at the commencement of the sweep is to make the sweep the timing control for the whole system. In this case the first part of the sweep is not used, wherefore all other functions of the radar system timed by the time base voltage must be delayed by a time which is the difference between the starting instant of the time base and the commencement of the usable part of it.

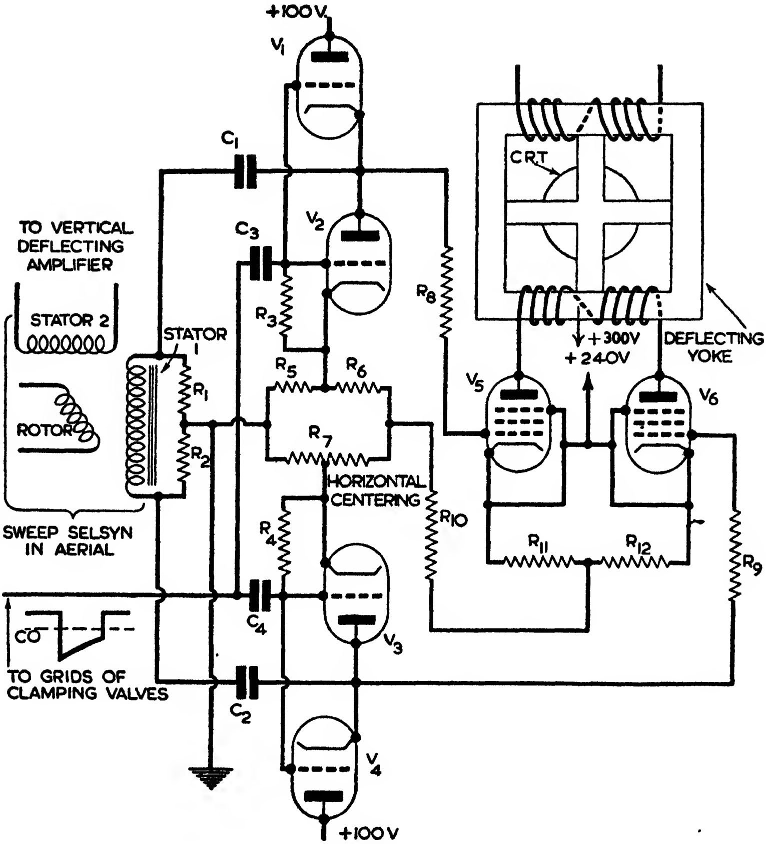


Fig. 315.—SELSYN TRANSFORMER CIRCUIT FOR PRODUCING ROTATING SWEEP FOR P.P.I. AS AN ALTERNATIVE TO ROTATING THE DEFLECTING COIL ON THE NECK OF THE C.R.T.

Note.—A centre tapping +300 V. is taken from the lower coil on the deflecting yoke.

Since the sweep is generated at the p.r.f.—say 800 times per second, and the aerial is rotated at a maximum speed of, say, 1 revolution per second, the outputs of the coils of the stator transformer are two waves of trapezoidal voltage, of which the amplitudes are modulated by two sine waves 90° out of phase.

The output from stator No. 1 controls the current in the horizontal deflecting coil and the output from stator No. 2

controls the current in the vertical deflecting coils. The resultant of the two magnetic fields produces the angular deflection. Thus a change in the angular position of the aerial produces an equal change in the resultant direction of the magnetic field, and the angular displacement of the radial time bases.

The vertical component of the magnetic field is the output of a circuit identical to that shown in Fig. 315 for the horizontal deflection component.

The output of stator 1 is applied to the grids of V_5 and V_6 in push-pull. The windings on the yoke are such that, if V_5 and V_6 conduct equally, their fields cancel. As the anode current of one valve rises for a signal on the grid, that of the other falls, and the result is a nett magnetic field in one or the other direction.

V_1 and V_2 clamp the grid of V_5 at a definite potential for the time constant between each sweep of the time base; V_3 and V_4 clamp the grid of V_6.

The grid of V_2 is connected to the cathode *viâ* the grid leak, so that during the time when no time base is applied, V_2 has zero bias, V_1 is biased by the drop across V_2.

If, at the end of a sweep time, the grid of V_5 returns to a potential higher than the normal value established for V_1, the bias on V_2 is increased, and V_1 then becomes a higher resistance, but the bias on V_2 remains at zero, so that the resistance of V_2 is nearly constant, making the potential of the anode and grid of V_5 return to the normal value.

Thus any change in the potential of V_5 is counteracted by a change in the relative resistances of V_1 and V_2 (which are established initially by resistances R_5 and R_6). V_3 and V_4 limit the potential of V_6 in a similar manner.

During the time base interval the clamping valves are held inoperative by a negative pulse on their grids, in order that the time base voltages can be applied to the grids of V_5 and V_6.

CHAPTER XXI

ANALYSIS OF SOME TYPICAL COMBINATIONS OF RADAR UNITS AS THEY MIGHT BE ARRANGED FOR RADAR SYSTEMS FOR VARIOUS APPLICATIONS

THE combinations of units described are not necessarily taken from any radar systems now in use. They are intended to illustrate the application of general principles upon which radar systems are laid out, and serve to introduce some of the many combinations of units that may be arranged to work together for various requirements.

General Warning Set for Detecting Targets at 100,000 Yards Maximum Range

Example 1. The application is a general warning set capable of detecting surface targets only at 100,000 yds. maximum range. A high degree of accuracy is not required since short-range accurate sets can be brought into service on receipt of an alarm. A power supply of 230 volts, 50 cycles, is available. Accommodation for the equipment is limited, but there are no restrictions regarding size and weight of the aerial. (*Note.* This might very well be part of some shore installation.) The set is to be simple to operate, easily maintained and economical. One " A " display only is required at the operating position, and no remote indicators are necessary. The indicator is to have two ranges, viz. 100,000 yds. and 20,000 yds., controlled by a switch. Provision for aircraft warning is not required.

Let it be assumed that in order to fulfil these conditions a carrier frequency of 100 megacycles is selected. This can very well be the case because (*a*) a high degree of accuracy is not required. (*b*) There are no restrictions regarding the size and weight of the aerial. Thus the advantage of a comparatively long wave can be secured in order to obtain the required range with minimum peak power. (*c*) The carrier frequency selected is well within the upper limits for the efficient operation of triode valves. (*d*) The peak power required might be fixed at, say, 15 kW.

Consequently facts (*c*) and (*d*), coupled with the other condi-

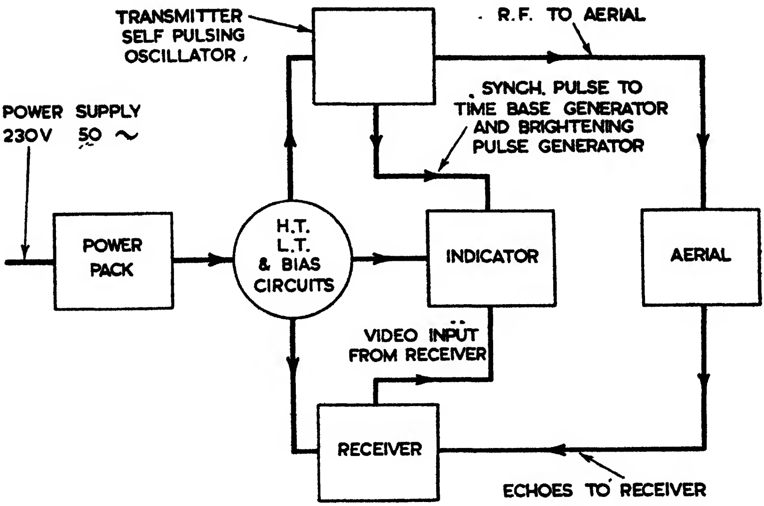


Fig. 316 (a). BLOCK DIAGRAM OF UNITS.

tions laid down, *i.e.*, frequency of the power supply (50 cycles per second),* limited accommodation, simplicity of operation, easy maintenance, overall economy and a wide accuracy tolerance may make it desirable to employ a self-pulsing oscillator.

R.F. Generator

This choice automatically fixes the source from which synchronising signals are to be derived.

The general arrangement of the system will therefore be as shown in Fig. 316 (a) and these units can be split up into sub-units with functions as shown in Fig. 316 (*b*). The pulsing system is simple and the power pack (not shown) is conventional. An alternative arrangement is shown in Fig. 316 (c).

Note. A brightening pulse to brighten the trace for the duration of each sweep is, in Fig. 316 (b), taken from the time base generator output. Differentiation of a sawtooth wave gives a square wave of the same duration. This is applied to the control grid of the C.R.T. as a brightening pulse. The self-pulsing transmitter might be as shown in Chapter XVIII. The power supply unit is not shown here: it is conventional

* Fifty per second would be a very low repetition rate. Consequently the main supply is not used to supply the synchronising signals for this system.

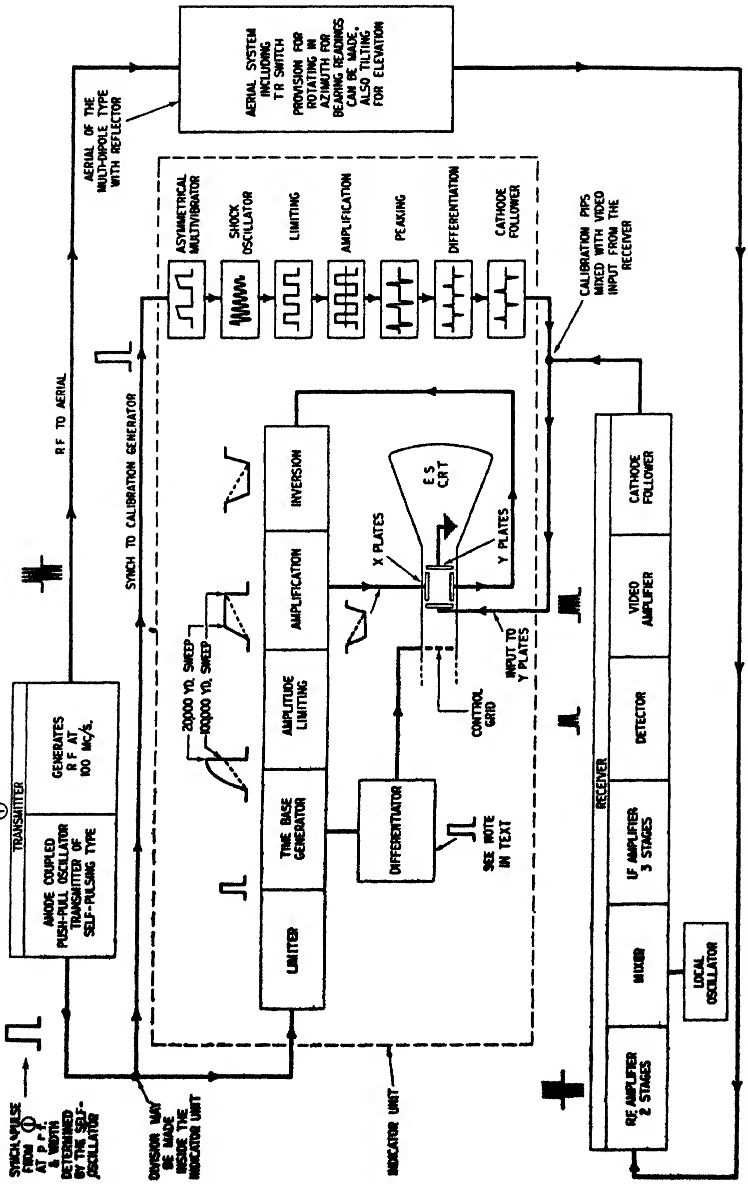


Fig. 316 (b).—MAJOR UNITS AND SUB-UNITS OF SIMPLE GENERAL WARNING SET. RANGE MEASUREMENT BY RANGE MARKS.

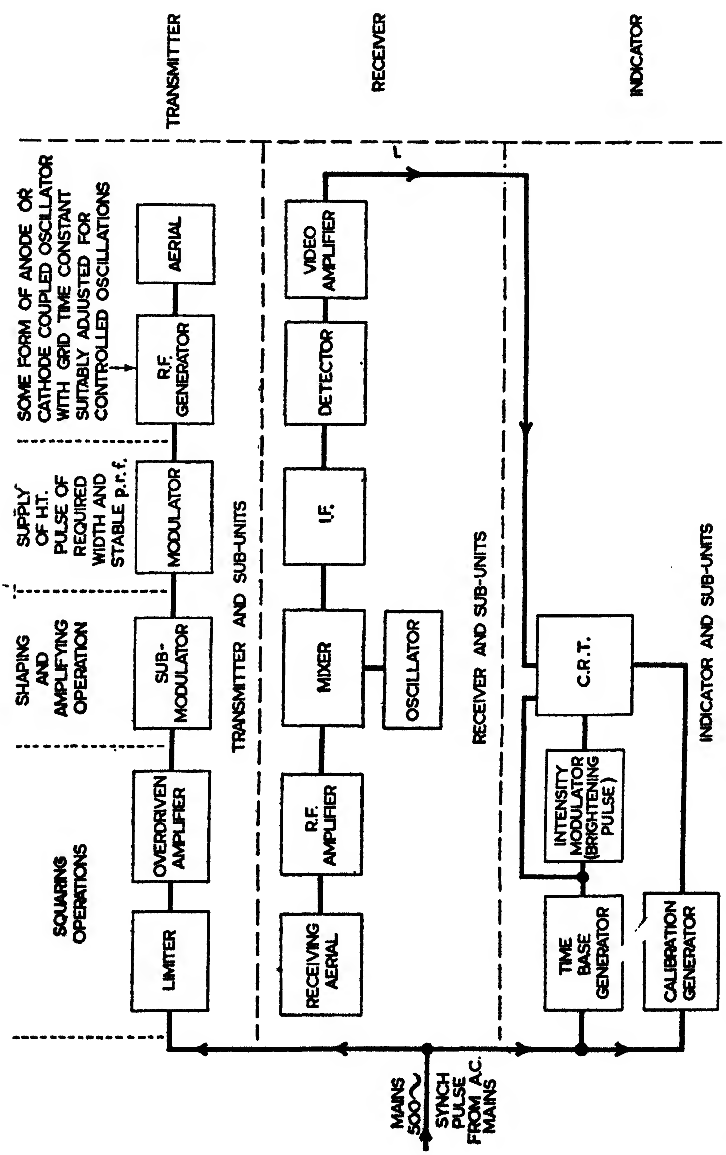


Fig. 316 (c).—POSSIBLE ALTERNATIVE ARRANGEMENT TO FIG. 316 (a) AND (b) WHEN MAINS OF 500 CYCLES ARE AVAILABLE.

and follows normal radio engineering practice. The receiver might be suppressed in this case by a pulse from the transmitter, arranged to provide negative bias to one or more valves in the I.F. stage for the duration of the transmitted pulse.

In Fig. 316 (c) the principal difference lies in the transmitter employed and the synchronising arrangements. The p.r.f. would probably be more stable, also there is good control over the shape of the R.F. envelope.

An Airborne Equipment for Observation of Surface Targets

Example 2. The lay-out shown in block schematic form for Example 2, Fig. 317, might be suitable for an airborne equipment for observation of surface targets. Whilst a high degree of accuracy is required, the controlling factors are space and weight. The range required is of the order of about fifty miles and the equipment is required to detect and discriminate between quite small targets.

In the example an "A" display, a "B" display and a selected range indicator are provided, mainly for the purpose of showing how these units can be worked together. In actual practice, a P.P.I. display would probably be substituted for the "B" and "C" displays for this application. The combination of the "A" and "B" displays and a selected range indicator ("C" display) is more suitable for fighter interception work. The operator searches for and finds the target on the "B" display under R/T control of a powerful ground station. When the operator of the "B" display has located, identified and held the target to which he has been directed, he selects that portion of the "B" display scale on which this particular echo appears and reproduces it on the selected range indicator at the pilot's position to enable him to direct the plane at the target and to bring it into visual range.

The Carrier Frequency

The overriding considerations of size, weight and accuracy influence the selection of a carrier frequency in the micro-wave range, since a high degree of accuracy can thus be obtained with a comparatively small aerial. The peak power required to obtain the required range at this high frequency might be perhaps of the order of 50 kW.

A narrow pulse width of about 2 microseconds is chosen to secure accuracy and also to keep the duty cycle as small as

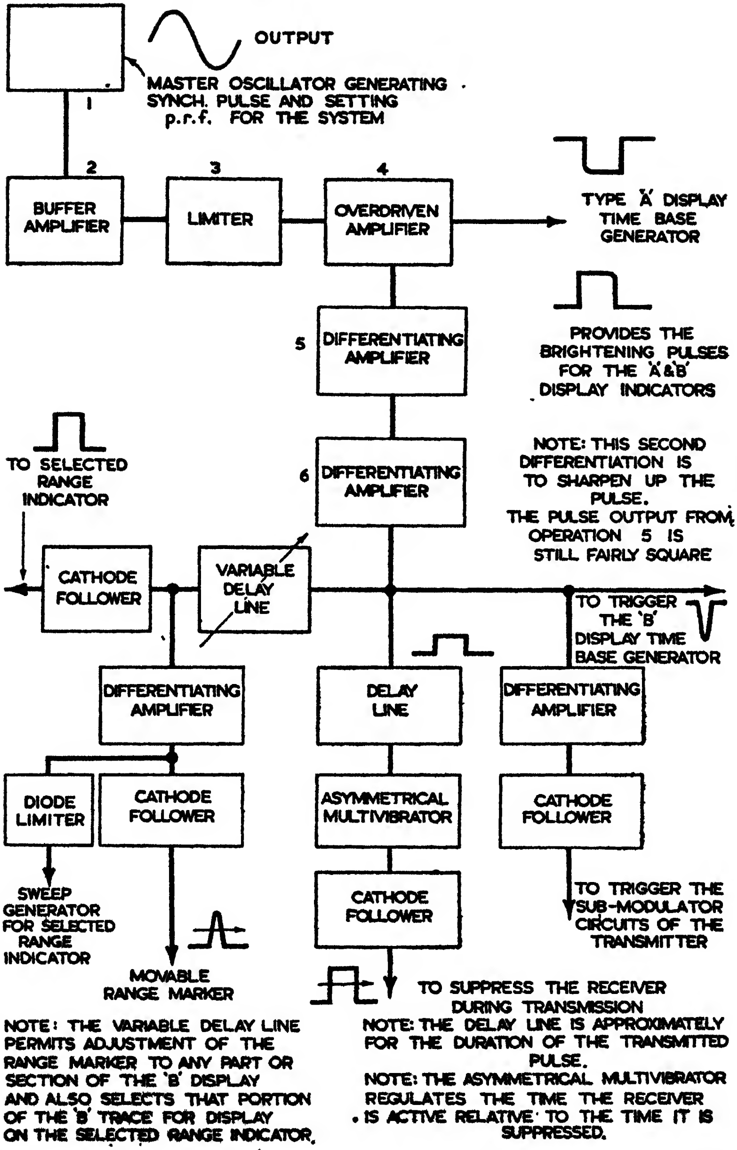


Fig. 317.—SCHEMATIC FOR A SYSTEM EMPLOYING AN INDEPENDENT SYNCH. GENERATOR, AND FOR DISPLAY PURPOSES AN "A," A "B" AND A SELECTED RANGE DISPLAY RESPECTIVELY.

possible. This also enables the required peak power to be developed from a valve of low rating.

Since the pulse width is narrow, and small targets are to be located, definition must be good and, therefore, a fairly high repetition rate will be used, 800 per second, in order that when the aerial is rotated continuously, small targets may be flooded with a reasonable amount of transmitted energy at each revolution and so return a good echo.

Since the carrier frequency selected is in the micro-wave range, the R.F. generator must be a magnetron. Let it be assumed that the repetition rate stability of a spark gap modulator is considered to be inadequate for the degree of accuracy required, and that it is necessary to provide Synch. pulses at a very stable repetition rate of 800 per second. In this case an independent Synch. pulse generator might be used and the pulse for modulating the R.F. generator would then be developed in a succession of shaping and amplifying circuits by sine wave input from the Synch. generator.

The magnetron may be built into the wave-guide which feeds the aerial and is powered by a pulse from the modulator output *viâ* a pulse transformer matched into a coaxial transmission line.

Since the signal-to-noise ratio for the carrier frequency employed is likely to be low, R.F. amplification of the received signals is not worth while; consequently, a crystal enclosed in a resonant cavity would be used as a mixer and the local oscillator might be a reflex klystron.

In these circumstances the mixer, the local oscillator and the first two or three stages of the I.F. amplifier would probably be located as near to the aerial as possible in the receiver branch of the wave-guide feeder system. The first two or three stages of the I.F. amplifier now act as an L.F. pre-amplifier and transmit a signal of reasonable magnitude to the input of the receiver proper, which in the above circumstances would be in the third or fourth stage of the I.F. amplifier proper incorporated in the receiver unit.

Arrangement of the System

A study of Figs. 318 and 319 shows that the following provisions have been made :—

(*a*) A square pulse differentiated to start the operation of the time base generator in the " A " display indicator.

(*b*) A square pulse differentiated to start the brightening pulse generator for the " A " display indicator.

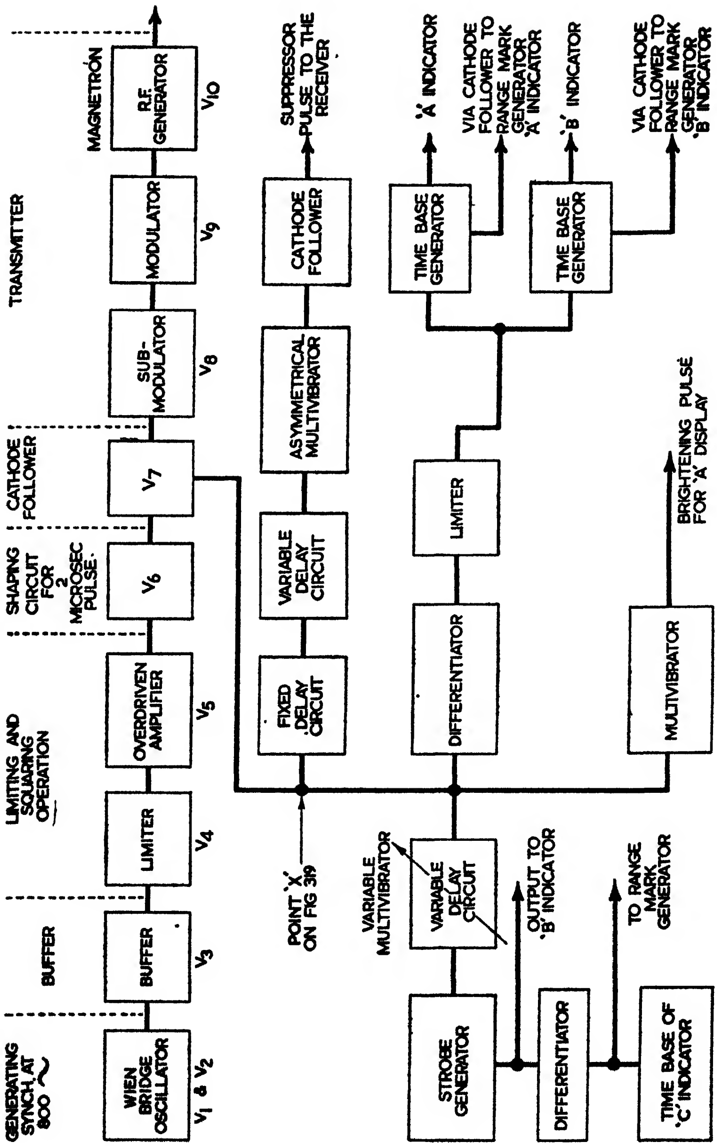


Fig. 318.—An alternative schematic for pulsing system to that shown in fig. 317.

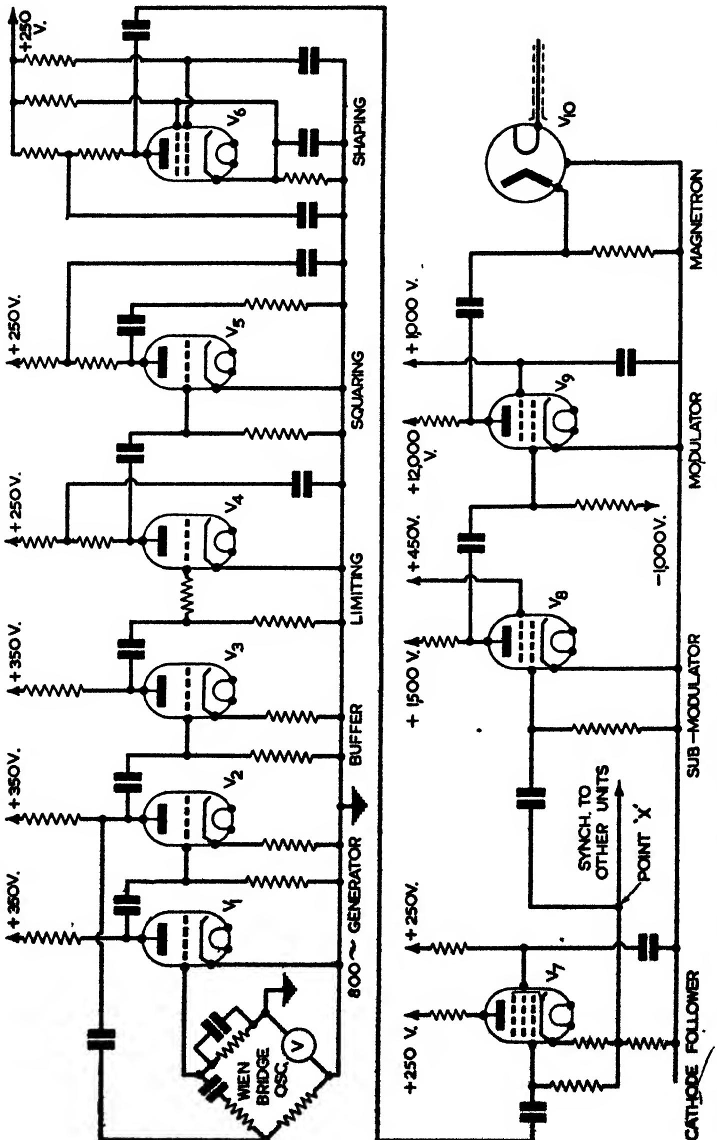


Fig. 319.—Transmitter system synchronised by 800-cycle generator with provision for synchronising other units of the system from the same source.

(*c*) A trigger pulse to start the time base generator for the "B" indicator display (square wave differentiated).

(*d*) A time-control trigger pulse to start the initial stage of the modulating pulse for the R.F. generator, *i.e.*, the sub-modulator.

(*e*) A pulse to render the receiver operative only after the end of each period of transmission, thereby preventing paralysis and enabling the receiver to find itself in a condition to receive echoes immediately after the end of each transmission. This makes for detection of nearby targets.

Note. After the start of the transmitted pulse a positive-going pulse to make the receiver operative is delayed by an interval of time at least equal to the width of the transmitted pulse. The pulse which operates the receiver is formed by an asymmetrical multivibrator which is started up by the falling edge of the pulse from the delay circuit, the width of which is equal (in time) to 2 microseconds.

(*f*) A variable time delay is provided by the variable width (in time) of a square pulse generated by a variable multivibrator or similar device.

This square pulse of variable width, which is the output from para. (*f*), is used to perform the following functions :—

To delay the start of a strobe generator the output of which produces two effects simultaneously, (*a*) a strobe or marker under the control of the operator marks off on the "B" display the portion of the trace on which appears the echo to be followed by the pilot. (*b*) The same control acts as a selector to confine the display on the pilot's indicator within the limits set by the operator on the "B" trace by means of the strobe.

Typical Set for use on Merchant Ships in Peace-time

Example 3. The block diagram for this system and the expanded block diagrams for the sub-units are typical of a radar set suitable for use on ships of the Merchant Navy in time of peace.

The general requirements would appear to be those of a medium-range warning set for the detection of distant objects, combined with good definition and accurate range and bearing measurement of nearby objects.

The performance specification might read : with fair visibility,* moderate sea† and a minimum aerial height of 25 ft.

* Specifies attenuation and refraction tolerances.
† Specifies "clutter" tolerance.

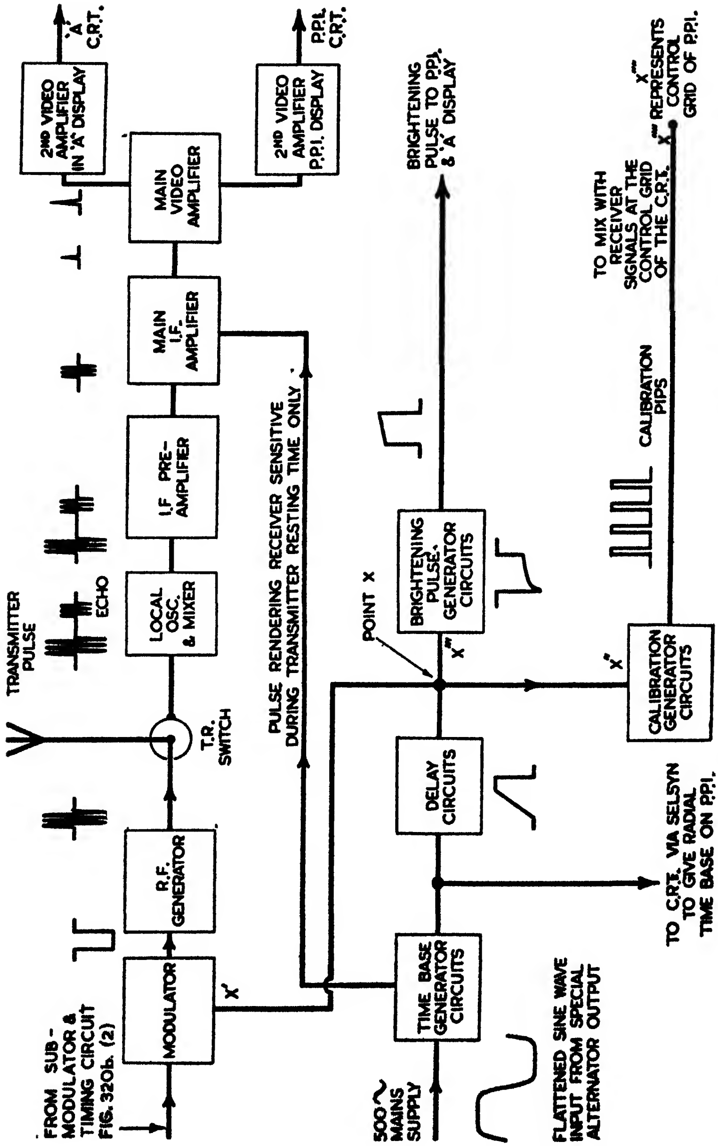


Fig. 320 (a).—Basic layout for typical marine set.

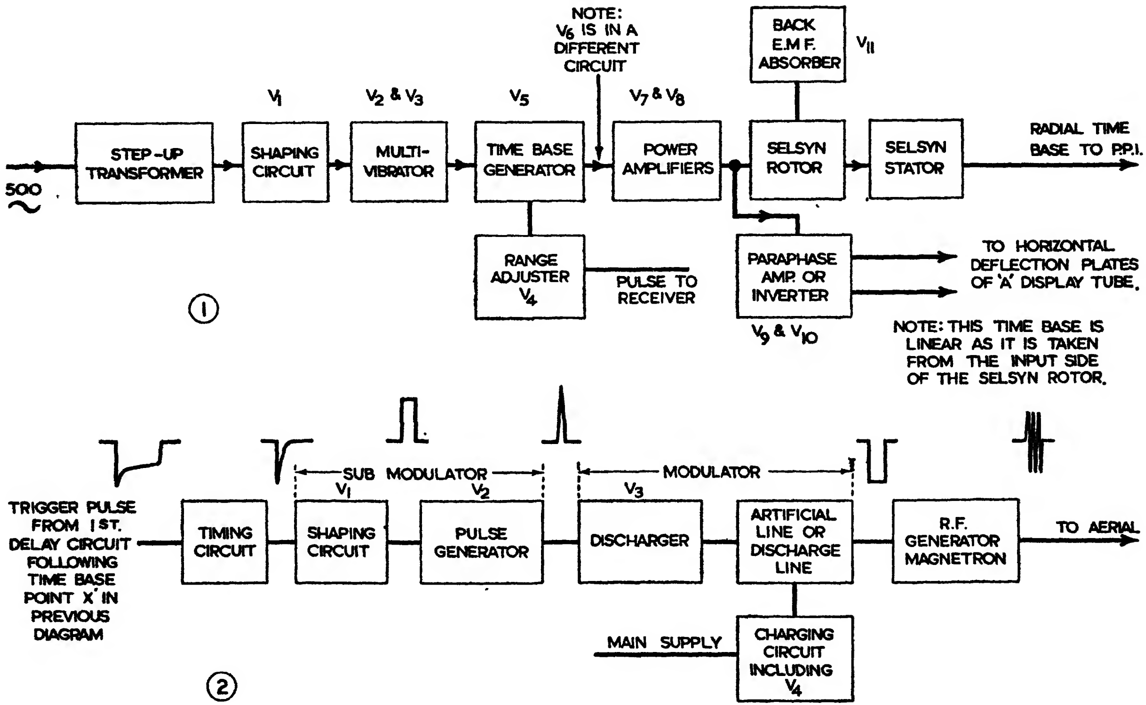


Fig. 320 (b).—EXPANDED BLOCK DIAGRAM OF PART OF FIG. 320 (a) (CONNECT WITH FIG. 323). (1) Block diagram of time base circuits. (2) Block diagram of transmitter.

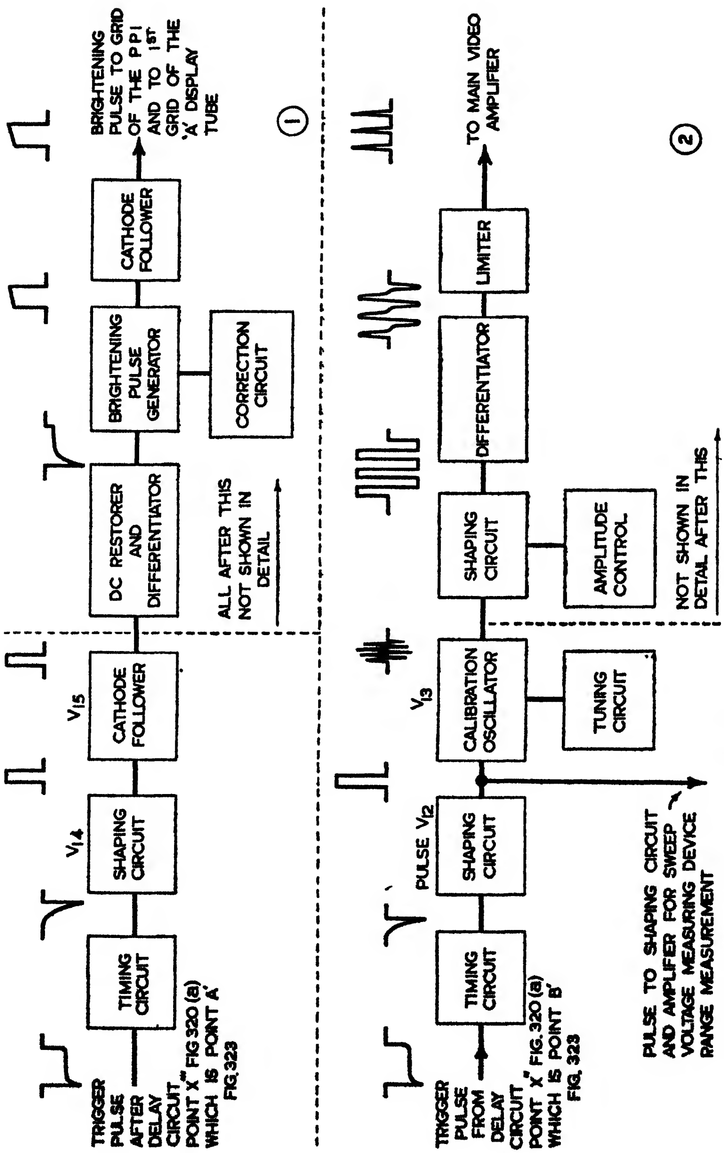


Fig. 320 (c).—EXPANDED BLOCK DIAGRAM OF ANOTHER PART OF FIG. 320 (a) (CONNECT WITH FIG. 323). (1) Block diagram of brightening pulse circuits. (2) Block diagram of calibration circuits.

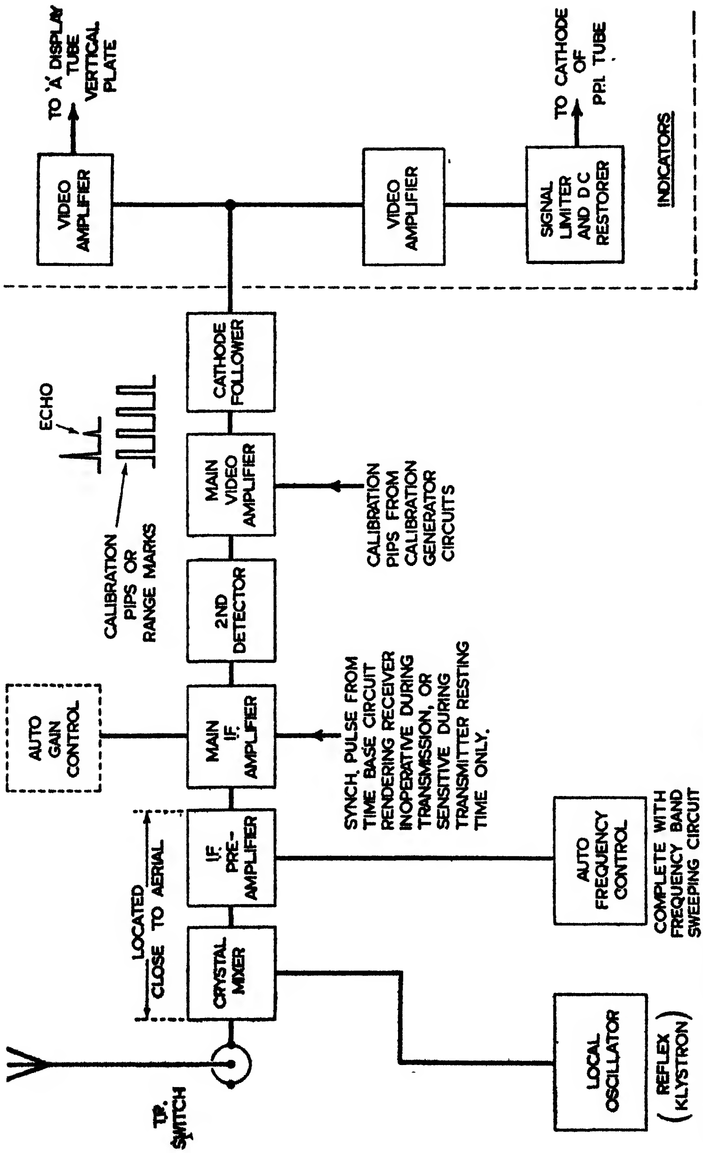


Fig. 320 (d).—EXPANDED BLOCK DIAGRAM OF THE RECEIVER CIRCUIT OF FIG. 320 (a) (CONNECT WITH FIG. 323).
Block diagram of the receiver circuit.

above the waterline, very small craft should be detected at from 6,000 yds. or less, and larger craft at 14,000 yds. and over. Low-flying aircraft may sometimes be detected.

Minimum range for target detection—250 yds.

Definition. At a range of about 1,000 yds. or more it should be possible to distinguish one small target from another if on the same bearing but differing in range by 200 yds.

Bearing Discrimination. Two buoy targets at a range of 5,000 yds. but 200 yds. apart may be distinguished as two separate targets.

Accuracy (range) plus or minus from 3 per cent. to about 16 per cent. for minimum and maximum ranges respectively (interpolation), but approximately 1 per cent. when more accurate methods of range measurement are employed. Bearing accuracy plus or minus two seconds.

The Display

The display is provided by a P.P.I. indicator at the operating position and two " A " display indicators, one of which may be remotely situated.

In this particular case a carrier frequency is employed corresponding to a wavelength in free space of about 3 cm. The R.F. generator is therefore a magnetron.

The pulse width is about $\frac{3}{4}$ microsecond and the peak power is 40 kW.

The repetition rate is 500 per second.

In order to fulfil the conditions laid down in the performance specification, it is essential that the R.F. energy should be radiated in the form of a very narrow beam.

To obtain the narrow beam required and at the same time keep the aerial dimensions within reasonable limits, it is necessary to employ a very high frequency carrier.

The aerial dimensions must be as small as possible, because it has to be mechanically strong, offer minimum wind resistance and at the same time it must be capable of continuous rotation at a speed of about 22 r.p.m.

The pulse width (in time) is made short in order to facilitate as much as possible the reception of echoes from nearby targets. Three-quarters microsecond is proportional to about 123 yds. of range, but this is not the only deciding factor for minimum range. The uneven surface presented by a moderate sea will in itself produce echoes which tend to mask those returned by nearby targets.*

* In radar terminology, " clutter."

The repetition rate is fixed at 500 c.p.s., having due regard to the pulse width, the speed of rotation of the aerial and the peak power required.

To avoid the mechanical difficulties encountered in mounting and rotating at 22 r.p.m. a deflecting coil on the neck of the electromagnetic tube used for the P.P.I. display, the time base current is passed through the rotor of the aerial Selsyn, to obtain a rotating radial time base from the fixed Selsyn stator mounted on the neck of the tube, in place of the normal deflecting coil.

This device introduces some distortion at the commencement of the output from the time base generator, which means that the time base proper must commence at some time after the start of the generator in order to avoid the distorted part. This requirement introduces a somewhat complicated system of time delays.

Time Delay

In this set several time delays are introduced. All are, however, concerned with the synchronising function.

This is brought about by the fact that the start of the time base is not linear. It is distorted due to the effect of the capacity of the Selsyn windings and connecting cables.

In consequence of this distortion the first part of the time base is not used, and for this reason a time delay must be introduced between the start of the time base generator and the start of the brightening pulse generator, the calibration generator and the firing of the transmitter. This time delay must be equal to the interval between the start of the time base generator and the end of the initial distorted portion of the time base.

This, however, is not all; a further delay must be introduced between the start of the time base generator and the start of the brightening pulse and calibration generators, because the trigatron which fires the transmitter is itself fired by the *trailing edge* of the trigger pulse, and this introduces the necessity for a further time delay.

Thus one delay circuit is introduced between the start of the time base generator and the Synch. pulse that fires the brightening pulse generator, the calibration generator and the transmitter, and a further delay is introduced into the circuits of the brightening pulse generator and the calibration gene-

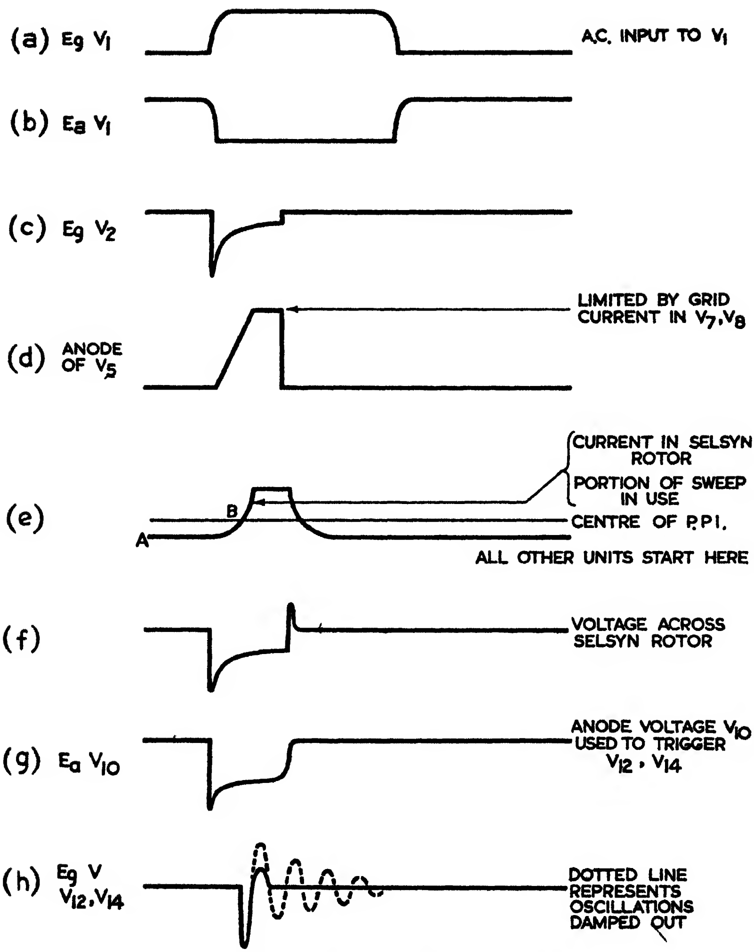


Fig. 321.—IDEALISED WAVEFORMS OF TIME BASE CIRCUITS OF FIG. 323 V_1, V_2, V_5, V_{10}, V_{12} AND V_{14}

rator to synchronise them with the start of the transmitter pulse.

This delay is equal to the time elapsing between receipt of the Synch. pulse at the transmitter and the firing of the trigatron, *i.e.*, approximately the width of the transmitter pulse.

The output from the A.C. 500-cycle generator is not a pure

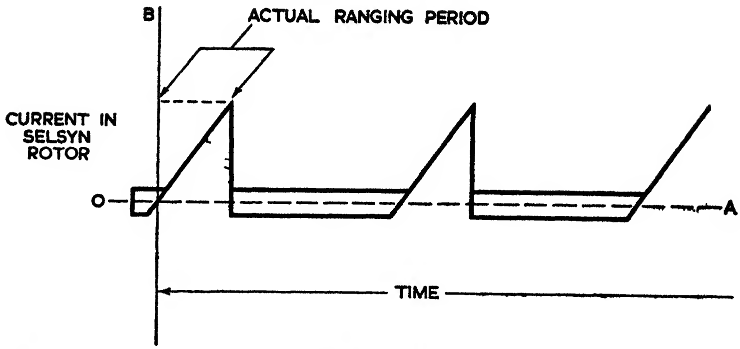


Fig. 322.—IDEALISED WAVEFORM OF TIME BASE CURRENT IN SELSYN.

sine wave, being somewhat flattened, as shown in Fig. 320 (a) * and Fig. 321.

A voltage from this source is applied *viâ* a transformer to the input of the time base generator (Fig. 323 (a)).

The first operation is a squaring operation, which is carried out by a valve amplifier biased at zero (see Fig. 323 (a)). The negative portion of the input drives the valve well below cut off. On the positive half-cycle grid limiting by means of a grid stopper prevents the grid from swinging very positive, due to the flow of grid current. The output from the anode is therefore approximately a square wave (Fig. 320 (a) and 321 (b)).

The next stage is a multivibrator which is triggered every 1/500 second by the A.C. mains. An output is taken from the grid of the right-hand section as a negative wave. This is called the gate wave. The width of this gate is determined by the time constant of the grid circuit and can be varied by the position of the range switch, indicated by a variable resistance and marked R.S. (Figs. 323 (a) and 321 (c)).

The Time Base

The negative-going gate wave is applied in parallel to the grids of V_4 and also V_5 of the time base generator which it drives beyond cut off. The anode current ceases and one of the condensers (shown as a variable condenser in the diagram for simplicity), selected by the range switch charges to the supply H.T. voltage.

* *Note.* Figs. 320 (a) and 323 are complementary. Fig. 321 should be referred to for waveforms of time base circuits and Selsyn circuits. Fig. 322 shows details of D.C. and A.C. in rotor of Selsyn.

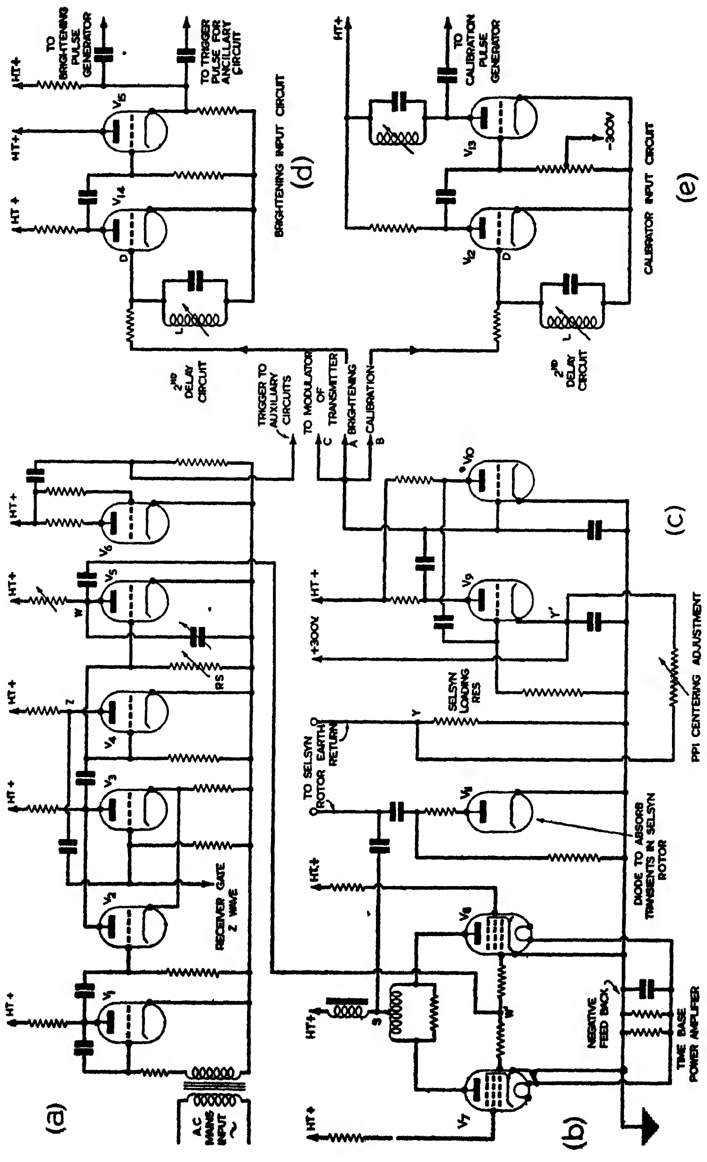


Fig. 323.—CIRCUIT DETAILS OF BLOCK DIAGRAM, FIG 320 (a), (b), (c) AND (d).
(a) Time base generator and receiver gate. (b) Time base amplifier. (c) Selsyn circuit and Synch. pulse distribution. (d) Brightening input circuit. (e) Calibrator input circuit.

This rise in anode voltage is nearly linear and forms the time base (see Fig. 321 (d)).

The output from the time base generator is applied to a power amplifier consisting of two valves in parallel with negative feedback. The output is taken to the rotor of a standard Selsyn generator (see Fig. 323 (b)).

For the operation of the time bases, current need only flow in V_7 and V_8 during the time base intervals; consequently, in order to avoid waste of power and variations of the D.C. component of the anode current for different ranges, the time width of the gate is shortened, so that the average anode current for the valves in the power amplifier is maintained reasonably constant. *Note.* The D.C. component would otherwise change with change of position of the range switch.

The output from the anode of V_4 (point Z), Fig. 323 (a), is used as a receiver gate to make the receiver active only for the duration of the resting time of the transmitter.

The output from the time base power amplifier (point S, 323 (b)), coupled to the rotor of the Selsyn generator, induces in the stator an A.C. voltage, the waveform following that of the voltage input to the rotor. The conditions are as shown in Figs. 321 and 322, in which a radial time base is produced in the stator of the Selsyn on the neck of the cathode ray tube, the angular position of the time base at any instant being determined by the angular displacement of the aerial from its reference position.

In Fig. 322, point A represents the average current, *i.e.*, current at zero deflection. The sweep of the P.P.I. beam begins at one side of the centre of the screen, crossing the centre at point B. The brightening pulse, calibration generators and the transmitter must all be synchronised, therefore, to start at point B (Fig. 321 (e)).

A resistance load is provided to maintain a nearly constant load on the Selsyn rotor regardless of the number of P.P.I.s used. Point Y (Fig. 323 (c)).

The waveform of the negative-going voltage across the Selsyn rotor load is almost exactly like that of the Selsyn rotor current (Fig. 321 (f)). A part of this output is applied to an amplifier the output of which is split, one half going to one of the horizontal deflecting plates of the "A" display, the other half is applied to the other plate *viâ* an inverter circuit (not shown), Fig. 323 (c).

The negative-going voltage from the Selsyn earth return

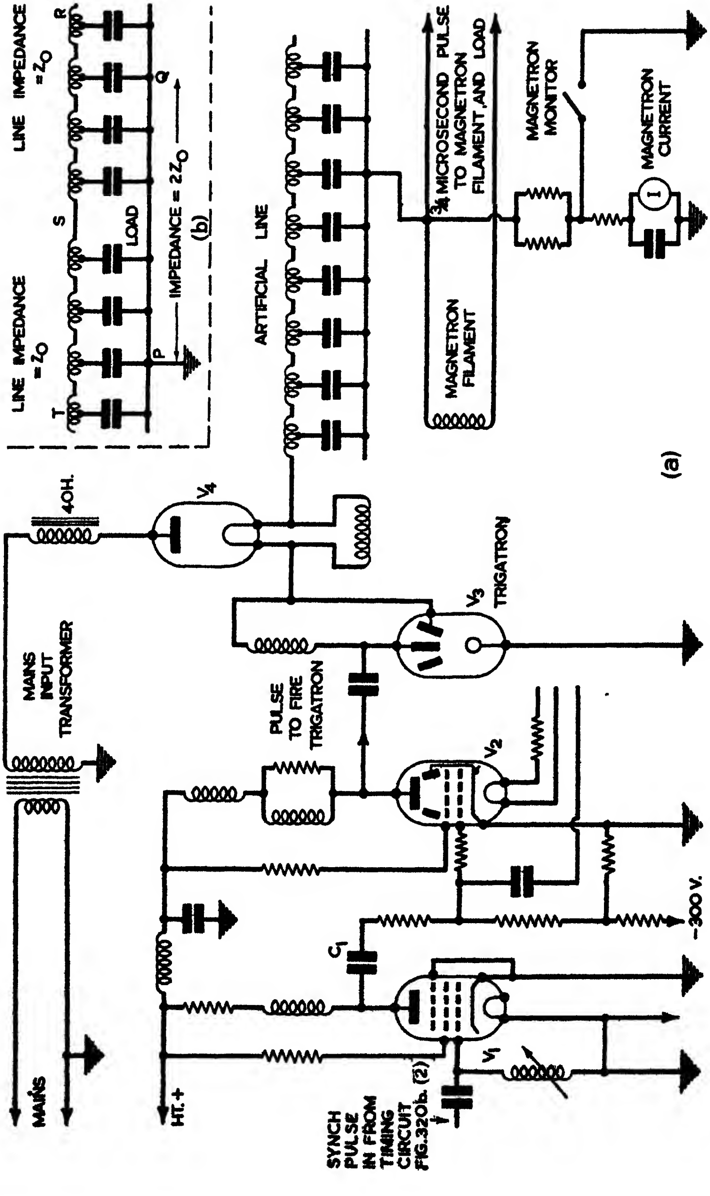


Fig. 324.—TRANSMITTER CIRCUITS OF FIG. 320 (b) (2).
(a) Expanded diagram of transmitter circuit. (b) Explanatory diagram of artificial line (see text, p. 483).

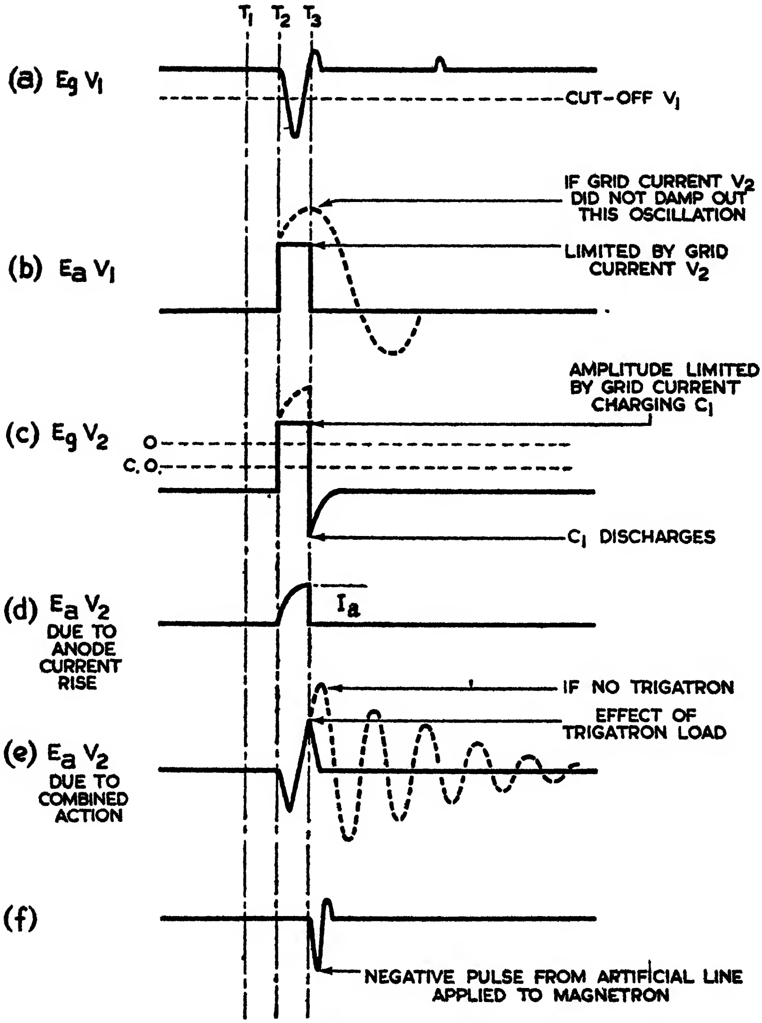


Fig. 325.—IDEALISED WAVEFORMS OF TRANSMITTER CIRCUIT SHOWN IN FIG. 324.

(point Y, Fig. 323 (b)), is applied to the *cathode* of V_9 (left-hand section), (Fig. 323 (c) Y′). This valve is normally biased close to cut off by connecting the cathode to a plus 300 voltage supply. As the negative-going wave is applied the condenser begins to charge until the valve begins to conduct. The anode voltage is approximately a square negative-going wave.

In order to make the slope of the leading edge steeper, and independent of the range used, a multivibrator circuit is introduced by the addition of V_{10}, so that at the instant the valve conducts it is suddenly run to saturation (V_9, V_{10}) Fig. 323 (c).

The output from this circuit is a distorted square wave, Fig. 321 (g), which is used to trigger the brightening pulse generator, the calibration generator and to fire the modulator, points A′ B′ C′, Fig. 323.

Some further delay occurs at the input circuit of the modulator (the voltage pulse that fires the trigatron being produced from the trailing edge of the voltage wave at its anode). Consequently it becomes essential to introduce a similar and additional time delay before producing the brightening pulse and the calibrator output. This is done by a tuned circuit in the first grid circuits of the first stage of the brightening pulse generator and the calibrator circuits, V_{14}, V_{12} respectively (Fig. 323 (d) and (e)).

The negative-going outputs from the delay circuits points D, Fig. 323 (d) and (e), are fed to the tuned input circuits of V_{14} and V_{12}, the frequency being adjusted by the variable inductance L. The time period of the oscillation is from 6 to 8 microseconds, which establishes the time delay for the firing of the trigatron, since it fires on the trailing edge of its firing pulse.

The Modulator

The first negative-going swing due to shock excitation drives the grid of V_1, Fig. 324,* well beyond cut off. The large inductance in the anode circuit, together with its stray capacity, forms an oscillating circuit with a somewhat longer time period than that of the grid. When the grid goes beyond cut off the anode current generates a shock of large amplitude. Just as this oscillation reaches its maximum amplitude the valve again conducts and the oscillation is damped out as shown in Fig. 325 (b). The anode output is approximately a square wave whose duration is determined by the constants in the grid circuit.

The square wave output from the anode of V_1 is applied to the grid of V_2, which is normally biased well beyond cut off. The grid is driven suddenly positive, as shown in Fig. 325 (c).

* Figs. 324 and 325 are complementary. Refer to Fig. 325 for waveforms. These two figures should be consulted simultaneously when reading text.

Grid current flows and partially charges the coupling condenser, thus limiting the amplitude of the positive swing.

At the end of the positive pulse, C_1, Fig. 324 (a), discharges, driving the grid very negative. The grid gradually rises again as the condenser discharges. Due to the inductance in the anode load, anode current rises gradually. When the grid is driven negative the anode current stops instantly and produces a shock oscillation of about 10 kilovolts and a period of about 2 microseconds, as shown in Fig. 325 (a). This voltage builds up to a peak in about ½ microsecond, and therefore produces no noticeable jitter in the trigatron firing. V_3 is a valve containing a triggered spark gap. The potential developed produces a spark discharge at the trigger electrodes. This provides the ions required to produce the main discharge.

When the trigatron fires the oscillation in the anode circuit of V_2 is damped out.

The inductance between the trigger electrodes and the main anode of the trigatron ensures that the two electrodes are at the same potential when the next trigger pulse arrives.

The artificial line which provides the high-voltage high-power negative-going pulse to operate the magnetron is charged from an 8-kW. power supply and discharged by the trigatron. The lines are charged by a high-voltage rectifier (not shown) through a 40 H. choke L_{40} to a voltage of approximately 2E, the power input being 13 to 14 kV. This is achieved by the method described in Chapter XVIII, and is a case of resonant charging.

The firing of the trigatron shorts the line and it takes about 1 microsecond for the line to discharge. During this time a negative pulse, fairly rectangular and about ¾ microsecond duration, appears across the load. The ratio of voltage reflected to that arriving at end of line $= \dfrac{Z - L}{Z + L}$ (see Fig. 324 (b)).

Z_0 = Characteristic impedance of each section.

$$Z \text{ (the load)} = 2 \text{ line } Z \;\therefore\; \text{load} = 2Z.$$

If the line is initially charged to 2E, the potential differences between points TP, S and P, S and Q and R and Q all equal to 2E.

$$\therefore\; E \text{ across the load} = 0.$$

When the trigatron fires, T and P are shorted and a wave of 2E travels down the line. When the wave arrives at SP it finds an impedance of $2Z_0$.

$$\therefore -2E\frac{3Z_0 - Z_0}{3Z_0 + Z_0} = -E.$$

Since S is now E volts lower than P and also E higher than Q, the voltage across the load is 2E. The voltage wave in the left-hand line will be reflected at the short-circuited end as + E and that on the right-hand side will be reflected at the open end as — E.

The left-hand wave reaches SP at the same instant that the right-hand wave reaches SQ and, at this instant, the voltage across the load drops to 0. The duration of the negative pulse across the load ($\frac{3}{4}$ microsecond) is the time required for the wave to travel down the line and back.

The function of the diode is to permit the artificial line to charge up to the value 2E. When the diode is conducting during the charging period its resistance is very small, and therefore there is very little damping introduced. At the end of the charging period, however, when the polarity of the anode of the diode is reversed, it becomes practically an open circuit, and the artificial line cannot discharge through it but is charged to 2E, as explained in earlier chapters.

In the wave-guide run from the magnetron as far as the rotating joint, the H_{01} mode is propagated; this is changed by transformer action to the $E_{0.1.}$ mode for propagation in the circular wave-guide. The $H_{1.1.}$ mode which the circular guide would normally transmit, in addition to the $E_{0\,1.}$ mode, is eliminated by the adjustment of the length of the stub end of the circular guide.

Automatic Frequency Control

In the receiver circuits automatic frequency control is provided, but automatic gain control is not incorporated in this outfit.

There are many ways of obtaining automatic frequency control, just as there are usually several alternative circuit arrangements for obtaining most of the special effects utilised in radar circuits. In general, however, the principle is common. Here there are two channels, one tuned above the middle frequency of the I.F. waveband and the other below it. If the intermediate frequency varies above or below the mid-point of the band the voltage across one of the branches rises and the other falls. This voltage difference is classified in a discriminator circuit, amplified and applied to the repeller of the klystron in such a manner as to correct the frequency drift.

Application of a Sweeping Circuit to A.F.C.

In this particular case, in order to make the A.F.C. more effective over a wider range, a sweeping circuit is introduced. This is operated by a multivibrator which sweeps through about 15 volts at a rate of one or two sweeps per second. This sweeping voltage is introduced for two reasons : (*a*) the A.F.C. circuit becomes effective over a much wider range, and (*b*) the local oscillator has to be brought close enough in to produce a signal in the discriminator in order that the A.F.C. circuit can function and time the oscillation correctly. By sweeping the reflector voltage, and therefore the frequency of the local oscillator, tuning may be further away from the correct value without reducing the signal to zero. Thus the chance of pull-in is much increased.

Secondly, it provides a means for checking the proper operation of the A.F.C. circuit. When the local oscillator is being held at the correct frequency, the whole of the A.F.C. system is in equilibrium, which is difficult to upset. When the local oscillator sweeps, the intermediate frequency changes in a similar manner. In general, the A.F.C. circuit must be operating correctly when the indicator meter is not sweeping.

This feature in connection with A.F.C. has been specially mentioned, because it is so very important to stress the point that any drift in the frequency of the intermediate frequency is immediately reflected as a loss of overall efficiency of the system. All drifts of the intermediate frequency are immediately reflected as a drop in the receiver output which may fall to zero. Since the efficiency of the entire system, all other things being equal, depends upon the ability of the receiver to make use of very weak signals, it is of the utmost importance that the I.F. should be maintained at a constant value.

Air-conditioning Unit

An air-conditioning unit is supplied with this equipment to produce continuously a flow of dry air through the wave-guide, in order to prevent moisture collecting and so seriously attenuating signals and R.F. power.

The unit is arranged to supply 1 cu. ft. of dry air per minute against any prevailing wind pressure.

In principle, a centrifugal air blower forces air through a silica gel dryer, feeding dry air into the wave-guide. Two absorber units are supplied, one being in operation whilst the other is drying out. Each unit can be operated for about ten hours before a drying-out process becomes necessary.

CHAPTER XXII

RADAR AERIAL SYSTEMS—AERIALS

THE aerial system of a radar equipment comprises the following units :—

(*a*) The aerial proper, *i.e.*, elements that perform the dual purpose of radiating R.F. energy and receiving reflected energy or echoes.

(*b*) The feeder system which joins the output from the R.F. generator to the input of the aerial proper.

(*c*) Some form of (T. and R.) electronic switching device for isolating the receiver when transmitting, and the transmitter during reception.

(*d*) In many cases a beam switching device for accuracy in determination of bearing.

(*e*) Mechanism for rotating the aerial and controlling its movements, according to operational requirements and the conditions imposed by the type of display employed. Also means for indicating at the operating position the bearing of the aerial relative to the ship's head at any instant of time.

The electrical characteristics and physical appearances of radar aerials may vary widely. Contrast Fig. 326 (a) with Fig. 326 (c).

Appearance and size are determined by the carrier frequency selected, the required performance and prevailing conditions.

Design Considerations

The general design of the aerial for a given equipment is controlled by considerations governing the overall efficiency and the degree of accuracy of the data to be obtained.

Since the propagation of electromagnetic waves, the basic principles of radiation, W/T aerials, feeder systems and arrays are all fully dealt with in many available text-books on tele-communications, aerial systems for radar will be dealt with on the assumption that principles common to both W/T and radar can be summarised in brief review.

The principles of radiation of electromagnetic energy are founded on the law that a *moving* electric field *creates* a magnetic field and a *moving* magnetic field creates an *electric* field.

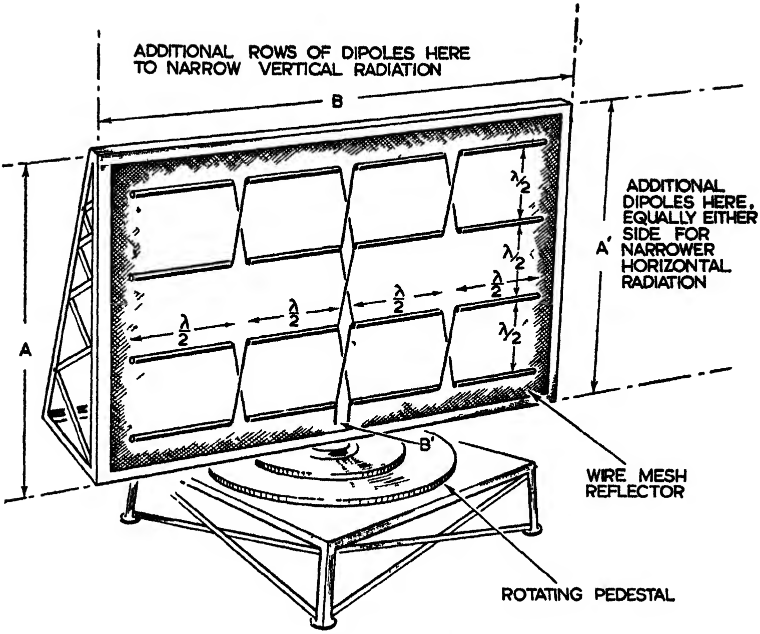


Fig 326 (a) —STACKED DIPOLE AERIAL ARRAY FOR 100 MEGACYCLES AND 20° BEAM

The frame can be rotated 360° in azimuth or tilted through a small angle of elevation

Thus an electromagnetic wave is characterised by a simultaneous appearance of magnetic and electric fields at right angles one to the other, mutually sustaining one another as they move through free space. These phenomena are covered by the statement that the created field at any instant is in phase, in time, with its parent field but is perpendicular to it in space.*

The chief consideration in connection with a radar aerial (apart from its overall efficiency as a radiator and collector of energy, and its size and weight), is its directive properties.

* Huygens views the propagation of a light wave in the following manner: each point of the wave front at any instant can be regarded as a new spherical wave, and at the next instant the new wave front will be made up of the fronts of these infinitely numerous secondary waves. If an infinite number of secondary waves is imagined, the new wave front will be a similar sphere concentric with the parent sphere and so forth (see Fig 326 (b)).

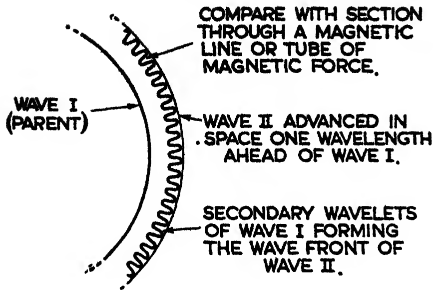


Fig. 326 (b). — ILLUSTRATING HUYGENS' CONCEPTION OF THE MECHANISM OF PROPAGATION OF A LIGHT WAVE.

To help in forming an equivalent picture the corrugations might represent vertical sections through magnetic tubes of force in the horizontal plane about the vertical electric field. (See footnote, p. 487.)

Requirements of the Aerial System

From the foregoing the aerial system for any radar set must conform to the following requirements :—

(*a*) Possess the required directional properties when functioning at the selected carrier frequency.

(*b*) Maximum overall efficiency for transmitting and receiving under conditions (*a*).

(*c*) Size and weight must be confined between any limits specified, and with due regard to mechanical problems involved in mounting and rotating the structure. It must also be rigid and robust.

All practical aerials are directive to some extent. In general, however, the term directive aerial refers to a radiating system which has been designed deliberately to concentrate its radiation in a relatively narrow beam.

R.F. energy can be reflected in much the same way as light energy, and under similar conditions, inasmuch as the dimensions of the reflector must be large compared to the wavelength of the energy to be reflected.

Directional Effects

Directional effects are obtained by the use of two or more aerials, so placed and phased, that radiation from the system adds in some preferred direction and cancels in others.

A modification of this is a system of radiating elements, disposed in such numbers and so positioned that a common reflector of suitable dimensions causes radiation to take place in the form of a beam.

Parabolic metallic reflecting surfaces may also be used to beam R.F. energy in a fashion similar to the beaming of light by a searchlight. This type of beam transmission is most practical when frequencies are sufficiently high (centimetre

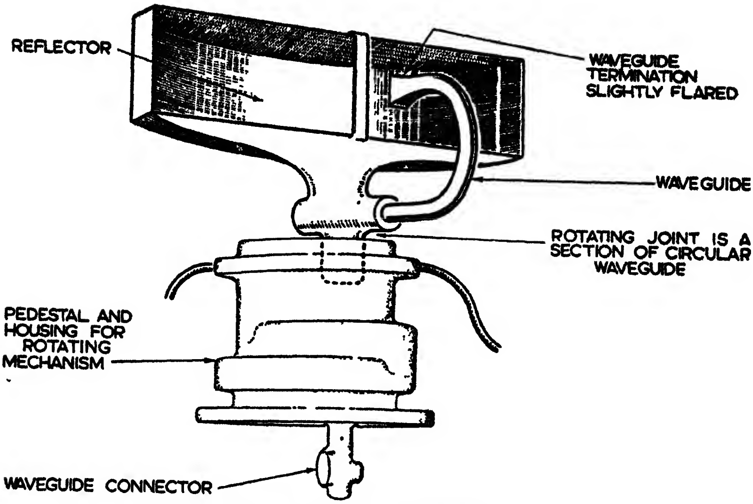


Fig. 326 (c).—PARABOLIC REFLECTOR AERIAL.

R.F. energy issuing from the flared end of the wave-guide floods the parabola and is reflected as a beam narrow in the horizontal plane, but fairly wide in the vertical plane. If the parabolic reflector is turned through 90° in the vertical direction, the beam is narrow in the vertical plane and broad in the horizontal plane.

waves) to permit the dimensions of the reflector to be comparatively small.

The study of the directive properties of a radiating unit or aerial proper necessarily involves a detailed examination of the electric field distribution. Accordingly, field strength may be plotted against bearing thus producing a Cartesian diagram. Alternatively, lengths proportional to field strength may be laid off along radii from 0° to 360° with the aerial as centre, each radius corresponding to a bearing as seen from the circumference. This constitutes a polar diagram. Of the two, the Cartesian diagram is probably the simplest and most widely used.

These diagrams may be drawn to represent the electric field conditions in the following circumstances :—

(*a*) When there are no ground reflections and the aerial is radiating into free space.

(*b*) When there are reflections from the ground wave and surrounding objects.

(*b*) may be derived from (*a*) by superimposing algebraically upon (*a*) the diagram of the distribution in space of the *reflected* energy.

One set of diagrams must be drawn for radiation in the horizontal direction and a second set (which may be widely different from the first set) must be drawn to represent radiation in the vertical direction.

Radiation Diagrams

It is important to bear in mind when studying radiation diagrams that the lobes shown do not define or bound the distance of transmission or range. A lobe is merely a picture or visual representation of the ratio field strength/bearing or field strength angular radius.

For convenience, the maximum field strength measured along the axis of a lobe is called 100 per cent. for any particular diagram. Field strengths shown on all other bearings by the boundaries of the lobe are to be regarded as percentages of the maximum.

Thus for another set of readings taken from the same aerial at twice the distance, the field strengths on all bearings would be less, but they would still all bear the same percentage relation to the maximum as for the first set of readings.

Thus, if a receiver calibrated to read field strength is transported round an aerial in a circle of diameter not less than 100 wavelengths, when radiation is taking place continuously, observed field strengths may be plotted against each bearing of the aerial to give a horizontal radiation diagram. The vertical diagram may be plotted by turning the aerial 90° about the axis of maximum radiation and proceeding as before.

If a portable transmitter is substituted for the portable receiver, diagrams can be plotted to show how well the aerial can locate in azimuth or elevation the direction from which signals are transmitted.

Field strength is measured in terms of volts per metre, *i.e.*, 1 volt per metre is equivalent to a potential of 1 volt induced in an aerial 1 metre in height. This unit is too large for practical purposes and the sub-unit of millivolts per metre, or microvolts per metre, is generally employed. At the higher frequencies the electric field is usually greater at high altitudes than near the ground, because of the cancellation caused by ground reflection in the latter case.

In practice, since the aerial is never entirely isolated from

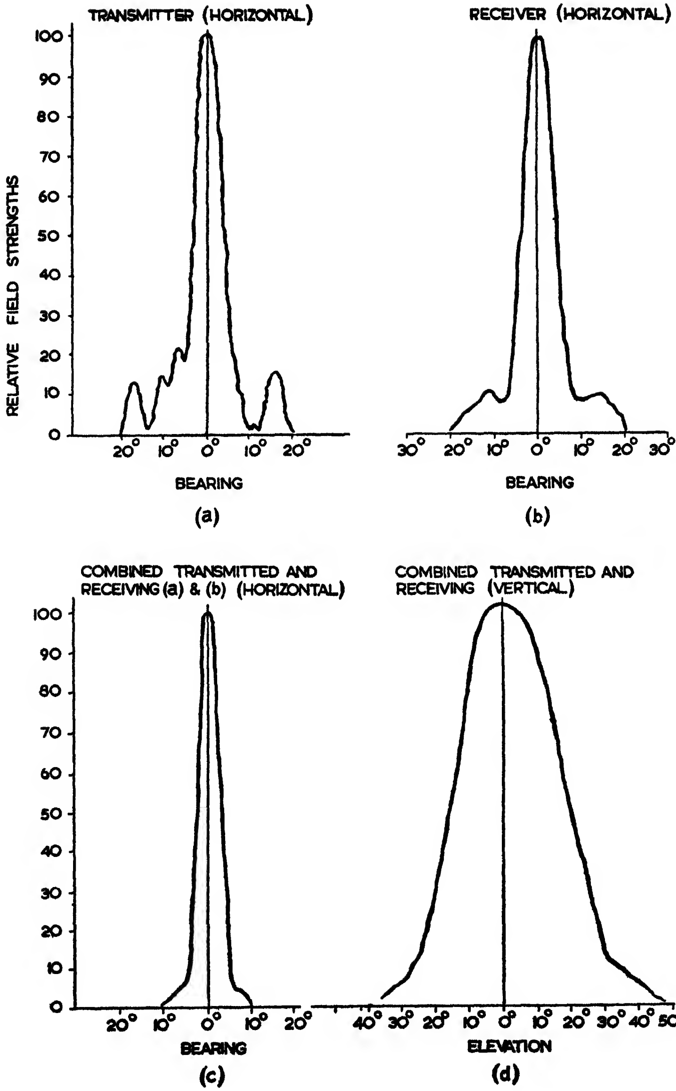


Fig. 327.—TYPICAL RADIATION DIAGRAMS.

These diagrams were constructed from observations made when separate aerials were employed for transmitting and receiving. In this case, the combined diagram reflects the directivity for the entire system.

surrounding objects, the velocity of the energy wave in the aerial is never as great as its velocity in free space, thus at frequencies above 30 megacycles the electrical length of the aerial is approximately 0·95 its physical length.

The following formula gives the physical length of a half-wave aerial for a given frequency :—

$$l = \frac{492 \times 0{\cdot}95}{f} = \frac{468}{f}$$

where l = length in feet and f = frequency in megacycles.

The position of an aerial in space determines the polarisation of the radiated wave. Thus an aerial which is vertical to earth radiates a vertically polarised wave.*

For low frequencies the polarities remain unchanged whilst the radiated field travels through space. At high frequencies polarisation varies, sometimes quite rapidly, because the wave splits into several components which follow different paths. When the recombined vectors are not parallel, the point traced by the point of the resultant vector may be circular or elliptical. A radiated field of this nature is said to be a circular or elliptically polarised field.

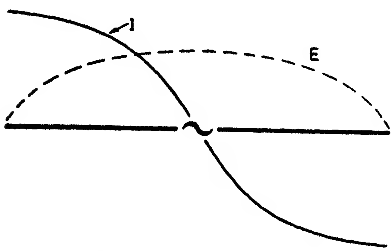


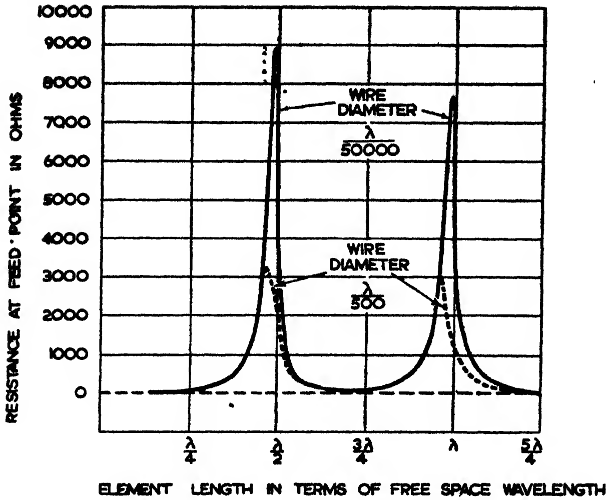
Fig. 328.—CURRENT AND VOLTAGE DISTRIBUTION FOR A $\frac{\lambda}{2}$ AERIAL.

When aerials are close to the ground, vertically polarised waves give a stronger signal close to the earth than do horizontally polarised waves, but when transmitting and receiving aerials are at least one wavelength above ground, the two types of polarisation give approximately the same field intensities near the earth's surface.

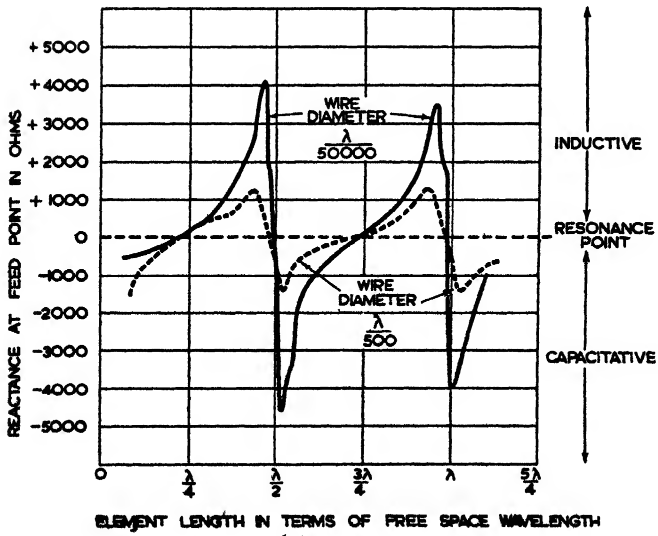
When the transmitting aerial is several wavelengths above earth, horizontally polarised waves result in a stronger signal close to the earth than is possible with vertical polarisation.

Power gain is a term used to express the power increase of

* In the case of non-directional propagation, the motion of the particles of the medium is in all directions in the wave front, *i.e.*, in all directions perpendicular to the direction of propagation of the waves. When propagation of electromagnetic energy is restricted to a given direction by the directional properties and orientation of an aerial, the radiation is said to be polarised in the direction of the electric field or E vector.



(a)



(b)

Fig. 329.—RESISTANCE AND REACTANCE AT FEED POINT DETERMINED BY DIMENSIONS.

(a) Resistance $\frac{\lambda}{2}$ and λ dipoles plotted for two diameters.

(b) Variation of reactance at feed point with length, plotted for two diameters.

an aerial over a standard aerial used for comparison. The standard aerial is usually a half-wave aerial and comparative measurements are usually taken in the optimum direction.

Aerial Impedance

The distribution of current and voltage in a half-wave aerial is shown in Fig. 328. Since the aerial input impedance is $Z = \frac{E}{I} = \sqrt{R^2 + X^2}$, the impedance varies all along the aerial. It is a minimum at the centre and a maximum at the ends. Resistance at feed point is determined by dimensions of dipole (Fig. 329 (a)).

The impedance of a $\frac{\lambda}{2}$ aerial erected in free space (a dipole) is generally taken as 73 ohms at the centre and as 2,500 ohms (allowing for losses) at the ends. Intermediate points have intermediate values of impedance.

Fig. 329 (b) is a plot of the input reactance of centre-fed aerials of any length up to $\frac{5\lambda}{4}$. Comparison of the thick and thin curves illustrates clearly the effect of the diameter of the wire.

The following facts emerge from a study of the curves :—

(1) The aerial may be capacitative or inductive, according to its length.

(2) Abrupt changes of impedance always recur near the $\frac{\lambda}{2}$ point or multiples thereof.

(3) The point of intersection of the reactance curve with the zero base line indicates the resonant lengths of the aerial.

(4) Since the curves are plotted in terms of "free space" length, the effect of reduced wave motion velocity along the aerial is shown by the curves.

(5) Thus a half-wave aerial is resonant only when it is less than the "free space" half wavelength. This foreshortening is caused by increased capacitance associated with the elements.

(6) The aerial employing

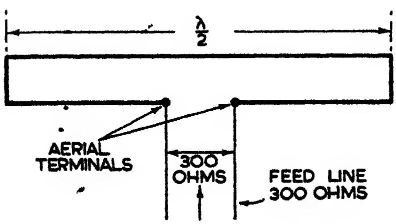


Fig. 330.—FOLDED DIPOLE OR DOUBLET MATCHING TO A $\frac{\lambda}{2}$ DIPOLE (HIGH IMPEDANCE INPUT).

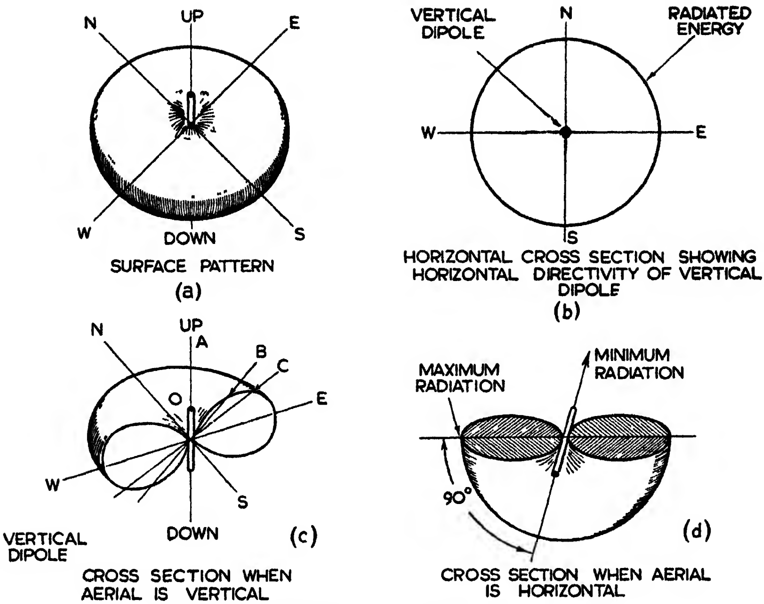


Fig. 331.—RADIATION PATTERN OF DIPOLES.

Comparing (b) and (d), horizontal radiation of (b) is similar to the vertical radiation of (d).

the larger diameter wire is foreshortened more than the one using the smaller diameter wire because of the larger capacity involved.

The input impedance of an aerial is affected by the presence of nearby conductors.

Radiation Resistance

Radiation resistance is defined as the ratio of the total power radiated to the square of the effective value of the maximum current in the radiating system. For a half-wave aerial this value is approximately 73 ohms and equals the input resistance of the centre of the aerial. For a λ aerial at its current maximum the radiation resistance is approximately 36·6 ohms.

The radiation resistance is affected by the height of the aerial above ground and by its proximity to other nearby objects.

Other smaller losses are I^2R, Corona* and insulator losses.

Since the input impedance of a single dipole is comparatively

* Sometimes loosely referred to as " brushing."

low, the feeder should be matched to this value. This is not always convenient or desirable. In order to obtain a higher impedance at the feed point two or more dipoles may be connected as shown in Fig. 330. The input impedance is now increased because the *total* length of the radiating element is approximately one wavelength. The aerial is fed at high voltage and with a low value of current. This half-wave doublet consists of two parallel closely spaced $\frac{\lambda}{2}$ wires connected together at the ends and fed in phase.

With the aerial in free space the input impedance is about 300 ohms. Thus it is possible to connect an open-wire transmission line of this impedance directly to the terminals of the doublet without a matching transformer.

Fig. 331 shows the pattern of the energy radiated by a half-wave dipole, (a), (b) and (c) with the dipole in the vertical position, and (d) in the horizontal position. From current and voltage distribution along a $\frac{\lambda}{2}$ dipole, it can be seen that a vertical $\frac{\lambda}{2}$ dipole radiates equally in all directions in the horizontal plane. Theoretically, in free space the vertical dipole has no radiation along the direct line of its axis, but there is considerable radiation at other angles measured to the line of the axis. Due to varying vertical directivity of a vertical dipole a field strength pattern taken in a horizontal plane must specify the vertical angle of radiation for which the pattern applies. This is demonstrated in Fig. 331 (c). The radiation along OA is zero, but at the angle AOB there is appreciable radiation, increasing still more at AOC, etc.

If a horizontal dipole is used, the free space pattern shown at Fig. 331 (b) applies if it is rotated 90°. Fig. 331 (d) shows the cross-section for this pattern when cut in half. The maximum radiation takes place vertically in a line perpendicular to the aerial. As the angle decreases from 90° the radiation also

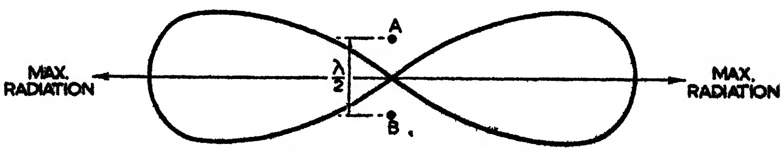


Fig. 332.—Horizontal radiation pattern of two aerials spaced half a wave apart and fed in phase.

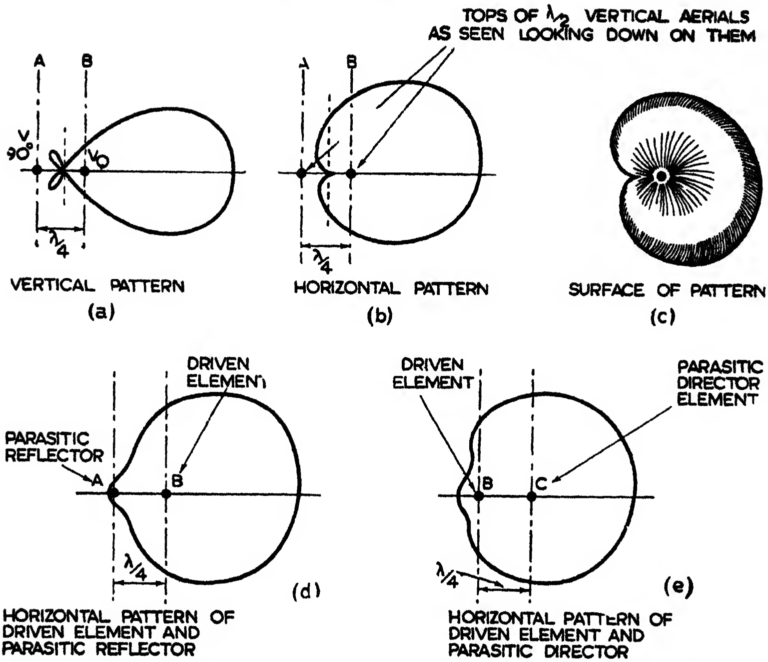


Fig. 333.—DIRECTIONAL RADIATION RESULTING FROM COMBINED FIELDS OF SPACED AERIAL ELEMENTS.

(a) Vertical pattern resulting from $\frac{\lambda}{4}$ spacing (fore and aft) when fed 90° out of phase.
(b) Horizontal pattern as in (a) looking down on the aerial.
(c) Surface of radiation about (a), (b) as it might appear in three dimensions.
(d) Horizontal pattern of driven element and parasitic reflector.
(e) Horizontal pattern of driven element and parasitic director.

decreases. Thus the horizontal dipole has a decided unidirectional characteristic in the horizontal plane.

Where two aerials, A and B, a half-wave apart are erected in phase, their radiation is concentrated along a line at right angles to their plane (see Fig. 332).

Analysis. The currents in aerials A and B are equal in value and in phase.

∴ Fields of A and B are similar in polarity, phase and amplitude.

The delay time of field A to reach B $= \frac{\lambda}{2}$, or 180° or π.

∴ A field cancels B field.

The same condition applies to energy radiated from A.

∴ Radiation in the plane of A and B is practically zero.

But at any point at right angles to the plane of A and B the radiated energy is in phase and additive.

This array gives a bi-directional field pattern and radiation is in the plane at right angles to the array.

Let A and B be placed as in Fig. 333 (a) $\frac{\lambda}{4}$ apart and let them be fed with equal currents having a phase angle of 90°. The radiation pattern is now unidirectional, see Fig. 333 (a).

The time interval is now such that the radiated wave from A reinforces B in the forward direction (to the right). But waves leaving B for A find the wave leaving A 180° out of phase and complete cancellation takes place to the left of A.

At other angles the waves add to give intermediate values, see Fig. 333 (b) and (c).

If an aerial slightly longer than a $\frac{\lambda}{2}$ and not connected to a power source is placed parallel to but slightly less than $\frac{\lambda}{4}$ directly behind a driven $\frac{\lambda}{2}$ aerial it will act as a parasitic reflector. It absorbs power and re-radiates it with such a phase relation to the original radiation that the fields of the two aerials add in one direction and subtract in the other, producing the pattern shown in Fig. 333 (d).

A parasitic element shorter than $\frac{\lambda}{2}$, placed parallel to and slightly less than $\frac{\lambda}{4}$ ahead of a $\frac{\lambda}{2}$ aerial is a director. It absorbs power and re-radiates it with such a phase relation that the fields of the two aerials again add in the direction of the director, see Fig. 333 (e).

The field of the driven aerial induces in the reflector a voltage and field which opposes the inducing field. For a properly designed reflector the induced field is comparable in magnitude to the inducing field. In practice the parasitic reflector is assumed to set up about 85 per cent. of the field of the driven aerial. Therefore only a small amount of energy from the driven aerial can travel beyond the reflector.

The time delay is such that the field from the reflector reaching the aerial is in phase and reinforces the aerial field.

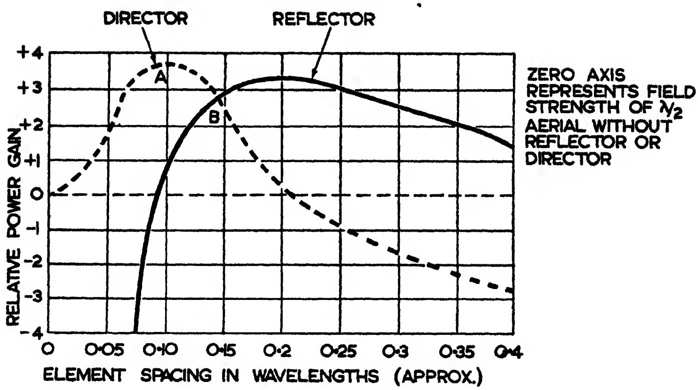


Fig. 334 (a).—RELATIVE POWER GAINS OF PARASITIC ARRAYS FOR DIFFERENT DIRECTOR AND REFLECTOR SPACINGS.

A wide variety of fields can be obtained with only one parasitic element. Two variables can change this pattern :—

(*a*) The length or timing of the parasitic element.

(*b*) Spacing between driven element and parasitic element.

The parasitic element should be spaced $\frac{\lambda}{4}$ from the driven element to provide the required cancellation and reinforcement of the field of the driven aerial.

If the physical spacing is made less than 90° $\left(\frac{\lambda}{4}\right)$ for reasons of economy in space or mechanical efficiency, the required time delay must be provided electrically.

This adjustment depends upon the fact that current lags or leads on the voltage induced in the parasitic aerial according to whether it is longer or shorter than resonant length so that when off resonance it may behave inductively or capacitatively.

If therefore the parasitic reflector is placed $\frac{\lambda}{8}$ behind the driven element the current in it must be made to lag 45°.

To do this the resonant frequency of the reflecting element must be made slightly lower than that of the driven element, so that it is sufficiently inductive to produce a 45° lag.

This means that the reflector must be physically slightly *longer* than the aerial.

Similar reasoning applies to the spacing and length of the director which is placed in front of the aerial. If the spacing

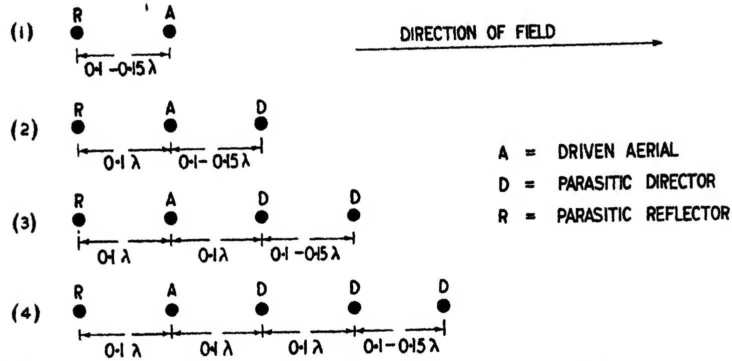


Fig. 334 (b).—EXAMPLES OF MULTI-ELEMENT PARASITIC AERIAL ARRAYS (YAGI).

is less than $\frac{\lambda}{4}$ the length of the director must be slightly *less* than that of the aerial. It then acts as a capacitative reactance and causes the current in the director to lead. Its length must therefore be adjusted to provide the lead which will cause its field to reinforce in the forward direction.

Fig. 334 (a) shows the power gain for director spacings in wavelengths between a half-wave director and its driven half-wave radiator. Curve B shows the power gain for half-wave spacings between the reflector and the driven half-wave radiator.

From a consideration of the two curves, A for the director and B for the reflector spacings, it will be seen that a reflector provides only a very small amount of gain when placed at 0·10 from the driven element and that the gain falls off sharply at spacings of about 0·015. The director, however, may have useful gain at much smaller spacings.

Several parasitic elements can be used in conjunction with a driven aerial to increase further its directivity and power gain. Various practical combinations are shown in Fig. 334 (b).

The following table shows the theoretical power gain of the alternative arrays shown in Fig. 334 (b).

No. elements	Power gain
2	2·5
3	3·6
4	5·0
5	6·4

The advantages are apparent when a unidirectional beam with minimum front to back ratio is required

Fig. 335 shows the type of pattern produced by multi-element parasitic arrays or Yagi aerials.

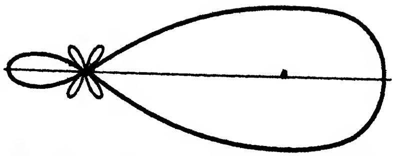


Fig. 335.—FIELD PATTERN OF YAGI ARRAY.

Phased Arrays

The two most widely used systems of phased arrays are :—*

(*a*) Colinear.
(*b*) Broadside.

The colinear array consists of two or more $\frac{\lambda}{2}$ radiators placed end to end and excited in phase.

Maximum radiation is perpendicular to the axis of the elements and is bi-directional when the aerials are horizontal and when no reflector is used. The vertical colinear aerial is known as a Franklin aerial.

The broadside array consists of two or more elements placed back to back or in parallel with one another.

The minimum radiation is broadside to the plane of the elements and is also bi-directional.

Another system sometimes used is the end fire array. This consists of two or more $\frac{\lambda}{2}$ radiators placed parallel to each other but excited out of phase. The exact phase difference depends upon the separation and the desired pattern. The maximum radiation is in the plane of the wires through the centres of the elements. Radiation from the end fire array is either bi-directional or unidirectional, depending upon the separation and phase difference of the elements.

Colinear Arrays

The simplest form consists of two horizontal $\frac{\lambda}{2}$ aerials erected in line with their free ends directly opposite and fed in phase, see Fig. 336 (a).

* Initially, consideration is focused upon the shape of the beam formed by the various arrays when no reflector is present, and the manner in which the beam can be varied according to the number and disposition of the dipole elements forming the array. Without reflectors, the beams are bi-directional. They do affect the beam formation, but apart from making the beam uni-directional they also contribute to the energy which the uni directional beam gains.

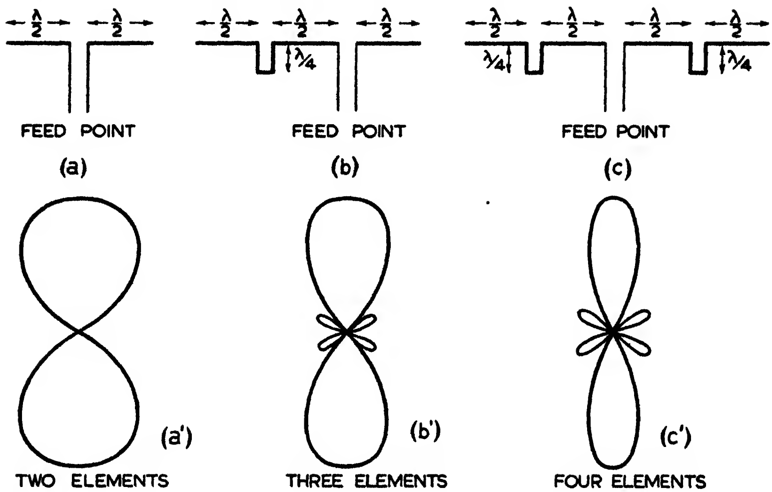


Fig. 336.—HORIZONTAL FIELD PATTERNS FOR HORIZONTAL COLINEAR ARRAYS.

The resulting radiation is in a direction at right angles to the plane of the aerial conductors. Two-, three- and four-element arrays together with their field patterns are shown in Fig. 336 (a, a′), (b, b′) and (c, c′).

Note. The *beam* becomes sharper as the number of elements placed end to end is increased.

Two $\frac{\lambda}{2}$ sections in phase give a power gain approximately 1·6 greater than a single $\frac{\lambda}{2}$ aerial. Three sections give a gain of 2·6 and four sections approximately 3·2.

When a colinear system is oriented horizontally it has most of its directivity in the horizontal plane. When vertically erected, most directivity is in the vertical plane.

A colinear array may be made unidirectional by use of reflecting wires or a small reflecting screen.

Broadside Arrays

If colinear elements are stacked above and below another set of similar elements, the result is a broadside array. It is sometimes used with a sheet-metal or screen reflector. High-power gains may be obtained with this arrangement, depending upon the spacing, number and tuning of the elements.

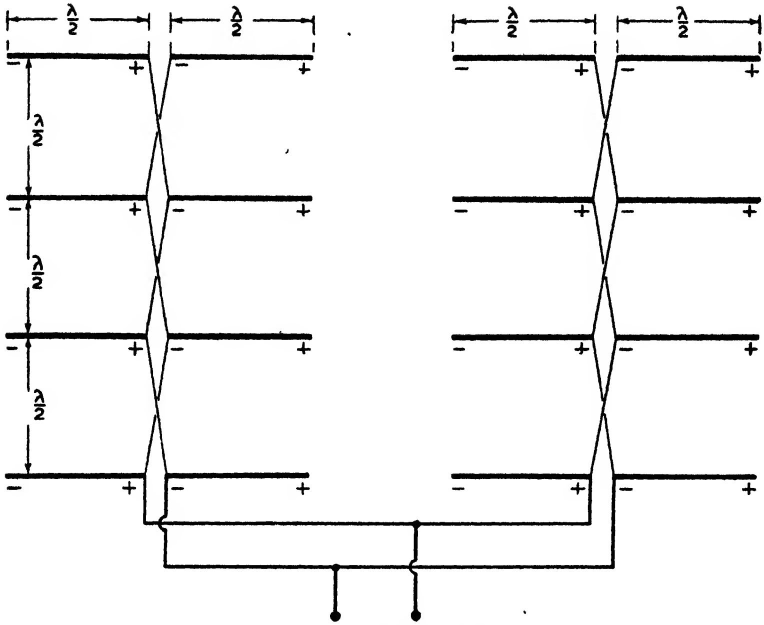


Fig. 337.—BROADSIDE ARRAY.

Fig. 337 shows a broadside array consisting of a number of $\frac{\lambda}{2}$ sections arranged to concentrate energy into a horizontally narrow beam and at very low angles above the horizon.

This low-angle radiation is obtained by arranging several in-phase $\frac{\lambda}{2}$ aerials one above the other and separating them by $\frac{\lambda}{2}$. This method of stacking dipoles practically eliminates vertical radiation.

Equivalent polarisation and voltages on all adjacent elements are separated by $\frac{\lambda}{2}$. This separation is accomplished by spacing the aerials $\frac{\lambda}{2}$ apart vertically and by transposing the transmission lines between aerial sections.

Without a reflector this aerial would radiate a narrow beam of energy in both directions at right angles to the plane of the conductors.

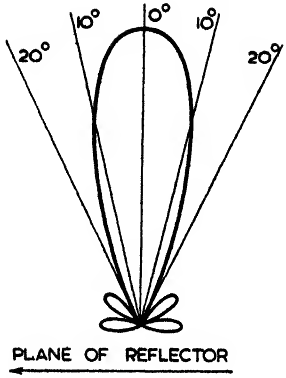


Fig. 338.—FIELD PATTERN FOR BROADSIDE ARRAY WITH REFLECTOR.

Unidirectional operation can be obtained by erecting a similar array of parasitic reflectors from $\frac{\lambda}{10}$ to $\frac{\lambda}{4}$ behind the driven aerial. Better results are obtained by using a metal sheet or screen reflector.

Fig. 338 shows typical energy distribution from an array as in Fig. 337 but with a reflector added. The beam is about 20° wide.

If a narrower beam is required, additional units may be added either horizontally or vertically, depending on the direction in which it is desired to restrict radiation.

The pattern may be made narrower in the vertical plane by stacking many elements one above the other. Similarly, the horizontal pattern may be made narrower by adding more elements alongside one another.

These arrays can be used with the elements placed either vertically or horizontally.

Ground Reflection

The total radiation from an aerial is made up of two components. One component leaves the aerial directly, the other is a ground reflection which appears to come from an underground image of the real aerial. The image is considered to be as far below the ground as the aerial is above it. Also the aerial need not be placed at the surface to produce the image. The image concept holds equally well for aerials above the surface and for aerials in front of large flat sheets of conducting material. The image aerial assumes a variety of forms depending upon the type of soil adjacent to the aerial (Fig. 339).

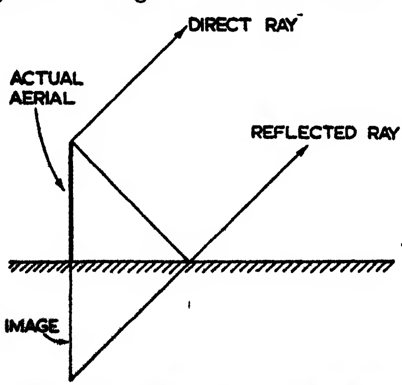


Fig. 339.—REFLECTED RAY FROM AERIAL IMAGE.

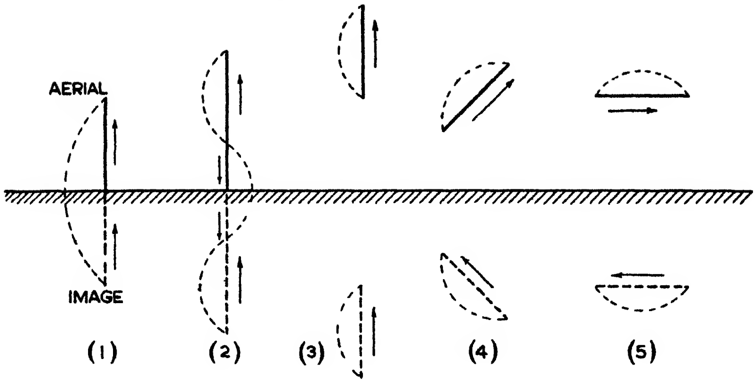


Fig 340 —CURRENT DISTRIBUTION IN REAL AND IMAGE AERIALS

Fig. 340 shows current distribution for real aerials along with their images. It is particularly noticeable that the currents in the horizontal aerial and its image are flowing in opposite directions and are therefore 180° out of phase ((4) and (5)). The current in the vertical aerial and its image are flowing in the same direction and are in phase ((1), (2) and (3)).

Hence the resultant field strength may be either twice the strength from the real aerial alone or zero field strength or some intermediate value depending upon the direction and location of the point at which the field strength is measured.

Fig. 341, vertical polar diagrams, are used to show the result of ground reflection for aerials of various heights.

Note. These graphs are not plans of the radiation patterns of vertical aerials. They represent multiplying factors arising as a result of ground reflection and are applied to the theoretical free space pattern in order to ascertain the resultant form of the energy distribution.

As seen from the graphs in Fig. 341, the reinforcement of fields in some areas and their subtractions or cancellations in other areas when applied to the free space pattern, produce a non-uniform field pattern (see Fig. 342).

Thus where reinforcement occurs, lobes are produced, and where cancellations take place, gaps are in evidence.

The main factors that determine the angle of elevation of the lobes and gaps in the vertical polar diagram are :—

(*a*) Wavelength.
(*b*) Height of aerial above earth.

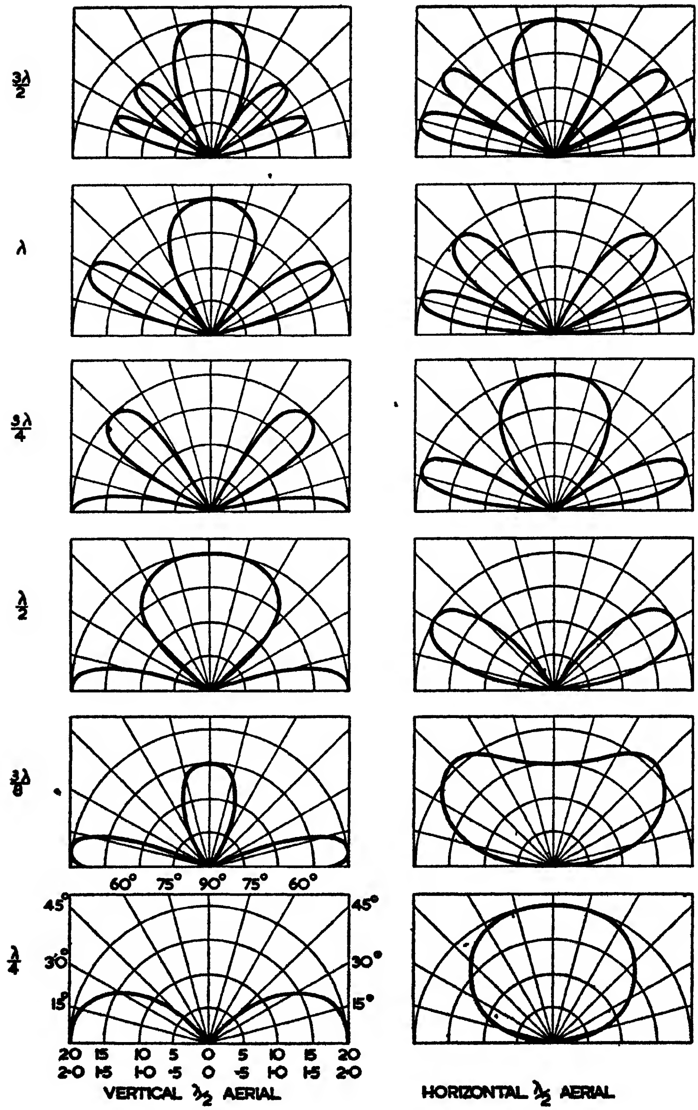


Fig. 341.—VARIATIONS OF GAPS AND LOBES FOR $\frac{\lambda}{2}$ AERIALS OF DIFFERENT HEIGHTS (APPROX.), *i.e.*, PATTERNS OF ENERGY DISTRIBUTION.

Distances in wavelengths indicate heights of the centres of the aerials above a perfectly conducting earth.

If the height of the aerial is large compared with the wavelength, the first lobe is at a very low angle.

Vertical Pattern of a Directional Array

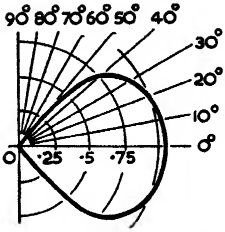
If the array is located horizontally near the earth, all values of the free space pattern must be multiplied by the factor shown in Fig. 341. The resultant vertical pattern developed from Figs. 342 (a) and 342 (b) is shown at Fig. 342 (c).

The fundamental principles of aerial design for use at micro-wavelengths are the same as those for use at lower frequencies.

Micro-wave multi-element parasitic arrays occupy very small space. Similarly, parabolic and horn-type reflectors which would be too bulky at lower frequencies become practicable at micro-wavelengths because of their physical size and desirable because of their high efficiencies.

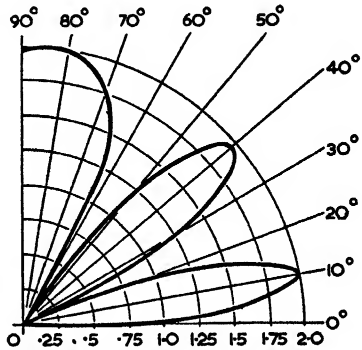
Micro-waves have characteristics very similar to those of light waves and therefore the parabolic reflector is an obvious application at these frequencies (Fig. 343 (a)).

If a dipole is placed at



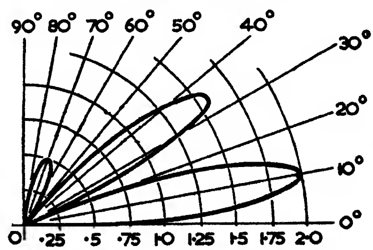
(a)

VERTICAL PATTERN OF HORIZONTAL ARRAY IN FREE SPACE



(b)

GROUND REFLECTION FACTOR HORIZONTAL AERIAL 1½ λ HIGH



(c)

DIRECTIONAL PATTERN IN A PLANE PERPENDICULAR TO THE EARTH OF AN ARRAY OF HORIZONTAL AERIALS 1½ λ ABOVE EARTH i.e. THE RESULTANT OF SUPERIMPOSING PATTERN (a) ON PATTERN (b).

Fig. 342.—PATTERN OF ENERGY DISTRIBUTION (c) DUE TO THE COMBINED FIELDS OF (a) AND (b).

This illustrates the manner in which the lobe or free space pattern (a) is modified by the ground reflection pattern (b) to a resultant as shown at (c).

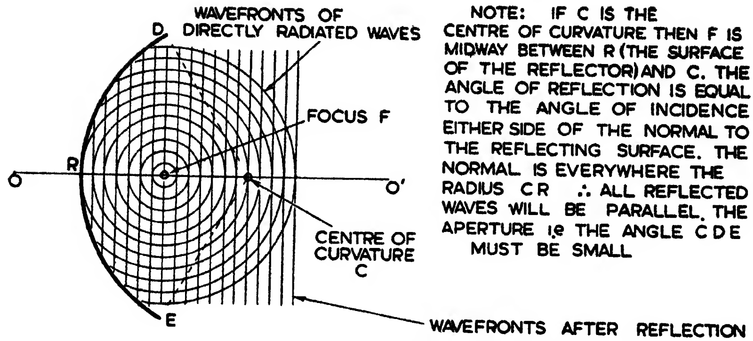


Fig 343 (a) —CONCAVE REFLECTOR.

the focal point A in Fig. 343 (b) the concave reflector concentrates the radiation from the dipole into a beam.

A parabolic reflector converts spherical waves radiated by the dipole into vertical lines that represent the wavefront after reflection. The equation of the parabola is $y^2 = 4ax$, where a = distance between focus and vertex.

Fig. 343 (b) shows the cross-sectional view of a rotatable paraboloid excited by a vertical aerial located at the focal point inside the paraboloid.

A hemispherical shield is used to direct all radiation back towards the parabolic surface.

This results in :—

(*a*) Elimination of direct radiation.
(*b*) Sharper beam.
(*c*) Power is conserved.

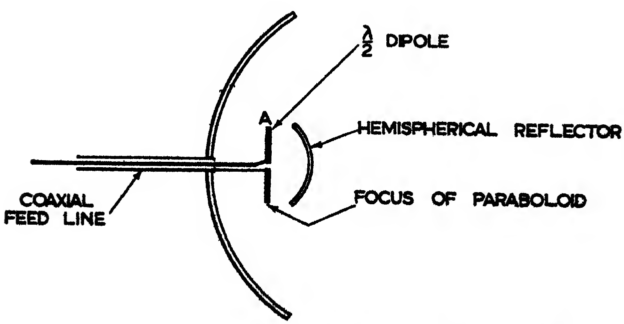


Fig. 343 (b). PARABOLOID EXCITED BY REFLECTED ENERGY FROM A DIPOLE.

The radiation pattern (typical) is shown in Fig. 344 (a) and (b). Very narrow beams are possible with this type of reflector.

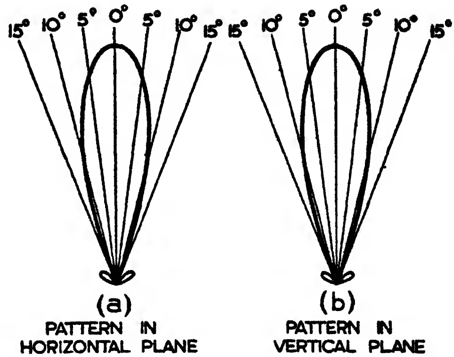


Fig. 344.—TYPICAL RADIATION PATTERN FOR RADIATOR AS IN FIG. 343 (b).

Cylindrical paraboloids with open or closed ends are also employed, see Fig. 345. These reflectors have a curvature in one plane, usually horizontally. They are excited normally by an aerial placed parallel to the cylindrical surface and located at the axis of the paraboloid. The focus should be well within the mouth of the device for maximum efficiency as an interceptor and reflector. Fig. 326 (c) illustrates a typical radiation system in which the electromagnetic energy from the guide termination is used to flood a reflector (" cheese " type).

Radiation into Free Space from Termination of a Wave-guide Feeder System.

The only practical solution to the problem of the transference of electromagnetic energy from one point of a system to another, by means of *guided* waves of low, high and ultra-high frequencies, has been the use of transmission lines of open-wire or coaxial type. The advent of micro-wave transmission and reception technique has opened up the possibility of the use of other types of transmission systems for guided waves at these frequencies.

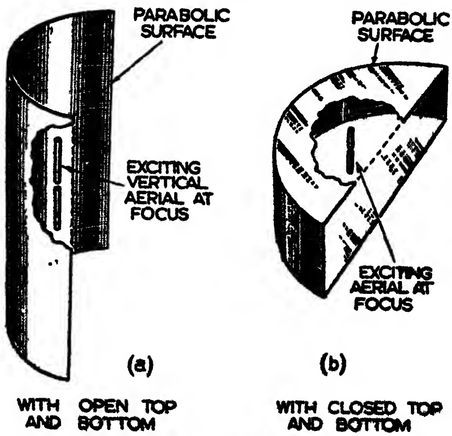


Fig. 345.—REFLECTOR AERIALS FLOODED BY RADIATION FROM DIPOLES.

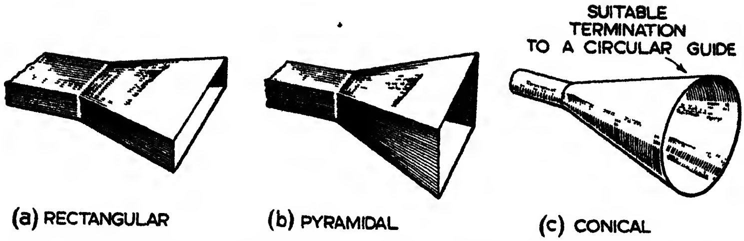


Fig 346.—TYPICAL TERMINATIONS FOR RECTANGULAR AND CIRCULAR WAVE-GUIDES

Of these new types of transmission systems, the one of most practical importance is the hollow feeder. This consists of a hollow tube of conducting material of suitable cross-sectional shape and dimensions, circular or rectangular, through which electromagnetic energy can be " piped " * in the form of wave-trains (see Appendix II).

There is no go-and-return circuit in the ordinary sense of current theory. Whereas in the case of ordinary current-carrying conductors, the electromagnetic energy is regarded as being transported by the current within the conductor itself, the main bulk of the energy transported in a hollow feeder resides in the hollow interior of the tube, and the only currents involved are the circulatory flow of electrons induced on the inner surface of the tube by the oscillatory fields adjacent to this surface or boundary.

On reaching the open end of a tubular feeder, waves are radiated into free space in the form of a directional beam containing a main lobe centred on the axis of the feeder, guide or tube, and a number of subsidiary side lobes. The angular width of the main beam, in planes parallel and perpendicular to the emergent electric field, will depend upon the dimensions of the aperture of the tube in terms of wavelength.

It can be shown that narrow beam width is associated with large tube apertures. Improved directivity of the beam can, therefore, be obtained by flaring the ends of the tube.

Thus a rectangular tube can be flared in the form of a sectoral horn, to produce a beam which is sharp in the plane of the longer sides and comparatively broad in the plane of the shorter sides. The actual beam angles obtained depend

* When wave-guides are employed, the form of construction for transferring electromagnetic energy in radar terminology is frequently referred to as " plumbing."

upon the length and angle of flare of the horn. Similar results hold for conical horns at the ends of circular tubes.

Thus there is a remarkable similarity between the laws governing the directivity of aerial arrays and those which determine the beam angle of radiation from flared wave-guides.

Likewise, it can be shown that the receiving pattern or collecting properties of flared horns, relative to electromagnetic energy travelling in free space and at suitable frequency, follows very closely the transmitting directivity pattern.

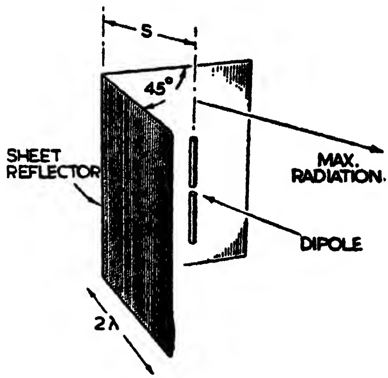


Fig. 347.—CORNER REFLECTOR ERECTED BY DIPOLE.

Horn-type Radiators

Horn-type radiators are frequently used and are very practical. The dimensions must be large compared with wave-length. This requirement is not difficult to meet for ultra-high frequencies (micro-waves), and they have the advantage of being usable over a wide frequency band. The operation of the device is comparable in some respects to that of an acoustic horn. The throat, however, of an acoustic horn usually has dimensions much smaller than the sound wavelength for which it is used, whereas the electromagnetic horn must have throat dimensions comparable to the wavelength being used.

Horn-type radiators may be conveniently used with wave-guides, since they serve the dual purpose of matching the impedance of the wave-guide to free air and at the same time produce a directive pattern.

The application of a horn is not confined to wave-guide operation, since they may be fed by a coaxial or other type of transmission line.

An electromagnetic horn may take a variety of shapes, see Fig. 346 (a) to (c). The shape of the horn and dimensions of the mouth measured in wavelengths determine the field pattern shape for a given magnitude and the phase distribution of the field produced across the mouth of the horn. In general, the longer the opening of the horn, the more directive is the resulting field pattern. An opening of approximately 5λ produces a radiated major lobe of about 30°.

Corner Reflector

The corner reflector aerial consists of two flat sheets which meet at an angle to form a corner, see Fig. 347. This type of reflector is normally driven by a $\frac{\lambda}{2}$ dipole which bisects the corner with maximum radiation in the horizontal plane.

This type has a somewhat larger gain than the parabolic type. The gain is greatest when the distance S is λ and the angle between the sides is 45°. The sheets are made the same height as the dipole which excites the corner and about 2λ on a side. The side length is not critical, but the dipole should bisect the angle accurately.

CHAPTER XXIII

RADAR AERIAL SYSTEMS—FEEDER SYSTEMS

CHOICE of the feeder system to be employed is determined basically by the transmitting and receiving units employed, the radiating system and the layout.

An open-wire transmission line is simple, cheap to construct and maintain, and is very efficient at the lower frequencies. It suffers from the great disadvantage, however, that due to the necessary spacing of the conductors, radiation takes place from the feeder itself, which results in increased losses. The higher the frequency the greater are the losses.

The coaxial line has the main advantage over the open-wire line of being completely shielded. Since the energy is conveyed entirely within the outer conductor, none of it is lost through radiation.

At very high frequencies the saving in power is considerable. In addition, the space occupied by its installation is small and the method of mounting is very simple, since the outer surface is at earth potential.

The disadvantages of the coaxial line are :—

(*a*) Exact matching is complicated.

(*b*) The peak power is limited by the spacing of the conductors.

(*c*) Moisture must be excluded if the breakdown point is to be kept high, *i.e.*, if the peak power-carrying capacity of the cable is to be preserved.

(*d*) The connection to the aerial is usually balanced to earth, the coaxial line has one side earthed and this necessitates the use of a balance converter.

(*e*) The losses in a coaxial line become excessive at wavelengths shorter than 10 cm.

Transmission Line Construction

Fig. 348 is an example of open-wire transmission line construction.

For wavelengths of 10 cm. and less, wave-guides are usually employed. At a frequency of 3,000 megacycles the dimensions of a guide are such as to render its use practicable. The losses are much smaller in wave-guides and they have the additional advantage of simplicity in construction and the moisture problem is more easily solved.

System Using Separate Aerials

Fig. 349 shows a simple feeder system such as might be applied when separate aerials are used for transmitting and receiving. In this case no T.R. switch is required, and beam switching is not employed.

A coaxial line is used to eliminate possible radiation and prevent stray pick up of R.F. energy by receiver and indicator.

Note the following points :—

(1) Output coupling circuit tuned to transmitter by C.

(2) Slides A and A′ adjusted for an impedance of 70 ohms (the characteristic impedance of the transmission line).

(3) The open-wire transmission line from A and A′ to the coaxial line A″B is very short and has only a small effect on the impedance that is seen looking from the output coupling to the coaxial line.

(4) The line balance converter B is used to prevent R.F. energy appearing on the outside of the outer conductor of the coaxial line. It has no matching effect. (See Appendix I, "Transmission Lines.")

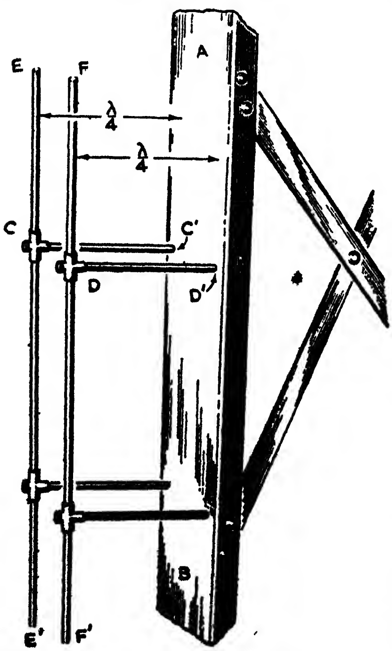


Fig. 348.—TRANSMISSION LINE CONSTRUCTION EMPLOYING SHORTED $\frac{\lambda}{4}$ SECTIONS AS INSULATORS.

Note. AB is a metal member of a self-supporting mast or tower.

CC′ and DD′ are $\frac{\lambda}{4}$ sections shorted by the metal member to which they are attached. EE′ and FF′ are open-wire transmission lines. This form of construction is only possible for one fixed frequency, *i.e.*, that frequency which makes the $\frac{\lambda}{4}$ sections resonant and, therefore, causes them to act as insulators.

Impedance at resonance approaches infinity. This construction is cheap and efficient but is possible only when one fixed frequency is employed.

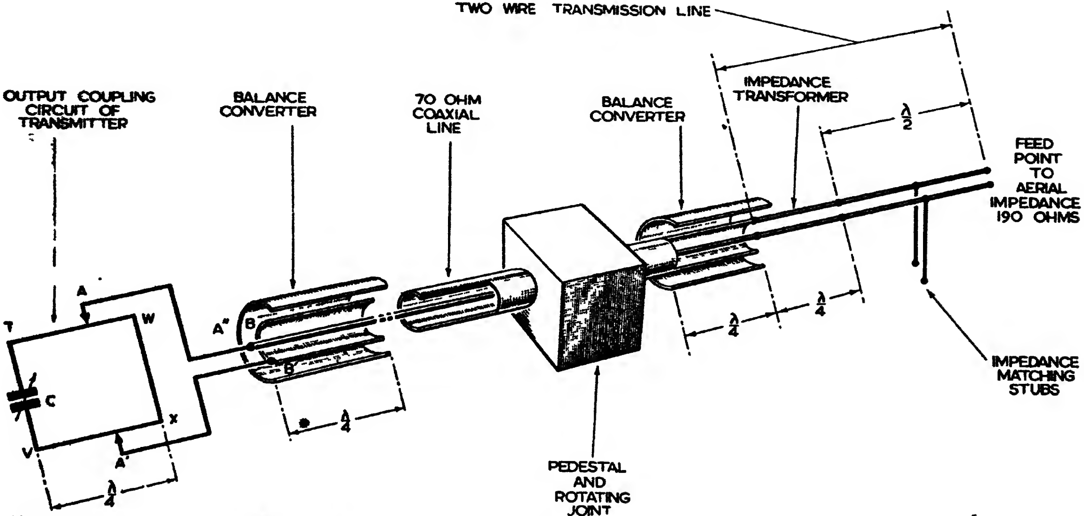


Fig. 349.—TYPICAL FEEDER SYSTEM FROM TRANSMITTER TO AERIAL WHEN SEPARATE AERIALS ARE USED FOR TRANSMISSION AND RECEPTION.

No T.R. switch is incorporated and no beam switching is employed. TVXW is the secondary or coupled circuit of a push-pull tuned anode, triode transmitter.

$$\text{Note the Ratio of } \frac{\text{Power loss in unmatched line}}{\text{Power loss in matched line}} = \frac{1 + R^2}{2R}$$

where R = Standing Wave Ratio or S.W.R.

Since the ratio of power lost in an unmatched line to that lost in a matched line becomes smaller as the reflected wave vanishes, matching to Z_0 numerically is not in itself sufficient to ensure maximum efficiency. Reflections due to reactive components must also be tuned out. Hence the impedance matching stubs.

(5) The rotating joint is generally built into the aerial pedestal.

(6) The characteristic impedance of the coaxial line is 70 ohms.

(7) The impedance of the aerial is approximately 190 ohms at the feed point.

(8) The impedance transformer makes an approximate match between coaxial line and aerial.

(9) The characteristic impedance of the transformer which is $\frac{\lambda}{4}$ section of open-wire line is given by :—

$$Z_0 = \sqrt{70 \times 190} \text{ ohms} - 115 \text{ ohms (approx.)}.$$

(10) To reduce standing waves to a minimum, a closer match is necessary than is possible with the transformer alone.

(11) Matching stubs are used for this purpose.

(12) At the junction of the coaxial line (unbalanced) with the two-wire impedance transformer (balanced) another line converter is required. No converter is required at the other end of the impedance transformer since both lines are balanced. The receiving aerial follows similar practice.

Aerial System for Micro-Wave Transmission

Fig. 350 is the schematic diagram of an aerial system suitable for installation in an aeroplane employed on surface reconnaissance. Size and weight are of first-rate importance. In order to fulfil the radar requirements for this particular application, the installation has the following specification :—

Carrier frequency, 3,000 megacycles.
Pulse width, 2 microseconds.
p.r.f., 800.
Display, P.P.I.
Aerial rotation, 20 r.p.m.
Average power input, 250 watts.
Peak power, 50 kW.
Coaxial transmission line (not wave-guide in this instance) since a dipole is to be fed.

The aerial system includes :—

(*a*) Coaxial transmission line.
(*b*) T.R. switch.
(*c*) The radiating elements which consist of a dipole with reflector and a parabolic reflector together with rotating and tilting mechanisms.
(*d*) Transmission and reception by common aerial.

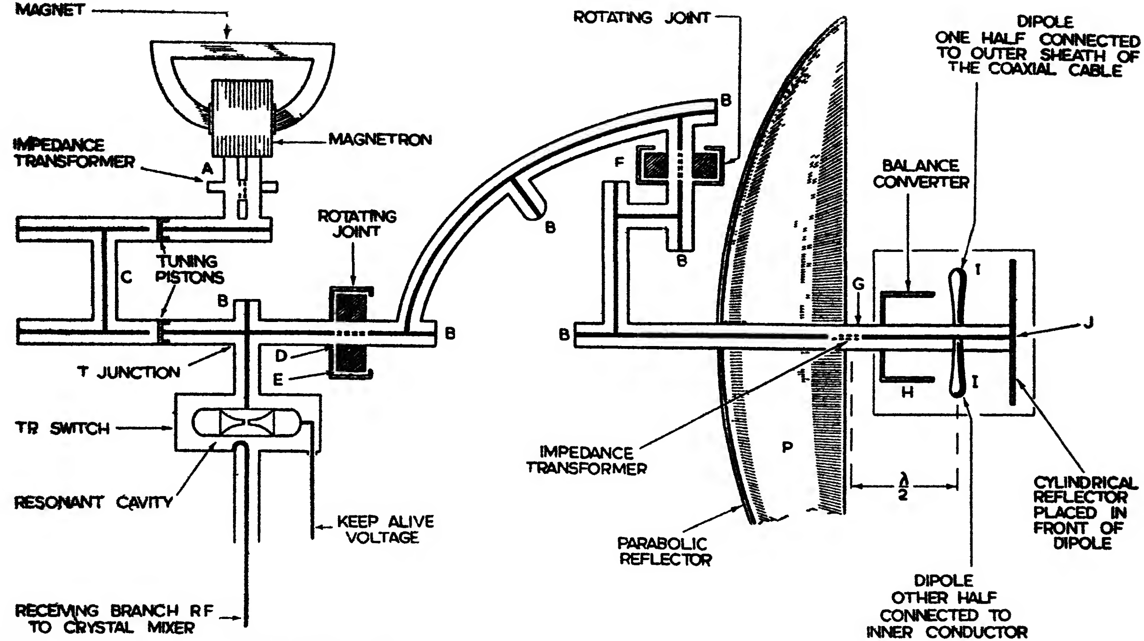


Fig 350 —SCHEMATIC OF AN AERIAL SYSTEM TO FEED A DIPOLE BY MEANS OF A COAXIAL CABLE AT MICRO-WAVE FREQUENCY

R F. energy from the dipole is reflected by J to flood the radiating paraboloid P Stubs marked B support the centre conductor and insulate it from the outer conductor The pistons tune the $\frac{\lambda}{2}$ transformer section C, the magnetron output and the line respectively

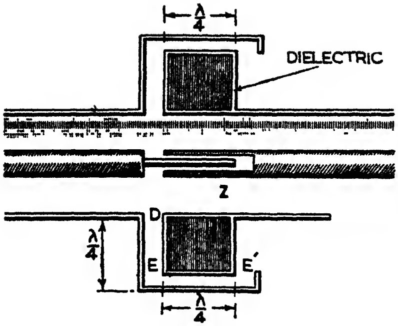


Fig. 351.—DETAIL OF ROTATING JOINT

The coaxial R.F. transmission line has a characteristic impedance of 50 ohms. The magnetron is matched during transmission to the line by an impedance transformer A built into the output connection of the magnetron.

The rotating joints at E and F permit rotation in azimuth and tilting for elevation (see Fig. 350). The two coaxial collars which form part of the line act as chokes (see Appendix II). The distance D—E is electrically $\frac{\lambda}{4}$ but physically shorter because of the increased capacity due to dielectric. In Fig. 351 the short circuit at D is reflected to E as a high impedance. Since E to E′ is also $\frac{\lambda}{4}$ the high impedance at E appears as a short circuit at E′. Therefore the R.F. energy sees a continuous conductor at D and one section of the line can be rotated freely relative to the other section without appreciable loss of energy. The inner conductor appears continuous at point Z because of the low capacitative reactance between the mating ends.

At its termination the outer conductor is short circuited at the cylindrical reflector J to provide mechanical support for the line (see Fig. 352). The dipole I I′ is mounted $\frac{\lambda}{4}$ from J.

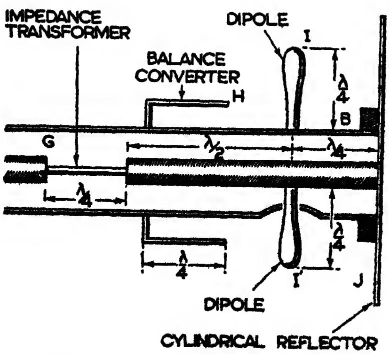


Fig. 352.—DETAIL OF AERIAL.

I′ is connected to the inner line *viâ* the aperture. A line balance converter H is used to isolate the earthed outer surface of the coaxial line from element I′.

The feed point impedance of the dipole is approximately 60 ohms. This impedance is reflected to transformer G by the $\frac{\lambda}{2}$ section of line connected to the dipole.

This transformer G is $\frac{\lambda}{4}$ section of line with changed characteristic impedance to match 50-ohm coaxial line to the 60-ohm aerial feed point.

T.R. switches are dealt with collectively in the next chapter.

General Characteristics of Wave-Guide Feeder Systems

The attenuation-frequency characteristic of a wave-guide is vastly superior to that of a transmission line, whilst a coaxial line will transmit 3,000 Mc/s waves with an attenuation of ·5 db. per metre; a wave-guide designed for a corresponding frequency shows an attentuation factor of only ·05 db. per metre.

Like a transmission line, a wave-guide presents an impedance to a source of electromagnetic energy which feeds it. An intrinsic property of a wave-guide of uniform cross-section, analogous to the surge impedance of a transmission line, is the wave impedance (Z_g). However, Z_g differs for different types and modes of waves propagated, *e.g.*, for E waves $Z_g = 120\pi A \frac{\lambda}{\lambda_g}$ ohms and for H waves $Z_g = 120\pi A \frac{\lambda_g}{\lambda}$ ohms, where λ = wavelength in free space, λ_g = wavelength within the guide and A is the area of the guide in sq. cms.

A resistive film of impedance Z_g placed across the guide will absorb all the travelling energy in a wave of the appropriate mode, and since no portion of the wave is reflected, no standing waves appear along the guide.

If the termination is of value other than Z_g, partial absorption and partial reflection of the wave takes place, and standing waves are generated along the guide. Similarly, if the load impedance is numerically equal to Z_g, but not a pure resistance, the reactive elements will introduce a phase difference resulting in reflection and standing waves along the guide.

Matching adjustments are made, as for transmission lines, to reduce standing waves and consequent loss of power to a minimum.

Short-circuit and open-circuit terminations set up standing waves in a guide by reflection. Also resonant lengths of guide can be adjusted to provide transformations similar to those obtained with quarter- and half-wave lines.

It is possible to design joints between different portions of a guide system where no mechanical connection is possible (*e.g.*, between a rotating and stationary section of a guide), in which the mechanically open circuit appears as an electrical short circuit to the wave. This is effected by the resonant properties of the region of space between the section of wave-guide.

The wave impedance of a wave-guide of non-uniform cross-section contains, in general, both resistive and reactive elements; thus obstructions of any kind, placed in the path of waves travelling in a guide of uniform cross-section, may introduce reactive terms in the impedance presented to the travelling waves. The nature of these reactances depends upon the geometry of the obstacles, together with their position and orientation in the guide.

The interposition of any obstacle within a wave-guide will, in general, necessitate the introduction of a *tuning* adjustment calculated to annul the reactive component presented by the obstacle and, therefore, to reduce the S.W.R.

In the design of practical systems of feeders, it is necessary to introduce bends in the guide. In general, the effect of a bend or discontinuity is to cause reflection due to the corner not looking like Z_g. By rounding off the corner suitably, the impedance can be made to look like Z_g and the wave passes round the corner without appreciable loss.

Wave-guide Aerial System

Examples of micro-wave aerial systems with wave-guide feeders are shown in Figs. 353, 354 and 355.

Fig. 353 shows a complete aerial system using both circular and rectangular wave-guides.

(1) The magnetron output is applied *viâ* coaxial line and small coupling loop to the wave-guide.

(2) The shorting plunger may be adjusted to cause a voltage maximum in the wave-guide $\frac{\lambda}{4}$ away, and propagation in the H_{01} mode.

(3) Maximum output is obtained by the employment of a matching stub to tune out the reactive elements.

(4) The wave-guide is excited in the H_{01} mode.*

* For transmission lines, the plane of polarisation of the waves at any point in the field is defined as the plane containing the E vector and the direction of propagation.

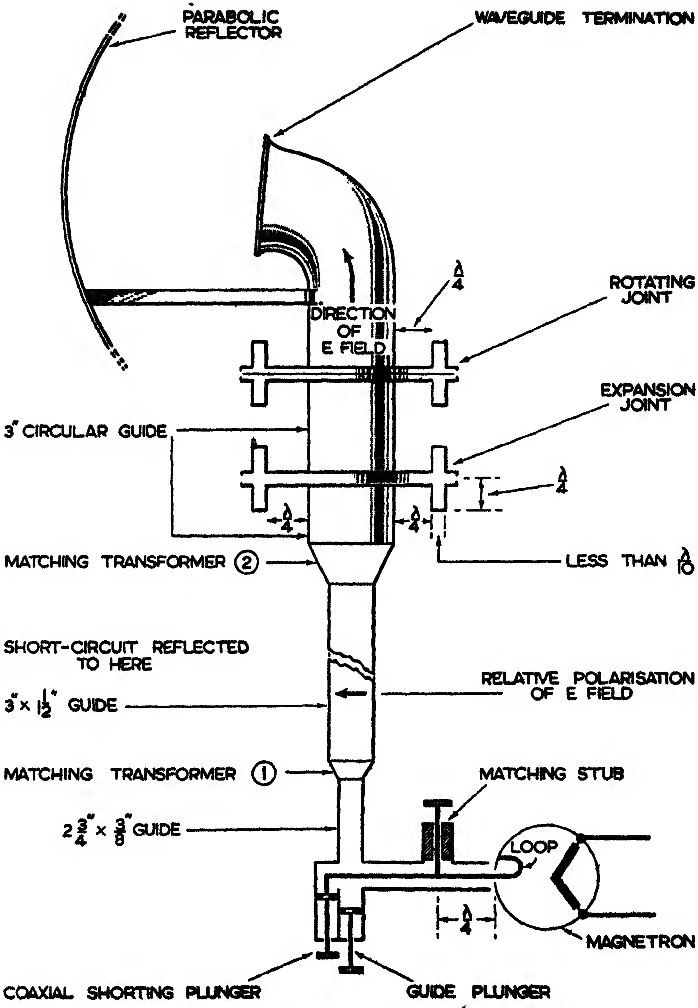


Fig. 353.—WAVE-GUIDE AERIAL SYSTEM.

Electromagnetic energy from the wave-guide termination floods the paraboloid. Output of the magnetron is coupled by coaxial line to the guide, and the three plungers can be adjusted for minimum standing wave ratio. Matching transformer (1) matches the impedance of 3 in. rectangular section to the initial 2¾ in. section. Transformer (2) matches the 3 in. rectangular section to the initial 3-in. circular section. H_{01} mode of the rectangular guide is converted to a symmetrical E_{01} mode in the fixed portion of the circular guide.

(5) The guide plunger also adjusts the standing wave ratio in the guide and so varies the coupling.

(6) Near the rotating joint the rectangular guide is changed to a 3-in. circular guide excited in the E_{01} mode.

(7) Note the expansion joint and choke. (For theory and details of the choke, see Appendix II.) The gap is variable and therefore the side slots are $\frac{\lambda}{4}$ deep and less than $\frac{\lambda}{10}$ in width, so that they form resonant circuits which present an open circuit at the outer edges of the flanges. This open circuit is reflected to the guide $\frac{\lambda}{4}$ away as a short circuit bridging the gap electrically. This short circuit is maintained by the side slots as the expansion joint opens or closes.

(8) The rotating joint is similar in principle to (7) except that the gap is fixed and the choke is simpler. Two flat flanges are attached either side of the break with fixed spacing of less than $\frac{\lambda}{10}$. The flanges form the equivalent of two sides of a $\frac{\lambda}{4}$ line open circuited at the outer end. The energy in the guide sees this high impedance reflected as a short circuit which confines the energy within the guide.

(9) The 3-in. guide continues to the aerial and directs the energy at a parabolic reflector.

(10) The electric field is transposed in the circular guide since it is propagated there in the E_{01} mode. Thus in passing through the rotating joint the energy undergoes a change in relative polarity (see Appendix II).

(11) If a circular guide is rotated 90° relative to fixed rectangular guide, the E and H vectors may be thought of as rotating with it. Consequently, at 90° there is no relative change in polarisation as determined by the direction of the E vector, but at other angular positions between 0° and 90°, polarisation varies between vertical and horizontal in the circular guide. This is one of the difficulties encountered at the circular joint. However, by transforming the H_{01} mode in the rectangular guide to an E_{01} mode in the fixed circular guide, a symmetrical pattern is obtained. This permits the circular guide above the joint to be rotated, through any angle relative to the fixed guide below the joint, without affecting the pattern in either one or the other. Since it is obviously

undesirable and inefficient to propagate two modes simultaneously in the circular guide, special precautions must be taken to eliminate the H and other undesirable modes from the fixed circular guide section, either by suppression using obstacles, or by dimensioning the guide so that they cannot exist.

Constant vertical polarisation can alternatively be maintained by using a coaxial line (short section) as the rotating joint.

In Fig. 354 the rectangular guide is terminated in a slightly larger section and shorted at the end. Two tuning plugs adjust for impedance match to the enlarged section.

The coaxial line has its inner conductor extended into the rectangular guide as a voltage probe and into the circular guide as one-half of the dipole fastened to it.

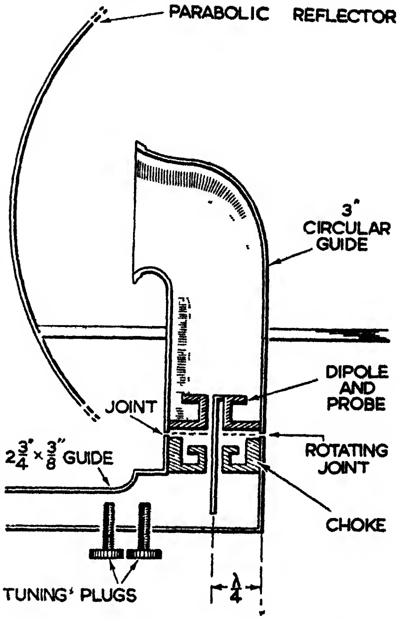


Fig 354—ROTATING JOINT IN WHICH A DIPOLE, COAXIAL CABLE AND PROBE ARE USED TO MAINTAIN UNIFORM POLARISATION DURING ROTATION.

The choke slot is bent back on itself and has a total length of $\frac{\lambda}{2}$, so that the short at the end of the slot is reflected as a short across the gap of the joint.

Polarisation is maintained for all angles by the position of the dipole in the guide, since the dipole moves with the upper guide and the probe is stationary relative to the axis.

Rectangular Guide System

Fig. 355 illustrates a system suitable for 3 cm.

(1) The magnetron is manufactured with the inner conductor of a coaxial line in place. This conductor is formed into a coupling loop within the magnetron and into an expanded voltage probe in the guide. The magnetron plugs into the side of a rectangular guide through a collar which forms the

outer conductor of the coaxial cable. The length of the inner conductor is made an odd number of $\frac{\lambda}{4}$.

(2) A tuning plunger in the end of the rectangular guide is adjusted for maximum energy transfer, *i.e.*, minimum S.W.R.

(3) The guide conveys energy to a cavity rotating joint where the mode of excitation is changed from H_{01} to E_{01}.

(4) The conversion is accomplished by means of a large-diameter tuning plug in the end of the cavity. This plug rotates the electric field through 90°.

(5) A second plug in the other end of the cavity couples it to the rectangular guide feeding the aerial.

(6) A metal plate is fastened along an equipotential line across the open end of the guide to support a parasitic dipole aerial.

(7) The metal plate has a second plate fastened to it at right angles, forming a reflector for the R.F. energy from the dipole

(8) A dipole is placed $\frac{\lambda}{4}$ in front of the reflector plate and is excited parasitically by the direct and the reflected fields issuing from the end of the wave-guide.

(9) The dipole now radiates energy at the parabolic mirror and this radiates the R.F. energy into space as a narrow beam.

(10) The main advantage of this system is that the feed enters from the back of the parabolic reflector and therefore large metal masses in front of the reflector are avoided.

The complete aerial system for the radiator shown in Fig. 326 (c), p. 489, is illustrated in Fig. 356. The technical specification of the associated equipment is :—

Wavelength	3·20 cm.
Pulse width	$\frac{3}{4}$ m. sec.
p.r.f.	500 c.p.s.
Peak power	40 kW.

In this instance the H_{01} mode is propagated in the rectangular guide and is changed to E_{01} in the circular guide at the rotating joint. The E_{01} mode in the circular guide is changed back to H_{01} in the final rectangular section, a transformer being used to match the termination or outlet.

Two stub wave transformers are used to change from H_1 mode to E_0 mode for the circular guide. Since the circular guide diameter is $1\frac{1}{8}$ in. at 3 cm., this guide would transmit two modes, *i.e.*, an H_1 mode (guide wavelength 4·24 cm.) and

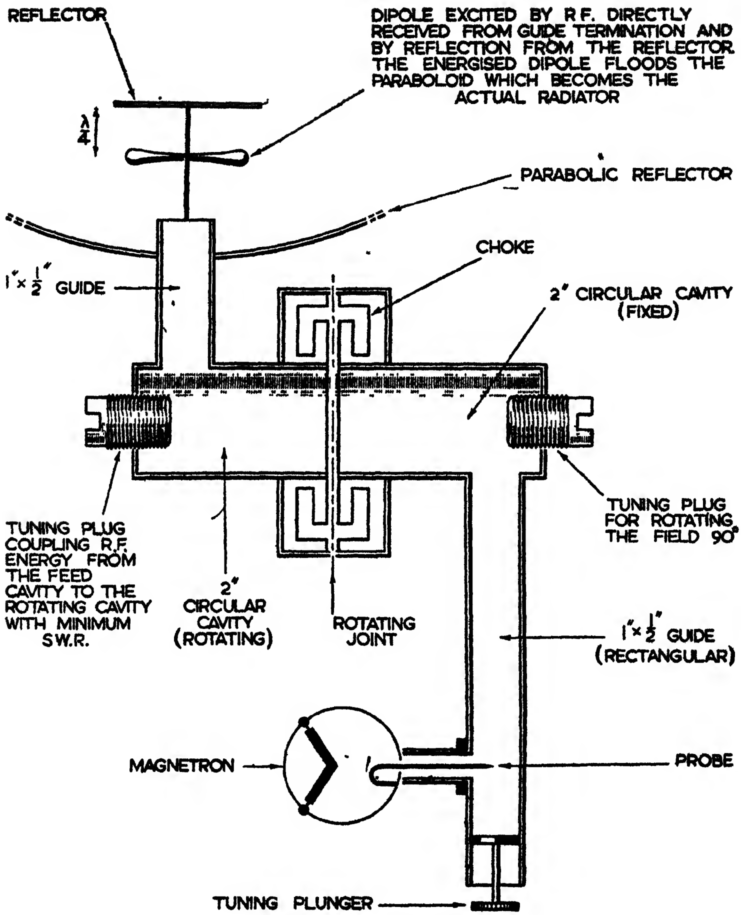


Fig. 355.—RECTANGULAR GUIDE AERIAL FEED EMPLOYING RESONANT CAVITIES AT THE ROTATING JOINT.

The H_{01} field of the rectangular guide is turned through 90° in the fixed cavity so that an E_{01} mode is produced there. This system also employs a parasitically-fed dipole.

an E_0 mode (guide wavelength 6·21 cm.). The undesirable H_1 mode is eliminated by proper adjustment of the stub lengths at the circular guide distant end. The stubs to mid-line of ends of the rectangular guides are $\frac{3\lambda}{4}$ of a guide wavelength for

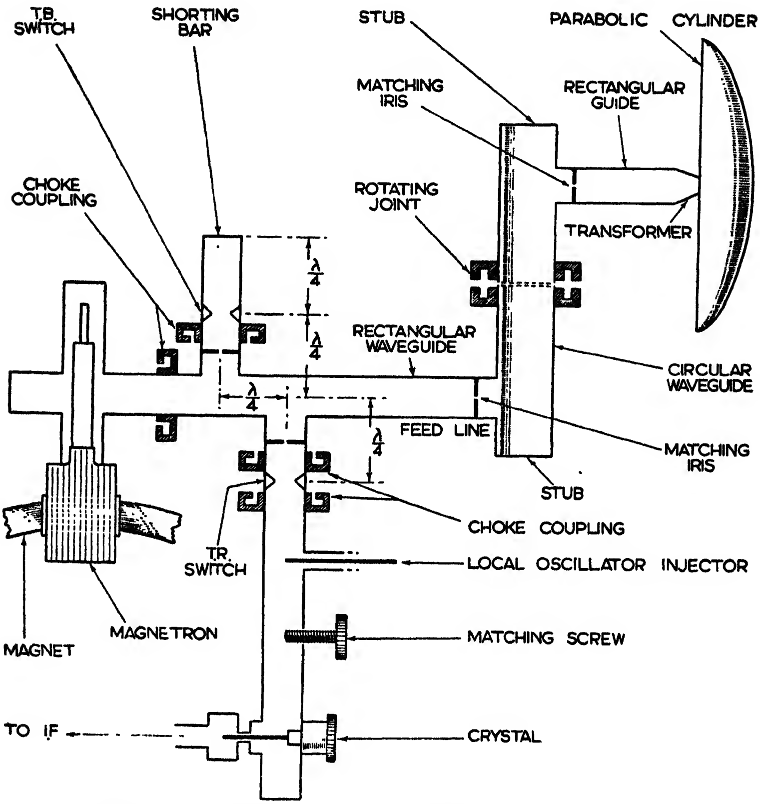


Fig. 356.—COMPLETE AERIAL SYSTEM FOR THE SET SHOWN IN FIG. 320 (a), EXAMPLE 3, pp. 469/470.

Note. The two distinct types of waves which may exist in a wave-guide are :—

(1) H waves in which the E vector exists at every point in a direction perpendicular to the direction of propagation, but has no axial component, whilst the H vector has components in directions both along and perpendicular to the direction of propagation.

(2) E waves in which the H vector has no axial component but the E vector has.

In guides of suitable dimensions, both E and H waves can exist in a doubly infinite series of modes. (The analogy with the singly infinite series of modes represented by a fundamental note and its harmonics may be mentioned, but it does not bear any very close comparison.)

In general, the higher the order of the wave mode, the higher is the " cut-off " or critical frequency.

H_1 mode, consequently a high impedance is presented to this mode. Also the overall length of the circular section is made an odd number of $\frac{\lambda}{4}$ guide wavelengths for the H_1 mode. If this length was resonant for the H_1 mode, an appreciable amount of undesirable H_1 mode would be set up in the circular guide.

The length of the stubs is approximately $\frac{\lambda}{2}$ wave-guide lengths for the E_0 mode and, therefore, the entrance to the circular guide appears as a low impedance point to the E_0 mode.

Inductive irises in the rectangular guide match out residual reflections and give a very low standing wave ratio.

The upper section of rectangular guide after leaving the circular section is twisted through 90° and terminated in a flared horn to which it is matched by inductive irises.

Inductive irises are also used to match the T.R. switch to the guide.

The matching screw probe between local oscillator and crystal is to prevent standing waves due to R.F. energy reflected by the crystal.

Matching is performed by varying the depth of penetration of the matching screw into the guide.

CHAPTER XXIV

RADAR AERIAL SYSTEMS—DUPLEXING SYSTEMS

T.R. AND T.B. SWITCHES

WHENEVER a single aerial is used for both transmitting and receiving (duplex), the problem arises of ensuring that maximum use is made of the available energy.

Selection of suitable switching arrangements is controlled by the following considerations :—

(*a*) Protection of the receiver input circuits.

(*b*) The mixer input circuit for micro-wave operation is subject to damage more easily than the comparatively sturdy R.F. amplifiers which form the input circuit at lower frequencies. Special protective measures must therefore be taken for micro-wave operation.

(*c*) Apart from protection of the receiver from actual damage, the T.R. switching must be sufficiently adequate to prevent (in conjunction with a gating wave), receiver paralysis.

(*d*) Efficiency requires that losses of transmitter energy in the receiver circuit or receiver energy in the transmitter circuit should be avoided as much as possible.

Principle of T.R. Switching

The principle of T.R. switching can be applied to wave-guides equally well as to open and coaxial transmission lines. In the former case $\frac{\lambda}{4}$ and $\frac{\lambda}{2}$ sections of wave-guides are substituted for similar sections of line and resonant cavities take the place of resonant $\frac{\lambda}{4}$ sections.

Fig. 357 serves to illustrate the principle and simplest form of T.R. switching.

The switching action is accomplished by the spark gap in conjunction with the $\frac{\lambda}{2}$ and $\frac{\lambda}{4}$ lines with which it is associated.

The spark gap itself makes a reasonably good switch, because it is an open circuit until the voltage across it rises sufficiently to break it down by arc over.

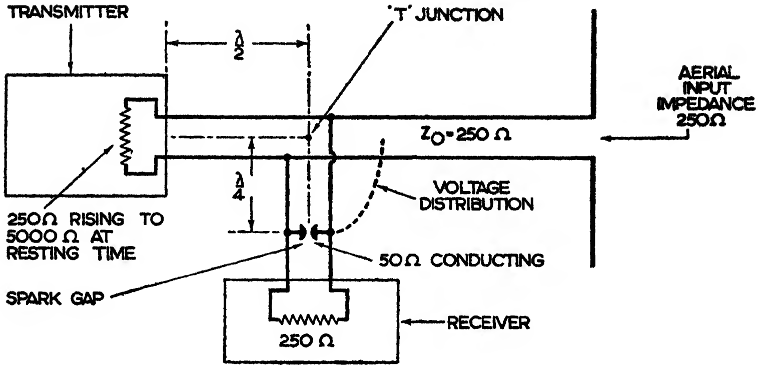


Fig 357.—OPEN WIRE T.R. SWITCH.

The principle is based upon the properties of $\frac{\lambda}{4}$ and $\frac{\lambda}{2}$ sections of transmitting line. A $\frac{\lambda}{4}$ section terminated in a high impedance at one end reflects a low impedance at the other end A $\frac{\lambda}{2}$ section reflects to the open end the impedance with which it is terminated at the closed end. (See Appendix I)

The arc is formed by ionisation of the gas or vapour between the electrodes. Once the process of ionisation has been started, it can be maintained by a very low voltage, and the resistance of the gap in this condition resembles a short circuit.

The ionised gap voltage is independent of the applied power, so that its resistance in the arc-over condition varies inversely with the power consumed in the gap.

Air at atmospheric pressure requires about 30,000 volts per inch of gap to start the arc, but the running voltage is about 50 volts. Both voltages depend upon the pressure and the gas or vapour used.

In order to simplify matters, it is assumed that the characteristic impedance of the transmission line, the feed-point resistance of the aerial, the input impedance of the receiver and the output impedance of the transmitter are all 250 ohms. It is also assumed that the transmitter output impedance during resting time *rises* to 5,000 ohms. The resistance of the gap conducting may be approximately 50 ohms.

(1) The transmitting pulse divides at the T junction, part goes to the aerial and part to the spark gap, which it breaks down.

(2) As a result, 50 ohms is placed across the 250-ohm line $\frac{\lambda}{4}$ from the T joint.

(3) As seen from the T junction, the $\frac{\lambda}{4}$ line terminated in 50 ohms appears to have an input impedance of $Z = \frac{250^2}{50} =$ 1,250 ohms.*

(4) Most of the transmitting energy now goes to the aerial, since it terminates the line in Z_0, *i.e.*, 250 ohms, as against the alternative path of 1,250 ohms impedance.

(5) In effect, just sufficient energy goes to the spark gap branch to keep the spark alive, the balance goes to the aerial.

At the end of the transmitted pulse the gap is de-ionised, and received signals arriving at the T junction see a low impedance path to the receiver, that path being terminated in Z_0. The path to the transmitter from the T junction which is a $\frac{\lambda}{2}$ section is terminated in 5,000 ohms during resting time and the received signals therefore see this high impedance repeated at the T junction.†

Note. If the transmitter is such that its impedance decreases below 250 ohms during resting time instead of increasing, the path from the T junction to the transmitter is then made a $\frac{\lambda}{4}$ length so that a low impedance at the transmitter reflects a high impedance at the T joint and effectively blocks any considerable loss of received signal.

The output impedance of the transmitter does not always change sufficiently for a simple resonant line to block efficiently received signals from the transmitter. In this case a second spark gap is added to block the transmitter branch at the T junction.

* For $\frac{\lambda}{4}$ sections $Z_0^2 = Z_a \times Z_b$ where Z_a = the impedance at one end of the line and Z_b = the impedance at the other end. Z_0 = characteristic impedance of the line.

$\therefore$ If Z_a represents the impedance of the spark gap load, the impedance at the T junction $\frac{\lambda}{4}$ away is $Z_b = \frac{Z_0^2}{Z_a}$. See Appendix I.

† A $\frac{\lambda}{2}$ line repeats at one end the impedance at the other termination.

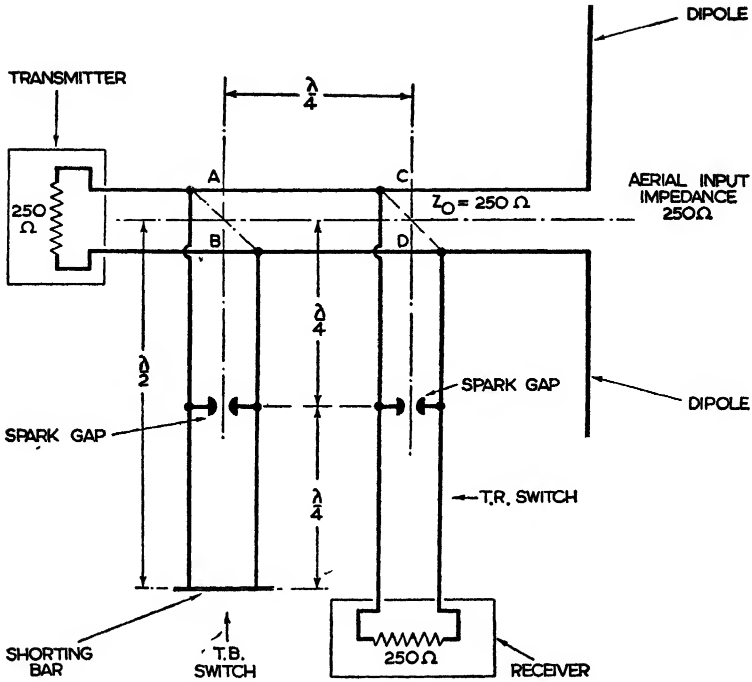


Fig. 358.—OPEN LINE T.R. SWITCH SUPPLEMENTED BY T.B. SWITCH.

For Transmission.—T.B. switch is in parallel with T.R. and has no effect on the general operation.

For Receiving.—T.B. line becomes a shorted $\frac{\lambda}{2}$ line reflecting a short at AB. The short at AB reflects a very high impedance across CD $\frac{\lambda}{4}$ away, thus blocking the transmitter path to received signals and offering a low impedance path to the receiver.

In this case a $\frac{\lambda}{2}$ section repeats a short circuit to the main feeder and this short-circuit is reflected a $\frac{\lambda}{4}$ length away as a high impedance.

The T.B. Switch

Fig. 358 repeats the T.R. system of Fig. 357 with the addition of a second switch which can be distinguished by the term T.B.

The T.B. switch works on a principle similar to that which has already been described.

(1) The transmitted pulse causes the T.B. switch to arc over as well as the T.R. switch and the outgoing signal sees a

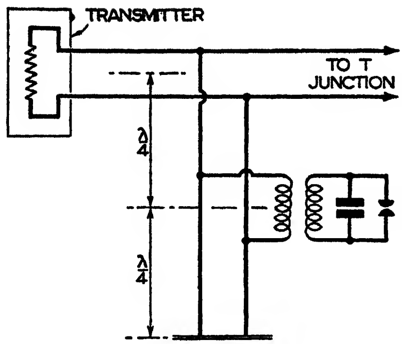


Fig 359.—T B SWITCH OPERATED BY TRANSFORMER ACTION

The equivalent circuit is shown here for explanatory purposes only.

high impedance at the entrance to the T.B. line as well as to the T.R. line. During transmission both T.B. and T.R. switches each use a small amount of the transmitted energy to keep their spark gaps alive.

(2) During the resting period the T.B. switch becomes a shorted $\frac{\lambda}{2}$ line due to the action of the shorting bar since the path across the spark gap is now an open circuit.

(3) The T.B. line therefore reflects a short circuit across the main feeder at AB.

(4) At $\frac{\lambda}{4}$ along the main feeder, from AB, *i.e.*, the T joint for the T.R. line, received signals see a high impedance looking towards the transmitter and a low impedance to the receiver path terminated as it is in Z_0.

(5) The function of the T.B. switch is therefore to create a high impedance across the transmitter path at the T.R. switch junction during the resting time of the transmitter.

Application of Transformer Action

Both T.B. and T.R. switches require some power to operate them, and this detracts from one of their functions, which is to increase efficiency. This loss can be greatly reduced by utilising transformer action, see Fig. 359.

Let the spark gap be placed in the secondary circuit of a transformer, the pulse signals being applied to the primary. Also let the secondary be tuned as a parallel resonance circuit to obtain a very high secondary resistance when the gap is not conducting. This high secondary resistance which is in effect equivalent to an open circuit, is reflected as a high impedance in the primary until the spark gap breaks down. When this occurs the low resistance shunted across the tuned secondary is transformed to a very low resistance in the primary. The advantages of this arrangement are :—

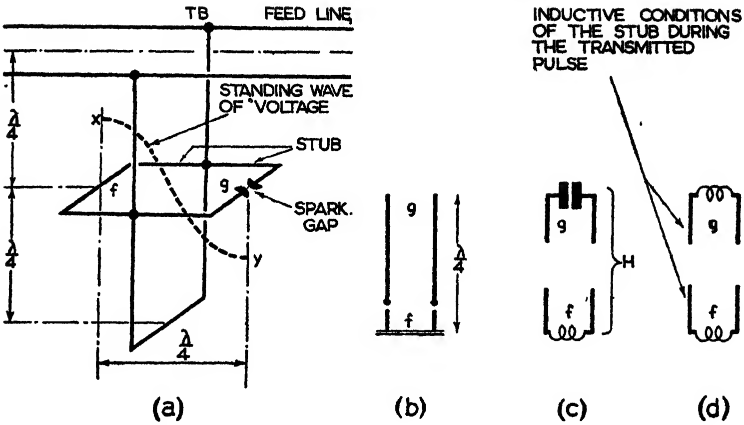


Fig. 360.—RESONANT LINE TRANSFORMER T.B.

xy = curve of voltage distribution across the parallel resonant circuit during transmission.

(b), (c) and (d) illustrate in steps the action of the stub, each half, g and f, being less than $\frac{\lambda}{4}$.

H shows the development of a parallel resonant circuit and (d) emphasises the existence of the inductive reaction due to the fact that f and g are respectively less than $\frac{\lambda}{4}$ and are, therefore, inductive at the resonant frequency.

(1) The voltage in the primary is stepped up in the secondary, causing the spark gap to break down at a relatively early stage.

(2) The low resistance of the conducting gap is placed across the tuned secondary.

(3) This value of resistance is stepped down in the primary to a still lower resistance.

(4) Assuming the gap resistance to be 50 ohms and a step-down ratio of 10 : 1, the primary has an impedance of 5 ohms only.

(5) This 5 ohms across the T.B. line is reflected by $\frac{\lambda}{4}$ to the main feeder as $\frac{250^2}{5} = 1{,}250$ ohms across the entrance to the receiver path and the power taken by the gap is only $\frac{1}{10}$th of that required when placed directly across the T.B. line.

Development of Resonant Line Transformer

R.F.-type transformers are not suitable at radar carrier frequencies, therefore resonant lines are used to produce the same result (see Fig. 360 (a)).

Fig. 360 (a) shows the application of $\frac{\lambda}{4}$ lines for this purpose.

The stub, Fig. 360 (b), can be considered as two sections of transmission line, one of which is terminated in a short circuit f and the other in an open circuit g. The switch line is less than $\frac{\lambda}{4}$, which gives it an input impedance that is inductive (c). The open line also less than $\frac{\lambda}{4}$ exhibits capacitative reaction (c). The two are in parallel, and since the total length is $\frac{\lambda}{4}$, their reactances are equal, thus forming a parallel resonance circuit of very high impedance (H). During the resting time, with the spark gap extinguished, the stub high impedance has little or no effect in bridging the T.B. line.

When energy is applied to the stub, a standing wave of voltage that is maximum across the gap, is set up along its length. The received signals are normally not large enough to break down the gap. The transmitted pulse, however, is large enough to do this, and in consequence places a low resistance across the open end of the stub by causing the gap to conduct.

The stub now consists of two lines in parallel across the T.B. line, each of which is inductive (d). The result of this and transformer action is to place a very low inductance across the T.B. line $\frac{\lambda}{4}$ from the feed line. This low impedance is then reflected as a very high impedance to the feed line. The arrangement is similar in its action to the equivalent circuit shown in Fig. 359 and limits the energy necessary to operate the gap.

Coaxial Line T.R. Switch

The effectiveness of a T.R. switch is dependent on the resistance of the gap as compared with the characteristic impedance of the transmission line system employed. Consequently, resonant transformers must be used with all coaxial lines, since the impedance of such lines is about 60 ohms as compared with a gap resistance of 30 to 50 ohms.

Fig. 361 (a) shows a simple form of T.R. system (coaxial).

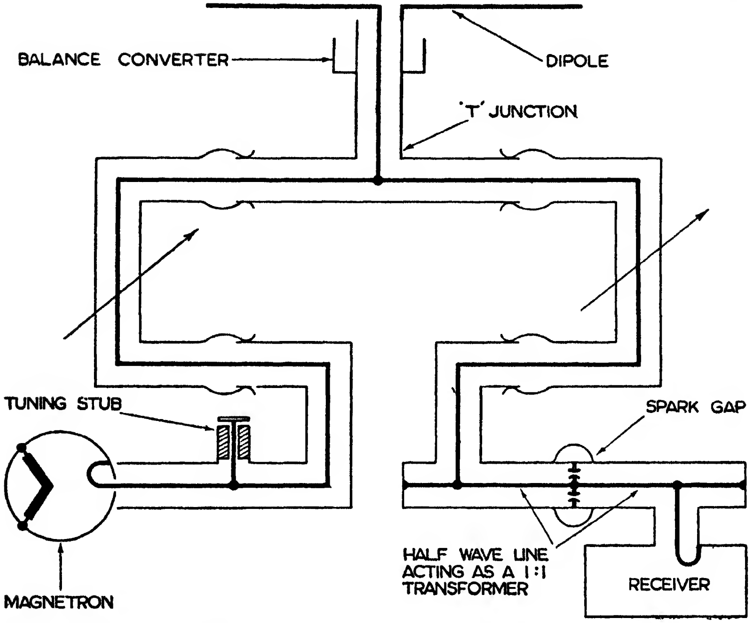


Fig. 361 (a).—COAXIAL T.R. SWITCH.

(1) The magnetron is matched to the line during transmission by the tuning stub.

(2) The length of the transmitter branch is such that the impedance seen at the T junction looking toward the magnetron is high when it is not generating.*

(3) In consequence of this received signals take the low impedance path to the receiver.

(4) The receiver feed line has a 1 : 1 transformer inserted (*i.e.*, $\frac{\lambda}{2}$ line shorted at both ends).

(5) This $\frac{\lambda}{2}$ line has similar properties to those of a tuned circuit. Its input impedance is zero at the shorted ends and maximum at the centre.

(6) The magnitude of this impedance depends upon the Q of the $\frac{\lambda}{2}$ line and the connected load (receiver).

* The input impedance to a magnetron changes from a high value during resting time to a much lower value during its period of operation.

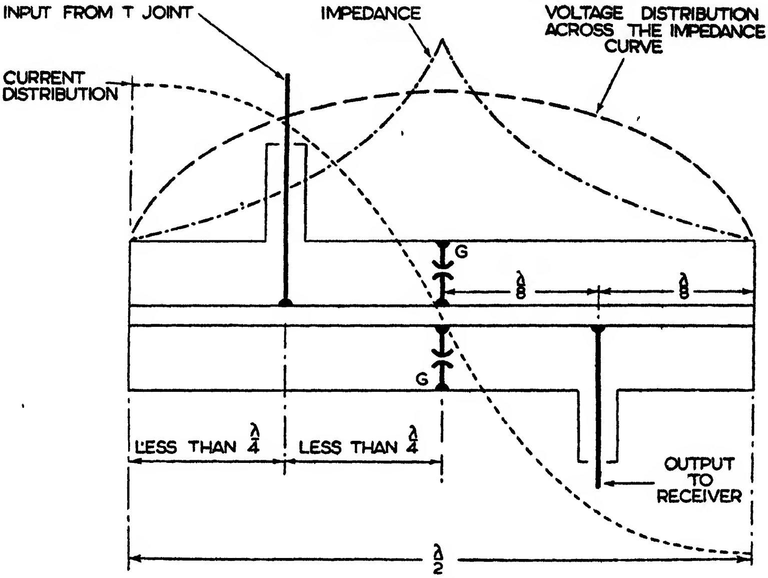


Fig. 361 (b).—DETAILS OF THE HALF-WAVE LINE T.R. SWITCH, WITH VOLTAGE, CURRENT AND IMPEDANCE DISTRIBUTION CURVES.

(7) The coaxial lines on either side of the $\frac{\lambda}{2}$ line are connected to it at points which match the coaxial line characteristic impedance.

(8) The receiver input circuit terminates the feed line in its proper impedance Z_0.

(9) The receiver branch of the system is matched throughout.

(10) The $\frac{\lambda}{2}$ line has spark gaps placed between inner and outer conductors at the middle of its length.

T.R. Action

The transmitted pulse passes to the T junction (Fig. 361 (a)) and causes a large voltage at the centre of the $\frac{\lambda}{2}$ line which breaks down the gap (Fig. 361 (b)). This is inherent in a shorted $\frac{\lambda}{2}$ line.

The lowered gap resistance shorts the $\frac{\lambda}{2}$ line so that the input terminal of the $\frac{\lambda}{2}$ transformer is connected to two shorted lines each less than $\frac{\lambda}{4}$ long, and since a shorted line less than $\frac{\lambda}{4}$ looks from the open end like an inductance, and, acting on the principle outlined on page 532, a very high total impedance value is thus presented to the receiver branch. The receiver feed line branch is adjusted in length to present a low impedance at the $\frac{\lambda}{2}$ line when unshorted, as at resting time, and a high impedance at the T joint (transmitter branch).

The transformer action of the $\frac{\lambda}{2}$ line reduces the effective gap resistance, thereby lowering the power necessary to operate the T.R. switch and the fraction of the transmitted signal reaching the receiver can be reduced to some small unimportant value.

At micro-wave frequencies transformer coaxial lines are reduced to such small dimensions that they become, in effect, resonant cavities,* and are usually so constructed. The use of a cavity resonator (Appendix III) gives a higher Q and by the use of this device greatly improved overall operation and efficiency is obtained.

T.R. Spark Gaps

Spark gaps used in radar circuits vary from a simple gap formed by two electrodes placed across a transmission line to one which is enclosed in an evacuated glass envelope with special features to improve operation.

The operating requirements of the spark gap, if it is to be efficient, are :—

(*a*) High resistance until the arc is formed.

(*b*) Low resistance through the arc.

(*c*) Rapid extinction of the arc when the pulse ends, to reduce loss and permit signals from nearby targets to reach the receiver as soon after the end of each transmitted pulse as possible.

The simple gap in air usually has resistance values that are

* A resonant cavity may be regarded as a small section of shorted wave-guide across which a standing wave is developed due to resonance effects. It possesses properties similar to those of a parallel resonant circuit.

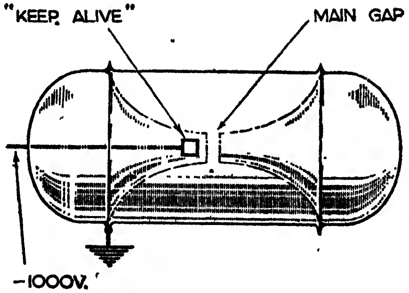


Fig. 362.—T.R. GAS GAP WITH "KEEP ALIVE."

too high for use with any but an open transmission line. Its air resistance is about 30–50 ohms when the arc has been formed. Also the de-ionisation time is about 10 microseconds, equivalent in recovery time to a target at about 16,000 yds., during which the gap acts as an increasing resistance across the transmission line.

Thus in a T.R. system using an air gap, the received signals reaching the receiver through the gap have only about half their proper magnitude after 3 microseconds (equivalent target range, 492 yds.). This delay is known as recovery time. Recovery time limits the minimum range at which echoes from nearby targets can be received. In the case cited the minimum range at which a target could be detected would be about 500 yards assuming a high signal to noise ratio and an absence of clutter.

The required voltage to break down a gap initially and the running voltage to maintain the arc can be lowered by reducing the pressure of the gas surrounding the electrodes. T.R. gaps are therefore used in which the spark gap is enclosed in a glass envelope which is partially evacuated.

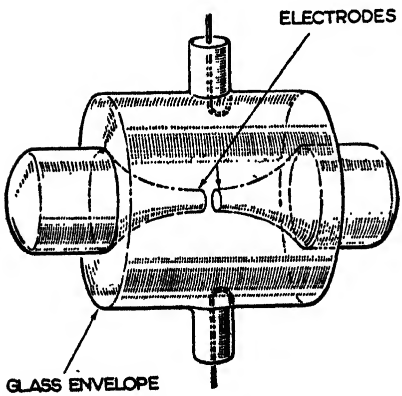


Fig. 363.—CAVITY T.R. GAS GAP WITH COAXIAL INPUT AND OUTPUT LINES AND PICK-UP LOOPS.

The arc is formed by conduction through a gas or vapour so that complete evacuation is not possible. Thus there is an optimum pressure for best results (Fig. 362).

The recovery or de-ionisation time can be reduced by introducing water vapour. At a pressure of 1 mm. of mercury, recovery time is reduced to ·5 microsecond (82 yds.), the lower limit of minimum range for this arrangement.

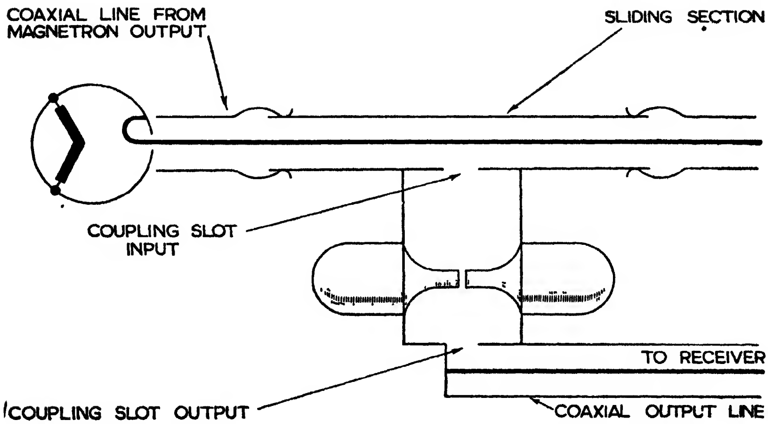


Fig. 364.—Slot-coupled cavity and coaxial line.

T.R. Gas Gaps

T.R. gas gaps at micro-wave frequencies are built to fit and be part of the resonant cavity or transformer. The high Q of the cavity and the presence of vapour reduce the power needed to maintain the gap and also the power of the transmitted pulse which reaches the receiver (see Figs. 362, 363).

The speed with which ionisation takes place can be increased by placing a third electrode within one of the main electrodes. This is known as the " keep alive," and has a potential of about 1,000 volts with respect to the main gap.

A glow discharge is maintained by the " keep alive " and one electrode of the main gap to provide a constant supply of ions and to form quickly an arc across the main gap when the transmitted pulse is applied.

The negative voltage of the " keep alive " tends to prevent stray ions from reaching the main gap and so to produce noises in the receiver.

The life of a T.R. gas gap is controlled by two main factors :

(1) The most common failure is due to build up of metal particles torn from the electrodes of the gap and spattered on the inside of the envelope. These particles act as small conducting areas, lower the Q and waste power. After a time these particles form a detuning wall within the cavity and prevent the device from functioning.

(2) The second cause of failure is due to adsorption of gas by the metal electrodes. This results in reducing the pressure

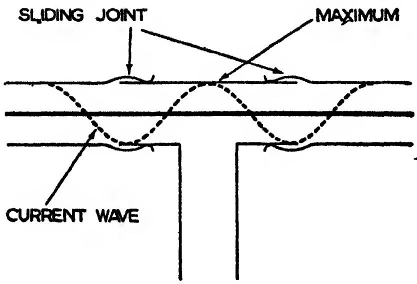


Fig. 365.—CURRENT WAVE REFLECTED BY MAGNETRON.

to a point where it becomes difficult to break down the gap in consequence, and extremely strong signals are therefore fed to the receiver.

For these reasons all T.R. gas gap units must be checked periodically for efficient operation.

Cavity T.R. Gas Gap

Gas gaps mounted in cavities may be excited with a transverse magnetic field and in the E_0 mode to produce a strong electric field across the gap, and therefore to cause an arc with minimum applied signal voltage.

Fig. 363 shows a method of excitation by terminating a coaxial feed line within the cavity, forming a coupling loop with the inner conductor. The coupling loop is a low impedance across the line so that current through the loop is large and a strong magnetic field is set up. The loop is positioned in the cavity so that its magnetic field reinforces the magnetic field of the cavity. The degree of coupling is controlled by rotating the loop.

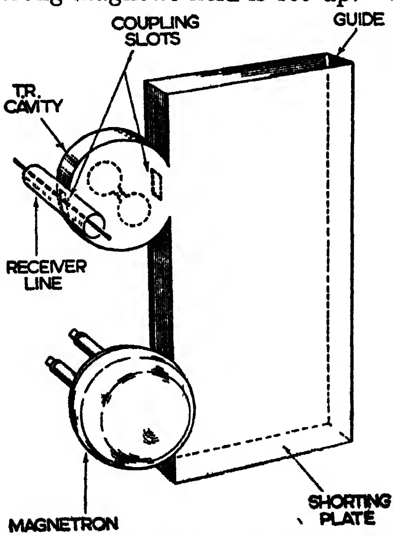


Fig. 366.—WAVE-GUIDE WITH CAVITY T.R. GAS GAP, SLOT-COUPLED TO A COAXIAL RECEIVER FEEDER.

Received signals which excite resonance conditions, but are insufficiently strong to break down the gap and detune the resonator, are removed from the cavity by a similar loop placed on the opposite side of the gap from the input loop.

Slot-Coupled Cavity

A second method of feeding the cavity from a coaxial line is to use slots which couple the field of

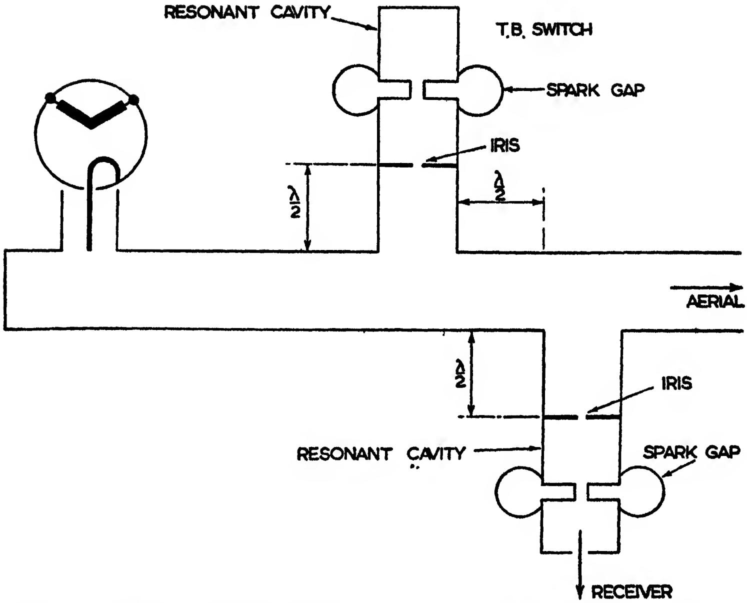


Fig. 367.—CAVITY RESONATORS APPLIED AS T.B. AND T.R. SWITCHES TO BRANCH LINES OF A WAVE-GUIDE.

the line to that of the cavity, see Fig. 364.

During transmission, energy is coupled into the cavity and produces a large voltage across the gap. The breakdown of the gap shorts the centre of the cavity. The resonant field built up in the cavity is then very weak and the effect is that of a short across the transmitting feed-line slot, so that very little energy is coupled into the receiver line.

At the end of the transmitter pulse the magnetron impedance changes, so that received signals are a mismatch at the magnetron and are reflected to the receiver branch. The position of the slot is adjusted by a sliding joint and must be set to absorb all the receiver line energy from the aerial feed line, to prevent standing waves between the T.R. switch and the aerial (Fig. 365).* In other words, in order to prevent a mismatch with the receiver line.

* The condition of resonance is to be avoided in feeder systems, since this indicates a mismatch and therefore a loss of power. The standing wave ratio or S.W.R. must therefore be measured in order to determine optimum matching conditions, *i.e.*, the condition in which the S.W.R. maximum/minimum approaches infinity or in other words the condition where no standing wave exists at all and resonance is entirely absent: $Z_{matching} = \sqrt{Z_{line} \times Z_{load}}$.

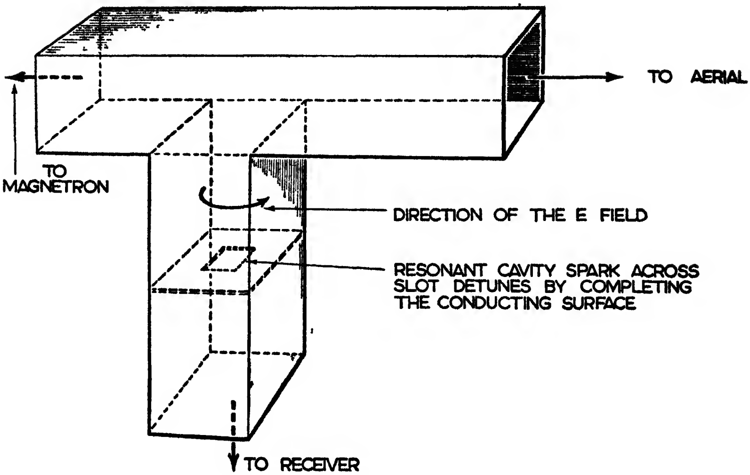


Fig. 368.—SLOT-TYPE T R SWITCH FOR USE IN WAVE-GUIDE FFEDER SYSTEMS

T.R. switches in the form of resonant cavities are also applied to wave-guides.

Wave-guide Using Cavity T.R. Gas Gap

In one method, which uses a coaxial line T.R. system, the wave-guide can be coupled to a coaxial line, but when large losses are likely to occur from the use of this method, resonant cavities are coupled directly to the wave-guide. Fig. 366 illustrates this arrangement.

The transmitted pulse of energy travelling up the guide spills into the cavity through a slot. The cavity builds up a strong electric field across the gap, breaking it down and detuning the cavity.

The impedance seen at the slot by the wave in the guide is decreased to an approximate short circuit, which effectively seals the opening and the pulse of R.F. energy passes on to the aerial without loss other than the small amount of electromagnetic energy which is absorbed to maintain resonance in the cavity.

Signals received during resting time are reflected from the magnetron and shorting plate. Since the slot is located at a maximum of the standing wave produced by reflection the cavity is suitably linked with the received signal energy. Received signals transferred to the cavity are not strong enough to break down the gap and so disrupt the resonant properties of the cavity; consequently they are coupled *viâ* the cavity resonator into the receiver coaxial line with maximum energy transfer.

Applied to Branch Lines

The cavity resonator can also be applied as an electromagnetic switch to the branches of wave-guide systems (see Fig. 367).

In order to ensure maximum efficiency, a T.B. switch (also a cavity resonator) is generally included.

During transmission time the energy travels down the wave-guide to the T.B. switch, where part is directed to the gas gap; An iris* is placed across the guide branch $\frac{\lambda}{2}$ from the main guide to pass the energy into the cavity. When the gap breaks down a short circuit appears across the iris. This short is reflected back to the main guide $\frac{\lambda}{2}$ away to close the mouth of the T.B. branch. Most of the energy is therefore directed to the aerial.

On reaching the T.R. branch a similar effect is produced by the T.R. switch. Both openings being effectively closed (short circuits acting as a wall), maximum energy is transferred to the aerial.

During resting time the T.B. gap is not broken down by the received signals but the open circuit created in this manner is reflected to the main guide, and at $\frac{\lambda}{4}$, that is, at the transmitter end of the T.R. branch it is reflected again as a short circuit or wall across the main guide. The effect of this is to turn back received signals from the transmitter and T.B. branch and direct them into the T.R. branch, where they pass through the resonant cavity to the receiver in the manner described above.

Resonant Slots

Resonant slots which act as spark gaps are sometimes used instead of irises and spark gaps. The action of the resonant slot is dealt with in Appendix II.

* An iris may be regarded as an aperture similar in form, etc., to that which is generally used in conjunction with the lens system of a camera.

The dimension of the slot in the direction of the electric field is small enough for the transmitted pulse to cause an arc. The arc completes the conducting surface of the slot, providing a short circuit path with similar results to those produced by a spark gap and resonant cavity, Fig. 368.

Summary

The general principle of the electronic switching devices employed in connection with transmission lines, coaxial lines or wave-guides are common to all three. They are based on the use of a spark gap which either shorts a resonant section of $\frac{\lambda}{4}$ and reflects a high impedance, or it may short a $\frac{\lambda}{2}$ section and reflect a low impedance to the main guide at some point $\frac{\lambda}{4}$ away from the channel which it is desired to protect. In wave-guides, resonant cavities perform the function of creating an antinode of voltage across the gap during transmission, just as the stubs in Fig. 360 and the $\frac{\lambda}{2}$ line in Fig. 361 (a) and (b) perform this function for transmission lines and coaxial lines respectively.

CHAPTER XXV

RADAR AERIAL SYSTEMS—BEAM SWITCHING

In Chapter VII it was pointed out that the general shape of the major lobe, excepting very narrow beams, is such that the change in magnitude of the target signal, per degree change in azimuth of the aerial, is comparatively small in the region of the axis, *i.e.*, around the point of maximum signal strength (see Fig. 369).

In this region, the magnitude of the signal may change so little per degree change in azimuth that accuracy in setting the aerial to bear directly on target by change in signal strength cannot be ensured. The accuracy of angular measurement depends upon the sensitivity of the aerial to any change in the direction from which echo signals are received.

Beam switching is one solution of this problem.

Principle of Beam Switching

The method depends upon the principle illustrated in Fig. 370. If two aerials A and B are placed so that when energised alternately their transmitting patterns intersect at some point less than the maximum signal point (*i.e.*, approximately 85 per cent. of the maximum) the receiving patterns will be similarly related. Thus the signals received at A and B from the target T will be equal in magnitude, *but only when T lies along the axis*

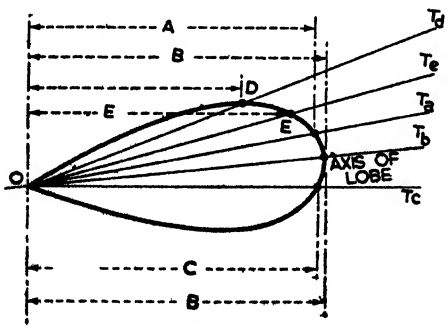


Fig. 369.—SIGNALS FROM TARGETS AT T_a, T_b AND T_c ARE PROPORTIONAL TO ABC.

The change per degree in azimuth T_b/T_a or T_b/T_c is quite appreciable, but the change in signal strength A/B or C/B is relatively small. The change of signal strength per degree azimuth in the region of T_d/T_e is very much greater, viz., D/E. In beam switching, therefore, the region about E (the ½ power point) is the part of the lobe employed for comparison purposes.

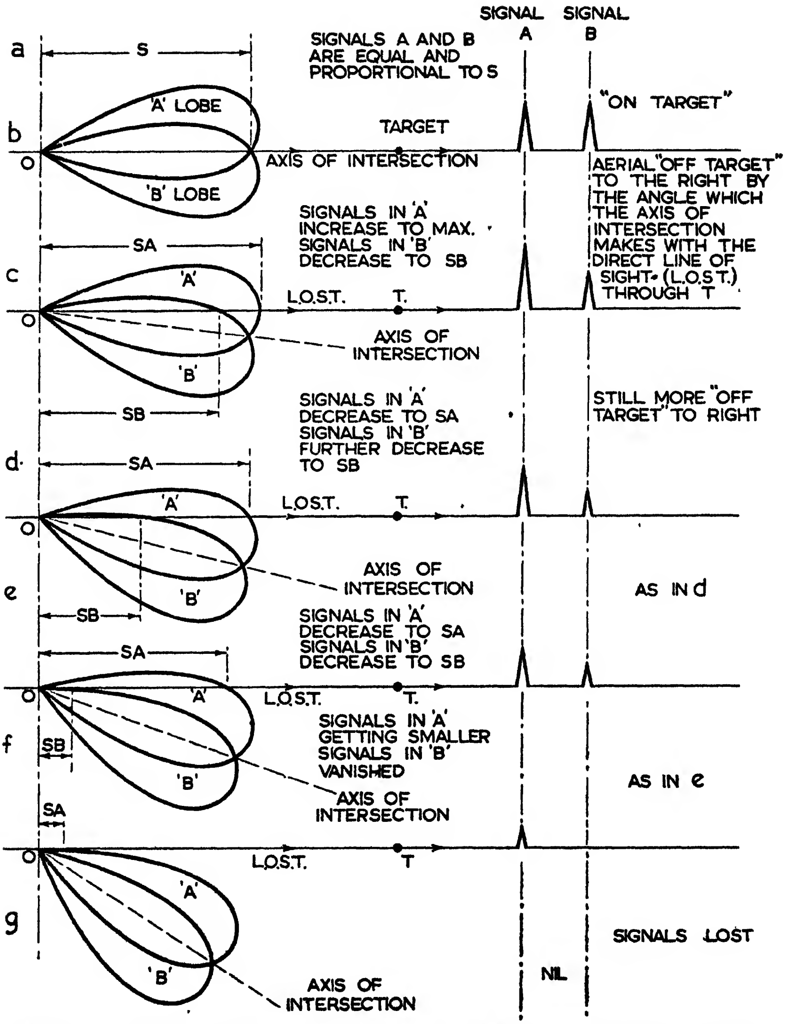


Fig 370.—SIGNAL STRENGTH VARIATIONS IN THE A AND B AERIALS AS THEY ARE SWUNG TOGETHER CLOCKWISE FROM "ON TARGET" POSITION (a) TO "OFF TARGET" POSITION (g).

If the aerials A and B are swung together anti-clockwise, the signal in the B aerial increases and A diminishes. In order to restore equality, *i.e.*, to bring the aerial "on target," the operator must rotate the aerial in a clockwise direction. O = observer, T = Target and L.O.S.T. = Line of sight.

of intersection. If T occupies any other position relative to the cross-over, the signals received at A and B will be ***unequal*** or will disappear altogether.

In using this method of aiming, the operator does not therefore adjust for a single maximum signal, but for equality of signals from the right- and left-hand patterns respectively.

This method is preferred because any change in position of the target in azimuth to either side of the cross-over, or dead-on position, must produce immediately an increased signal in one aerial and a decreased signal in the other aerial.

This differential action, which takes place in the immediate vicinity of the cross-over, permits very sharp discrimination and accurate adjustment.

Note. If the target moves to either side of the cross-over point (which is the same thing as saying " when the aerial swings the cross-over point to either side of the target "), the signals at first increase in that aerial which is swinging towards the target and decrease in the one which is swinging away from it. Let aerial A swing towards the target and aerial B swing away from it as would be the case if the aerials were rotated together in a clockwise direction for a fixed target.

Signals increase in aerial A and reach a maximum when the axis of the A lobe passes through the target. The signals in aerial B will by then have become small (see Fig. 370 (c)).

As the aerials continue to rotate, the A signals also begin to decrease and B will fade out altogether (Fig. 370 (d), (e) and (f)).

With continued rotation, the A signals will also fade out and the target is lost until the next revolution (Fig. 370 (g)).

Single aerials which utilise beam switching must include provision for alternately switching from one radiation pattern or lobe to the other, and some means for comparing the signals received from the lobes of each pattern in turn must be included in the display.

The simplest method would, of course, be to use separate aerials, receivers and a common indicator, but since a simple means of electronic pattern switching has been devised, it is more economical to use one aerial and one receiver to give the same results.

How Beam Switching is Carried Out

Taking a simple arrangement first in order to demonstrate the principle, if separate aerials and a single receiver were used the arrangement would be as shown in Fig. 371, S_1 and S_2

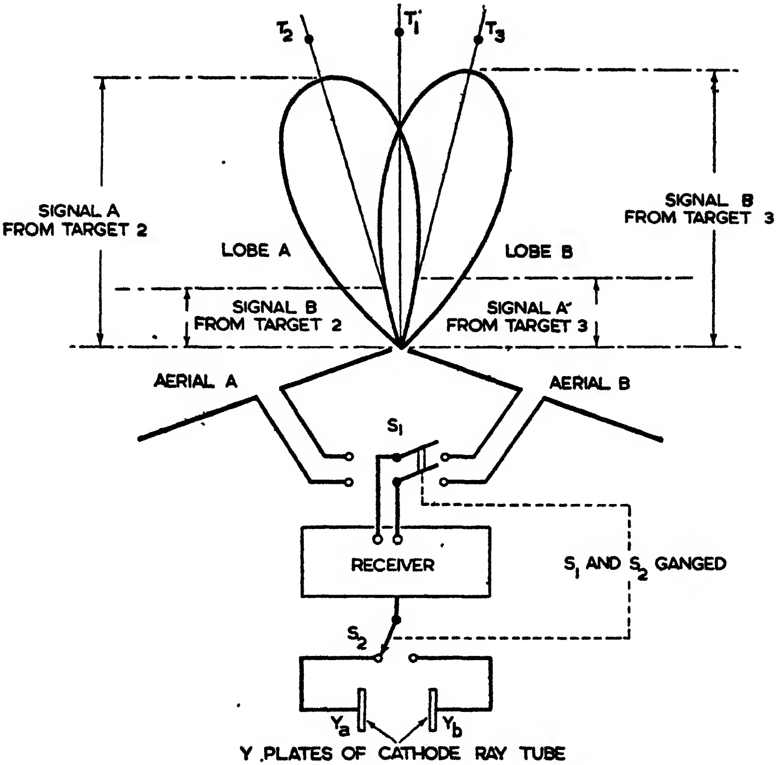


Fig. 371.—SCHEMATIC OF MANUAL SWITCHING SYSTEM USING TWO AERIALS AND ONE RECEIVER TO DEMONSTRATE THE PRINCIPLE OF AIMING BY THE BEAM SWITCHING METHOD.

With S_1 in either position, signals will be received from T_2 and T_3. In the A position, T_2 will give the largest signal in the A aerial; in the B position T_3 gives the largest signal in the B aerial. Equal signals will be received from T_1 at S_1 and they will be in phase.

being ganged so that the change over from aerial A to aerial B simultaneously changes the receiver inputs to plates Y_a and Y_b respectively of the cathode ray tube. The tube must have a screen exhibiting considerable persistence, and the change over should take place at a rate greater than the persistence of vision, so that the illusion is created of the picture from output A appearing on the screen at the same time as the output from aerial B, to enable comparison of the two signals to be made (see Fig. 372).

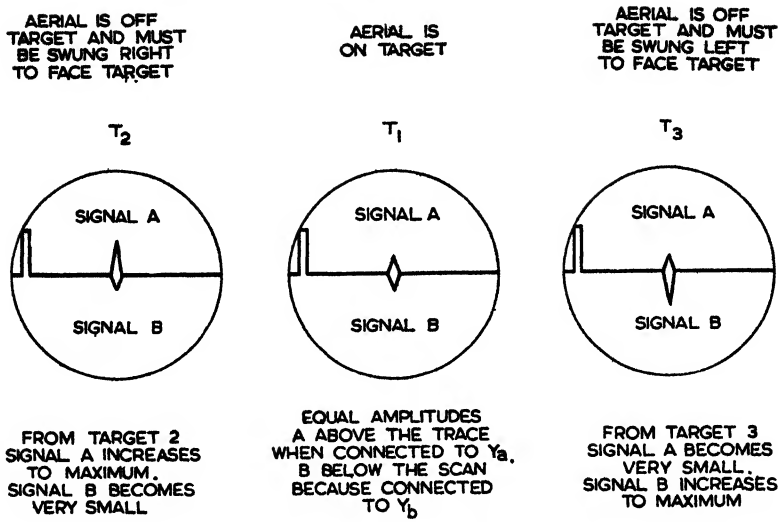


Fig. 372.—ILLUSTRATING THE DISPLAY FOR ONE FORM OF C.R.T. INPUT WHEN BEAM SWITCHING IS EMPLOYED.

To produce this effect, the aerials (Fig. 371) are supposed to be fixed and T_1 is supposed to move first to the left to T_2 and secondly to the right to T_3. Obviously this effect is the reverse of that obtained when the target is fixed and the aerials rotated about the line of sight anti-clockwise.

A single aerial can be used for all purposes by phasing one half of the aerial against the other half to produce a pattern whose axis is alternately deflected right and left.

In Fig. 373 an aerial four dipoles wide is used to receive echo signals and to measure azimuth by beam switching.

The aerial is divided into left and right halves, A and B respectively, connected by an external phasing section and feed line. The aerial is assumed to be fixed and T_1 moves right and left to T_2 and T_3 respectively, as in Fig. 373.

The signal from T_1 travels the same distance to A and to B, the induced voltages are therefore in phase.

The A and B voltages meet and combine over the phasing section. Since this junction is also equidistant from A and B, their voltages are still in phase and add to produce a resultant voltage input to the receiver.

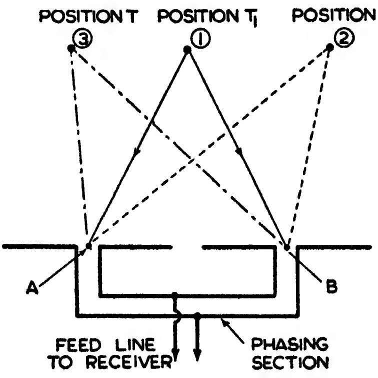
If the target is off centre to the right as at (2), the echo signal from T_2 reaches B before A. The maximum voltage occurs at A some time after it has occurred at B. The phase relationship at the junction is the same as at the dipoles, therefore the effective voltage input to the receiver is less from Target 2 than from Target 1 because of this phase difference between the A and B inputs. The target at T_3 will produce similar results to those from Target 2 but the relative phases of A and B will be interchanged.

Thus with the feed line connected to the centre of the phasing sections, the maximum signal is received only from position 1.

The difference in phase of signals received from positions other than (1) can be offset by varying the distance that each voltage travels before reaching the junction to the receiver feed line. In other words, the difference in phase between the voltages received at A and B when off target can be offset by a phase bias introduced into the phasing section in favour of the lagging voltage.

Phase Delay

In Fig. 374, let the target be located at T_2 with the receiver feed line connected to the left of the centre of the phasing section so that the voltage from B travels further over the phasing section to reach the junction than the voltage from A.



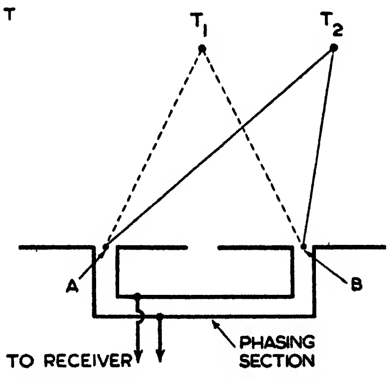


Fig. 373 (*left*).—EQUAL SIGNALS RECEIVED FROM T_1 WHEN THE FEED POINT IS AT THE ELECTRICAL CENTRE OF THE PHASING SYSTEM.

Differential signals are received at A and B from T_3 or T_2 respectively, and they are not in phase.

Fig. 374 (*right*).—EQUAL SIGNALS IN A AND B AERIALS FROM T_1, BUT THE COMBINED SIGNAL IS LESS THAN THAT FROM T_2 AT THE JUNCTION WHEN THE FEED POINT TO THE RECEIVER IS OFF CENTRE IN THE PHASING SECTION.

Maximum combined signal at the junction is now received from T_2, the phase difference in the A and B signals being offset by a bias in the phasing section. Since the maximum signal is now received from T_2, the effect of the phasing bias is to incline the receiving pattern to the right.

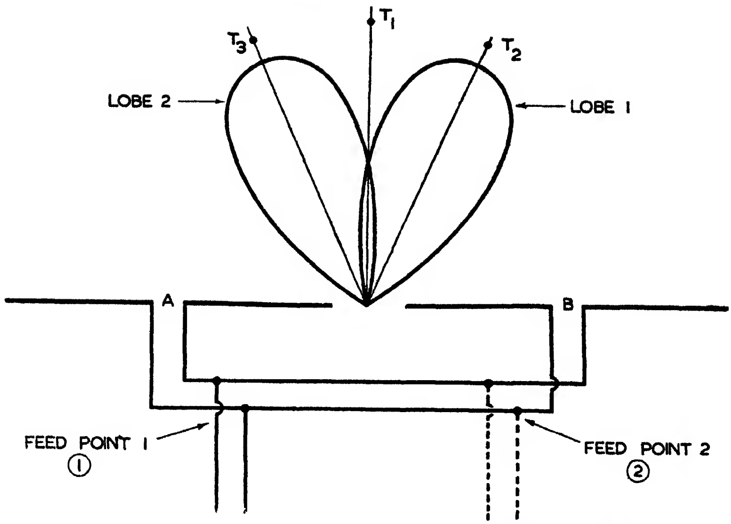


Fig. 375.—ILLUSTRATING THE CONDITIONS WHEN A PHASE DELAY IS INTRODUCED INTO THE RIGHT HALF OF THE PHASING SECTION (DOTTED LINE).

Right- and left-hand off-centre feed connections offset the differential and out-of-phase signals in the A and B aerials received from T_2 and T_3, to produce at the junction a maximum combined signal from targets at T_2 and T_3 respectively. Due to phase bias in both right- and left-hand sections, the equal A and B signals received from T_1 are out of phase at the junction, resulting in a combined signal which is less than those received from T_2 or T_3.

If the additional distance travelled by the voltage from B, over the phasing section, is equal to the additional distance travelled in space by the echo received at A, the A and B voltages are in phase at the receiver junction.

If with this biased phasing section the target is now moved to T_1 and the aerial remains fixed, the A and B signals received at their respective dipoles are in phase, but because of the bias in the phasing section, B lags A and the combined voltage at the receiver input is *not* a maximum.

(*a*) Thus with a balanced phasing section the signal input to the receiver is only maximum when the target is equidistant from A and B. (Fig. 373 (left).)

(*b*) With a biased phasing section the signal input is not at a maximum at the equidistant target position, but only at that point where the additional distance travelled in space cancels out the bias on the phasing section. (Fig 374.)

Fig. 375 illustrates the conditions when a phase delay is also introduced into the right half of the phasing section (dotted line). As in Fig. 374, the distance to aerial B from the full line feed point is greater than from aerial A by an amount less than $\frac{\lambda}{2}$. This phase delay, as we have seen in Fig. 374, produces a maximum combined signal when the target is at position T_2, which means that the receiving pattern must be a lobe as shown for lobe 1, Fig. 375. If now an additional feed line is connected (as shown dotted), to produce a similar phase delay in the A to B path, a mirror image of lobe 1 is obtained as shown by lobe 2, the reasons being similar to those already considered in connection with lobe 1.

When the delay path difference is greater than $\frac{\lambda}{2}$ it can be shown by similar reasoning that the lobes or patterns will be bent towards each respective feed point rather than away from it as in Fig. 375, *i.e.*, the cross-over point is moved along the axis of intersection.

From the foregoing it can be seen that beam switching can therefore be carried out by alternately switching the feed point for right- and left-hand bias along the phasing section.

Obviously the problem of shifting the feed point connection from one side of the phasing system to the other is difficult to solve mechanically, but it can be carried out electrically. For example, a phase difference can be introduced by switching a suitable reactance across one of the branch feed lines.

Reactance-phased Aerial

Fig. 376 shows an aerial fed by a coaxial line which branches symmetrically to each half of the array. A half-wave line is placed across each branch to act as a 1 : 1 transformer.

The inner conductor of each line is terminated in the small plates of condensers A and B. The other plate, C, is a half disc rotated by a motor so as to engage capacitatively with each half plate for approximately one-half of a revolution.

With C in the position shown, A is isolated and presents an open circuit across the coaxial line. This is reflected down the left branch line and has no effect. B and C form a small condenser which terminates the right coaxial line. This capacity is reflected to the right branch and is shunted across the line to *increase the effective length of the line.** The result

* See Appendix I, "Transmission Lines."

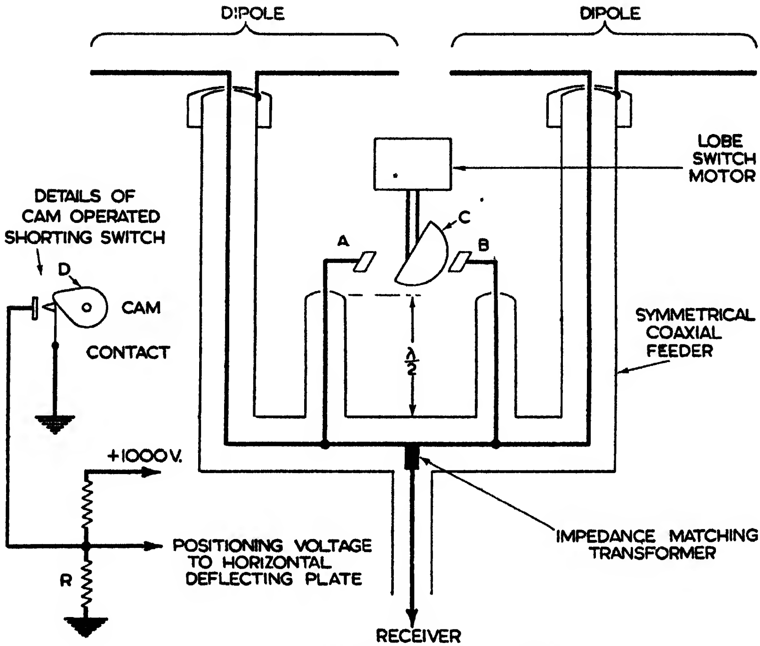


Fig. 376.—REACTANCE-PHASED AERIAL FOR BEAM SWITCHING

Note. The cam design controls the period for which the positioning voltage is shorted.

is to introduce a lag in the echo signals received by the right half of the aerial.

As the motor rotates this lag is placed alternately on the right and left branches and causes the lobe pattern to change as demonstrated by Fig. 375 and accompanying text.

The signals can be separated on the indicator screen by means of a variable position voltage controlled by a cam-operated switch. The cam is rotated by the motor at the same rate as the disc C. The cam is shaped to close during the time the lobe is bent to the left, and to leave it open when the lobe is to the right. When closed it shorts a resistance, so that the positioning voltage in the indicator is zero.

The output of the receiver, consisting of pulses from both lobes, is applied to the vertical deflecting plates of the C.R.T. The horizontal deflecting plates receive a sawtooth sweep voltage synchronised to the p.r.f., as well as the positioning voltage from the rotating switch D.

With C near plate B, the lobe is bent to the right and D is open. The positioning voltage is positive and causes the sweep trace to appear to the right of the centre of the indicator and to the left when switch D is closed. The positioning voltage is zero when the switch is closed and the sweep trace appears to the left of centre (Fig. 378). The separation or spread between the echo pips resulting from the receiver input can be adjusted by varying the positioning resistance R.

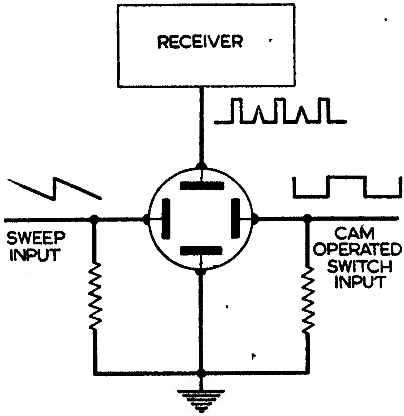


Fig. 377.—AMPLITUDE COMPARISON.
Schematic connections for separating and identifying for amplitude comparison, echoes received with phasing bias switched alternately to right- and left-hand sections.

The devices just described cause the receiving pattern to swing right and left alternately and provide the necessary means for a comparative display.

Another aerial system is available which produces both lobes simultaneously and which uses electronic switching. The main advantage is to eliminate moving switch contacts and noise, arising from dirt and incorrect adjustment.

Phasing is accomplished by an adjustable length line which connects the inner dipoles (Fig. 379).

The total voltage reaching the receiver is the vector sum of the voltages induced in A and B after they have travelled over the phasing system. Relative phase can therefore be controlled by varying the phasing section (Fig. 380).

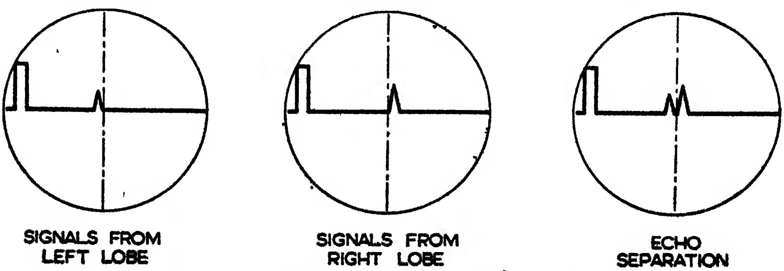


Fig. 378.—ECHO SEPARATION ON INDICATOR SCREEN.
Typical display employed for comparison of left and right switching.

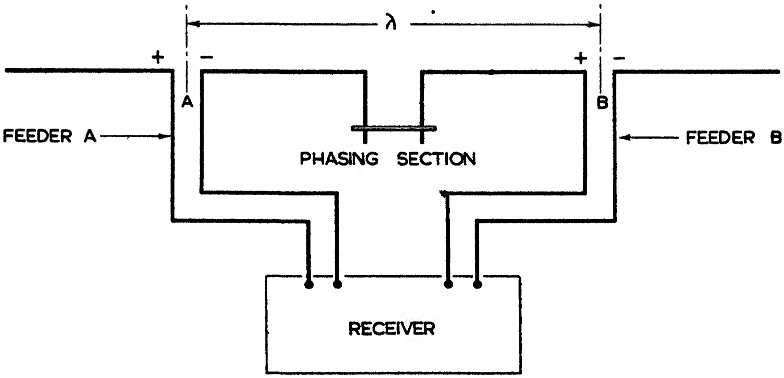


Fig. 379.—INTERNALLY-PHASED BEAM SWITCHING SYSTEM.

Signals received from a direction perpendicular to the face of the aerial induce in-phase voltages in A and B. Thus the induced voltage on the right feed point of A and the left feed point of B are 180° out of phase (see Fig. 380).

If the phasing section is zero length the induced voltage at B must travel one wavelength to reach A where it is 540° out of phase with the induced voltage at A and their algebraic sum is zero.

For a $\frac{\lambda}{2}$ phasing section, the total signal path length becomes $1\frac{1}{2}\lambda$ and the voltage from B arrives in phase with that at A combining to give a maximum total voltage. This means, of course, that when the phasing section is an odd multiple of $\frac{\lambda}{2}$ the phase difference is such that the maximum signal is produced in some direction other than straight ahead, *i.e.*, the axis of the lobe swings through some angle relative to the line of sight dead ahead.

Thus an echo arriving from a direction to the right induces a voltage in A which lags B (see Fig. 380). But the voltage induced in A also travels to B to produce a second voltage input to the receiver. Consequently the original lag produced by out-of-phase induction is still further aggravated by the delay which must occur in travelling from A to B over the connection between these two points. The total lag being 360° lag due to direct reception from the right-hand target plus 540° lag due to the length of path from A to B. This makes an A signal lag 900°.

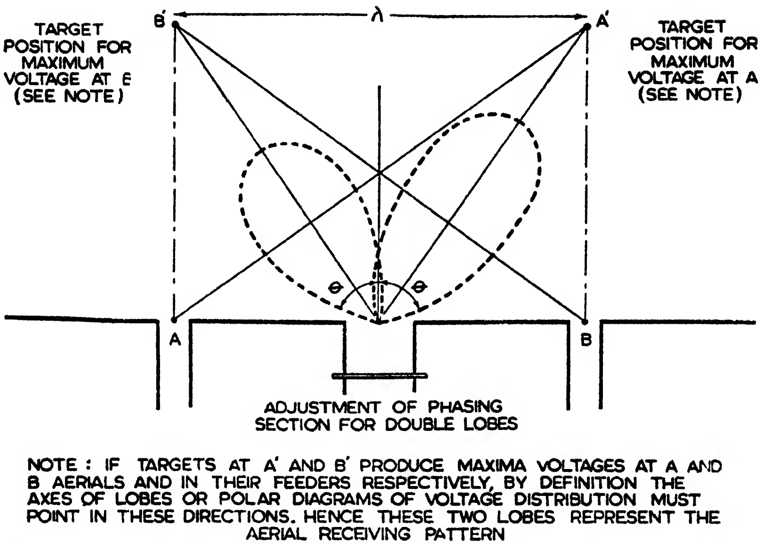


Fig. 380.—DEVELOPMENT OF RECEIVING PATTERN OR LOBES FOR THE INTERNALLY-PHASED BEAM SWITCHING SYSTEM OF FIG. 379.

Similarly, when an echo arriving from a direction to the right (see Fig. 380) induces a voltage in A which lags B, a voltage also travels from B to A and a correction can be made for this by introducing a delay in the path B–A sufficient to cancel out the delay of the echo received direct at A compared with the voltage reaching A from B.

But when induced voltages at A travel to B, this second delaying correction which must be made to the phasing section to bring B signals into phase with A (as detailed in the previous paragraph) increases still further the original delay or phase difference between the echo signals reaching B *viâ* A and the connecting path, and the echo signals which B receives directly from the target.

The only way to correct the phase difference between echo signals received directly at B and those received at B *viâ* A and the phasing system is to move the target to the left (see Fig. 380). This new position of the target then makes an angle to the left of the aerial centre, which is equal to the right-hand angular displacement of the target from the centre of the aerial. Thus a target at B′ produces a maximum signal voltage in the B aerial and the B feeder. Conversely, it can

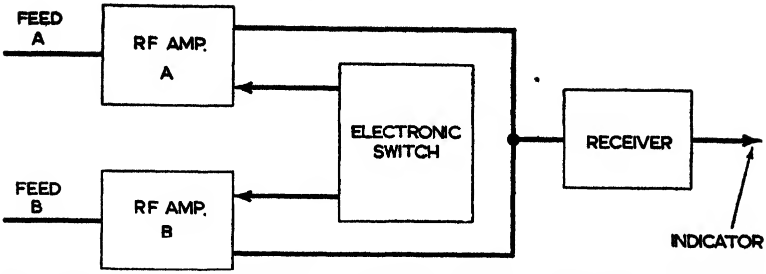


Fig. 381.—BLOCK DIAGRAM OF RECEIVER WITH ELECTRONIC SWITCHING DEVICE FOR USE IN CONJUNCTION WITH BEAM SWITCHING

This arrangement displays alternately the output from amplifier A and amplifier B for comparison of amplitudes on the indicator display.

be shown that a target at A′ produces similar results in the A aerial and the A feeder.

This means that the use of the internal phasing section can produce two lobes which are symmetrical about the perpendicular to the aerial.

If the total length of path from A to B is between 1 and $1\frac{1}{2}\lambda$, the lobe is bent away from the side of the aerial to which the feed line is attached. If the length of the path is between $1\frac{1}{2}$ and 2λ the lobe is bent towards the side to which the feed line is attached.

In order to make use of the two lobes and their signals which reach the receiver over separate feed lines, the receiver has two input channels, and an electronic switch selects them alternately. The outputs of the two channels are combined and applied to a conventional superheterodyne receiver (Fig. 381).

The electronic switch may produce two square-wave signals each at a frequency of, say, 1,000 c.p.s. according to the p.r.f., these signals being applied to the grids of the R.F. amplifier as blocking voltages.

Since they are in antiphase, R.F. amplifier A conducts when B is non-conducting. Every half cycle of the switching voltage this condition is reversed.

Thus the signal applied to the single-channel amplifier is alternately a signal from the right and left feed lines.

A Typical Switching System

Fig. 382 is the diagram of a typical switching system. V_1 and V_2 are duplicate amplifiers for feed-line signals A and B.

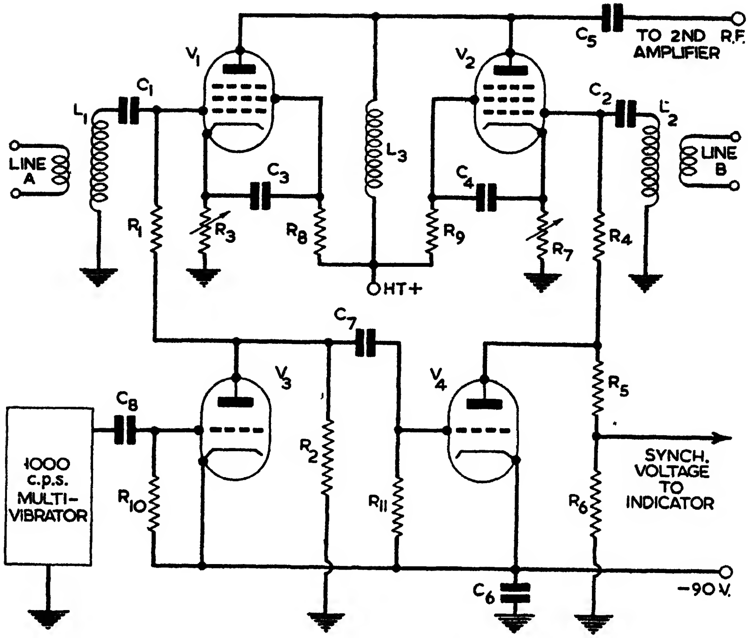


Fig. 382.—CIRCUIT DIAGRAM OF ELECTRONIC SWITCH FOR USE WITH INTERNAL BEAM SWITCHING SYSTEM.

Note cathode of V_3 is 90V. below earth, and anode at earth potential is 90V. positive to cathode.

The transformer is used on the input circuit to preserve a balanced feed-line system.

The output of the 1,000 c.p.s. multivibrator is an approximately square wave with amplitude sufficient to overdrive V_3. V_3 and V_4 improve the square waveform output from the multivibrator.

The anode voltage of V_3 and V_4 varies between earth and a negative value by the application of a negative supply to the cathode. Thus when the output of the multivibrator swings positive, V_3 conducts and anode current flows to earth *viâ* R_2. The anode is therefore negative relative to earth. This negative voltage change across the anode load is coupled to V_4, which cuts off. The anode of V_4 is now at earth potential. When the output of the multivibrator swings negative, V_3 is cut off and its anode rises to earth potential. V_4 now conducts and its anode voltage becomes very negative ; there-

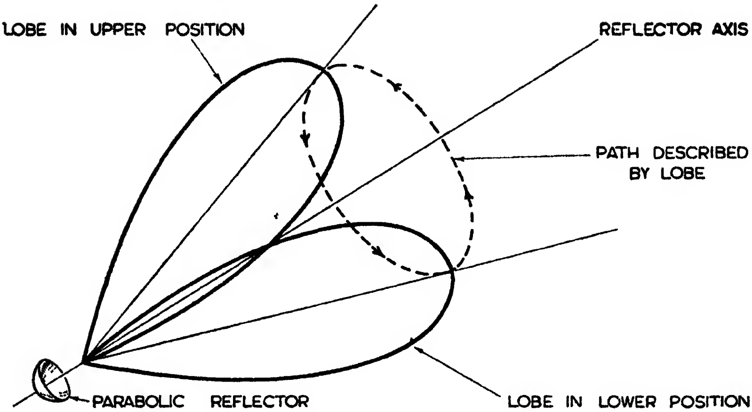


Fig. 383.—CONICAL SCANNING.

fore, when the anode of V_3 is at earth potential, the anode of V_4 is negative, and *vice versa*.

The grids of the R.F. amplifier are directly coupled to the anodes of the switch amplifier so that the anode potential of V_3 is a bias for V_1 and that of V_4 is a bias for V_2. V_1 operates as an amplifier with cathode bias when V_3 is cut off, and is cut off when V_3 conducts. V_2 operates in the same manner with V_4.

Since V_4 inverts the signal from V_3, V_1 and V_2 alternate in operation. The signal developed across L_3 is alternatively the signal from feed lines A and B.

The gains and anode currents of both amplifiers must be made equal, in order that the switching system may work properly. Otherwise the indicated equal signal position of the aerial will not correspond to the correct on-target position, and a square-wave modulation will appear in the output.

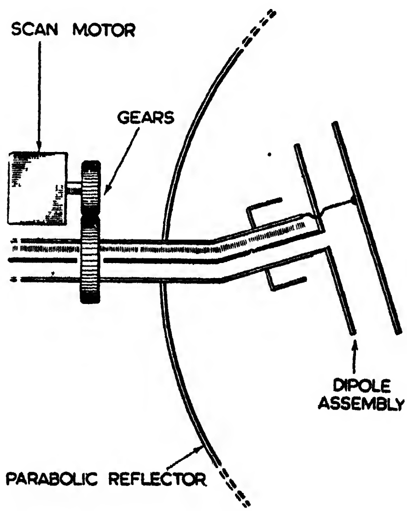


Fig. 384.—OFF-CENTRE DIPOLE FOR CONICAL SCANNING.

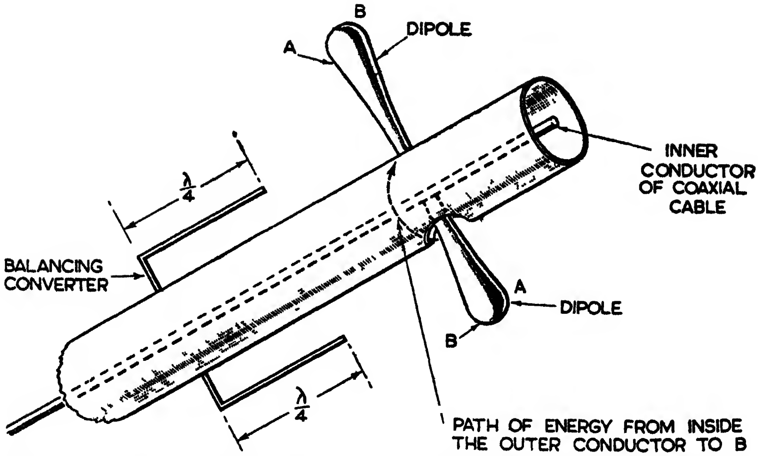


Fig. 385 —CONSTRUCTION OF MECHANICALLY-BALANCED DIPOLE FOR CONICAL SCANNING

This is done by balancing V_1 and V_2 with the individual gain controls R_8 and R_7.

The indicator requires a square wave to mix with the time base sweep voltage for spreading the echo pips. This signal must be synchronised to the switching action, and is obtained by taking the voltage drop across R_6 in the anode circuit of V_4. The receiver output is applied to the vertical deflecting plates of the indicator.

Conical Scanning

The principle of beam switching can be applied to give accurate azimuth and elevation, simultaneously, when applied to aerial systems using parabolic reflectors. This is conical scanning, and is achieved by the production of an off-centre lobe which is rotated about the axis of the reflector (see Fig. 383). The lobe axis describes a cone in space about the axis of the reflector.

The echo signal received from a target which lies on the axis of the reflector has the same amplitude for all positions of the lobe. If the target moves away from the reflector axis, the received signal varies sinusoidally (approximately), with the rotation of the lobe.

As the axis of the lobe nears the target the signal increases, and as the axis of the lobe moves away the signal decreases.

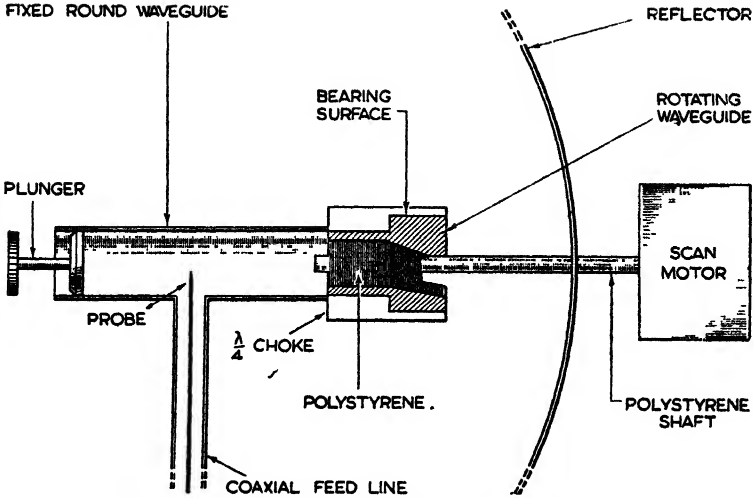


Fig. 386.—CONICAL SCANNING WITH WAVE-GUIDE FEEDER.

The *relative phase* of the signal variation therefore indicates the *direction* of the target from the reflector axis. The *magnitude* of the signal variation indicates the displacement of the target from the reflector axis. The circuits which supply the indicator can be used to indicate the relative position of the target, or to track the target automatically as well as to indicate position.

The simplest method is shown in Fig. 384. Here the coaxial line is bent away from the focal point of the reflector. Rotation may be at the rate of 20–60 r.p.m. The apparent source of energy is the electrical centre of the dipole assembly, since it is off centre with respect to the reflector, the lobe also will be off centre. The lobe rotates with the centre of the dipole. The dipole is not, however, balanced mechanically about the axis of rotation in this arrangement, and this is a disadvantage from the mechanical engineer's point of view.

Mechanically Balanced Arrangements

An alternative balanced arrangement is shown in Fig. 385. The path for energy flow to B is made longer by causing the energy to flow from the inside of the outer conductor through the hole and around the outside of the coaxial line to element B.

This manner of feeding, the choke, and the fact that the elements are differently shaped, gives an uneven current distribution. The effect of this is to cause the electrical centre of the dipole to move from the physical centre, and the energy is thus reflected from the paraboloid at a slight angle to the axis. When the element is rotated the point on the reflector at which the energy is directed describes a circle round the centre of the reflector and the re-radiated energy describes a cone.

The $\frac{\lambda}{4}$ balancing section mounted on the outer conductor of the coaxial line also prevents standing waves on the transmission line.

It is possible to produce a simple system of conical scanning for application to a round wave-guide which can be balanced mechanically (see Fig. 386).

The inner conductor extends into the guide to act as a coupling probe, and the plunger is used to adjust the degree of coupling. A "filled" round wave-guide is fitted into the other end of the main wave-guide.

The outer end of the filling is of the proper size to match the impedance of the fixed wave-guide to that of the rotating wave-guide.

The rotating wave-guide is bent off centre relative to the fixed wave-guide in order to produce a beam shift by supplying energy off centre to the parabolic reflector.

The conical scan is produced by driving the offset rotating wave-guide through a shaft made of plastic material coupled to the scan motor. A small hole in the outer end of the "filling" helps to match the rotating wave-guide to the paraboloid and to free space.

The system can be mechanically balanced by properly distributing the weight of the metal plug in which the hole is bored to form the rotating wave-guide.

Radiation through the rotating joint between the fixed wave-guide and the metal plug is prevented by a groove $\frac{\lambda}{4}$ deep which acts as an R.F. choke.

CHAPTER XXVI

RADAR AERIAL SYSTEMS—POSITION CONTROL AND INDICATION

In the operation of radar systems it is often necessary to have the angular position of a shaft follow accurately the motion of another shaft some distance away.

There are two major cases :—

(*a*) When it is desired to transmit data or information between two points, *i.e.*, the reading of a meter dial, the position of an aerial in azimuth or elevation, temperature readings, etc.

(*b*) When the controlled shaft must follow but also has to deliver much more power than the controlling shaft.

In case (*a*) the amount of energy and torque required is comparatively small, since the only requirement is that a light shaft at some remote position shall follow accurately a master or controlled shaft.

In case (*b*) conditions are different, *i.e.*, an aerial having considerable mass, and mounted on a substantial support, must follow accurately, despite windage and its own inertia, a light shaft revolved by a handwheel.

In effect, the problem in this case is to couple an electrically driven motor to the aerial shaft to supply the necessary torque, and at the same time to control the energy delivered by the motor to the aerial shaft, so that the torque always has the correct magnitude and direction for the aerial shaft to follow accurately the hand-driven shaft.

These functions are accomplished, in part, by use of a device which is basically a Magslip. A similar device frequently used is often referred to by its trade name as a "Selsyn." An alternative device which uses a different principle, but operates in a somewhat similar manner to that of a Magslip or Selsyn, is the "M"-type system.

All these devices perform their functions electrically. In case (*a*) angular mechanical motion is converted into electrical energy at one end of the system and is transmitted by wires from the point of origin to the remote position, where it is received and converted back to an angular mechanical motion, having variations in magnitude and direction which correspond to those at the transmitting end.

In case (*b*) the electrical state at the origin or transmitting end is made proportional to angular mechanical movement and compared with the electrical state at the remote or receiving end. The electrical difference, due to any change in relative position, which corresponds to mechanical difference, *i.e.*, the error, is then amplified and used to control the motor which rotates the aerial. This means that any movement of the handwheel changing the relative angular positions of the controlling and controlled shaft is immediately noted and compared by the remote Magslip or Selsyn, which thereupon generates a voltage proportionate to the difference or error.

This voltage is amplified and applied to the motor which rotates the aerial in the right sense, thus correcting the error and causing the aerial to follow the handwheel continuously. When the error has been rectified in this way the Selsyn ceases to generate, and the aerial comes to rest in a position corresponding to that of the handwheel.

Indicator and Rotating System

These two applications of the Magslip or Selsyn are illustrated in Fig. 387.

In Fig. 387, G_1 is a device which converts changes of position of the aerial shaft into proportional currents. These currents are transmitted by cable to M_1 which converts them back to shaft rotation. In this case the shaft rotates the light pointer of the indicator device. Consequently for every position of the aerial the direction of its beam in azimuth can be given by a dial reading at the operating position. Similar arrangements can be applied to elevation indication or any similar form of data transmission.

In the aerial rotating system, G_2 converts the position of the handwheel into an electrical indication which is transmitted to CT_1.

CT_1 compares instantaneously the relative positions of the aerial shaft with the position of the handwheel as conveyed over the connecting cable. As a result of this comparison, CT_1 develops a voltage proportional to the difference between the respective shaft positions, *i.e.*, an error voltage.

This error voltage is applied to a control amplifier which adjusts a power supply to a special motor with such polarity and magnitude that the resulting rotation reduces the error to zero.

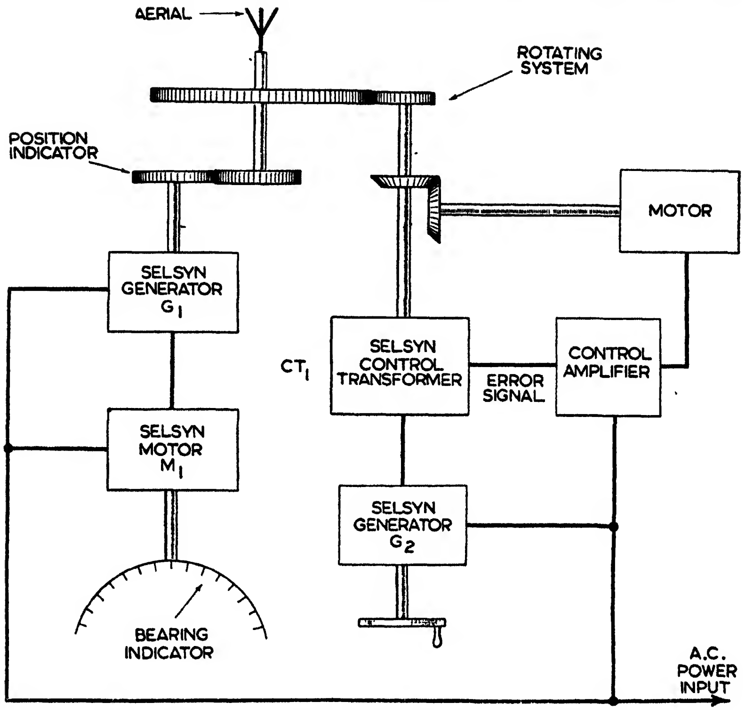


Fig. 387.—SERVO INDICATOR AND HANDWHEEL ROTATING SYSTEM.

When the aerial position corresponds to that of the hand-wheel, there is no error voltage, consequently the amplifier output falls to zero and the aerial comes to rest.

Briefly, therefore, the amplifier, under control of the Selsyn system, supplies variable power to the aerial-drive motor, which consequently causes the aerial to follow the handwheel.

The principal method of providing the electro-mechanical conversion is by means of the Magslip or Selsyn ; in simple cases a potentiometer sometimes suffices, in which case it is arranged to supply the correcting voltage.

When a target is to be examined the aerial must be trained accurately upon it. Normally the inertia of the moving system and associated electrical controls would militate against satisfactory operation by introducing hunting, causing thereby a tendency for the system to hunt about the line of sight.

In order to eliminate hunting and harmful mechanical effects which it would produce to the mechanical system, an anti-hunt device must be incorporated.

If zero over-travel is required, the drive must be reversed when approaching correspondence. Under transient conditions the motor armature carries either an accelerating or braking current. An early solution, together with a definition of hunting, can be found on p. 578. In terms of a power amplifier this is equivalent to negative feedback.

The Magslip or Selsyn

By definition a Magslip or Selsyn is a device used for the electrical transmission of an angular position.

This device takes the form of a small alternator. The rotor is turned by a shaft mounted on ball bearings and has a single coil of wire wound on an iron core which is laminated. The coil ends are brought out to a pair of slip rings.

The stator is fixed and may act as part of the frame. It is laminated and has uniformly placed slots into which are wound three separate windings. The coils are placed 120° apart round the stator and are distributed in several pairs of slots. The corresponding end of each coil may be connected to a common point, with the other ends brought out to terminals, or alternatively, the windings may be connected in series, each point of junction being brought to a terminal.

In other words, the device consists of a three-phase stator, connected in star, and a single-phase rotor.

If an A.C. input is applied to the rotor of the Selsyn shown in Fig. 388 and the rotor turned relative to the stator as shown, voltages are induced in the three-phase winding in the manner indicated in the diagram. Assuming the induced voltages in coils 1 and 2 is 45 volts (in the position shown for the rotor

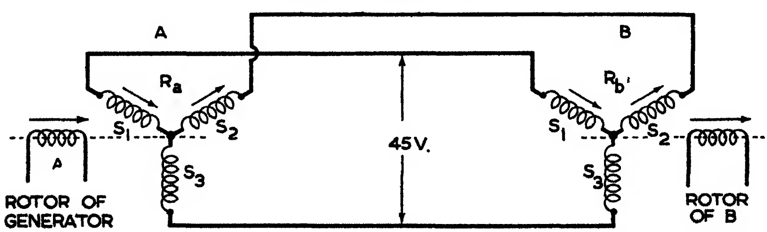


Fig. 388.—ACTION OF SELSYN SYSTEM. A AND B ROTORS BOTH IN THE HORIZONTAL POSITION.

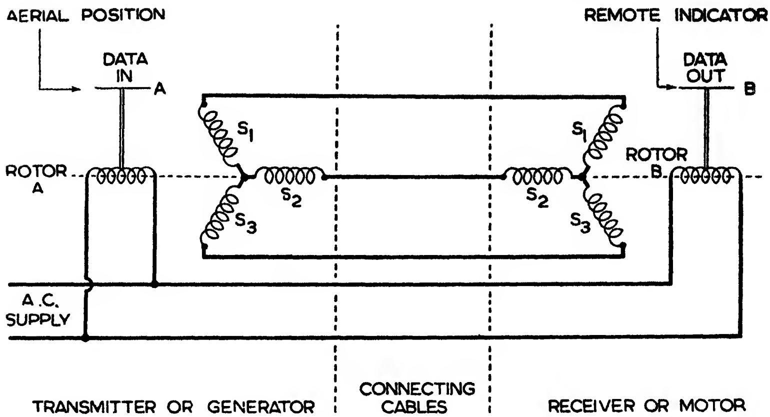


Fig. 389.—SIMPLE SELSYN SYSTEM.

these induced voltages are equal), the voltage distribution in stator A is:—

between S_1 and S_2 is 45 volts + 45 volts = 90 volts
" S_2 " S_3 " 45 " + 0 " = 45 "
" S_3 " S_1 " 0 " + 45 " = 45 "

By applying these line voltages to the proper terminals of a second Selsyn (B) the stator coils of B set up magnetic fields in directions as shown. The combination gives a field in the centre of the B stator which has the same direction, but, because of losses, of somewhat less magnitude, as the original field produced in stator A by the rotor of A.

Rotor B now sets itself along the axis of the magnetic field produced in the B stator by the currents from stator A, and thus reproduces exactly the angle taken up by rotor A.

Any change in position of rotor A causes a redistribution of the voltages in the A stator field and a similar redistribution of currents in the B stator, so that the B rotor follows the A rotor in every position.

The field induced in rotor B must always be in phase with the field of rotor A if the two Selsyns are to keep in step. This is accomplished by connecting the B rotor to the same supply service as rotor A.

Briefly then, if the rotors of A and B are energised by the same A.C. source, the induced currents in the A stator are repeated at the B stator, wherefore by Lenz's law the B rotor resets itself for equilibrium in the field of the B stator. The equilibrium position must be identical with that taken up by the A rotor.

Selsyn Generator and Selsyn Motor

Fig. 389 is a diagram showing Selsyns A and B connected in such a manner that any change in the position of rotor A is followed by a similar change for rotor B. It is essential with this type of Selsyn that frictional drag should be reduced to the absolute minimum. A Selsyn, designed for a freely turning rotor, would probably overheat if the rotor were held fixed, so that the angle between it and the stator field became permanently greater than 20°.

Selsyn A (the transmitter) is generally termed the generator, and Selsyn B (the receiver), the motor. Electrically, Selsyn A is identical with Selsyn B. Physically, they differ somewhat, since the motor has mounted on its shaft an oscillation damper. This consists of a lead ring, mounted within friction plates, in a sleeve secured to the rotor shaft. The ring has considerable inertia, consequently it exerts a damping effect on any tendency to oscillate since it is unable immediately to follow any change of state.

Note. A Selsyn does not necessarily rotate continuously, since it is not always desirable continuously to rotate the aerial. It may, therefore, only change its position a few degrees, as for instance when the operator is tracking some particular target. Voltages and currents in the wires of the three-phase system connecting the A and B stators have their maxima at the same time, but vary in magnitude according to the rotor position. Therefore the A.C. voltages and currents in Selsyn stators are single-phase.

In Fig. 387 the simple Selsyn system is used to transmit the aerial position. If the aerial is also capable of being tilted a second Selsyn system is employed to indicate elevation.

When very small angular changes are to be transmitted accurately, two Selsyn systems may be used. One directly coupled to the shaft from which position information is derived, and the other coupled through a gear ratio of, for example, 36 : 1, so that the second Selsyn generator makes 36 revolutions to one of the aerial shaft.

The aerial position is then given on two dials. The coarse reading is given by the first Selsyn system, and the fine reading by the second. The fine-reading pointer makes one complete revolution for each 10° angular movement of the aerial. The

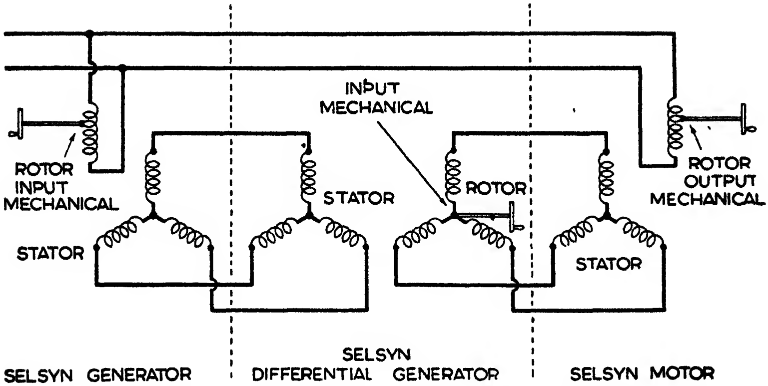


Fig. 390.—CONTROL USING DIFFERENTIAL SELSYN SYSTEM.

disadvantage is that the two Selsyns may be turned out of step when the power is off, so that the geared-up system may synchronise at some multiple of 10°, out of step with the parent system, when power is again applied

A combination coarse and fine system has been developed which is completely self-synchronising.

The Differential Selsyn

The differential Selsyn is similar to the basic Selsyn, except that the rotor has a three-phase winding instead of the single-phase winding of the basic Selsyn rotor.

The position of the rotor field flux with respect to the stator is now determined by two factors. The mechanical position of the rotor shaft and the *electrical* disposition of the field flux *within the rotor itself* (as determined by the three rotor input voltages). Like the basic Selsyn, the differential Selsyn may be a generator or a motor, the only difference being a damping device on the motor shaft. The differential Selsyn is used when the receiving Selsyn is to be controlled by two or more transmitting Selsyns.

Fig. 390 shows a control system which includes a differential Selsyn. The position of the rotors of the generator and of the differential generator are mechanically controlled by the connected apparatus and the algebraic sum of the resulting electrical outputs of these two sources determines the position of the rotor of the motor.

The stator of the differential generator is energised by the output from the stator of the Selsyn generator, so that the

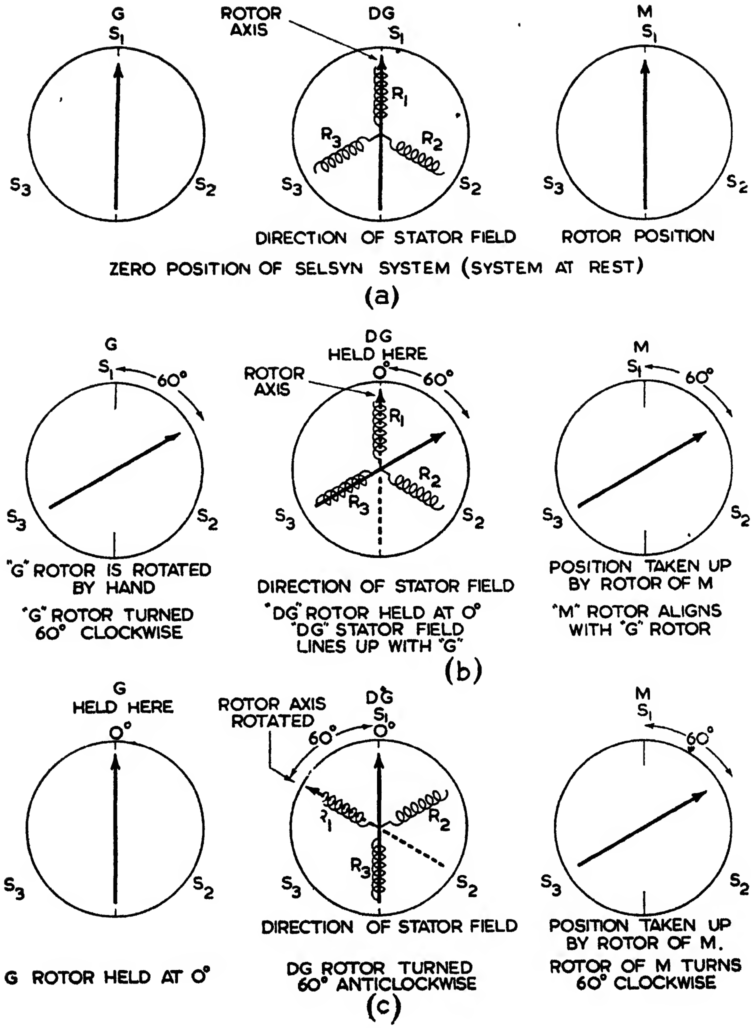
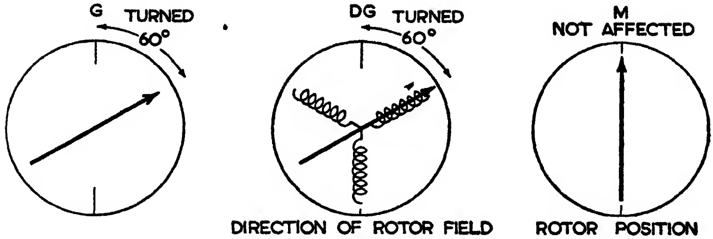


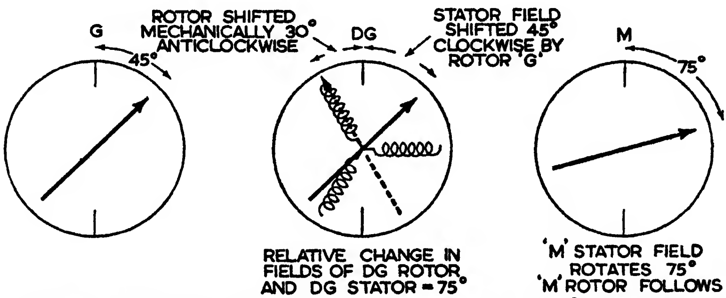
Fig. 391.—SKELETON DIAGRAMS TO SHOW THE ACTION OF THE DIFFERENTIAL SELSYN.

(a)–(c) demonstrate that the position taken up by the M rotor is determined by the angular displacement of the stator and rotor fields of DG. The diagrams (d)–(f), on the facing page, show that this displacement is determined by the final position of the rotors of G and DG respectively.



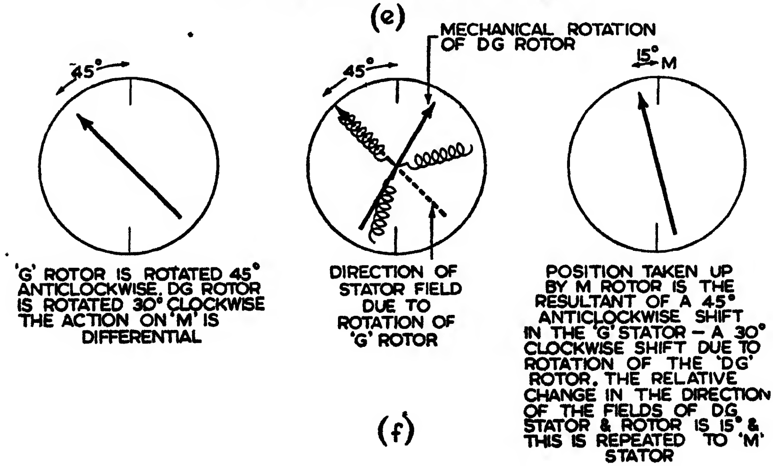
G ROTOR AND DG ROTOR BOTH TURNED 60° CLOCKWISE. 'G' ROTOR ROTATES THE FIELD OF 'G' STATOR & 'DG' STATOR. IF DG ROTOR IS ALSO ROTATED THERE IS NO RELATIVE CHANGE BETWEEN DG ROTOR & DG STATOR ∴ NO VOLTAGE INDUCED IN DG ROTOR AND NO CHANGE IN M STATOR

(d)



G ROTOR TURNED 45° CLOCKWISE. DG ROTOR TURNED 30° ANTICLOCKWISE. M STATOR FIELD RECEIVES 75° SHIFT.

(e)



(f)

axes of the stator fields of these two machines are always in the same direction as determined by the position of the generator rotor. The voltages induced in the *rotor* of the differential Selsyn determine the direction of the stator field of the motor and hence the equilibrium position of its rotor.

If the system is at rest and all rotors are free to turn, all the rotors will be at the same position, see Fig. 391 (a).

Let the rotor of generator G be turned 60° clockwise and let the rotor of DG be held fixed, then

(*a*) This change in position of the energised rotor of G causes the stator field of G to follow and the stator field of G is rotated 60°.

(*b*) By normal Selsyn action the field in the stator of DG which is electrically connected to G stator is also rotated 60°.

(*c*) Normally, the rotor of DG would follow the change of field of the DG stator since there is also a change in the field of the DG rotor but it is held, so that it is not free to move.

(*d*) Thus the currents in the stator of the Selsyn motor (connected electrically to the rotor of DG) will be identical with those in the DG stator and

(*e*) By normal Selsyn action the M rotor will take up a position similar to that which would have been taken up by the rotor of DG had it been free to move (*i.e.*, 60° clockwise rotation) (Fig. 391 (b)).

Obviously, the same effect can be produced by holding the rotor of G fixed and rotating the rotor of DG anti-clockwise. The relative change in position between rotor G, and rotor DG being the same in either case (Fig. 391 (c)).

If rotor G and rotor DG are both turned 60° clockwise (Fig. 391 (d)), there is no change in the voltages induced in rotor DG because there has been no relative shift between the axis of field of the DG rotor and the direction of its stator field. Consequently rotor M remains at zero and does not change position, no change having taken place in its stator field *viâ* the rotor of DG.

Thus the effect of the differential Selsyn in the control system is to cause rotor M to turn through an angle which is equal to the relative shift between the axis of the field of the DG rotor and the direction of the DG stator field.

To clear up this point finally :—

Let rotor G be turned 45° *clockwise* and rotor DG be turned 30° *anti-clockwise.* As a result the DG stator field is shifted 45° clockwise and since the rotor has been moved 30° anti-

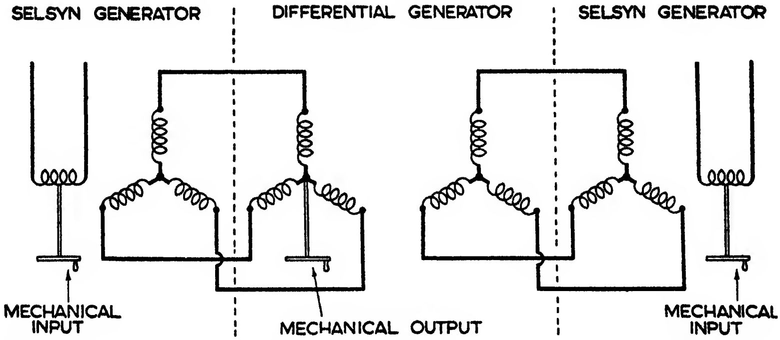


Fig. 392.—INDICATOR SYSTEM WITH TWO CONTROLLING SHAFTS AND ONE CONTROLLED SHAFT USING DIFFERENTIAL SELSYN MOTOR.

clockwise there is a relative total shift between DG rotor and DG stator of 75°. This causes rotor M to turn 75° clockwise (Fig. 391 (e)).

Now let rotor G be turned 45° anti-clockwise and let rotor DG be rotated 30° clockwise instead of 30° anti-clockwise as before. The fields of the two motors now subtract, for the reason that the rotor of G, acting electrically *viâ* DG, swings the rotor of M 45° clockwise, whilst the mechanical displacement of the DG rotor swings the M rotor 30° anti-clockwise (Fig. 391 (f)).

The *relative* shift of the DG stator field and rotor is now only 15° *anti-clockwise,* consequently rotor M rotates 15° anti-clockwise.

Note. The effect of the differential Selsyn in the circuit can be reversed by exciting the DG rotor from the Selsyn generator (G) in place of the stator. In this case the DG rotor must be turned anti-clockwise to neutralise an anti-clockwise motion of the G rotor.

Thus in example (1), in which G was rotated 45° anti-clockwise and DG 30° clockwise, the *relative* shift would for the modified connection be 15° anti-clockwise instead of 75° clockwise as before.

Application of the Differential Selsyn

A control or indicator system in which the output is taken from the shaft of the differential Selsyn may be arranged by replacing the motor in Fig. 390 with a generator and by making the differential a motor (Fig. 392).

The two end Selsyns have rotor shafts positioned mechani-

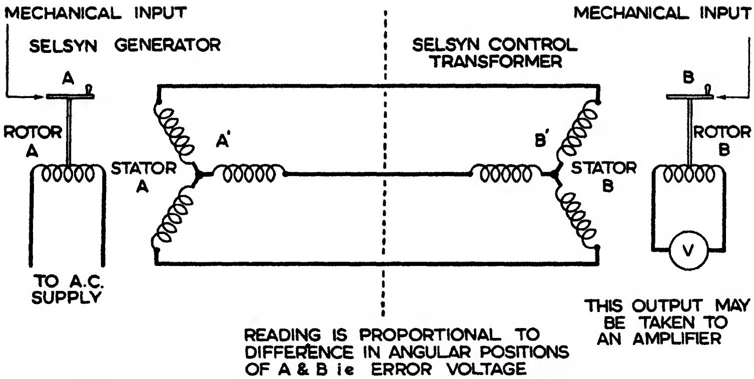


Fig. 393.—PRINCIPLE OF THE SELSYN CONTROL TRANSFORMER.

cally by other units The rotor shaft of the differential Selsyn now transmits the sum or difference of the angular positions of the other two rotor shafts relative to a fixed position or reference.

The differential Selsyns in Figs. 390 and 392 are supplied with power from a Selsyn generator. Thus Selsyn generators used with differential Selsyns are generally of heavier design to prevent overheating.

In the case of complex systems involving several differential Selsyns, condensers may be placed across the leads to the differential Selsyn rotor to counteract lagging currents. Where lines between Selsyns are long, this practice reduces voltage drop and power losses which otherwise may become appreciable as part of the total power consumed

Selsyn Control Transformer

Selsyn generators and motors have H section rotor cores. Selsyn control transformers have cylindrical rotor cores. Generator rotors and stators act as primaries and secondaries. When this device is used the rotor A of the transformer is held fixed in some manner. The rotor winding, B, instead of receiving an input from an A.C. line, supplies to an amplifier an alternating voltage, which is proportional to the angle between stator field A′ as determined by the position of rotor A and the axis of the rotor B of B′.

Thus a Selsyn acting as a transformer can be used to produce an error voltage for control purposes proportional to the angular displacement of rotor B relative to rotor A (Fig. 393).

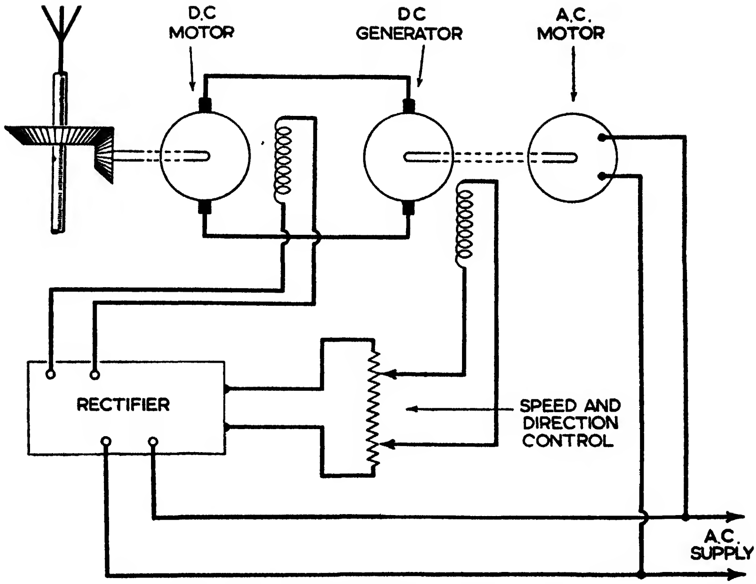


Fig. 394.—SIMPLE RADAR AERIAL DRIVE.

Motors and Driving Systems

It is frequently necessary to drive heavy aerial equipment at varying speeds, in either direction of rotation, and in most cases to maintain accurate control over their motion.

The ordinary single-phase or three-phase motor is essentially a constant speed device, therefore special D.C. motors must generally be used for this purpose.

The speed of a D.C. motor can be varied by armature voltage or by varying the voltage across the field coils. The first method gives the most stable results and the most desirable range of speed, but normally entails a large resistance in series with the armature and large power loss. The second method can be carried out by a physically small resistance and reduced power loss, but the speed range is limited from a certain minimum value upwards. At the higher speeds the field becomes so weak that speed is relatively unstable.

One alternative is to couple a D.C. motor to a D.C. generator running at constant speed. The field of the generator can be varied in magnitude and direction by means of a rheostat and reversing switch. In this way the motor armature is supplied with a smoothly variable input which can be varied from zero to full load value. The motor field is excited by a constant D.C. supply.

The advantage of this method is that by means of the variation of a small field current a smooth, flexible and stable control can be exercised over the speed and direction of rotation of a D.C. motor.

A simple drive for rotating, under control, a radar aerial, is shown in Fig. 394. The D.C. generator, direct-coupled to an A.C. motor, has its field connected to a potentiometer in such a way that the magnitude and polarity of the voltage applied to the armature of the D.C. motor can be varied, and the aerial can be rotated in either direction, and at any speed from 0 to maximum.

Aerial Control System

The basic system has a number of limiting factors. As a rule it is desired to control the direction in which the aerial points and the aerial may be required to rotate continuously for normal searching, or to be turned only a few degrees in azimuth to determine the bearing of a particular target.

Also the driving system must be capable of supplying sufficient power to make a large aerial rotate in step with the controls in spite of varying wind pressure.

Such a system is shown in Fig. 395. The field supply for the aerial driving motor and the D.C. generator is obtained from a small self-excited D.C. generator, driven by an A.C. motor.

The D.C. generator field supply from the exciter is regulated by a resistance bridge arrangement. When the bridge is balanced the voltage supply to the field of the D.C. generator is zero, and the aerial driving motor is stopped.

Any change in the resistance of the balancing arms results in field current to the D.C. generator, the resulting rotation depending upon the magnitude and direction of the change taking place in the electrical equilibrium of the bridge.

The shorting contacts are operated by a lever arm L geared to a differential Selsyn motor. Thus the speed and direction of rotation of the aerial are determined by the lever arm, which controls *viâ* the shorting contacts and the bridge, the magnitude and direction of the voltage impressed on the main D.C. generator field.

If the handwheel is turned when the aerial is in the stopped position, G_2 energises the stator of the differential Selsyn and

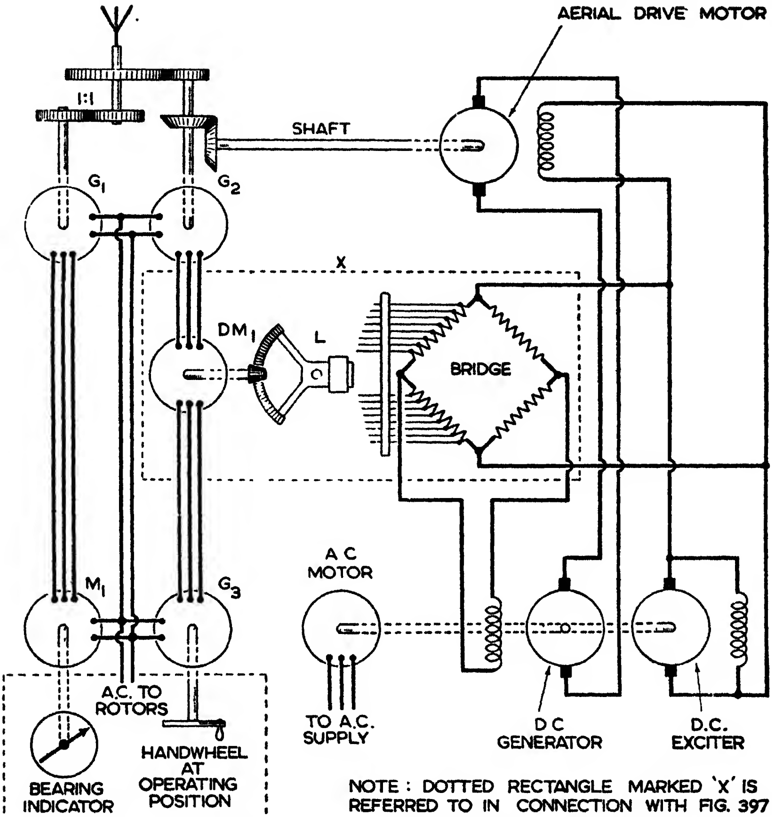


Fig. 395.—AERIAL CONTROL SYSTEM.

the rotor DM_1 turns, causing a deflection of the lever arm L. This deflection raises the shorting contacts momentarily to change to a new shorting position. The D.C. generator field is energised, and the drive motor turns the aerial in the same direction as the handwheel was operated.

Since the rotor of G_2 is geared to the aerial drive moto-, the rotation of this motor changes the output from the stator of G_2. This causes a rotation of the field of the stator of DM_1 which tends to bring the rotor of the machine to the neutral position.

Thus if the handwheel is turned only 10° the aerial turns 10° and stops, but if the handwheel is turned continuously the aerial will turn continuously, because the rotor of DM_1 is kept continuously out of neutral by some constant small angle. The effect of this small angle or error is relatively large if the handwheel is turned quickly, so that the aerial will also follow quickly. When the error is small the aerial turns slowly.

The system G_1M_1 forms a remote indicating system. If required, G_1 may feed signals to several Selsyn motors to indicate the aerial position at several points simultaneously.

When hunting* occurs, the voltages in the stator of G_2 cause the lever arm to be displaced from its neutral position, in a direction opposite to the displacement which causes the rotation, and the whole system will tend to oscillate about the desired position. This oscillation is not important to the operation of the radar set, but the resulting mechanical vibrations tend to cause breakdown of the mechanical system.

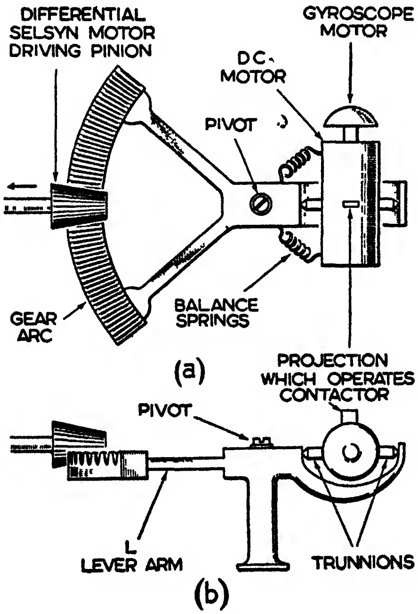


Fig. 396.—Gyro system for prevention of hunting.

(a) General arrangement.
(b) Details in elevation.

Gyro System

The undesirable effects of hunting can be obviated by the addition of a gyro mounted in trunnions on the lever arm (see Fig. 396 (a) and (b)). The position of the gyro relative to L is fixed in the static condition by a pair of balancing springs.

If the lever arm L is turned about its pivot, the gyro tends to precess or rotate about its axis through the trunnions, thereby eliminating hunting and increasing the speed of response of the controls.

For example, if the handwheel is turned suddenly, the rotor of DM_1 immediately tends to a

* Hunting, or search for equilibrium, is brought about by interaction between the mechanical and electrical momentums of the system. (See textbooks on A.C. motors.)

new position. The sudden motion of the lever arm unbalances the Wheatstone bridge, and the gyro at the same time precesses about its trunnions, causing a further motion of (a) in the same direction. Thus any change in control causes high acceleration.

As the aerial rotates and the error in the field position within DM_1 decreases, L turns back towards its neutral position. The gyro is now pulled to its normal position by its balancing springs, and then precesses in the opposite direction as the lever arm moves towards neutral.

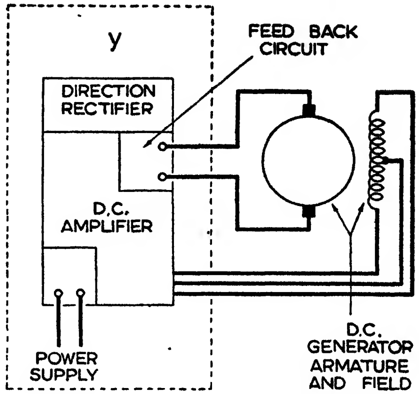


Fig. 397.—MODIFICATION TO FIG. 395 FOR ELECTRONIC CONTROL.

Remove x in Fig. 395 and replace by y above.

This tilting reduces the state of unbalance of the bridge and causes the aerial motor to slow down as the aerial approaches the desired position. The tendency to overshoot, with consequent oscillation or mechanical vibration, is therefore greatly reduced.

Thus the gyro action accelerates the aerial rotation at starting up and decelerates it as the final position is approached.

The rotation control function may be performed electronically by replacing the differential Selsyn, gyroscope, contactor and bridge with a Selsyn control transformer and an amplifier, which converts the Selsyn A.C. error voltage output to a direct current, of sufficient power to control the D.C. generator field. (Fig. 397).

The Amplidyne

The Amplidyne drive is similar to that described in Fig. 395 with the exception that an Amplidyne (trade name) is used in place of the D.C. generator.

The main difference is that the field of the Amplidyne requires much smaller control power for the same values of output power. This device may be compared to an electromechanical power amplifier.

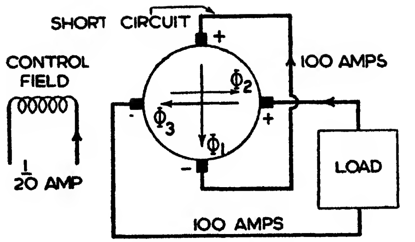


Fig. 398.—ARMATURE FLUX, CONTROL FLUX AND LOAD REACTION FLUX IN CURVATURE OF A SHORT-CIRCUITED D.C. SHUNT-WOUND MOTOR HYPOTHETICALLY LOADED.

Load and short circuit in parallel but their sources of e.m.f. are in quadrature.

If on load the armature current of a D.C. shunt-wound generator is, say, 100 amps. and the field excitation current 5 amps., it may be necessary to reduce the field excitation current to as low as $\frac{1}{20}$ amp. (if the armature current is not to exceed 100 amps.) *when the load is removed and the armature short circuited*, assuming that the speed of rotation is constant.

In this case the magnitude and direction of the *armature reaction flux* are the same as in the loaded generator, but the *control flux** is very small. The armature reaction flux Φ_1 and the control flux Φ_2 are still in quadrature.

Voltages are therefore generated in the armature conductors as they rotate through these two fluxes, but because of the location and direction of these two fields, the maximum voltage caused by the cutting of the reaction flux appears across the armature at right angles to the voltage developed by the excitation flux.

Consequently, by placing a second set of brushes at right angles to the short-circuited brushes, sufficient voltage is available to supply another 100 amperes to an external load, in addition to the 100 amperes flowing through the short-circuited armature path.

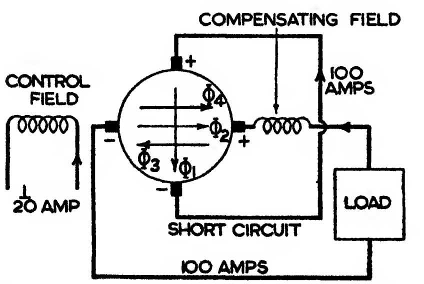


Fig. 399.—ADDITION OF COMPENSATING WINDING TO NEUTRALISE THE LOAD REACTION FLUX SHOWN IN FIG. 398.

Since the control flux has only to build up to a low value, and since the resistance of the short-circuited armature is very small, full load current may be obtained in an

* The control flux is that field strength due to the small magnetising field current of $\frac{1}{20}$th amp. assumed to be sufficient to maintain a current of 100 amps. when the load proper is switched off and a short circuit is substituted.

exceptionally short time. Thus changes in the control field current are amplified almost instantaneously by the Amplidyne arrangement (see Fig. 398).

The direction of the second load current just discussed is such that it must produce a second armature reaction flux Φ_3 in direct opposition to the original control flux Φ_2 (Fig. 399).

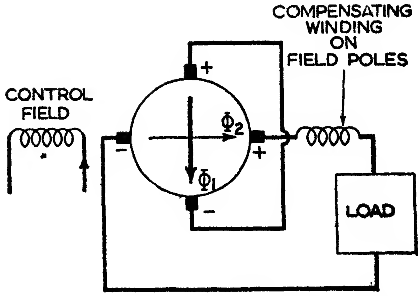


Fig. 400.—EFFECTIVE MAGNETIC FIELDS IN AMPLIDYNE.

As result of modifications shown in Figs. 398 and 399.

The load armature reaction flux will be much greater than the control flux and would prevent the correct field current from controlling the output.

It is essential that the small control flux Φ_2 should not be affected by this second armature reaction flux Φ_3 if it is to retain control over the output.

Therefore a series compensating winding through which the load current flows is wound on the control field poles. The number of turns of this winding are adjusted so that the compensating flux Φ_4 exactly cancels the second or load armature reaction flux Φ_3 for all values of load current in the operating range (Fig. 400). In this case the effective magnetic fields are as shown in Fig. 400.

If the compensating flux is slightly under the value for complete compensation, the machine has reduced power gain and acts as a "negative feedback," feedback being taken from output current, *i.e.*, the internal impedance of the Amplidyne is increased by under-compensation. Over-compensation creates the effect of regenerative reaction and operation may easily become unstable.

Since residual magnetism along the axis of the control field would have a large effect on the Amplidyne output, it is necessary to demagnetise the core material. This is accomplished by attaching a low permeability magnet to the end of the armature. This magnet, with a separate field winding, revolves and generates a small A.C. voltage which is applied to two sets of opposed windings on the field poles.

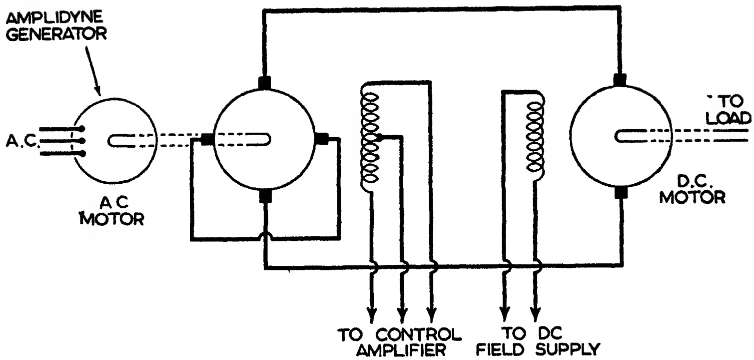


Fig. 401.—BASIC AMPLIDYNE AERIAL DRIVING SYSTEM.

Thus the generated A.C. neutralises any residual magnetism when the control field current is zero.

In summation, it may be noted that the production of the short-circuit current and the associated armature reaction flux by a small control field represents a first stage of amplification, which can be regarded as corresponding to current amplification.

The use of this large current, and the flux it produces, to induce sufficient voltage to drive an equally large current through the external load circuit, represents a second stage, which may be regarded as voltage amplification.

The compensating winding can be regarded as a feedback circuit, when exact compensation corresponds to zero feedback. The power gain of the Amplidyne may range from 3,000 to 10,000, and perhaps higher in certain machines.

The additional power appearing at the output is of course supplied by the additional mechanical power which must be applied to the shaft by the motor.

Basic Amplidyne Drive

The Amplidyne drive as commonly used consists of the basic system shown in Fig. 401.

The drive is usually an A.C. motor. The control field of the Amplidyne is shown as a split winding, to enable it to be supplied by a control amplifier having separate outputs for each polarity of the applied signal. The series-compensating winding is omitted from the drawing for the sake of clarity.

The field of the D.C. motor can be supplied by a rectifier or permanent magnet generator. In order to prevent demagnetisation when permanent magnets are employed, compensating windings are connected in series with the armature, and wound on the faces of the field poles to neutralise the effect of the armature reaction flux.

The amplifier is controlled by comparing an error voltage from a Selsyn control transformer with an A.C. reference voltage furnished by the A.C. line supply. In the diagram, Fig. 402, it is assumed that there is no phase shift, and therefore both A.C. input voltages are in phase.

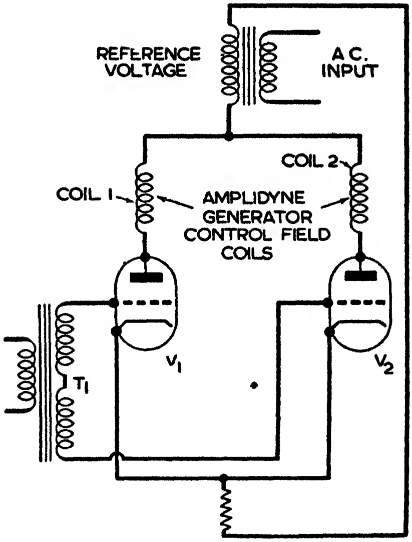


Fig. 402.—BASIC AMPLIDYNE CONTROL AMPLIFIER FOR AMPLIDYNE AERIAL DRIVING SYSTEM.

The anodes of both valves are positive, but the grid V_2 is below cut off and the grid of V_1 above cut off. In this condition the output is that of a half-cycle rectifier supplying coil 1; coil 2 is inoperative.

If the phase of the voltage input to the grid transformer T_1 is shifted 180° while the phase of the A.C. line remains the same as before the circuit acts as a half-wave rectifier for coil 2 and coil 1 is inoperative. If the two coils are both wound on the Amplidyne field poles but in opposite directions, a 180° shift in grid voltage changes the direction of the control field flux and the polarity of the Amplidyne output. The magnitude of the Amplidyne output is thus controlled by the magnitude of the A.C. error voltage and the polarity by the phase of the error voltage.

Alternative Arrangements

Numerous alternative arrangements incorporating various refinements are available for aerial control systems; these include aided tracking, used in cases where the motion of the target is such that its rate of change of bearing is constant.

Instead of following the target by turning a handwheel, a small D.C. motor is used to drive the rotor of the Selsyn generator. The speed of the motor is adjusted by a potentiometer.

The error voltage output from the Selsyn control transformer can now be applied to an Amplidyne control amplifier, so that the aerial follows the target continuously as long as the rate of change of bearing and elevation is constant. The occasions when this form of tracking can be used are, however, somewhat infrequent.

The device can be usefully employed for continuous search in conjunction with a P.P.I.

A case in point is when an "A" scan indicator and a P.P.I. are operated together. A slewing motor can be adjusted to rotate the aerial at the required speed for P.P.I. operation until it is desired to follow a target of particular interest more closely. The slewing motor is then disconnected, the target being followed for detailed examination on the "A" scan by operating the handwheel by hand.

True bearing correction is incorporated with some equipments when a gyro compass is installed. In this case it is usual to provide another complete Selsyn indicator system which transmits a true bearing from the gyro compass.

The information from both indicator systems (*i.e.*, the bearing of the aerial relative to the ship's head and the true bearing as transmitted from the gyro compass) is then placed on a special bearing indicator dial, the outer dial being fixed. If the aerial controls the position of a pointer and the inner dial is rotated by a Selsyn operated from the ship's gyro compass, the reading on the dial opposite the pointer gives true bearing.

Two-phase induction motors may be employed to rotate the aerial system and hydraulic systems were also in use before electrical systems were developed. A D.C. motor used in conjunction with a thyratron has also been employed.

The fundamental requirements, however, are the same in all cases. A smooth drive with good acceleration and a large reserve of power during rotation towards the required position. On approach to the desired position the aerial should be retarded so that it comes to rest in its final position with zero velocity, and excessive mechanical vibration of the system does not occur.

APPENDIX I

THEORY OF TRANSMISSION LINES

TRANSMISSION lines and sections of transmission lines are employed for the following purposes :—

(*a*) To guide electrical energy from one point to another.

(*b*) To *guide* and *match* the energy output from one unit to the input of another unit, *e.g.*, to match transmitter outputs to an aerial.

(*c*) To establish proper phase relationships between the elements of an aerial array and to perform transforming functions.

(*d*) To act as circuit elements, *i.e.*, to function as a tank or oscillating circuit in place of coils and condensers.

(*e*) In certain circumstances to act as metallic insulation and balance converters.

(*f*) To form part of an electronic switching system and to invert a load.

Thus (*a*) is a guiding function carried out by virtue of conductivity.

(*b*) Combines the function (*a*) with matching functions which are carried out by virtue of the characteristic or surge impedance of the line, the value of which depends upon its geometry.

(*c*), (*d*), (*e*) and (*f*), and sometimes (*b*), are functions carried out by virtue of the properties possessed by sections of transmission line bearing definite relationship in linear length to the wavelength of the applied R.F.

A transmission line may consist of two wires, of any length suitably insulated from one another; alternatively a coaxial cable, in which the inner conductor is entirely surrounded and shielded by the outer conductor. Either wire may be thought of as composed of a number of very small lengths joined in series. Since inductance and resistance is associated with each of these elementary lengths, in the case of a line of uniform cross-section, inductance and resistance are distributed evenly throughout its length. Similarly, it can be seen that there will also be a distributed capacity effect throughout the length of the transmission line, due to the capacity effects between the elementary lengths in opposite wires.

In approaching the subject of transmission lines it must be realised that any element of voltage initially applied to one end takes *some* time to reach the far end, no matter how short the line may be. In the case of open transmission lines the velocity of propagation

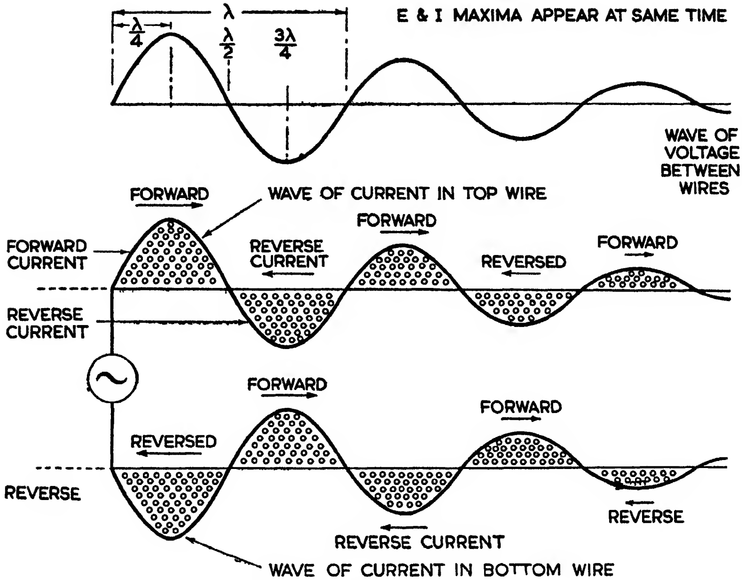


Fig. 1.—DISTRIBUTION OF CURRENT AND VOLTAGE IN A LINE OF INFINITE LENGTH.

The lines act as base lines for current curves. The perpendicular distance from wire to curve is a measure of the electron motion in the wire at a point. The arrows represent the direction of electron motion for each half wave.

along the line is generally taken as 186,000 miles per second; this figure may, however, be considerably reduced for coaxial cables.

In order to avoid more than one variable, it is convenient in the first place to consider the distribution of voltage and current (or more strictly speaking the electric and magnetic fields) along a line of infinite length and uniform cross-sectional area, when the frequency of the input is varied. It is assumed also that losses due to leakage *viâ* insulators are so small as to be negligible in effect upon the analysis of the general case.

The voltage and current distribution in an *infinitely* long two-wire transmission line (for any *finite* frequency) may be shown as in Fig. 1.

Analysis of Fig. 1. Analysis of Voltage and Current along a Two-wire Transmission Line System

The conditions for this analysis are :—

(1) That the generator is supposed to have been in operation for some time.

(2) That at some instant, when the voltages between the terminals

of the generator are at zero in the A.C. cycle, the voltages and currents along the line are deemed to be frozen for examination purposes.

From consideration of Fig. 1, then, the following facts emerge :—

(*a*) The voltage and current are everywhere in phase.

(*b*) Both voltage and current reverse at regular intervals along the line.

(*c*) The amplitude of the voltage and current waves decreases along the line.

(*d*) Energy hypothetically supplied by the generator to the infinitely long line never reaches the infinitely distant end. It must, therefore, all be consumed in the line itself.

Let an observer now stand at some fixed point along the line, and let transmission continue in the normal way. If the waves of voltage and current could be seen, they would appear to sweep past the observer as alternative hills and dales. This means that regularly spaced reversed currents and voltages sweep along the line with the same velocity.

All other things being equal, let the frequency of the generator input be changed. Let frequencies of 60 cycles, 1,000 cycles, 30 megacycles and 3,000 megacycles be applied in turn.

If the velocity is taken at 186,000 miles per sec. or 300,000,000 metres per second, and the two-way transmission line is infinitely long, the positive peak of a 60-cycle wave of voltage or current will have travelled 3,100 miles down the line before the generator applies the next positive peak. Therefore the voltage, in the line for continuous transmission, reverses in polarity at intervals of 1,550 miles along the infinitely long line until the energy is completely absorbed at some infinitely great distance.

If a frequency of 1,000 cycles is applied to the same line, voltage reversals take place at intervals of 93 miles whilst at 3,000 megacycles, reversals occur at intervals of about 3·86 inches.

Thus a line of 300 miles would be electrically short for a frequency of 60 cycles, but electrically long for 3,000 cycles. Similarly, a line of 1 metre would be long for a frequency of 3,000 megacycles but short for 3 megacycles.

Thus when a line is electrically short, the impedance presented to the generator is mainly that of the load. When the line is electrically long, however, and if the load is not of a certain critical value, the voltages necessary to produce a given current or to deliver a given power, may differ greatly from that which can be accounted for by the impedance of the load in series with the resistance of the line.

A two-wire transmission line has resistance, capacity, inductance and a leakage or conductivity coefficient.

Ignoring the leakage factor, which is small as compared with capacity and inductive reactions, a transmission line may be represented as in Fig. 2 (a), (b) and (c), subject to the conditions that the

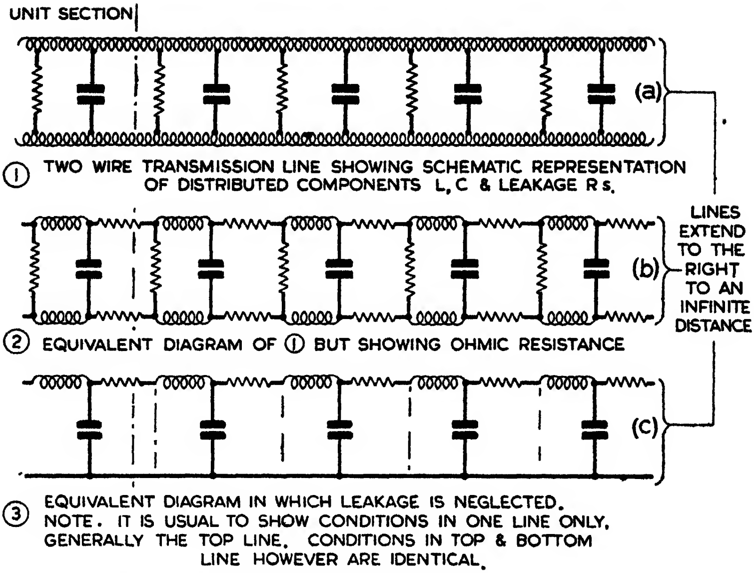


Fig. 2.—THE INFINITE LINE (DEVELOPMENT OF SCHEMATIC REPRESENTATION).

line shall be divided into many sections, and that the capacity, inductance and resistance lumped together, in each section, shall be as small as possible per section. This is essential if the distributed resistance, inductance and capacity of a transmission line is to be simulated.

A two-wire transmission line may, therefore, be specified by stating :—

(*a*) Its resistance per mile.
(*b*) Its inductance in millihenrys per mile.
(*c*) Its capacity in microfarads per mile.

Smaller units for short lines can be adopted when necessary.

When leakage is taken into consideration it is expressed in terms of conductance, and as micromhos per mile. It will, however, be ignored in this analysis.

The properties of a transmission line made up from a number of sections of lumped resistance, capacity and inductance can now be considered. The leakage factor can be neglected, since it is very small compared with X_C and X_L at high frequencies. In fact, R (the ohmic resistance) itself is frequently overshadowed by these other values.

If a voltage is applied to the line shown in Fig. 3 (a), some current

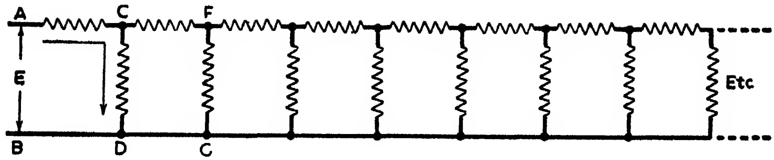


Fig. 3 (a).—DEVELOPMENT OF THE CHARACTERISTIC IMPEDANCE OF A LINE.

Note. Resistances only have been employed in order to simplify the explanation, but this in no wise affects the general principle. If X_L, X_C and R are substituted in their proper proportions the calculation merely becomes more complex, but it does not alter the fundamental principle that the continuous addition of similar sections causes the impedance, as seen at AB, gradually to approach a finite value, *i.e.*, the characteristic or surge impedance of the line.

must flow and the ratio of the voltage to the current (Z) is the impedance, where $Z = \frac{E}{I}$.

Conclusions to be Drawn

The impedance presented at the input of a transmission line is not merely the resistance of the line in series with the impedance of the load. The effect of series inductance and shunt capacity on the line itself may be large as compared with the resistance, and may even overshadow the load value when looking in* at the line from the input terminals.

The impedance for a single elementary section AB of the line shown in Fig. 3 (a) can be calculated by the use of series parallel impedance formulæ, provided the impedance across CD is known. But since this elementary section is merely one small part of a longer line, another similar section FG is across terminals CD. Again the impedance at AB of the two sections can be calculated provided the impedance of the third section is known. This process of adding one section after another can be carried on and on.

With the addition of each section, the impedance at AB has a *new and lower value.*

After many sections have been added in this way, each successive addition has less and less effect on the impedance at AB. Thus, if sections are added endlessly the line becomes infinitely long and a certain value of impedance across AB is approached so nearly that it may be regarded as finite. This value is the characteristic impedance for the line.

* *Note.* The terms "looking in" and "looking out" are met with constantly. This arises because the generator "sees" an impedance represented by the line and the load which may be quite different from the impedance seen from the load end, which includes the line impedance and the impedance of the generator.

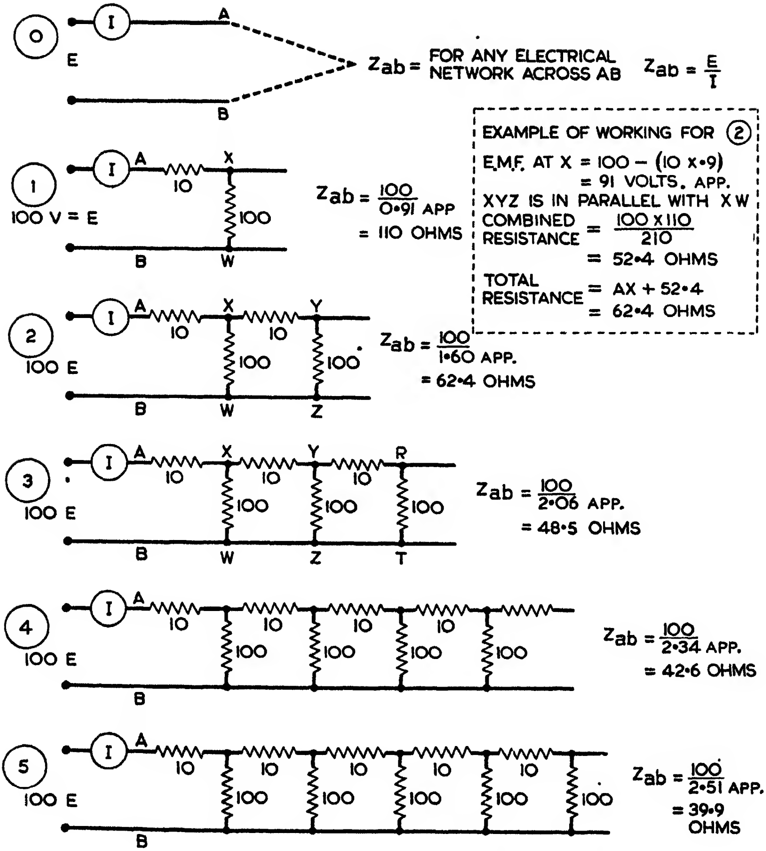


Fig. 3 (b).—NUMERICAL EXAMPLE TO ILLUSTRATE THE MEANING OF CHARACTERISTIC IMPEDANCE. (ALL VALUES APPROXIMATE.)

Note. As additional units are added, the total impedance approaches closer to a finite value of 37·0 ohms. The convergence is such that after about the twelfth addition the series is still converging upon 37. Subsequent changes are so small in effect that 37 ohms can be taken as a finite value. It now follows that if a line is terminated in a resistive load equal to Z_0 for the line, it must behave as though it had infinite length, *i.e.*, the applied energy will be wholly asborbed in the line and load. This corresponds to the ideal condition of an infinite line. A maximum amount of energy is transferred to the load under these conditions.

The Characteristic Impedance

Let it be assumed that the sections of Fig. 4 (b) extend to the right with an infinite number of sections. In these circumstances the impedance across AB is Z_0 (its characteristic impedance).

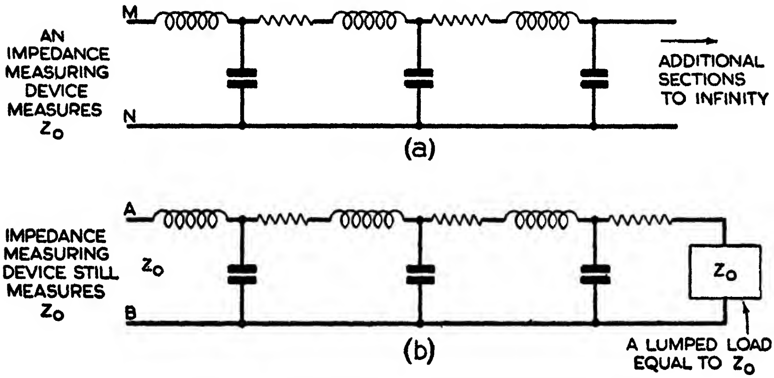


Fig. 4.—COMPARISON OF INFINITE LINE AND A FINITE LINE TERMINATED IN AN IMPEDANCE EQUAL TO Z_0.

(a) An infinitely long line severed at some point MN near the generator end is still infinitely long.

(b) A line terminated in a load, having an impedance equal to the characteristic impedance of the line, may be thought of as the severed portion AB plus MN to infinity, MN being as represented by the load impedance which is equal to Z_0. The characteristics of (a) and (b) are thus identical.

If the line is cut at MN, Fig. 4 (a), an infinite number of sections still extends from MN to the right since the line is endless in that direction. Consequently, from the results obtained in the previous paragraph, the impedance appearing across MN must also be Z_0, since the conditions are similar.

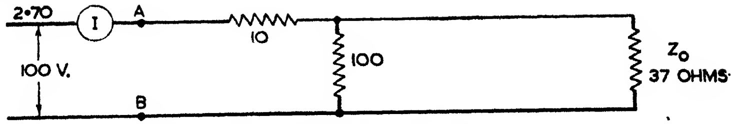
The impedance across the input of a theoretically infinite line has a very valuable use, for if a *load equal to this impedance* is placed across the output end of any short or convenient length, an impedance appears at the *input* terminals of the shortened line of the same numerical value as that of the load. In other words, a line loaded to meet this condition, no matter how short it may be, behaves as though it was a line extending to infinity, and the maximum transfer of energy to the load can take place.

Only one value of load impedance for any particular geometrical construction can cause a line to behave in this way.

This value of load impedance must be numerically equal to the characteristic impedance of the line which is independent of its length, but is a function of its *geometry*.

A Numerical Example

Fig. 3 (b) gives a numerical example of the manner in which (for demonstration) small sections of line loading can be added successively to cause the impedance looking in at AB to converge towards some finite value which is the characteristic impedance. It is



$Z_{ab} = Z_0 = \frac{100}{2 \cdot 70} = 37$ OHMS

Fig. 5.—This figure should be compared with fig. 6.

The line is terminated in Z_0 and will, therefore, behave as a line of infinite length.

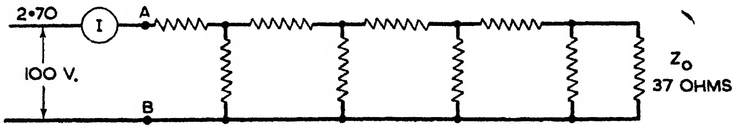


Fig. 6.—Compare with fig. 5.

The load along the line is distributed, but since it is terminated in Z_0 it will still behave as a line of infinite length. In both cases, a maximum amount of the applied energy will be absorbed in the line and load.

readily seen that this finite figure is determined by the geometrical construction of the network.

In (1) a section of parallel wires AX, BW, having a 10-ohm resistance in series and 100-ohm resistance across the wires has an impedance of 110 ohms. Assuming the simple case where D.C. is applied to pure resistance let (2), an identical section XYZ now be added, *i.e.*, 110-ohm section in parallel with XW. The resistance across AB = 62·4 ohms and the total current is increased from 0·91 to 1·60 amps. to supply the current flowing in the second section. Addition of the third section YRT places 110 ohms in parallel with YZ of the network, giving a new impedance of 48·5 ohms, etc. As this process is repeated the resistance across AB is continuously lowered and the line current is increased, but the *change* per section in both current and resistance grows less as more sections are added. After the addition of about twelve sections, the impedance across AB is reduced to nearly 37·0 ohms. Several hundred sections can now be added without any appreciable further reduction in the impedance at AB. For all practical purposes therefore, the converging impedance value of Z_{ab} may be considered to be finite at 37·0 ohms, which is Z_0 the characteristic impedance of the line.

Referring to the example just given, when the process of adding successive sections has caused the input impedance across AB to converge to the numerical value of Z_0, the line current is 2·70 amps. Nothing can now be done at the output end of the line to make the current increase, provided the line is very long and made up of resistance only.

Thus if the first section has an impedance of value equal to Z_0 (in this case 37 ohms) placed across its output instead of a second section (see Fig. 5), the impedance appearing across AB is still 37 ohms. This must be so since the geometry of the *part* is the same as that for any length of line. In fact, 37 ohms placed across any length of this line confers upon it an *input* impedance equal to its characteristic impedance, *i.e.*, 37 ohms.

This means that if the impedance appearing across the input terminals of a line of infinite length is Z_0, then *any length* of the same line, when terminated by a load with a resistive impedance equal to Z_0, will exhibit a similar input impedance.

In the numerical example, D.C. and pure resistance only were considered. Whilst the actual numerical values for A.C. and D.C. would be different, the fundamental principle is the same. The characteristic impedances of lines, in practice, generally range between 50 ohms and 600 ohms.

Sections of two-wire transmission lines exhibit a variety of properties according to :—

(*a*) Their loading.
(*b*) Their length.

The two extremes of (*a*) are open circuit (no load) and short circuit. Both conditions are employed to obtain certain characteristic effects.

Sections of less than a quarter wavelength behave differently from sections which are longer than a quarter and less than a half wavelength. Sections longer than a half wavelength but less than a full wavelength are obviously multiples and therefore repeat the behaviour of their respective sub-multiples which are all less than a half wavelength.

It is therefore necessary to investigate separately the properties of :—

Half wavelength sections on open circuit and on short circuit.

Quarter wavelength sections on open circuit and on short circuit.

Sections shorter than a quarter wave on open and closed circuit.

Sections longer than a quarter wave and less than a half wave on open and closed circuits.

Making eight cases in all.

Cases involving other terminations for the classifications mentioned above must be treated separately.

Firstly, therefore, consider the general conditions of voltage and current distribution for open and closed lines of finite length.

Open Circuits and Standing Waves

When energy is applied at the generator end of a two-wire transmission line which is open-circuited at the load end *time*

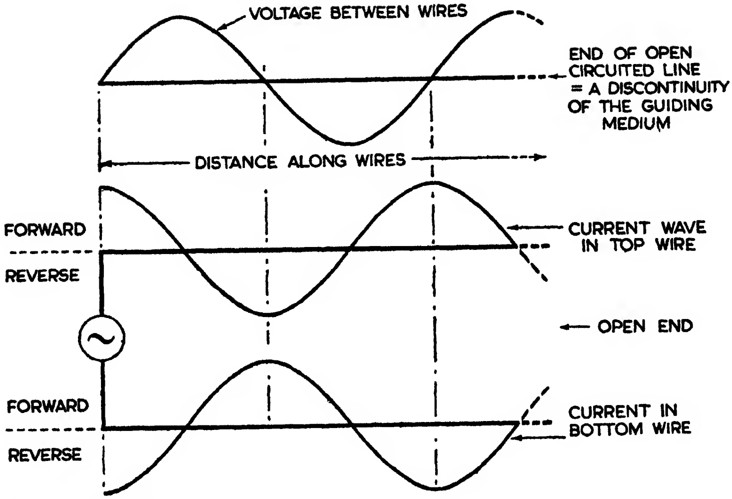


Fig. 7.—DURING THE INITIAL SURGE, VOLTAGE AND CURRENT WAVES GO DOWN THE LINE TOGETHER IN PHASE.

Upon reaching the open end, an abrupt change of phase of 90° takes place between the voltage and current waves reflected back to the source.

elapses before the energy applied at the source can reach the far end during which an initial surge occurs. This consists of a wave of voltage and a wave of current sweeping down the line, with their positive maxima together until the end of the line is reached. The *initial* conditions are, therefore, the same as for a line of infinite length.

When the current wave reaches the open-circuited end of the line it faces a discontinuity and must fall to zero, whereupon new voltage and current waves are set up which travel *back* along the line toward the input end. These reverse waves are, in effect, reflections caused by the high voltage generated when the current wave and accompanying magnetic field falls suddenly to zero on reaching the discontinuity.* Immediately the first reflection occurs at the discontinuity a phase change takes place between voltage and current. Since the voltage must be maximum and the current must be zero at the extreme end of an open line, the effect is that of a 90° displacement (see Fig. 7).

In general, whenever the conditions at the termination of a transmission line are such that the advancing energy sees any

* The effect when a guided electric wave meets a discontinuity may be compared with that which occurs when sound meets an unyielding obstruction in the normal to its surface. The wave is reflected backwards along its line of propagation.

change whatsoever in the value of the impedance of its guide, some of the energy is always reflected back along its line of propagation, its magnitude and phase being determined by the magnitude and nature of the impedance change which causes the reflection. Reflection is, in effect, the reaction to any change of state of the conducting path.

Whenever waves are reflected as above, the nett effect produced along the line, by the advancing wave from the generator and the waves reflected from the load, is the appearance of *standing* waves * of voltage and current displaced from each other all along the line by a quarter of a wavelength.

Thus, when an open-circuited line is connected to a generator, after the first initial surge, standing waves of voltage and current appear along the line separated a quarter of a wavelength apart. Also a current minimum or node is accompanied by a voltage maximum or anti-node, and *vice versa.* (See analogies given in Figs. 8 (a) and 8 (b).)

In a line which exhibits standing waves of current and voltage, the initial energy from the generator continually surges back and forth and is in fact carried in the electromagnetic fields about the line. Since energy is transformed in creating these fields the process is reversed and energy is returned to the circuit when they die away.

Thus, when the first wavefront of current surges down the line from the generator, the parts of the line carrying maximum current (anti-nodes of current) are surrounded by magnetic fields. The energy of these fields is proportional to $\frac{1}{2}LI^2$ and since, in the case of an open circuit, there is no load to absorb this energy, the energy of the magnetic field in the region at the end of the line is transferred to an electric field with a consequent rise of voltage at the open end of the line. Thus the reflected wave, which always appears with a discontinuity or some change in the impedance of the guide, is the reaction to the energy transformation that must take place in the region of the impedance change, *i.e.*, the change of the ratio $\frac{E}{I}$.

Energy absorbed by the line from the generator supplies the line losses due to heat and radiation, so that the energy stored in the electric and magnetic fields is maintained.

Actually, the current minima (nodes) do not fall to zero, since there must always be some line losses and a small in-phase current must flow to provide for them.

Indicating instruments seldom discriminate between positive and

* The term " standing wave " can be misinterpreted. When they exist, the positions along a line at which minima or nodes occur, are fixed. They do, in fact, stand still in so far as the positions at which they are found do not wander along the line. But the amplitudes between these nodes rise and fall at the frequency of the input, consequently maxima or anti-nodes also occur at fixed points along the line at exactly a quarter of a wavelength from the nodes.

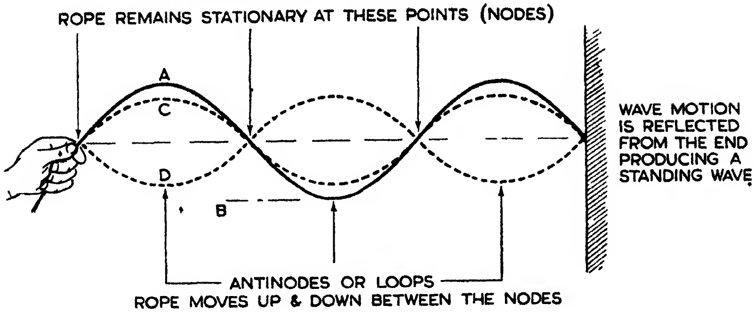

ROPE REMAINS STATIONARY AT THESE POINTS (NODES)
A
C
D
B
WAVE MOTION IS REFLECTED FROM THE END PRODUCING A STANDING WAVE
ANTINODES OR LOOPS
ROPE MOVES UP & DOWN BETWEEN THE NODES

Fig. 8 (a).

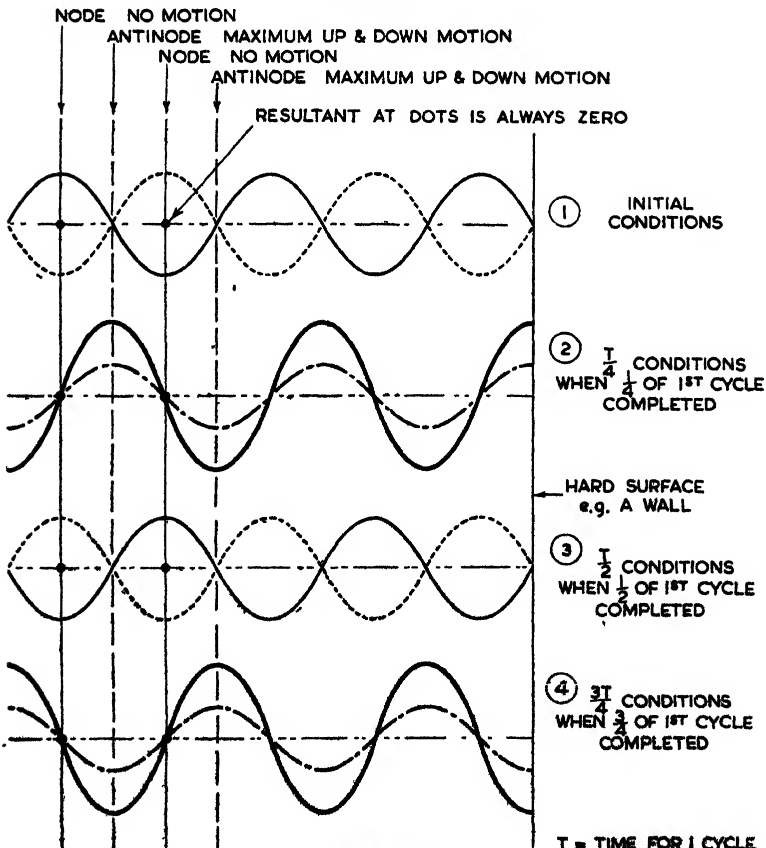

NODE NO MOTION
ANTINODE MAXIMUM UP & DOWN MOTION
NODE NO MOTION
ANTINODE MAXIMUM UP & DOWN MOTION
RESULTANT AT DOTS IS ALWAYS ZERO
1 INITIAL CONDITIONS
2 T/4 CONDITIONS WHEN 1/4 OF 1ST CYCLE COMPLETED
HARD SURFACE e.g. A WALL
3 T/2 CONDITIONS WHEN 1/2 OF 1ST CYCLE COMPLETED
4 3T/4 CONDITIONS WHEN 3/4 OF 1ST CYCLE COMPLETED
T = TIME FOR 1 CYCLE

Fig. 8 (b).

Fig. 8 (a).—ROPE ANALOGY FOR STANDING WAVES.

This analogy exemplifies standing waves on a rope. C and D indicate the manner of change of amplitude at the anti-nodes. An observer holding on to the rope at an anti-node would be jerked up and down between A and B at the frequency at which the rope is shaken. An observer holding on at a node, however, would not be moved.

Fig. 8 (b).—SOUND ANALOGY. REFLECTION OF SOUND WAVES FROM A HARD SURFACE AND FORMATION OF STANDING WAVES.

The thin continuous line represents the forward wave moving left to right. The dotted line is the reflected wave moving right to left. The dot-dash line in (2) and (4) shows reflected and forward waves together at $\frac{T}{4}$ and $\frac{3T}{4}$. The thick line shows the resultant at these instants.

At the instant (1) the forward and reflected waves are 180° out of phase. The resultant is zero. Amplitudes are equal.

At $\frac{T}{4}$ later—(2) the forward wave has travelled along the line $\frac{\lambda}{4}$ to the right and the reflected wave has travelled $\frac{\lambda}{4}$ to the left. The maxima of the two waves, therefore, coincide and the resultant amplitude is twice that of either wave.

At $\frac{T}{2}$—(3) the two waves are interchanged with respect to the configuration in (1), because they have each travelled $\frac{\lambda}{2}$ in opposite directions and the resultant amplitude is again zero.

At $\frac{3T}{4}$—(4) the two waves are interchanged with respect to the configuration in (2) and the resultant amplitude is double that of either. Maxima and minima are also reversed. After this the sequence is continually repeated.

---

negative anti-nodes, so it is convenient to show diagrammatically both waves of voltage and current on one side of the line. Thus the voltage wave between wires, and the magnitude of the current wave in the line are shown together in their correct phases, using the top conductor as the base or zero line for both voltage and current curves (see Fig. 9).

The Voltage and Current Relations in a Short-circuited Line

When a line is short circuited, the voltage and current relations are shown in Fig. 10. Since a short circuit is a condition of zero impedance, the current in the closed end of the line is a maximum, whilst the voltage is a minimum.

Reflection also occurs in a closed line for reasons similar to those for an open line, *i.e.*, the advancing wave sees a change of impedance of the guide (from Z_0 to approximately zero in this case).

No appreciable amount of the energy of the initial wave is absorbed by the short circuit. The high current in the shorted end of the line

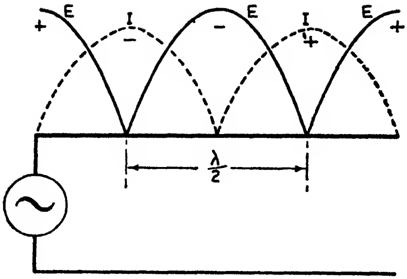


Fig. 9.—STANDING WAVES ON TRANSMISSION LINE CAUSED BY REFLECTION FROM THE OPEN END.

Voltage and current nodes and anti-nodes shown together using the top wire as a base. In the extreme case of an open line, anti-nodes of current coincide with nodes of voltage. If this line had been terminated in a resistive impedance equal to Z_0, there would have been no reflections and no standing waves.

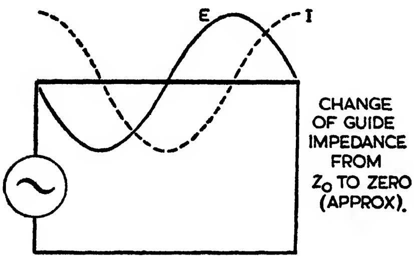


Fig. 10.—STANDING WAVES OF VOLTAGE AND CURRENT IN SHORT-CIRCUITED LINE CAUSED BY POTENTIAL ENERGY REFLECTED AT THE SHORT CIRCUIT.

Compare with Fig. 9 and note that positions of E and I maxima have been transposed.

represents potential energy $= \frac{1}{2}LI^2$. Consequently, because this energy exists across zero impedance, it cannot be absorbed and reflection must occur.

Comparison of Open and Closed Lines

It is important to note that the voltage and current relations in open and closed transmission lines are in reverse, as shown by comparative Figs. 11 (a) and 11 (b).

Examination of Figs 13 and 14, and comparison with Fig. 12,

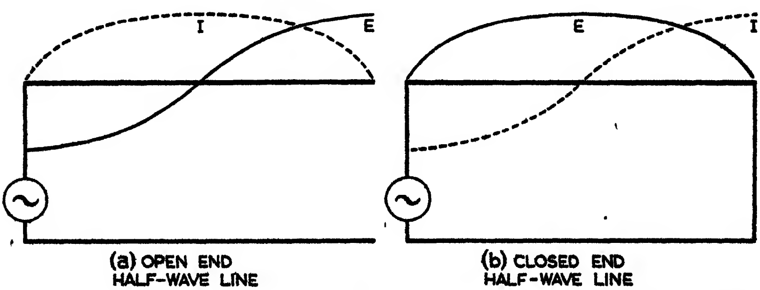


Fig. 11.—COMPARISON OF VOLTAGE AND CURRENT DISTRIBUTION FOR OPEN AND CLOSED HALF-WAVE LINES.

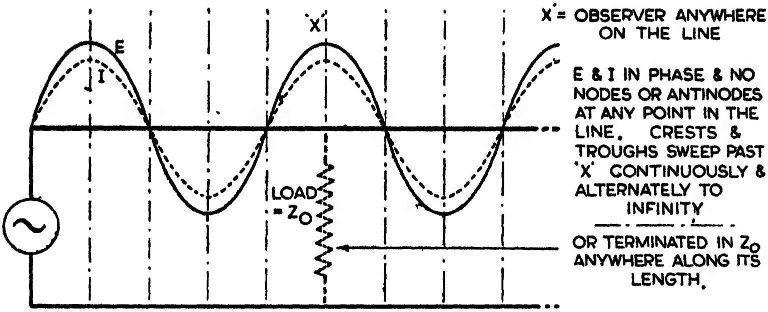


Fig. 12.—Voltage and current distribution along an infinite line or line terminated in Z_0 (resistive). Note that E and I are in phase because there is no reflected wave.

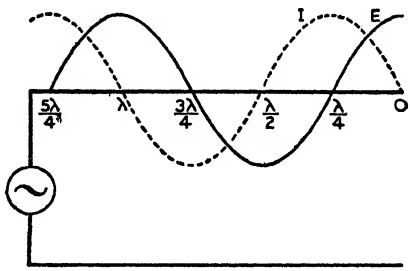
When dealing with finite lines it is *essential* to work from the load end back towards the source. This is because the conditions for voltage and current distribution along the line are fixed by the nature of the termination, *i.e.*, open circuit, closed circuit, terminated in Z_0 (resistive) or some other impedance.

Fig. 13.—Open-end line.

E and I 90° out of phase and standing waves with I nodes at $\frac{\lambda}{2}$, λ, $\frac{5\lambda}{4}$. Anti-nodes at $\frac{\lambda}{4}$, $\frac{3\lambda}{4}$, $\frac{5\lambda}{4}$.

Voltage nodes are at o, $\frac{\lambda}{2}$, λ.

Voltage anti-nodes at $\frac{\lambda}{4}$, $\frac{3\lambda}{4}$, $\frac{5\lambda}{4}$.



To determine the impedance presented to the generator, it is first necessary to fix the E and I nodes or anti-nodes at the *load* end. This is decided by the terminating condition as already explained in Figs. 9–12. These conditions are repeated for each wavelength and from load end to generator terminals.

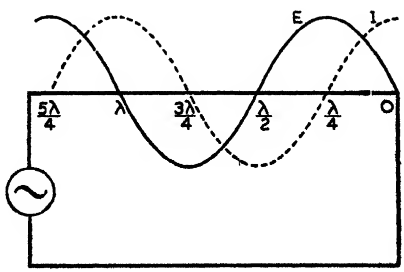


Fig. 14.—Closed-end line of same length as in Fig. 13, the applied frequency also being the same in both cases.

Compare E and I nodes and anti-nodes at load end and generator end with those for the open-end line of Fig. 13.

shows the relationship between voltage and current distribution in open lines, closed lines and infinitely long lines or lines terminated in Z_0 at various distances along the line measured in terms of wavelength or in terms of time for one cycle of the applied frequency.

The location of nodes and anti-nodes of voltage and current along a line must always be traced from the *output* end of the line, because they are created by reflected waves, the phase of which is determined by the line termination.

The requirement for the production of standing waves is that reflection must occur. Conversely, standing waves can only be suppressed by preventing reflection.

A non-resonant line is a line in which there are no standing waves. In this case all the energy passed down the line is absorbed by the load. The waves shown in Fig. 12 are travelling waves and there are no standing waves. For lines carrying radio frequencies the characteristic impedance is almost a pure resistance.

For a line to be resonant it may be either open or short circuited at the output end and cut according to whether it is open or short circuited, to some multiple of a quarter wavelength. If the length is not exactly some odd or even multiple of a quarter wavelength the line acts in a manner similar to either a capacitative or inductive reactance, according to whether it is greater or less than an odd quarter wave multiple. In fact at frequencies above or below the resonant frequency it behaves like a conventional tuned circuit.

A resonant transmission line as defined above is said to be resonant because it may assume characteristics similar to those of a resonant circuit comprised of lumped inductance and capacity. For example, it may exhibit the characteristics of a series resonant circuit or those of a parallel resonant circuit, according to whether it is cut to an odd or even number of exact wavelengths and also whether it is open or closed at the far end.

As for example :—

(1) *Series Resonance.*

(*a*) Resonant rise of voltage across circuit elements.

(*b*) Low impedance across the resonant circuit.

(2) *Parallel Resonance.*

(*a*) Voltage across the circuit never in excess of the applied voltage.

(*b*) High impedance across the resonant circuit.

Figs. 15 and 16 illustrate the relation of voltage, current and impedance for various lengths of both open-end and closed-end transmission lines. The impedance which the generator "sees" looking towards the termination for various lengths of line is drawn directly above each. The curves above the letters of various height indicate the nature of the impedance presented to the generator, and the circuit symbols indicate equivalent electric circuits or elements.

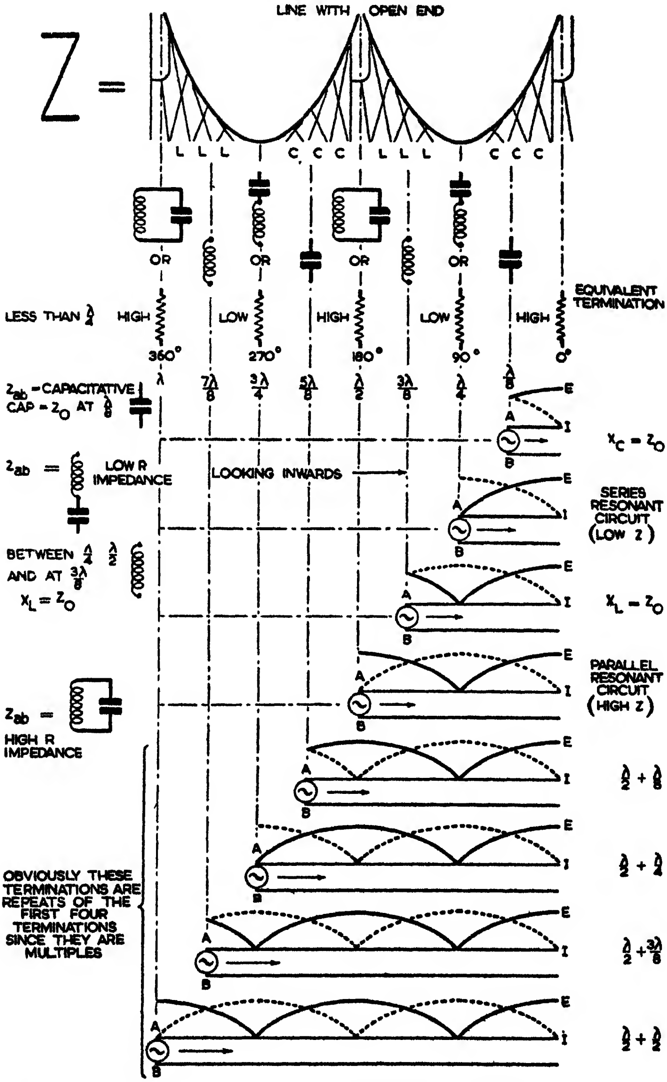


Fig. 15.—PROPERTIES OF OPEN-END TRANSMISSION LINES OF VARIOUS LENGTHS MEASURED IN TERMS OF MULTIPLES AND SUB-MULTIPLES OF $\frac{\lambda}{4}$.

Follow the dotted line upwards from generator end to read the nature of the impedance presented to the generator. Also check by procedure outlined under Fig. 13.

Standing waves of current and voltage are shown for each particular length of line reckoned from the load end.

Analysis of Open-end Lines

In conjunction with a detailed study of Fig. 15, note the following :—

(*a*) At all odd quarter-wave points, *i.e.*, $\frac{\lambda}{4}$, $\frac{3\lambda}{4}$, etc., measured from the *output end*, the current is maximum and the impedance is therefore minimum. The first voltage minimum occurs $\frac{\lambda}{4}$ from the open end.

(*b*) There is a resonant rise of voltage from the odd quarter-wave points towards the output end.

(*c*) Therefore, at all odd quarter-wave points an open-end transmission line acts like a series resonance circuit. The impedance is a very low resistance prevented only by the circuit losses from being zero. Consequently, an open line the total length of which is an odd multiple of $\frac{\lambda}{4}$ lengths or a single section $\frac{\lambda}{4}$ long presents a low resistance or series resonant characteristics to the generator at AB.

(*d*) At all even quarter-wave points, *i.e.*, $\frac{\lambda}{2}$, λ, etc., the voltage is maximum at AB (working back from the open end).

(*e*) The voltage for an even quarter-wave section never exceeds the applied voltage and there is a high impedance across both ends of the line, *i.e.*, it repeats at AB the conditions at the far end.

(*f*) Compared with the characteristics of an LC circuit, all even quarter-wave open-end lines act like parallel resonance circuits. The impedance presented to the generator is therefore an extremely high resistance.

(*g*) Open-end lines *less* than $\frac{\lambda}{4}$ and between $\frac{3\lambda}{4}$ and $\frac{\lambda}{2}$ act like a capacitative circuit, *i.e.*, the generatoı in fact sees a capacity at the open end of the line.

(*h*) Open-end lines longer than $\frac{\lambda}{4}$ and less than $\frac{\lambda}{2}$, or longer than $\frac{3\lambda}{4}$ and less than λ, act like an inductive circuit and the generator therefore sees an inductive termination.

(*i*) A $\frac{\lambda}{8}$ open line acts as a capacitative reactance numerically equal to Z_0.

(*j*) A $\frac{3\lambda}{8}$ open line acts as an inductive reactance numerically equal to Z_0.

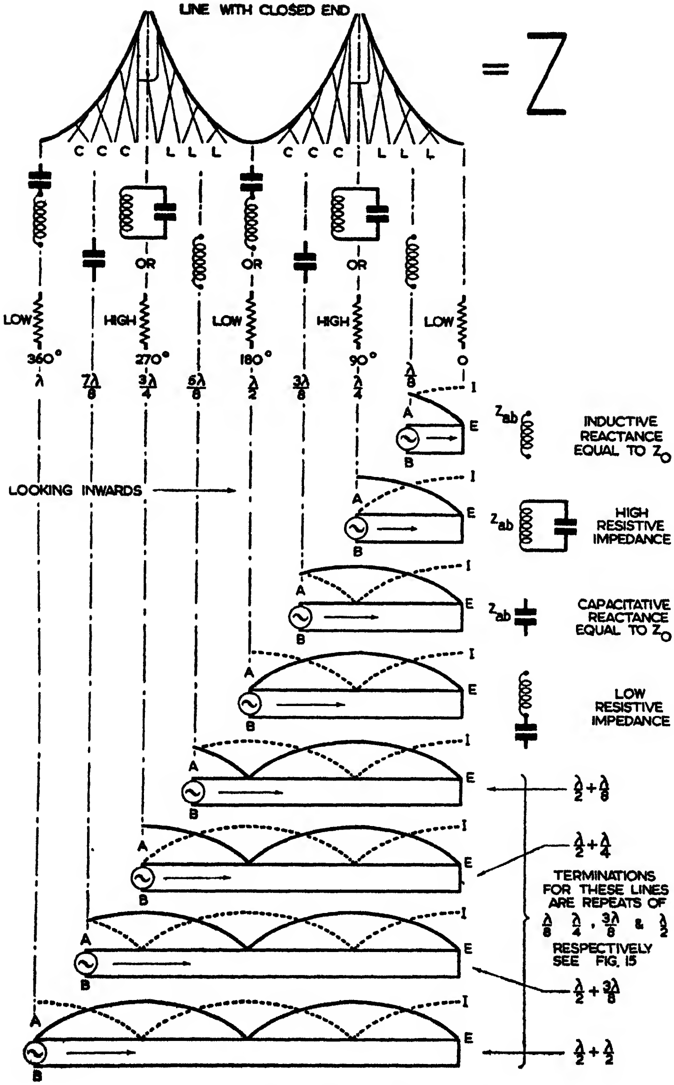
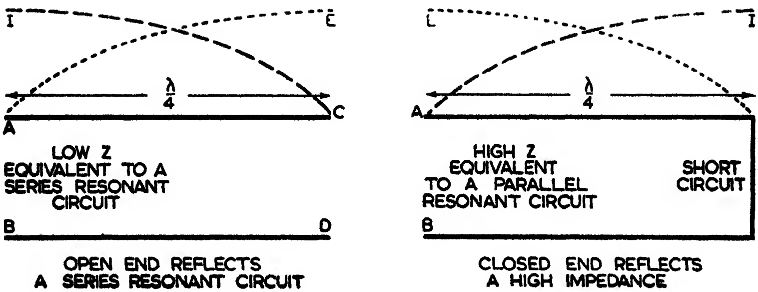
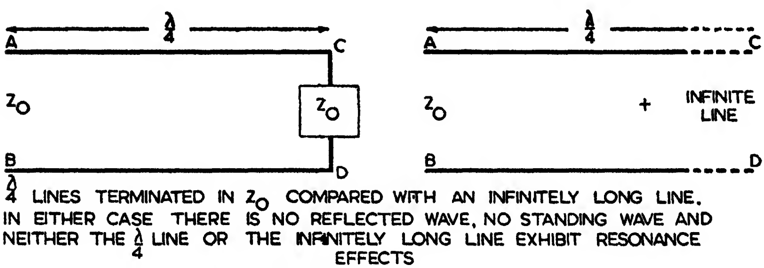


Fig. 16.—PROPERTIES OF CLOSED TRANSMISSION LINES OF VARIOUS LENGTHS MEASURED IN TERMS OF MULTIPLES AND SUB-MULTIPLES OF $\frac{\lambda}{4}$.

Read impedance presented to generator by same procedure as indicated for Fig. 15.



Figs. 17 *and* 18 — $\frac{\lambda}{4}$ LINES INVERT THE LOAD



Figs 19 *and* 20 — $\frac{\lambda}{4}$ LINES TERMINATED IN Z_0 (RESISTIVE) BEHAVE AS INFINITELY LONG LINES.

Analysis of Closed-end Lines

From a study of Fig. 16 note the following :—

(*a*) At odd quarter wavelengths, $\frac{\lambda}{4}$, $\frac{3\lambda}{4}$, etc., the voltage is high and the current low. The impedance is therefore high at the generator end and the first voltage maximum appears $\frac{\lambda}{4}$ from the shorted end.

(*b*) These conditions are similar to those in a parallel resonance circuit, consequently a shorted line of odd $\frac{\lambda}{4}$ length acts like a parallel resonance circuit. It also inverts the impedance of the termination at the generator end, *i.e.*, a short is reflected as a high resistance.

(*c*) At the even $\frac{\lambda}{4}$, $\frac{\lambda}{2}$, λ points the voltage is minimum, the current a maximum and the impedance is minimum as seen by the generator at AB. The first voltage maximum from the closed end is at $\frac{\lambda}{4}$ and the first minimum is $\frac{\lambda}{2}$.

(*d*) This resembles the effects produced by a series resonance LC circuit and can be compared with such. A half-wave closed-end line repeats at AB the low impedance at the far end of the line.

(*e*) A closed-end line *less* than $\frac{\lambda}{4}$ or between $\frac{3\lambda}{4}$ and $\frac{\lambda}{2}$ acts as an inductance and from $\frac{\lambda}{4}$ to $\frac{\lambda}{2}$ and between λ and $\frac{3\lambda}{4}$ it acts like a capacitative circuit.

(*f*) A $\frac{\lambda}{8}$ short-circuited line acts as an inductive reactance numerically equal to Z_0.

(*g*) $\frac{3\lambda}{8}$ short-circuited line acts as a capacity reactance numerically equal to Z_0.

Any line, even if cut to a particular fraction of a wavelength, loses its resonant characteristics when terminated in Z_0.

Since a $\frac{\lambda}{4}$ open-end line has a low value of E and a high value of I at AB, there is a low impedance at AB (Fig. 17). Conversely, since at AB a closed-end line has high E and a low I, there is a very high impedance at AB (Fig. 18).

If a $\frac{\lambda}{2}$ or a $\frac{\lambda}{4}$ line is terminated in Z_0 they become immediately

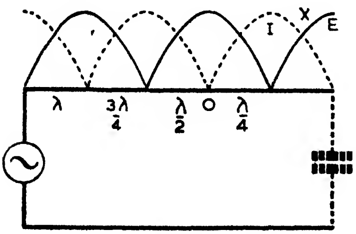


Fig. 21.—Reactance terminations.

Output end closed through a capacitative reactance equal numerically to Z_0. As capacity reactance becomes smaller (the termination conditions approach more nearly to a short circuit), X rises to resonant value nearer to the end of the line. In fact, the whole system of nodes and anti-nodes moves to the right. Check against Figs. 13 and 14.

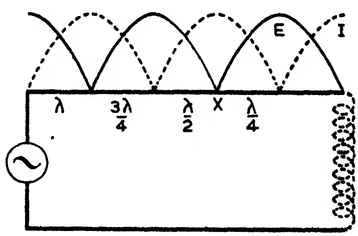


Fig. 22.—Reactance terminations.

Output end closed through an inductive reactance equal numerically to Z_0. Point X moves to the left as X_L increases and gradually approaches "open line" conditions. Check against Figs. 13 and 14. The effects in Figs. 21 and 22 are caused by the reactive components of the capacity and inductive loads respectively.

non-resonant, since there is then no reflected wave and no standing waves can be present. The impedance at AB $= Z_0$ under these conditions.

In Fig. 19 the $\frac{\lambda}{4}$ section terminated in Z_0 has no resonant properties whatever It behaves like an infinitely long line in this respect as in Fig. 20.

Thus it is shown that when a line is terminated in Z_0 its length has no particular significance.

A line terminated in a pure resistance equal to Z_0, normally, has no reflections present.

If, however, a line is terminated in a *reactance* numerically equal to Z_0, or equal to any other impedance, standing waves are not entirely eliminated. This follows from analyses of Figs. 15 and 16.

Fig. 21 shows the standing waves on a line terminated in a conventional capacity reactance numerically equal to Z_0. Fig. 22 shows the standing waves that exist on a line terminated in a conventional inductive reactance numerically equal to Z_0.

From Figs. 21 and 22 the following should be noted :—

With a capacity load the last *current* resonant point moves *closer* than $\frac{\lambda}{4}$ to the output end of the line.

For an inductive load $= Z_0$ the last resonant point of *voltage* is closer to the line termination than $\frac{\lambda}{4}$. Capacity termination produces a voltage and current distribution similar in character to that of the $\frac{\lambda}{8}$ open line, except that the curves are shifted towards the output end of the line by an amount that increases as the capacity reactance is reduced, *i.e.*, as the line approaches the closed-line condition.

With inductive termination the voltage and current distribution are essentially the same as for a *short-circuited* $\frac{\lambda}{8}$ line, except that the curves are shifted away from the output end by an amount that increases as the load reactance approaches infinity, *i.e.*, as the line approaches open-line condition.

Factors Controlling the Value of Z_0

For a two-wire transmission line it can be shown that $Z_0 = \frac{\sqrt{L}}{C}$.

An increase in the separation of the wires increases the inductance and decreases the capacity. Thus Z_0 is increased by increasing the spacing and reduced by the use of any dielectric that increases capacity.

When air is the dielectric between the wires, the formula

$Z_0 = 276 \log_{10} \frac{b}{a}$ can be applied. b is the spacing between centres of conductors and a = radius of the conductor. It is sufficiently accurate when Z_0 is a pure resistance.

Formula $Z_0 = 138 \log_{10} \frac{b}{a}$ where b is the inner diameter of the outer conductor and a is the outer diameter of the inner conductor can be applied to coaxial lines as an approximation.

A simple but by no means accurate method of measuring Z_0 is to terminate the line with a calibrated variable resistance. When the resistance is adjusted so that standing waves are no longer present, the line may be said to be terminated in Z_0.

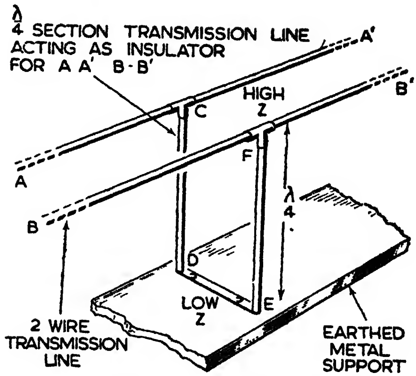


Fig. 23.—$\frac{\lambda}{4}$ SECTION AS A METALLIC INSULATOR AT RESONANT FREQUENCY.

$$\text{Resonant frequency} = \frac{300 \times 10^6}{\text{length of section} \times 4 \text{ (metres)}}.$$

Applications

Resonant lines may be employed as :—

(*a*) Metallic insulation.
(*b*) Wave filters and chokes.
(*c*) Reactances.
(*d*) Impedance-matching devices and transformers, etc.
(*e*) Phase shifters and inverters.
(*f*) Oscillator frequency controls.
(*g*) In converting from balanced lines to unbalanced lines.

As Metallic Insulators

When a shorted $\frac{\lambda}{4}$ line is excited at resonant frequency (the frequency which makes a particular length exactly equal to $\frac{\lambda}{4}$), standing waves appear. At the short circuit the voltage is zero and current a maximum. At the input end voltage is usually maximum and the current nearly zero. Therefore $\frac{E}{I}$ is large, which means that Z is large. In Fig. 23 the $\frac{\lambda}{4}$ section looks like an insulator to the lines AA^1 and BB^1 at resonant frequency. Such an arrangement constitutes a short circuit for direct current and, of course, it functions efficiently as an insulator only at resonant frequency, *i.e.*, at the frequency which makes CDEF an exact quarter wavelength.

Such an arrangement is a highly efficient electrical insulator at resonant frequency. It takes a negligible amount of energy from the line to make up losses caused by circulating current once it has absorbed enough energy from the line during the first few cycles, to set up the essential condition of resonance.

If the frequency varies too widely from the value for which the section is designed, it rapidly becomes a poor insulator, and begins to act as a condenser, or inductance, across the line, according to whether the frequency deviation is above or below the resonant frequency (see Fig. 16).

When properly designed and spaced, this type of insulator is more efficient than dielectric material, but at one frequency only.

Such an arrangement is mechanically strong and enables transmission lines to be readily fixed to steel towers, masts, etc.

$\frac{\lambda}{4}$ Line as Filter

By virtue of its impedance characteristics a quarter-wave line can be used as an efficient filter of even harmonics.

If interference on say 10 or 20 megacycles is experienced when a transmitter is operating on, say, 5 megacycles, the even harmonics may be suppressed by a $\frac{\lambda}{4}$ line filter.

In this case the short-circuited $\frac{\lambda}{4}$ line is made $\frac{\lambda}{4}$ for the fundamental frequency of 5 megacycles. At 10 megacycles this section becomes a half-wave line or a full-wave line at 20 megacycles. Since a half-wave line or a full-wave line offers zero impedance at its open end, the 10- and 20-megacycle harmonics are filtered to earth. The $\frac{\lambda}{4}$ section may be used anywhere along a non-resonant transmission line with similar effect.

Both open and closed $\frac{\lambda}{4}$ resonant lines may be used as wave filters. Fig. 26 shows how more than one $\frac{\lambda}{4}$ line may be fitted to eliminate unwanted frequencies.

In this case a $\frac{\lambda}{4}$ section B open at the output end is placed in series with the transmission line. This will offer low impedance in the transmission line to the fundamental frequency. For each odd harmonic such a line is an odd multiple of a $\frac{\lambda}{4}$ wavelength and

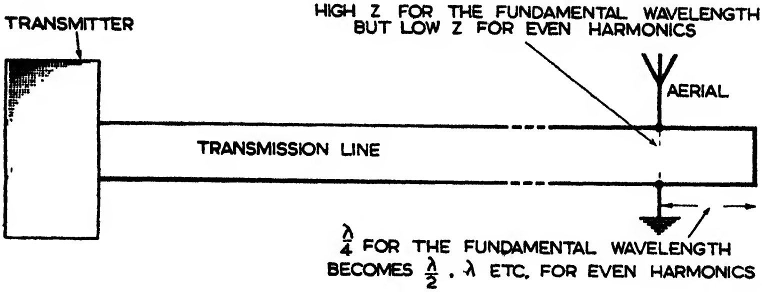


Fig. 24.—$\frac{\lambda}{4}$ STUB AS A HARMONIC FILTER.

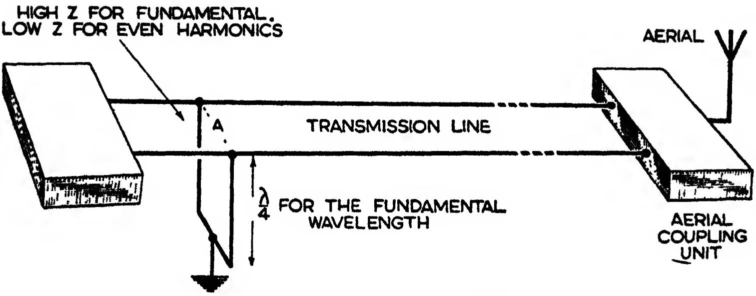


Fig. 25.—FILTER CONNECTED TO MAIN LINE.

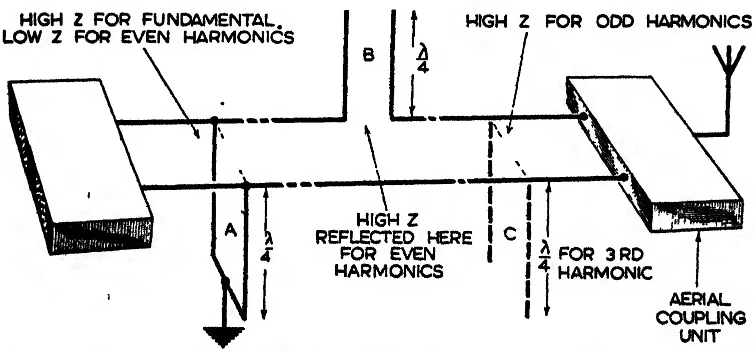


Fig. 26.—**VARIOUS ARRANGEMENTS OF FILTERS TO ELIMINATE EVEN HARMONICS.**

C demonstrates the impracticability of using the method for filtering out odd harmonics by means of an open stub.

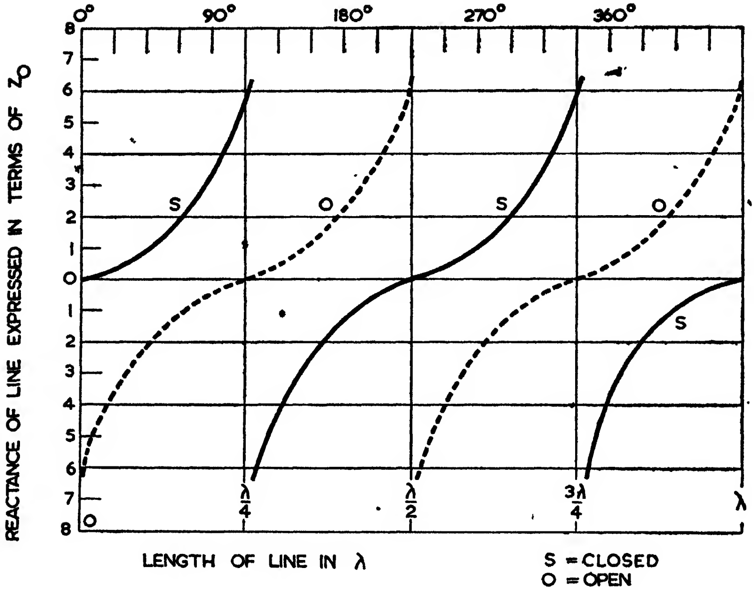


Fig. 27.—REACTANCE PLOTTED AGAINST ELECTRICAL LENGTH OF LINE MEASURED IN TERMS OF WAVELENGTH FOR OPEN AND CLOSED ENDS.

offers little impedance to the odd harmonics. Thus the $\frac{\lambda}{4}$ open filter B passes the fundamental and odd harmonics along the transmission line to the aerial coupling unit. As for the even harmonics, however, the length of the open line becomes a half wave or some multiple of a half wave so that line B offers a high impedance to the even harmonics and blocks their passage to the antenna coupling unit.

The joint action of the two quarter-wave lines is shown in the following table :—

			Harmonics			
$\frac{\lambda}{4}$ line	Connection	Fundamental	2nd	3rd	4th	5th
Open .	Series	Low Z	High Z	Low Z	High Z	Low Z
Shorted .	Shunt	High Z	Low Z	High Z	Low Z	High Z

Unfortunately, the method cannot be used to eliminate odd harmonics because any attempt to eliminate the odd harmonics must result in loss of the fundamental frequency.

For example, assume C in Fig. 26 is $\frac{\lambda}{4}$ at the third harmonic, viz., 15 megacycles. This frequency would be eliminated effectively before it reached the aerial. But the fundamental frequency to be transmitted would also be greatly attenuated, because a $\frac{\lambda}{4}$ line at 15 megacycles is a $\frac{\lambda}{12}$ line at 5 megacycles. Such a section would act as a condenser and offer a fairly low impedance to 5 megacycles, therefore, although the 15-megacycle frequency would have been eliminated, the 5-megacycle frequency would be considerably attenuated. Other types of filters are available, however, to meet this difficulty.

As a Reactance

Fig. 27 shows reactance is plotted against wavelength for closed and open lines. Note that reactance is zero for all multiples of $\frac{\lambda}{4}$.

The open-end line acts as a capacitative reactance at frequencies below the resonant frequency, causing the current to lead for electrical lengths shorter than $\frac{\lambda}{4}$ At $\frac{\lambda}{8}$ the capacitative reactance is equal to Z_0. Therefore an open line of less than $\frac{\lambda}{4}$ may be used to provide effects similar to that of a condenser termination, particularly when it is desired to avoid the use of actual condensers in the open air. At the exact $\frac{\lambda}{4}$ or odd multiples thereof, the open line has zero reactance. As the length of the line is increased beyond $\frac{\lambda}{4}$, alternatively if the frequency is increased above resonant frequency, the line exhibits inductive reaction, causing the current to lag. Finally, as a half wave or multiple half wavelength, the open end has infinite reactance.

The closed-end line acts as an inductive reactance at lengths shorter than $\frac{\lambda}{4}$ *i.e.*, for frequencies below the resonant frequency. At $\frac{\lambda}{8}$ it has X_L numerically equal to Z_0 and at exactly $\frac{\lambda}{4}$ or multiple thereof it has infinite reactance. Between $\frac{\lambda}{4}$ and $\frac{\lambda}{2}$ *i.e.*, at frequencies above the resonant frequency, the line exhibits capacity reactance and at exactly $\frac{\lambda}{2}$ or multiple thereof the line has zero reactance.

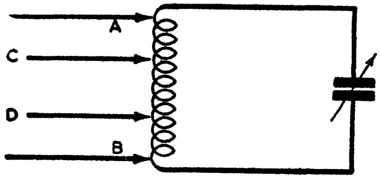


Fig 28.—EXAMPLE OF A PARALLEL RESONANT CIRCUIT COMPOSED OF A COIL AND A CONDENSER, EMPLOYED TO MATCH LINE CD TO AB.

Conditions for Matching

It has been shown that when the load impedance is not exactly matched to the line impedance, standing waves appear along the line. In order to obtain a good match, however, it is not sufficient that the numerical values of the load and line impedances should be equal.

Despite numerical equality, if reactive components are associated with the load, standing waves will still be present along the line.

The resistive and reactive components of a load and the phase-angle can be determined when the following information has been obtained :—

(*a*) The ratio $\frac{\text{voltage maximum}}{\text{voltage minimum}}$ (known as the standing wave ratio or S.W.R.).

(*b*) The positions of the voltage anti-nodes and nodes. (Refer Figs. 21 and 22 and relative text.)

For procedure in obtaining measurements see Appendix IV.

The effects of residual capacitative or inductive reaction of the load can be tuned out by the use of tuned lines.

If a line terminated in a resistance equal to Z_0 is checked for match over a wide frequency range, the load will look like a pure resistance at one frequency only, *i.e., the resonant frequency of the resistance.* Above or below this frequency the load will exhibit capacitative or inductive reactance.

In consequence of this, when it is desired to match line and load at some given frequency, other than the resonant frequency, a resistive load may look like a resistance paralleled by a condenser. Thus, in order

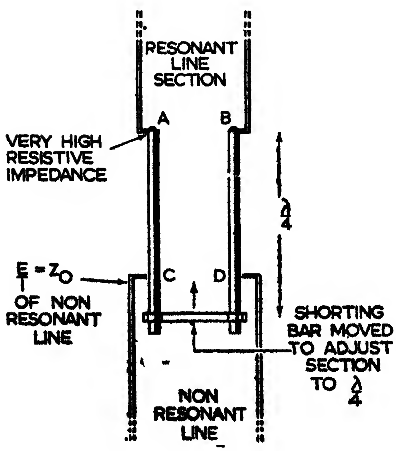


Fig. 29.—EXAMPLE OF SHORTED $\frac{\lambda}{4}$ SECTION EMPLOYED TO MATCH CD TO AB AS IN FIG. 28.

to eliminate the effects of capacity reaction, a suitable inductive section of line must be connected as an extension to the line, and in series with it *beyond* the load.

If this additional section is correctly dimensioned, it will resonate with the capacity component of the load, and the termination to the line then looks like a parallel tuned circuit.

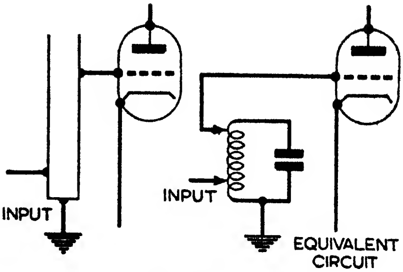


Fig. 30.—IMPEDANCE MATCHING.

Matching input to grid of valve by $\frac{\lambda}{4}$ line. This involves transformation from low impedance input from the source to a high impedance input to the grid.

Thus, instead of looking into a load with shunt capacity, the termination now looks like a resistive load shunted by a very high pure resistance. The latter has very little effect on the total resistance ; therefore the line is, to all intents and purposes, matched to its terminating load and all standing waves are eliminated. This point will be referred to again in a later paragraph headed " Final Matching by means of Stubs."

As Impedance Matching Devices

Circuits using coils and condensers frequently employ parallel resonance circuits as matching devices (see Fig. 28).

If a voltage at resonant frequency is applied across AB, the parallel circuit presents maximum impedance to the line equivalent to a pure resistance. If the voltage is applied across CD the circuit remains approximately in resonance but the impedance presented to the line is considerably lower.

The impedance of a shorted $\frac{\lambda}{4}$ transmission line is zero at the shorted end, increasing to a maximum toward the open end. If, for example, it is required to match the non-resonant line connected at CD to the resonant line AB, the arrangement shown

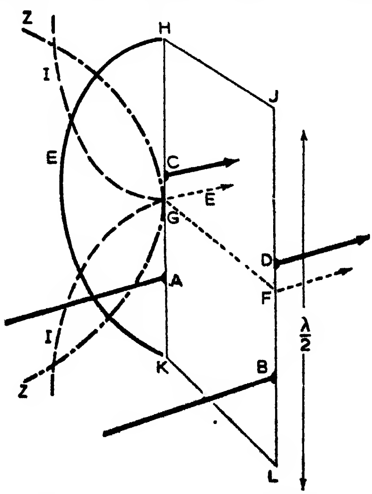


Fig. 31.—SHORTED HALF-WAVE MATCHING TRANSFORMER.

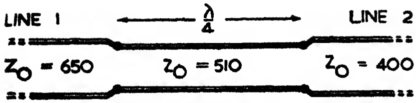


Fig. 32.—$\frac{\lambda}{4}$ MATCHING SECTION $Z_0 = \sqrt{Z \text{ LINE 1} \times Z \text{ LINE 2}}$.

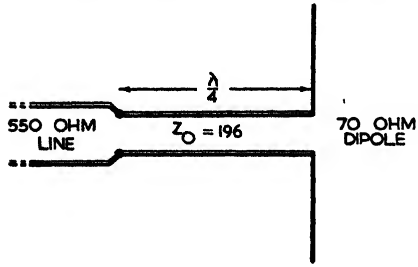


Fig. 33.—$\frac{\lambda}{4}$ MATCHING SECTION $Z_0 = \sqrt{Z \text{ LINE} \times Z \text{ CENTRE OF DIPOLE}}$.

in Fig. 29 can be employed.

The shorting bar is moved to adjust the matching section to $\frac{\lambda}{4}$ at the frequency of the resonant line AB. The tapping points CD are then moved closer to, or away from, the shorting bar, until the impedance at CD is equal to Z_0 of the non-resonant line.

The impedance presented to the resonant line is resistive since the matching section is equivalent to a parallel resonance circuit.

A second example of the use of a shorted $\frac{\lambda}{4}$ section is shown in Fig. 30. When used in this manner it is sometimes termed an impedance transformer. Here a relatively low impedance is transformed to a high impedance to suit the input grid conditions. The equivalent circuit is shown side by side for comparison purposes.

A half-wave section of line shorted at both ends may also be used as an impedance matching device.

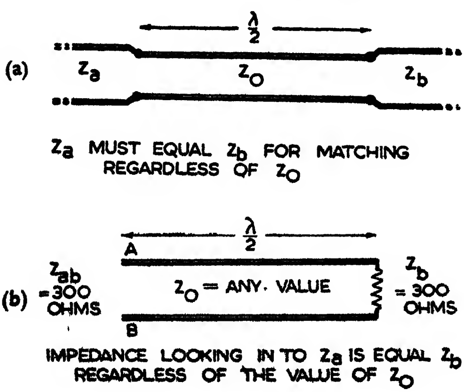


Fig. 34.—A HALF-WAVE MATCHING SECTION, Z_a INPUT = Z_b OUTPUT. IT REPEATS AT THE OUTPUT THE INPUT IMPEDANCE.

(a) Illustrates principle.
(b) Numerical example.

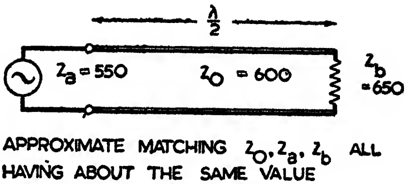


Fig. 35. — SPECIAL CASE OF MATCHING FOR A $\frac{\lambda}{2}$ SECTION.

Fig. 31 is a shorted half-wave matching section excited at AB with resonant current and voltage values as shown by the continuous and dotted curves. Z is shown by the dot-dash curve.

The input sees an impedance equal to $\frac{E}{I}$ at the input, *i.e.*, Z_{AB}, and the load "looking in" to CD sees a larger ratio, *i.e.*, Z_{AC}. Minimum impedance is at GF where voltage is highest and current lowest. Maximum impedance is at HJ and KL.

The upper half repeats the lower half, so that there are always two points on the frame having the same impedance. The current in one half, however, will be 180° out of phase with that in the other half.

$\frac{\lambda}{4}$ Open Line as Impedance Matching Devices

If $\frac{\lambda}{4}$ section is closed at one end by an impedance Z_b, the input impedance is

$$(1)\ Z\ \text{input} = \frac{(Z_0)^2}{Z\ \text{load}}$$

or

$$(2)\ Z_0 = \sqrt{Z_a \times Z_b}$$

where Z_a = input impedance and Z_b = output or load impedance.

When it is desired to couple two transmitting lines having different characteristic impedances, a $\frac{\lambda}{4}$ line having a characteristic impedance determined by $Z_0 = \sqrt{Z_a \times Z_b}$ can be used.

Let the line in Fig. 32 have $Z_0 = 650$ ohms and line 2 have $Z_0 = 400$ ohms. They can be coupled by a section where

$$Z_0\ (\tfrac{\lambda}{4}\ \text{section}) = \sqrt{Z_a \times Z_b}$$
$$= \sqrt{650 \times 400}$$
$$= 510\ \text{ohms.}$$

Line 1 sees an impedance of $\frac{(510)^2}{400} = 650$ ohms

Line 2 sees an impedance of $\frac{(510)^2}{600} = 400$ ohms.

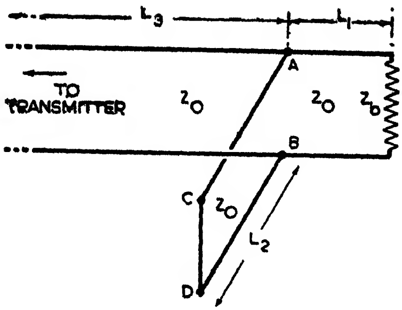


Fig. 36.—LOCATING POSITION AND LENGTH OF TUNING STUB TO TUNE OUT REACTIVE COMPONENT OF A LOAD, THEREBY REDUCING S.W.R. TO THE MINIMUM.

Thus line 1, looking towards line 2, sees no change in impedance, similarly line 2, looking toward line 1, is equally satisfied.

Matching line to aerial (see Fig. 33),

$$Z_0 = \sqrt{550 \times 70}$$
$$= 196 \text{ ohms approximately.}$$

Matching with Half-wave Sections (Non-shorted)

If a half-wave section is used instead of $\frac{\lambda}{4}$ section, $Z_a = Z_b$ regardless of Z_0. Therefore a half-wave section acts as a 1 : 1 transformer, repeating at one end the impedance seen at the other end.

From the above, and a consideration of Fig. 34, it can be seen that a half-wave non-shorted section cannot be used to match widely different impedances, but it can be used without regard to Z_0 as long as the two terminating impedances are alike.

The only exception is when Z_0, Z_a and Z_b are all approximately of the same value. The match then obtained is sufficiently close for most purposes (see Fig. 35).

Final Matching by Means of Stubs

In a previous paragraph it was shown that it frequently happens, when an aerial is fed by a fairly long transmission line with suitable characteristics, some standing waves still exist. Since standing waves represent losses, it is desirable, in such cases, to eliminate them entirely. For this purpose a matching stub is used.

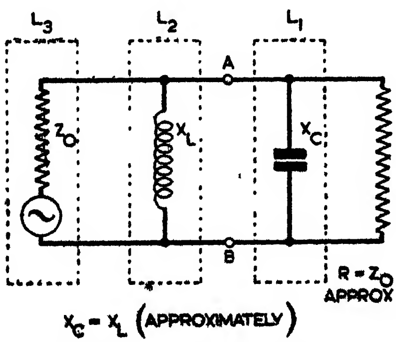


Fig. 37.—EQUIVALENT CIRCUIT OF FIG. 36.

The length of the stub, which is generally less than $\frac{\lambda}{4}$, and its position along the line may both vary in individual aerial systems.

Whilst it is possible to calculate the length and position of the stub with some degree of accuracy,

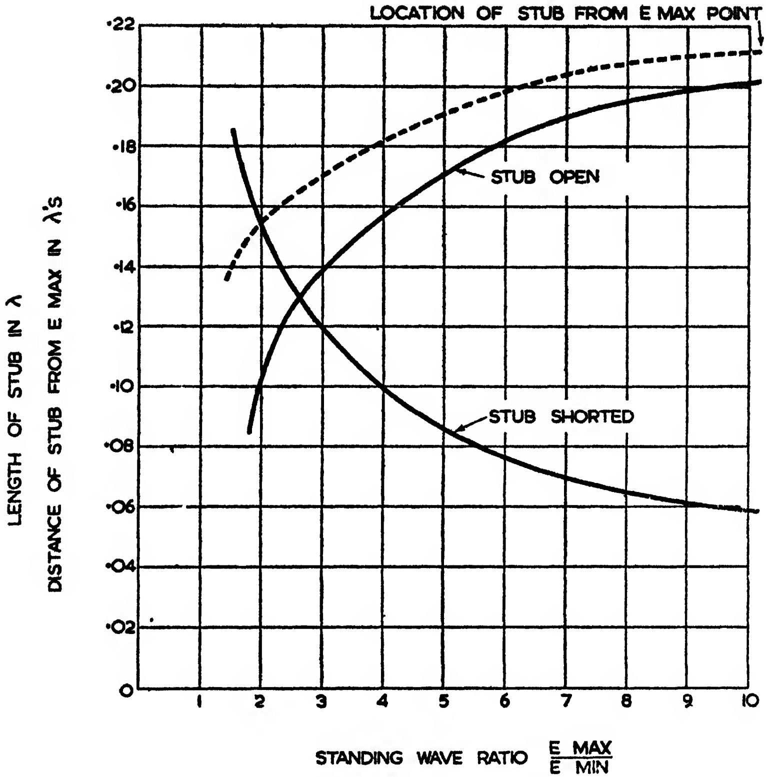


Fig. 38.—DETERMINATION OF STUB LENGTH AND POSITION ON THE LINE FOR CORRECT IMPEDANCE MATCHING.

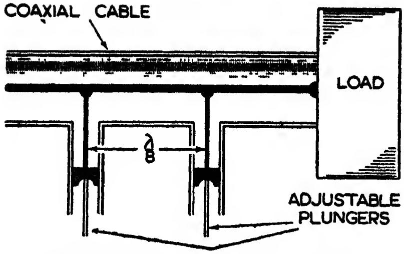
Length of stub in terms of λ and distance of stub from first E max. in terms of λ plotted against S.W.R.

the final adjustment is determined by trial for the complete elimination of all standing waves. The general principles of stub matching or tuning are, however, established :—

Open or shorted stubs may be used, but a shorted stub is generally preferred, because it reduces radiation losses, also it is easier to adjust and to support mechanically.

The problem is to locate the stub on the main line (see Fig. 36) with such spacing from the load end that the main line sees an impedance, looking into AB, which is in effect a pure resistance (or a parallel resonant circuit).

The impedance at AB is made up of L_1 and L_2 in parallel. L_1 is of such length that its impedance at AB, made up of the load impedance and the characteristic impedance Z_0 of L_1 in com-



STUB MATCHING FOR COAXIAL LINES

Fig. 39—MATCHING STUBS FOR COAXIAL LINES ADJUSTED BY TUNING PISTONS.

bination, has a resistive component equal to Z_0 plus some reactive component.

This reactive component shown as X_L (Fig. 37) is present at AB even if the load impedance is a pure resistance, since this impedance is not matched to L_1.

The stub (normally less than $\frac{\lambda}{4}$), acts almost as a pure capacity reactance. Therefore, the stub length can be adjusted at C and D to resonate with the reactance component due to L_1, the result is that a resistance of Z_0 remains across AB and the line L_2 is matched.

The equivalent circuit is shown in Fig. 37. The generator must be matched to the line, and line and generator can therefore be represented by L_3.

It is difficult to find experimentally the length L_1 which will present across AB a resistive component $= Z_0$. Also, any subsequent adjustment of length of L_2 would require a readjustment to L_1, L_1 and L_2 being complementary.

The method adopted, therefore, is to find :—

(1) The voltage maximum point nearest the aerial array where there is a voltage minimum, the standing wave ratio of which is $\frac{E_{max.}}{E_{min.}}$

By use of Fig. 38 both the stub location and stub length in wavelength can then be determined with reasonable accuracy.

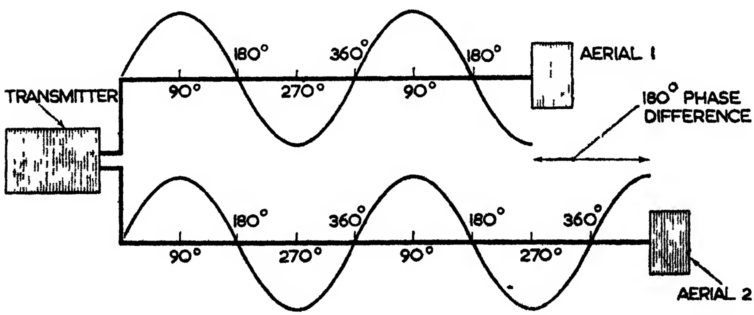


Fig. 40.—USE OF LINES OF UNEQUAL LENGTH TO OBTAIN A PHASE DIFFERENCE AT THEIR TERMINATIONS.

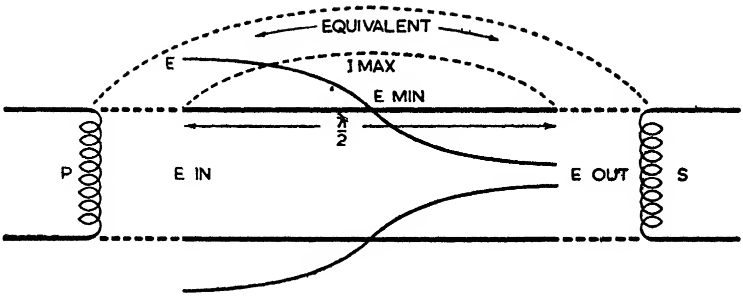


Fig. 41.—PHASE INVERSION BY $\frac{\lambda}{2}$ SECTION OPEN END WITH EQUIVALENT ELEMENTS, *i.e.*, PRIMARY AND SECONDARY OF A 1 : 1 TRANSFORMER.

Under some conditions it is easier to locate current maxima and minima. These serve equally well as long as it is noted that a current maximum coincides with a voltage minimum. (For method see Appendix IV.)

Stub Matching with Coaxial Lines

To avoid the problem of moving one stub along the line in search of the correct location, two stubs are used, their effective lengths being controlled by pistons. The second stub adds capacity or inductance to the line and thus varies the position of the standing waves. This produces the same effect as if the first stub was moved along the line. The two stubs are spaced either $\frac{\lambda}{8}$ or $\frac{3\lambda}{8}$ apart and near the load as shown in Fig. 39. Tuning is then carried out by piston adjustment.

As Phase Shifting and Inverting Device

Transmission lines may be used when it is desired to produce a phase difference between two aerials for example. This is done in some cases by employing paths of different lengths between transmitter and aerial.

This method is possible because there is a phase delay of 360° per wavelength, progressively, along a non-resonant line. For example, if one line is $\frac{\lambda}{8}$ longer than another the phase angle is 45°. Fig. 40 shows two aerials with

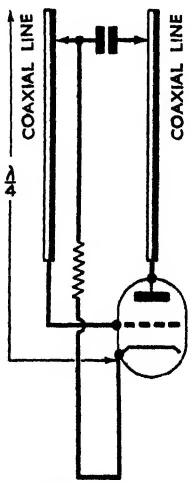


Fig. 42.—SIMPLE OSCILLATOR USING QUARTER-WAVE SECTIONS OF COAXIAL LINE TO CONTROL THE FREQUENCY.

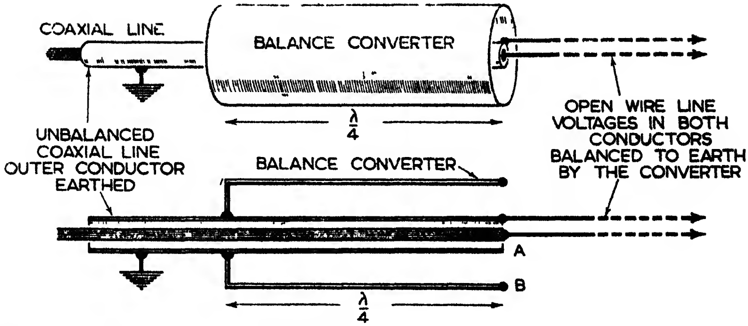


Fig. 43.—THE BALANCE CONVERTER OR BAZOOKA.

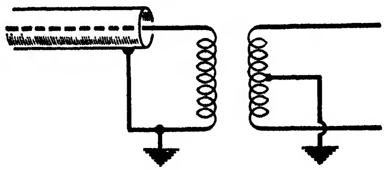


Fig. 44 (*right*)—EQUIVALENT CIRCUIT FOR FIG. 43.

180° phase difference since the line to aerial 2 is $\frac{\lambda}{2}$ longer than that to aerial 1. The accuracy of phasing in this manner depends upon the absence of standing waves.

A half-wave line can act as a phase inverter and a 1 : 1 transformer in which the secondary voltage is 180° out of phase with the primary voltage (Fig. 41).

As Frequency Control of an Oscillator

Since $\frac{\lambda}{4}$ short-circuited lines exhibit voltage and current distribution and, therefore, impedance conditions similar to those of a parallel resonance circuit, they may be substituted in place of conventional coils and condensers in oscillator circuits.

The advantages are :—

(*a*) High Q.
(*b*) Low loss.
(*c*) Oscillations at frequencies above those obtainable with tank circuits containing condensers and coils.

Since $Q = \frac{\omega L}{R}$ when R is low, Q is high, also good stability depends upon the value of Q.

The Q of a concentric line is at a maximum when the ratio of the inner diameter of the outer conductor to the outer diameter of the inner conductor is $= 4 ; 22$. The Q for a two-wire line is at maximum when the ratio of the spacing of the wires to the radius of 1 wire is 6 : 19. The Q of any line is also best at an effective electrical length of $\frac{\lambda}{4}$. This value of Q may be increased for ultra-high frequency by silver plating the lines.

In a single-valve oscillator the concentric line is frequently used. For push-pull circuits a balanced two-wire line or two concentric lines may also be used. The concentric line has the advantage that the inner conductor is shielded by the outer and this reduces radiation losses.

A simple form of line oscillator is shown in Fig. 42. It is tuned by the capacity-shorting bar and oscillates readily at the highest frequency at which the valve is capable.

As Line Balance Converter

The junction between a coaxial line unbalanced to earth and a pair of balanced lines cannot be made directly since there is a discontinuity in the electrical characteristics of the line.

This would result in an excessive standing wave ratio and loss by radiation because of unbalanced line currents. To effect a junction, a line balance converter is used.

The outer skin of the outside conductor of a coaxial line is at earth potential, whereas the outer skin of the inner conductor is well above earth potential and displays a high impedance to earth.

Both conductors of the balanced line, however, display the same potential and the same impedance to earth under ideal conditions.

The object of the converter is to produce a high impedance to earth between the outside of the inner conductor, at the point where it connects to one side of the balanced line, and thus to convert the end of the concentric line to a balanced line condition.

To achieve this a $\frac{\lambda}{4}$ shield is placed round the end of the coaxial line. This auxiliary shield is connected to the outer conductor of the coaxial line $\frac{\lambda}{4}$ from the end of the line, and is bonded firmly to the outside of the outer conductor of the coaxial line. Thus shield and conductor form a $\frac{\lambda}{4}$ section shorted at one end. Since a short-circuited $\frac{\lambda}{4}$ section displays a high impedance between A and B, point A is isolated from the earth, and B can be connected to one of the wires of the balanced line, both conductors of which have a high impedance to earth.

Fig. 44 shows the equivalent circuit. The action of this arrangement is, therefore, similar to that of a 1 : 1 transformer having a primary earthed at one end and the secondary earthed at the mid-point of the secondary winding.

APPENDIX II

THEORY OF WAVE-GUIDES

AT frequencies in the region of 3,000 megacycles, wave-guides become a practical proposition and are preferred to coaxial lines for energy transmission because circuit losses can be greatly reduced by their employment (attenuation for similar frequencies in a wave-guide is approximately $\frac{1}{50}$ of that for a coaxial line).

Whilst a tube of square, circular or elliptical cross-section is known to British technicians as a wave-guide,* the fact must not be lost sight of that a two-wire transmission line, a coaxial line, or any pair of conductors, may be regarded as performing a similar function in perhaps a somewhat different manner.

Actually, any surface which separates entirely, two different regions of electrical properties, can exert a guiding effect on electromagnetic waves. Such a surface may be one which separates a conductor from an insulator or one which separates two insulators.

By definition, therefore, the conductors of any electric circuit are wave-guides, since electric and magnetic fields may be propagated along them when a voltage is applied across the input terminals.

Examination of the general theory of energy transfer by transmission line was carried out, in Appendix I, by analysis of the voltage and current distribution along a two-wire transmission line.

It was pointed out that variations of the electric field between the wires, and of the magnetic field linked with the wires, were associated with waves of voltages and currents all along the line.

Thus, energy input from the generator may be regarded as being transported along a transmission line to the load by a series of transformations, alternating between an electric condition and a magnetic condition at intervals of $\frac{\lambda}{4}$ along the line.

It is, therefore, possible to think of the progress of electric energy along a line as an initial value of $\frac{1}{2}CV^2$ changing at $\frac{\lambda}{4}$ to a value of $\frac{1}{2}LI^2$, the latter value being numerically less than $\frac{1}{2}CV^2$ by an amount (I^2R + radiation and other losses), for each successive transformation.

This means that the operating conditions of a two-wire transmission line could have been examined by analysing the distribution of the electric and magnetic fields instead of by the conventional method applied.

* Definitions relating to wave-guides are standardised in Supplement No. 1 (1948) to British Standard 204 : 1943.

Analysis of Electromagnetic Fields

If the electromagnetic fields are regarded as the agents by which the input energy is conveyed to the load, the movements of the electrons, which also take place in the line relative to matter, may be thought of as manifestations of the changing electromagnetic fields which occur with each successive energy transformation.

This is only an inversion of the convention that electric fields appear when free electrons are displaced, and that magnetic fields demonstrate themselves when free electrons are in motion.

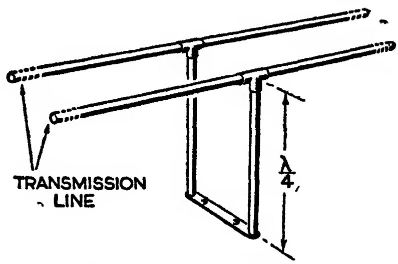


Fig. 1.—SECTION OF TRANSMISSION LINE SUPPORTED UPON $\frac{\lambda}{4}$ METALLIC INSULATOR.

If the first hypothesis is considered, it must be concluded that the function of the wires is to guide the energy from input to load, in which case the electrons constitute the link between the material of the wires and the electromagnetic fields.

Thus it is possible to think of a wave-guide as an improved form of transmission line, meaning that a two-wire transmission line may be regarded as an elementary and somewhat inefficient wave-guide.

If then the function of a two-wire transmission line is now deemed to be merged in the wave-guide as shown in Fig. 3 and, since energy is now to be guided by a single conveyer, analysis of the electromagnetic fields must take the place of the more conventional analysis of voltage and current distribution as applied to lines. In other words, the electromagnetic fields are the primary consideration, the movement of electrons being incidental thereto and generally of little importance in wave-guide analysis.

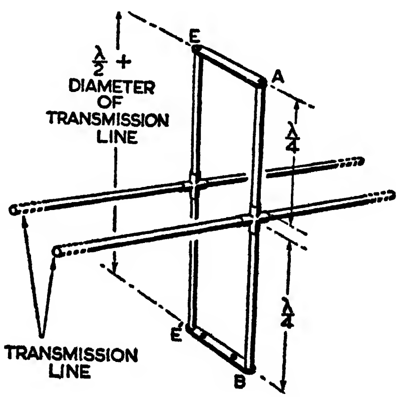


Fig. 2.—AS IN FIG. 1 BUT WITH A SECOND $\frac{\lambda}{4}$ SHORTED SECTION OF METALLIC INSULATOR INVERTED ABOVE THE FIRST SECTION.

Fig. 1 shows a shorted $\frac{\lambda}{4}$ section acting as an insulating support to a two-wire transmission line. Fig. 2 shows another $\frac{\lambda}{4}$

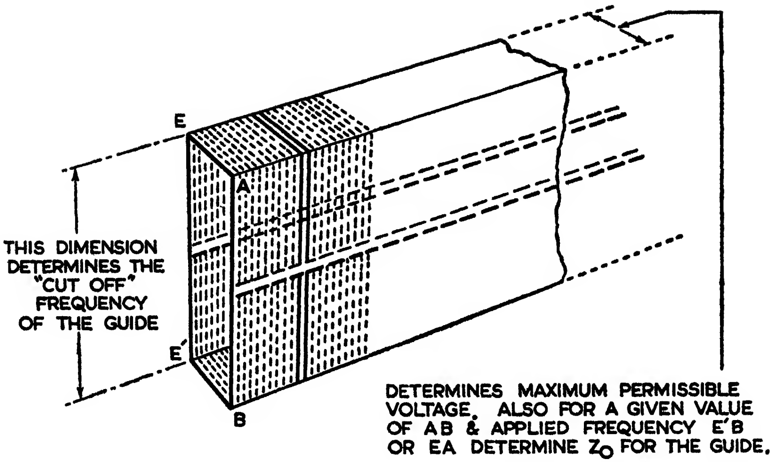


Fig. 3.—SECTION OF WAVE-GUIDE DEVELOPED FROM AN INFINITELY GREAT NUMBER OF ELEMENTARY FRAMES AS SHOWN IN FIG. 2, THE TWO PRINCIPAL DIMENSIONS BEING EA AND AB.

section inverted over the first section at the point of junction with the transmission line. The transmission line is now sandwiched between the two $\frac{\lambda}{4}$ sections and can be supported and insulated equally well by the top EA, or bottom section E^1B, or by both.

Fig. 6 (c) demonstrates that if a large number of these double $\frac{\lambda}{4}$ sections are assembled along a line as in Fig. 3 and welded together, a rectangular wave-guide is formed, wherefore each double $\frac{\lambda}{4}$ section with its sandwiched length of transmission line may be regarded as an elementary section of a wave-guide (Fig. 4 (1)).

From Figs. 4 (2) (3) and (4) it can be seen that the dimension AB of the wave-guide cannot be less than half a wavelength. In fact it must be a little more, in order completely to accommodate the transmission line function and at the same time to preserve the insulating properties of the resonant $\frac{\lambda}{4}$ sections. Clearly, any frequency lower than that which makes the AB dimension of the frame less than $\frac{\lambda}{2}$ will cause it to become an inductive shunt to earth.

The frequency therefore that makes the dimension AB $= \frac{\lambda}{2}$ is the " cut off " frequency for a particular wave-guide.

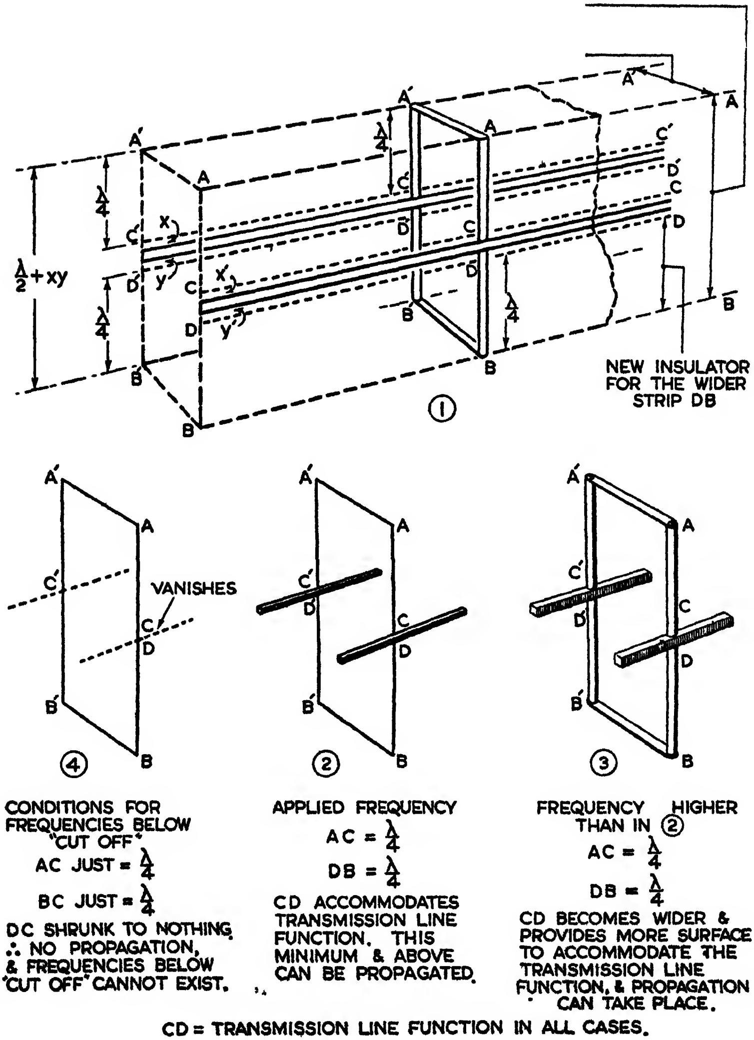


Fig. 4.—DIAGRAMS TO ILLUSTRATE THE FACTOR THAT GOVERNS THE FREQUENCY "CUT-OFF" FOR ANY GIVEN WAVE-GUIDE, *i.e*, THE DIMENSION AB.

Also illustrates the fact that any given guide can accommodate frequencies higher than the fundamental or "cut-off" frequency.

The wave-guide will, however, readily transmit frequencies higher than the "cut off" frequency, since the transmission line function can then be regarded as occupying the wider section CD, thereby shrinking the lengths A^1C^1 and AC, D^1B^1 and DB to $\frac{\lambda}{4}$ sections at the higher frequency, the nett result being that in order to make A^1C^1, AC and D^1B^1DB into $\frac{\lambda}{4}$ sections at the higher frequency, the transmission line function occupies the wider strip C^1D, CD.

It can be shown that the electric and magnetic fields of wave-guides are confined entirely to the interior of the guide and that voltages and current exist on the inner skin only. The external parts of the guide may be earthed therefore at any point along its length just like a coaxial line. This leads to the conception of boundaries which is dealt with in a later paragraph.

Conditions within the Wave-guide

In the characteristic field associated with the conductors of a normal transmission line, the electric and magnetic vectors (E and H) exist at every point of the field in directions mutually perpendicular and themselves perpendicular to the direction of propagation of the electromagnetic energy. This is not true of wave-guides. In wave-guides, two distinct types of wave may be distinguished, *i.e.*, the H wave and the E wave. These are dealt with in detail in later paragraphs. It is sufficient for the moment to appreciate the fact that energy may be propagated in a wave-guide in a doubly infinite series of modes analagous, to some small extent, to a singly infinite series of modes represented by a fundamental note and its harmonics. The modes are distinguished by the patterns of the lines of force traversing the fields, and their order characterised by the number of pattern repeats obtained in traversing the guide.

For a guide of given dimensions and a wave mode of given order there exists, as has been shown above, a critical frequency below which the mode cannot be propagated in the guide. Waves of lesser frequency would be very rapidly and completely attenuated.

For example, in wave-guides of rectangular cross-section a given wave cannot be propagated if the frequency "f" is such that :—

$$f^2 < c^2\left\{\left(\frac{m}{2a}\right)^2 + \left(\frac{n}{2b}\right)^2\right\}$$

where c = the velocity of light *in vacuo* and a and b equal the length and breadth of cross-section of the guide respectively; m and n are two positive whole numbers characterising the mode.

In general, the higher the order of the wave mode, the higher is the "cut off" or critical frequency.

The launching of a given type (E or H) wave in a wave-guide is usually accomplished by the insertion of an emitter probe within the guide. The probe is fed with micro-wave oscillations of the proper frequency, and must be so placed inside the guide as to coincide with the direction of the electric lines of force in the pattern of the mode which it is desired to excite.

This effect is very similar to that which could be obtained by placing a metal tube of suitable dimensions over a vertical aerial. If the axis of the tube is placed parallel to the aerial, the magnetic field is bounded entirely by the walls of the tube and has no component in the direction of the axis, but the electric field has maximum intensity with direction along the axis. Since the magnetic field is transverse, by definition, an E wave can be propagated. If the tube is placed so that the aerial lies transversely across its axis, the above conditions are reversed and electromagnetic energy will be propagated along the axis as an H wave.

Wave-guide Losses

Open transmission lines suffer losses under three headings :—

(*a*) Radiation losses. (*b*) Dielectric losses. (*c*) Copper losses.

The coaxial line eliminates radiation losses, but experiences dielectric losses from the spacers which are generally employed to support the inner conductor. This loss does not occur in the wave-guide. Also the larger conducting surface that can be provided by a wave-guide reduces the copper loss.

Fig. 5 compares the efficiency of wave-guides to coaxial lines for attenuation at increasing frequencies. The attenuation of the wave-guide is very great in the region of the " cut off " frequency, but it is very much less than that for the coaxial line at the frequency for which the guide is designed.

The interior surface of the wave-guide must be clean and is generally silverplated, in order to obtain higher surface conductivity.

The dimension EA or E^1B determines the maximum permissible value of the R.F. voltage, since this must be less than the arc-over value. A rough interior surface, dents, careless soldering all tend to reduce the AE dimension and hence the power-handling capacity of the wave-guide system.

For any frequency that the wave-guide can transmit, the power-handling capacity of a wave-guide is considerably greater than that of a coaxial line of the same diameter. This is shown in Fig. 5. The dimension determining maximum permissible voltage for the wave-guide is more than double that of a coaxial line having the same diameter.

Perhaps the most outstanding electrical advantage of the tubular feeder (circular or rectangular wave-guide) over the usual transmission line, lies in its vastly superior attenuation frequency characteristics. A guide constructed of perfectly conducting walls

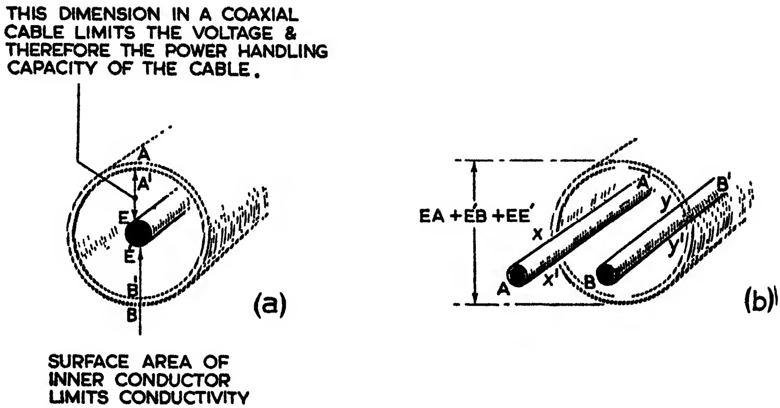


Fig. 5.—COMPARISON OF VOLTAGE-HANDLING CAPACITY AS DETERMINED BY ARC OVER CONDUCTIVE SURFACES FOR (a) COAXIAL LINE AND WAVE-GUIDE OF SIMILAR DIAMETER; (b) PHANTOM SURFACES PROVIDING THE TRANSMISSION LINE FUNCTION FOR A CIRCULAR WAVE-GUIDE.

Note. AA^1 and BB^1 are phantom sections of the guide. The width of this phantom section for any tube of given dimensions increases with frequency. Sections x to y and x^1 to y^1 automatically adjust themselves in this way to $\frac{\lambda}{2}$ for all frequencies higher than the fundamental or lowest frequency.

would have zero attenuation. The losses introduced by the imperfection of such a metal as copper as an electrical conductor are very small. Thus whilst a coaxial line will propagate 3,000 Mc/s waves with an attenuation of ·5 db per metre, a tubular feeder designed for the same frequency will have an attenuation factor of only ·03 db per metre. Q is of the order of 25,000.

Conditions for Detailed Analysis

Thus far it has been established that a frame having the dimension AB (see Fig. 3), a little greater than $\frac{\lambda}{2}$, can be thought of as an elementary section of a wave-guide. For analysis, the above assumption must be conditional upon the wave-guide being *closed at both ends*, therefore any comparison that is made between a transmission line and a wave-guide, using the "elementary frame" as a basis, must be made with a transmission line also closed at one end by the generator and shorted at the far end. In other words, the comparison must be based upon the common condition that maximum reflection is present, causing standing waves to appear. The reason for making this condition is to simplify the analysis by considering, firstly, corresponding conditions for standing waves and so to avoid the complications which are introduced by travelling waves and the complex phenomena by which they are progressed along the guide. Since under the above conditions the wave-guide is not delivering energy to a load, the manner in which energy is progressed along the wave-guide does not enter immediately into the discussion, thus enabling us to approach wave-guide technique along the similar and more familiar avenue of transmission-line technique. In making the analysis the following principles of wave-guide technique must be observed :—

Boundaries

(1) All fields must be continuous throughout the region in which the dielectric is the same. This means that the frequency everywhere along the guide must be the same.

(2) At the surface of a perfect conductor placed in an electromagnetic field that varies with time, the electric field is perpendicular to the surface, and the magnetic field is parallel to the surface, *i.e.*, an electric field which is equivalent to a voltage is short circuited when it exists across a perfect conductor. Also the magnetic field created by induced currents is equal and opposite to the exciting magnetic field, wherefore the field at the surface of this perfect conductor is zero.

Let the wave-guide section AA^1 CC^1 shown in side elevation at Fig. 6 (M) be closed at both ends and assume it to be excited in a mode producing a distribution of the electric and magnetic fields within the guide similar to that which would normally appear between shorted lines excited at the same frequency.

Also let line and guide be $\frac{3\lambda}{2}$ long at the excitation frequency. Then the equivalent transmission line can be represented by either of the closed sections A or F (they are identical).*

For purposes of comparison, therefore, let the wave-guide, section M, be represented by the transmission line A, supported at top and bottom by $\frac{\lambda}{4}$ sections, thus forming an elementary or skeleton wave-guide, as at C, and having certain characteristics similar to those of M. C is to be excited by some form of R.F. generator across ZY (not shown), at the same frequency as M.

Since C is short circuited at A^1B^1, and because the line is three half waves in length, reflection will occur at A^1B^1 and standing waves will be produced all along the line.

The voltage distributed along the line is shown at B and the

* There is one difference between the transmission line condition and wave-guide example, *i.e.*, the transmission line must be energised from one end which is closed through the generator, whereas the wave-guide is closed at both ends and energised by a probe, which is, in effect, an aerial or internal generator of electromagnetic waves within the guide.

current distribution is shown at G. Note that density and direction are both indicated.

The distribution of the electric field between the wires is shown at A, and is also represented by the curve B.

The distribution of the magnetic field linked with the wires is shown at F, and is represented by the curve G, which has as its base the top wire AA^1.

At C and H respectively, the electric and magnetic fields due to the standing waves of voltage and current are shown. The magnetic field is developed by combining the fields associated with the $\frac{\lambda}{4}$ sections with the fields of the transmission line wires $z\,B^1$ and $y\,A^1$. The reader can check this for himself by Fleming's rule.

K and L are end views of C and H respectively, looking towards the generator from A^1B^1, C^1D^1.

A key is supplied for reading these diagrams.

When analysing the development of the fields shown in C and H, the following previously mentioned rules must be borne in mind :—

(1) The electric field which is equivalent to a voltage is short circuited when it exists across a perfect conductor.

(2) Action and reaction are equal and opposite, thus currents induced by alternating electric fields set up in a perfect conductor are accompanied by magnetic fields which are equal and opposite.

Thus, in the case of the vertical components of the magnetic lines of force shown at F, they may not penetrate the walls of the wave-guide because of (2),* and for a similar reason the magnetic lines of force must follow a flux path parallel to the internal walls of the guide.

The electric and magnetic fields for the closed guide are 90° out of phase and at right angles to one another. This can be seen by comparing D with H and E with F.

Also along the guide lengthways there are magnetic components

* The walls of a practical guide are not, strictly, perfect conductors, but their resistance is such that they approximate very closely to that condition and to all intents and purposes act as though they were perfect conductors.

Fig. 6.—ANALYSING THE DEVELOPMENT OF THE ELECTRIC AND MAGNETIC FIELDS IN A WAVE-GUIDE (RECTANGULAR) FROM CONSIDERATIONS OF A TRANSMISSION LINE SUPPORTED BY AN INFINITELY LARGE NUMBER OF $\frac{\lambda}{2}$ FRAMES AND EXCITED BY AN EXTERNAL GENERATOR.

The length of line and wave-guide is $\frac{3\lambda}{2}$ for convenience. This length has no special significance. Both transmission line and guide are shorted at the far end for reasons given in text. The E vector is shown at (A) (C) (D) (E) and (K). The vertical component of the H vector is shown at (F), the component in direction of propagation at (J) and (L) and the resultant field at (H).

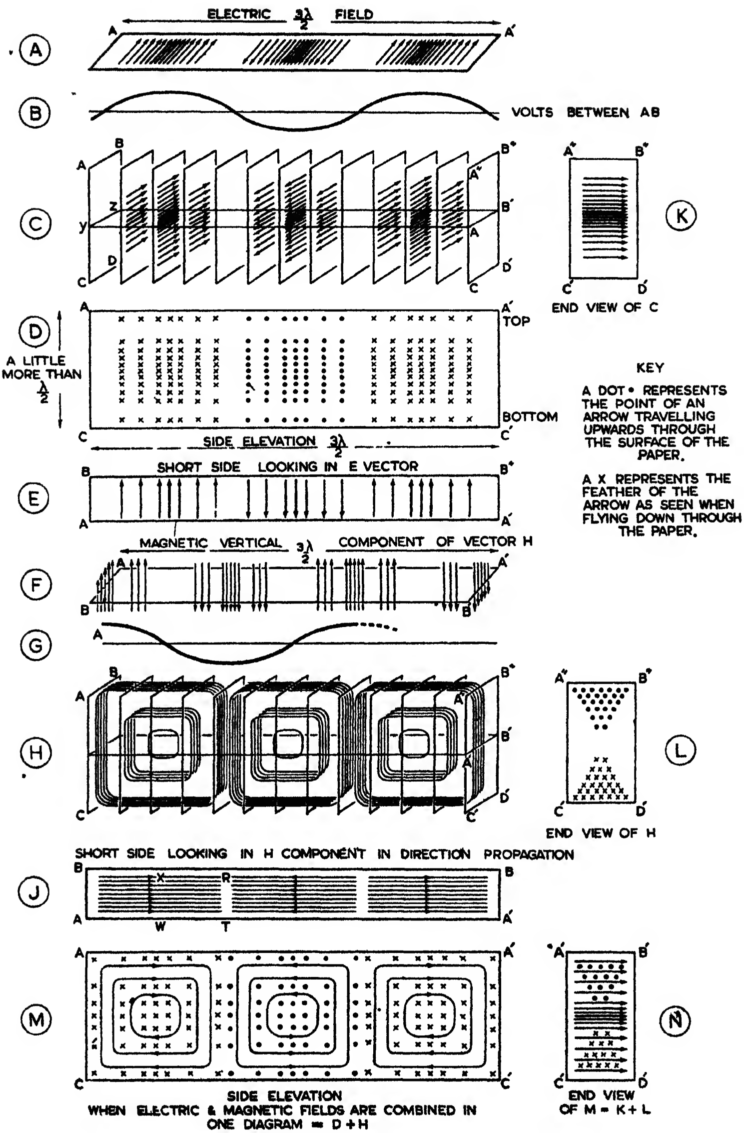

ELECTRIC $\frac{3\lambda}{2}$ FIELD
VOLTS BETWEEN AB
END VIEW OF C
A LITTLE MORE THAN $\frac{\lambda}{2}$
TOP
BOTTOM
SIDE ELEVATION $\frac{3\lambda}{2}$
KEY
A DOT • REPRESENTS THE POINT OF AN ARROW TRAVELLING UPWARDS THROUGH THE SURFACE OF THE PAPER.
A X REPRESENTS THE FEATHER OF THE ARROW AS SEEN WHEN FLYING DOWN THROUGH THE PAPER.
SHORT SIDE LOOKING IN E VECTOR
MAGNETIC VERTICAL $\frac{3\lambda}{2}$ COMPONENT OF VECTOR H
END VIEW OF H
SHORT SIDE LOOKING IN H COMPONENT IN DIRECTION PROPAGATION
SIDE ELEVATION
WHEN ELECTRIC & MAGNETIC FIELDS ARE COMBINED IN ONE DIAGRAM = D + H
END VIEW OF M = K + L

Fig. 6.

only (as in J). Across the guide, however, for the $\frac{\lambda}{2}$ dimension, there are electric and some magnetic components. These points are seen plainly on E and F respectively.

Distribution of the electric and magnetic waves that may exist in a wave-guide may take various forms, depending upon (*a*) the manner in which energy is introduced into the wave-guide, (*b*) the relations between wavelength and the dimensions of the guide.

The various field patterns of operation (or oscillation) are termed modes, and they are classified as E or H modes, according to the direction of their electric and magnetic components relative to the direction of propagation.

All H or TE waves have magnetic components only along the guide, the electric components being transverse across the guide.

All E or TM waves have electric components only along the guide with magnetic components transverse to it.

Identification of Modes

In this country the major characteristic used to distinguish one mode from another is that field, either E or H, which lies principally along the direction of propagation. In America, however, the major identification is stated by the field which is transverse or at right angles to the direction of propagation. Thus a TE mode (transverse electrical) corresponds to the British designation of an H mode; similarly a TM (transverse magnetic) mode corresponds to an E mode.

A mode is fully described by stating the identification letter or letters followed by two numbers. Thus H_{01} in the British system or TE_{01} in the American system.

In the British system, H indicates that the magnetic field lies along the direction of propagation and that the electric field is transverse or at right angles to it.

The first subscript number is the number of maxima of magnetic force which would be encountered on travelling across the guide (at right angles to the longer side).

The second number is the number of maxima of magnetic force which would be encountered on travelling across the guide (parallel to the longer sides).

In the American system, the first of the subscript numbers used in describing field configurations in a rectangular guide indicates the number of half-wave patterns of the transverse lines which exist along the short dimension of the guide through the centre of the cross-section. The second number indicates the number of transverse half-wave patterns that exist along the long dimension of the guide through the centre of the cross-section.

In both cases, when there is no change in the field intensity, a zero is used.

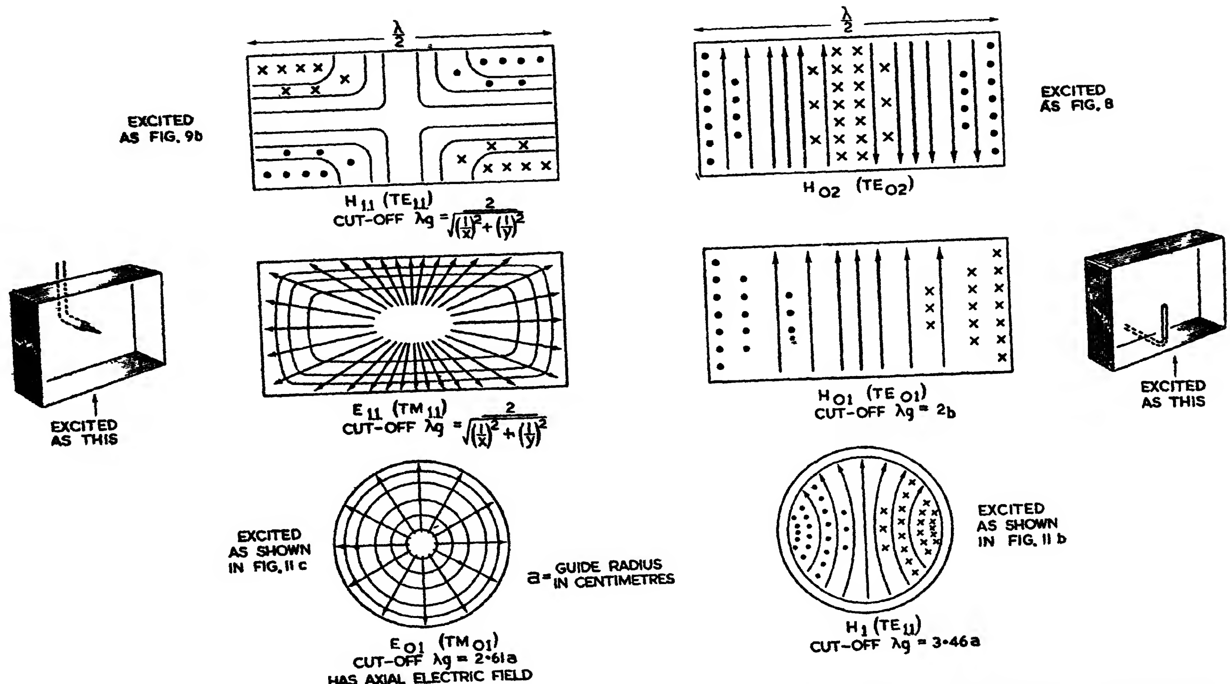


Fig. 7.—Classification of typical modes employed in rectangular and circular wave-guides illustrating methods for excitation of the mode and cut-off wavelength (in the guide).

Note. Other modes sometimes employed in rectangular guides E_{12} (TM_{12}) and H_{12} (TE_{12}).

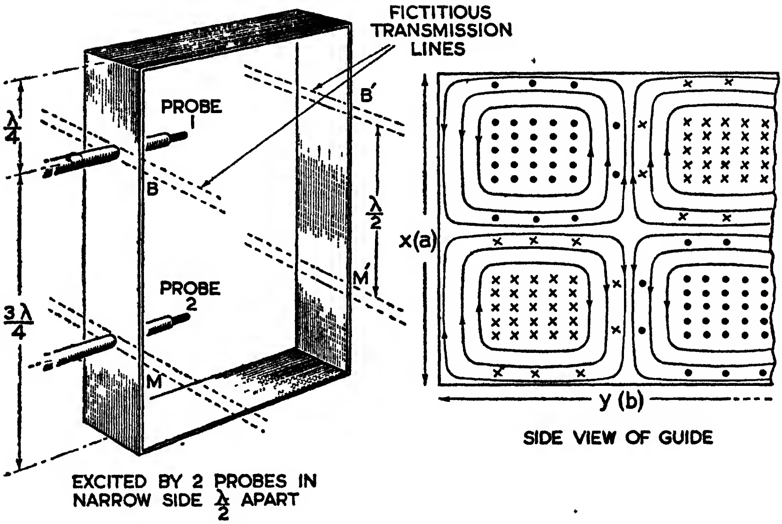


Fig. 8.—ANALYSIS OF A RECTANGULAR WAVE-GUIDE SHOWING METHOD OF EXCITATION FOR AN H_{02} (OR TE_{02}) WAVE OR PRINCIPAL MODE.

Thus H_{01} or TE_{01} indicates that the magnetic field lies along the length of the tube and that the electric field is therefore transverse. Travelling across the tube, Fig. 6, from A^1 to B^1, as in N, no change in intensity of the magnetic field is encountered.

For example, X to W in J, the field is uniform and again, from R to T in J it is also uniform, being zero everywhere along this line. This is again seen in N, although the number of lines shown along a straight path from A to B is different, there is no concentration such as may be encountered in going from C^1D^1 to A^1B^1 (in N). On this journey one definite maximum is encountered. Thus H_{01} or TE_{01} fully describes this mode.

When circular guides are employed, the same major classifications are employed but the first number indicates the number of whole-wave patterns of the lines encountered around the circumference of the guide, the second number indicates the number of maxima that exist along a diameter.

Development of Modes

Fig. 7 illustrates a number of other modes that can be developed for rectangular and circular wave-guides, together with an indication of the manner in which these modes may be excited and their individual cut-off frequencies.

Fig. 8 illustrates the development of a wave-guide for operation in the H_{02} mode. A two-wire transmission line is supported at B

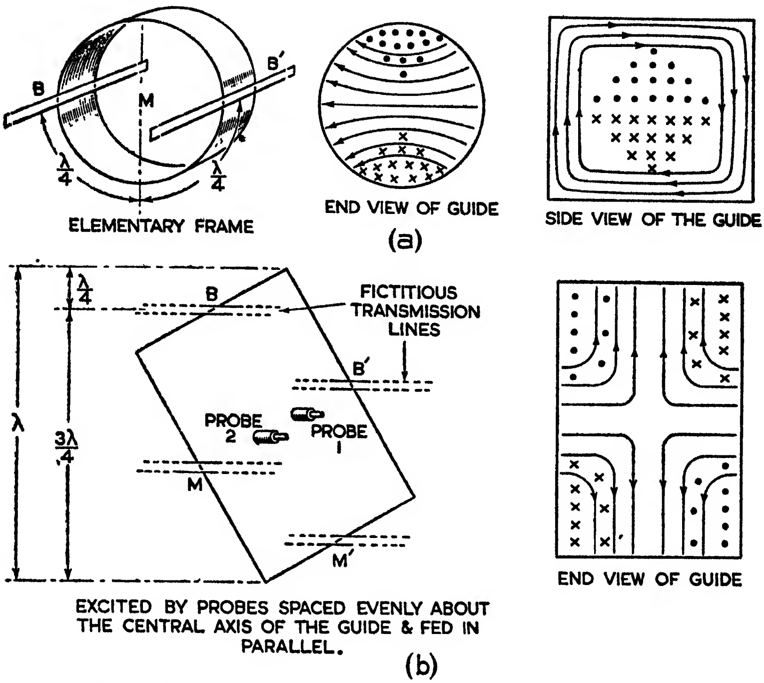


Fig. 9.—ANALYSIS OF EXCITATION CONDITIONS FOR H_1 AND $H_{1\,1}$ MODE IN CIRCULAR AND RECTANGULAR GUIDES.

(a) Analysis of circular guide for H_1 ($TE_{1\,1}$) mode.
(b) Analysis of rectangular wave-guide for $H_{1\,1}$ ($TE_{1.1}$) mode, also method of excitation.

and B^1 by a $\frac{\lambda}{4}$ insulator on the top and $\frac{3\lambda}{4}$ insulator on the bottom. If this arrangement is developed into a wave-guide by adding successive frames, the field pattern shown in Fig. 7 can be obtained.

For a guide of the same size as that in Fig. 8 (H_{02}), the cut-off frequency is doubled since one complete λ exists between top and bottom of the guide. Thus the guide may be considered to be made up from two transmission lines separated by $\frac{\lambda}{2}$ section and supported by a pair of $\frac{\lambda}{4}$ insulators at BB^1 and MM^1. Higher frequencies than cut off can be passed by this guide as the width of the transmission line elements at BB^1 and MM^1 can increase and so shorten both the insulators and separator to $\frac{\lambda}{4}$ and $\frac{\lambda}{2}$ respectively, at some higher frequency. Thus it will be seen that in general a wave-guide acts like a high pass filter.

Fig. 9 (a) shows the development of H_1 mode for a circular guide. Note the similarity between this mode and $H_{1.1.}$ mode, for a rectangular guide. The maximum wavelength that a circular guide can transmit is 1·70 times its diameter.

Fig. 9 (b) shows the development of $H_{1.1}$· mode for a rectangular wave-guide.

Another method of developing a guide is shown in Fig. 10. In this case the mode is E_0/TM_{01} for a circular guide. The distance X is extended to some multiple of $\frac{\lambda}{2}$ and standing waves are produced across it as well as along the side of the insulator, Fig 10 (2). If several frames are attached to small discs A and B, the pattern shown at 4 and 5 is obtained. Rectangular wave-guides for the $H_{1.1.}$ mode can be developed in a similar manner.

The most important modes used for transmission in circular wave-guides are H_0 and E_0.

H_0 has the advantage of decreasing attenuation with frequency increase, but H_1 can be used in a smaller tube for the same frequency.

The mode most frequently used in rectangular guides is probably H_{01}.

H_{01} (TE_{01}) has the following advantages :—

(1) It is the simplest form of possible wave.

(2) It uses the smallest tube for a given wavelength.

(3) Attenuation is less than for other modes.

(4) Most suitable for radiation from the open end of a guide.

Thus far, all comparisons between two-wire transmission lines and wave-guides have been made on the assumption that the transmission lines were shorted at both ends and the wave-guides were similarly closed. Under these conditions, in both cases, maximum reflection occurs and standing waves appear. With the closed transmission line, the energy input appears as antinodes of voltage and current along the line with a phase difference of 90°. The energy in the closed wave-guide may be thought of as oscillating between the electric and magnetic states, the time displacement of these two states being equivalent to a phase difference of 90°.

These conditions change when either the transmission line, or the guide, delivers energy to a load of impedance Z_0, capable of absorbing the energy input in such a manner that no reflection takes place.

Under these conditions the standing waves disappear in the direction in which the energy travels (towards the load), but can be considered as remaining on the insulating sections of the frame elements from which the wave-guide is assumed to have been developed. The general condition can be represented by moving the E line maxima along the guide until they coincide with the H

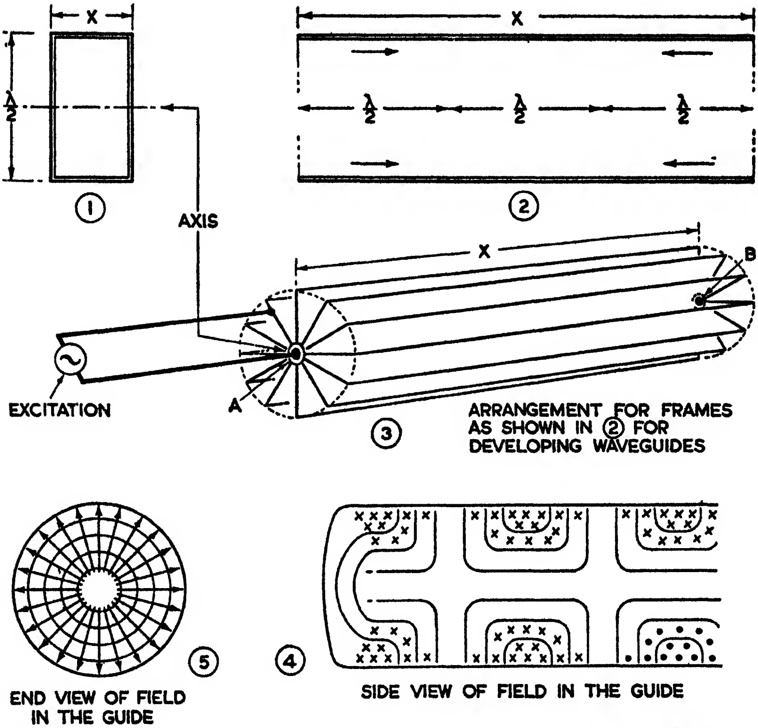


Fig. 10.—ANALYSIS OF A CIRCULAR GUIDE FOR E_0 (TM_{01}) MODE. METHOD OF EXCITING THE MODE IS ALSO SHOWN.

line maxima. This corresponds to the conditions surrounding a transmission line when voltage and current are in phase, *i.e.*, when the line is terminated in a pure resistance equal to Z_0 (characteristic impedance).

When the wave-guide is open at the far end it can, under favourable conditions of termination, radiate some of its energy into free space and, therefore, becomes loaded. In general, however, there remains a residual reactance, in consequence of which there is a certain standing wave ratio and phase angle both of which may be reduced by " tuning " by the application of principles similar to those employed with transmission lines. Unless the dimensions of the guide are such that only one mode can exist, several modes will be set up simultaneously in any guide when it is excited, although the desired mode is usually the strongest. When this occurs the wave may be purified by inserting suitable obstacles within the guide positioned to short out the undesired modes.

Pattern Modifications

Fig. 11 shows the pattern modifications for some of the modes when a guide is matched to a load. H_{01} mode for the closed guide (see Fig. 6) should be compared with H_{01} when energy is being transmitted to a load (Fig. 11).

When a wave-guide is terminated in a suitable load, groups of E and H components move down the guide towards the load in the form of electromagnetic waves. These waves are similar in composition to those radiated into free space by an aerial, but they are wholly within the guide. *Note.* The probe may be thought of as an aerial within the guide situated $\frac{\lambda}{4}$ away from the closed end or other surfaces which act as reflectors to reinforce the electromagnetic wave in the direction of its propagation.

Velocity and Propagation of Energy in Wave-guides

Whilst the energy is carried along the axis of the guide as a result of *group* movement, the electromagnetic waves themselves (group components) do not move down the guide in a straight line, they are progressed by a series of reflections from wall to wall of the guide in a manner similar to that in which the vertical components of electromagnetic waves in free space are progressed by reflection alternately from the ionosphere and the earth's surface.

Due primarily to the zig-zag motion of the electromagnetic wave along the guide, the field intensity variations (phase velocity) appear to occur with a velocity greater than that of light. This apparent violation of the laws of physics is explained by the hypothesis that the wavefront strikes the wall of the guide at an angle and is reflected at a similar angle.

In the time required for the wavefront to move the distance L (see Fig. 13) the point of reflection has moved the greater distance P. Thus the apparent speed (phase velocity) is greater than the true velocity of propagation.

The analogy of the progress of a caterpillar from point A to point B has been used with some success to explain these phenomena. The caterpillar moves along as a group, and the rate at which its whole body moves over the ground is its group velocity. In the process of moving little ripples or waves run up and down its back with a much greater velocity. This is its phase velocity.

This paragraph may be summarised as follows :—

The wavelength of the wave train pattern propagated in a wave-guide is greater than that of waves propagated in free space at the same frequency.

$\frac{1}{\lambda_g^2} = \frac{1}{\lambda_f} - \frac{1}{\lambda_0}$. where λ_g = wavelength in the wave-guide, λ_f = wavelength in free space and λ_0 = cut-off frequency. This implies

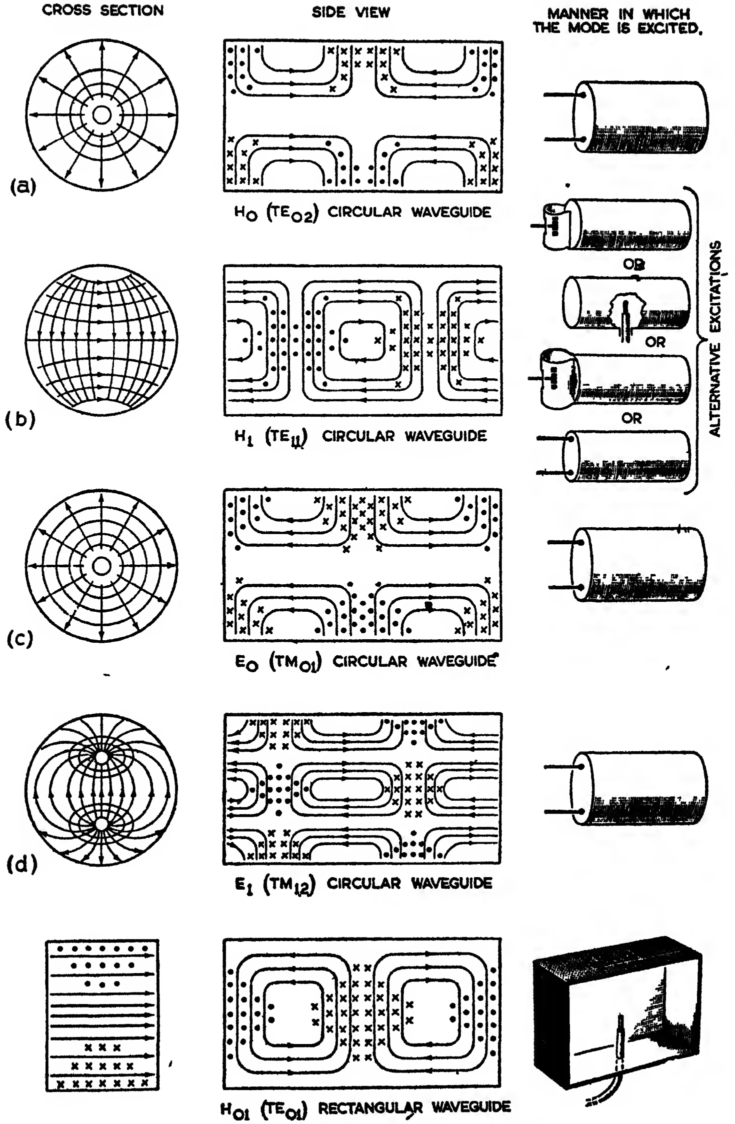
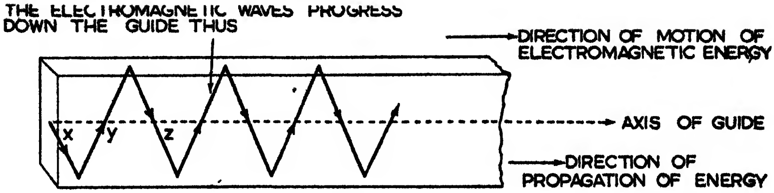


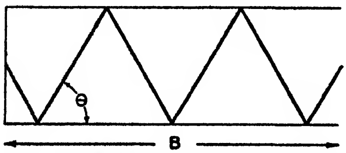
Fig. 11.—MODES IN WAVE-GUIDES (LOADED).

The waves are travelling down a guide terminated in Z_g. Note the change of pattern which takes place. The electric and magnetic fields now have their maxima together, but they are still in quadrature relative to the axis of the guide.



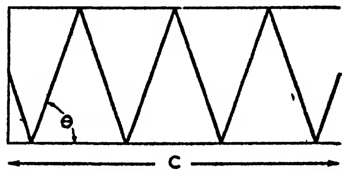
THE ELECTROMAGNETIC ENERGY MOVES DOWN THE GUIDE WITH VELOCITY V (LESS THAN LIGHT IN VACUO) ALONG ITS AXIS (DOTTED).
THE ELECTROMAGNETIC WAVE IN WHICH THE ENERGY RESIDES PROGRESSES DOWN THE GUIDE AS SHOWN BY x, y, z ETC. WITH A PHASE VELOCITY GREATER THAN THAT OF LIGHT IN VACUO.

(a)



THE FREQUENCY f_1 IS HIGHER AND λ IS SHORTER THAN IN (c). NOTE THE ANGLE θ IS SMALLER THAN IN (c) AND THERE ARE FEWER REFLECTIONS FOR EQUAL LENGTHS OF GUIDE.

(b)



THE FREQUENCY f_2 IS LOWER AND λ IS LONGER THAN IN (b). NOTE ANGLE θ IS LARGER AND THAT THERE ARE MORE REFLECTIONS REQUIRED TO COVER EQUAL LENGTHS OF GUIDE B AND C.

(c)

Fig. 12.—DIAGRAMS TO ILLUSTRATE THE PROPAGATION OF ELECTROMAGNETIC ENERGY, AND THE ELECTROMAGNETIC WAVES IN WHICH IT RESIDES, DOWN AN OPEN END GUIDE.

Note. The longer the wave the steeper the angle θ. The longer wave more nearly coincides in length with the y dimension of the guide, *i.e.*, at C.O. it will fill the guide. The shorter wave must expand before reaching the wave-guide walls or boundaries. This occupies time during which some movement along the guide takes place.

that the velocity with which the geometrical patterns are propagated in the direction of the axis of the guide (phase velocity = v) must be greater than that of light *in vacuo*. (This does not, however, conflict with the principle of relativity which states that no material object or "packet" of electromagnetic energy may travel with a speed greater than that of light *in vacuo*. *There is nothing to prevent a quite unsubstantial thing like a geometrical pattern from being propagated with any velocity whatsoever up to infinity.*) The energy residing in the electromagnetic field is not, however, transported at the phase velocity (v), but at the group velocity (u) and the relationship between the two is always $uv = c^2$. Hence, since $v > c$, u must be $< c$ in accordance with the principle of relativity.

Group velocity is dependent upon :—

(*a*) The exciting frequency.
(*b*) The guide dimensions.

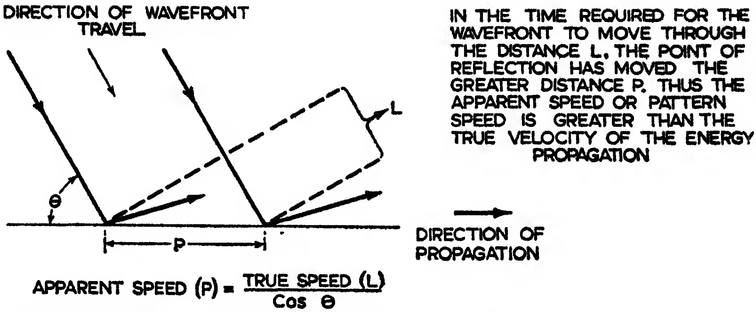


Fig. 13.—DIAGRAM TO ILLUSTRATE THE RELATIONSHIP OF PHASE VELOCITY TO GROUP VELOCITY.

For a given size of rectangular guide the group velocity increases for shorter and decreases for longer wavelengths. This is because the angle of incidence, and therefore reflection, becomes greater as the frequency decreases, which means that more reflections must take place for a given amount of forward travel for the longer wave and lower frequency than for the shorter wave and higher frequency. Consequently, relative to the speed of light, the group velocity of the longer waves is slower than that of the shorter waves for any given wave-guide.

The transit time for a group, therefore, between any two points in a wave-guide is greater than it would be in free space, since the electromagnetic waves do not take the shortest path, *i.e.*, a straight line. In consequence of this the group velocity or energy velocity is less than that of light, and the wavelength excited in the guide by any given frequency is longer than it would be if excited by the same frequency in free space.

All wave-guide dimensions referred to in wavelengths are therefore in terms of the wavelength excited in the guide for the given frequency, and *not* the wavelength that would exist in free space at that frequency.

The wavelength in the guide is defined as the distance along the axis of the guide between similar field patterns at any instant.

If the wavelength in the guide is λ_g cms., and the wavelength in free space is λ_0 cms., and the longest wavelength capable of travelling along the guide is λ_c cms., where the length of the side of the wave-guide perpendicular to the electric field is b cms.,

$$\text{then } \lambda_c = 2b$$

$$\text{and } \frac{1}{\lambda_g^2} = \frac{1}{\lambda_0} - \frac{1}{\lambda_c}.$$

Attenuation

In a rectangular wave-guide of given dimensions, the attenuation increases as the wavelength approaches cut-off value. This is due to increased losses resulting from the greater number of zig-zag reflections per metre length of guide. These losses are skin losses in the guide walls.

Excepting a special case where the electric field does not terminate at the guide walls, wavelengths much shorter than the optimum wavelength attenuate rapidly due to increased skin effect per reflection.

Excitation

There are three possible ways in which energy can be introduced or removed from a wave-guide :—

(*a*) Positioning a small loop of wire so that the H lines are linked with it for magnetic conduction.

(*b*) A probe (or aerial) placed parallel to the E lines.

(*c*) Linking of the internal fields with an external loop or probe through a slot in the wave-guide wall.

Inductive or Magnetic Coupling by Loop

Fig. 14 shows inductive coupling. The loop may be placed at positions (*a*), (*b*) or (*c*) as long as it links with the H lines where they are at a maximum.

To alter the degree of coupling the loop can be rotated relative to the directions of the H lines, or shielded or moved to another position where the H field is weaker. Fig. 14 (*d*) shows the details of the magnetic coupling loop and method of connection to a coaxial line.

Electric Coupling by Probe

Fig. 15 shows electric coupling by means of a probe. Coupling can be varied by situating the probe in a weaker or stronger field, or by reducing the length of the probe exposed to the action of the lines of electric force.

Fig. 16 shows a method of adjusting the end stub of the guide so that the distance between the coaxial line and the end of the guide is $\frac{\lambda}{4}$ for a given frequency by moving plunger A.

Plunger B performs two functions :—

(*a*) Acts as a shorting plate to form a $\frac{\lambda}{4}$ insulating support for the centre conductor.

(*b*) Provides a means of tuning the coaxial section so that the electric field of the coaxial line is at the proper position along the centre conductor to energise the wave-guide.

Coupling through Slot

Fig. 17 shows an example of coupling through slots or openings.

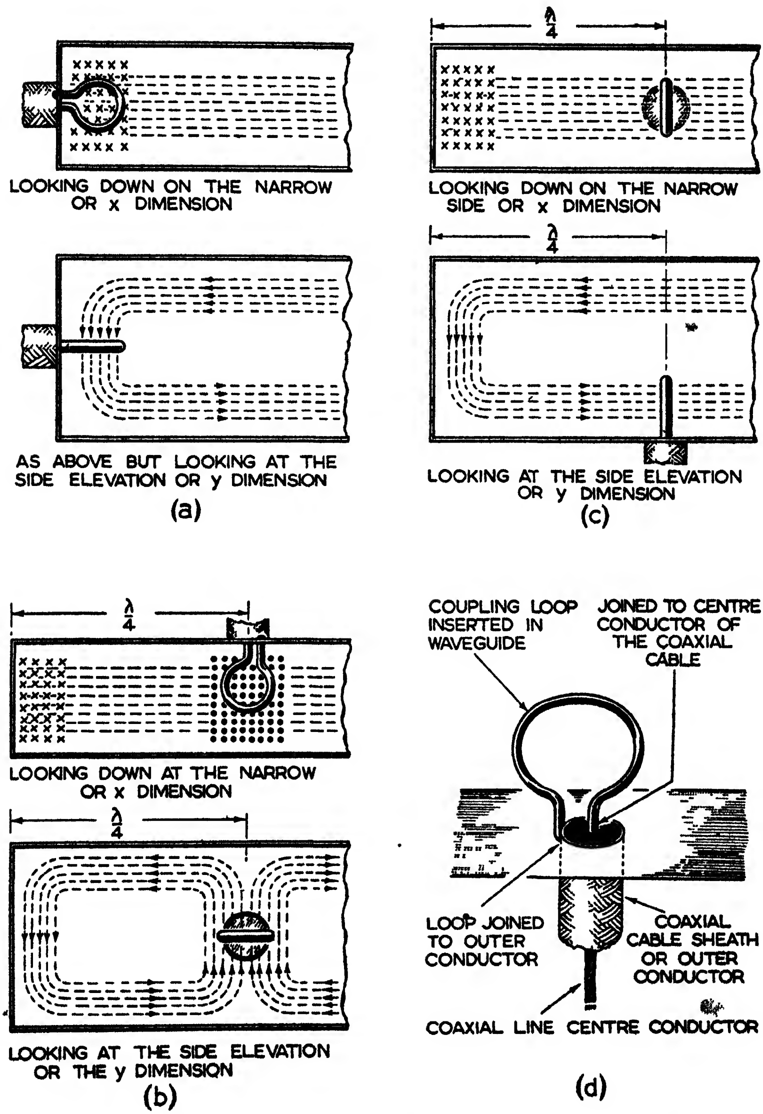


Fig. 14.—MAGNETIC COUPLING FROM COAXIAL LINE TO WAVE-GUIDE BY LOOP.

(a) Coupling located at end of the guide. (b) Coupling located in the broad side. (c) Coupling located in the narrow side. (d) Detail of loop coupling.

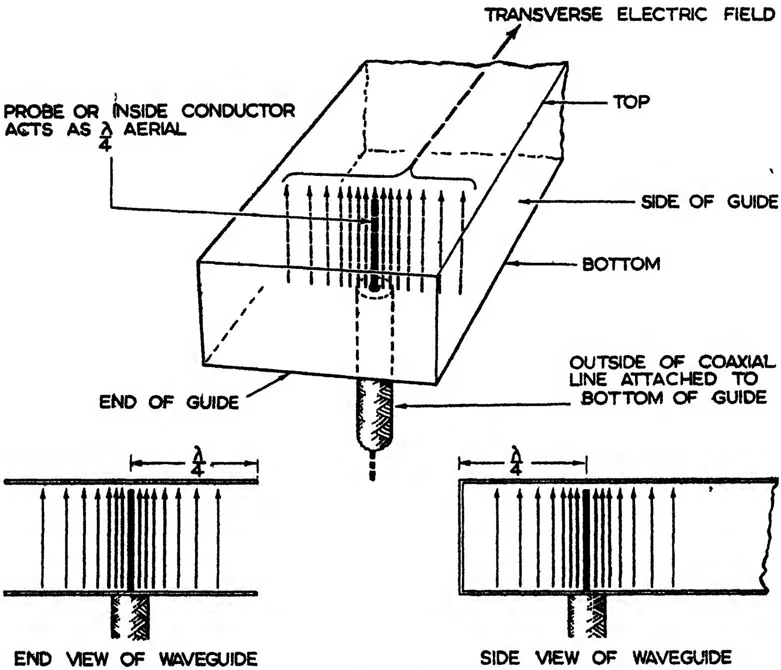


Fig. 15.—ELECTRIC COUPLING FROM COAXIAL LINE TO WAVE-GUIDE BY PROBE.

A variation of this method is that of shooting a stream of electrons through a hole in one side and out at the other.

Impedance of Wave-guides

As in the case with a transmitter line, a wave-guide presents an impedance to a source of electromagnetic energy that feeds it. An intrinsic property of a wave-guide of uniform cross-section, analogous to the surge impedance of a transmission line, is the wave impedance (Z_g). In contrast with the properties of the surge impedance, however, the wave impedance differs for different types and modes of waves propagated. For E waves $Z_g = 120\pi A\lambda_\phi/\lambda_g$ ohms and for H waves Z_g is $120\pi A\lambda_g/\lambda$ where λ is the wavelength in free space, λ_g = wavelength within the guide and A is the area of the guide in sq. cms.

A resistive film of impedance Z_g placed across the guide will absorb all the energy in a wave of the appropriate mode travelling down the guide. If the " termination " is other than Z_g, and not purely resistive, partial absorption and partial reflection of the wave takes place and standing waves are generated along the guide. Thus

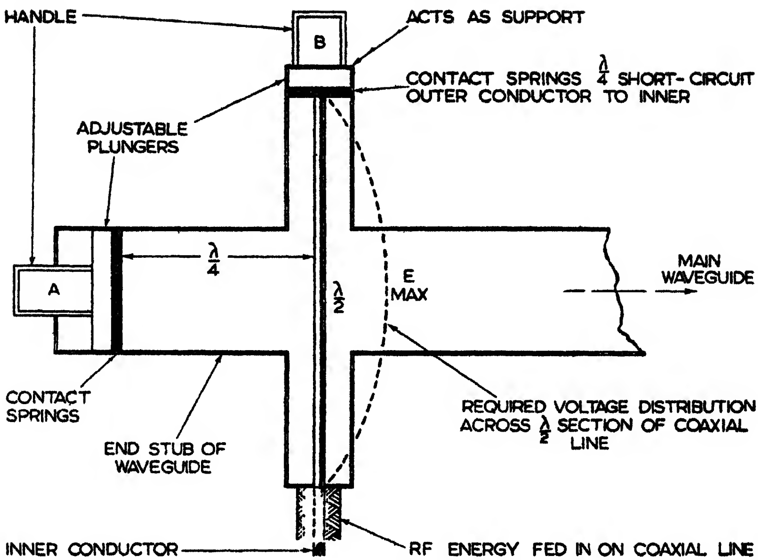


Fig. 16.—ELECTRICAL COUPLING BY COAXIAL LINE TO WAVE-GUIDE IN WHICH THE $\frac{\lambda}{2}$ SECTION SETS UP AN ELECTRIC FIELD TO EXCITE THE GUIDE.

a reactive termination numerically equal to Z_g results in a reflection which can be expressed as a phase angle (see Appendix IV).

Matching.—Adjustments are made as for transmission lines to reduce standing waves, and consequent loss of power to a minimum. Short circuits and open circuits set up standing waves in a guide by reflection, and resonant lengths of guide can be adjusted to pro-

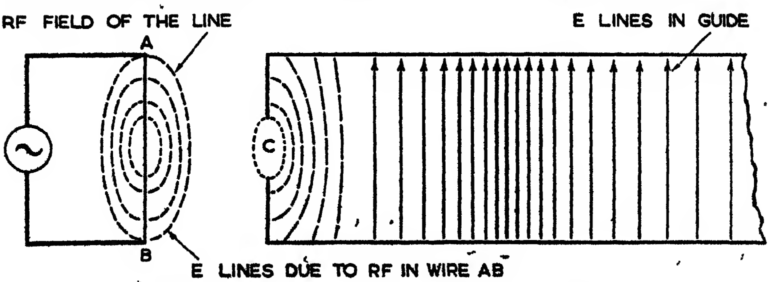


Fig. 17.—E LINE COUPLING TO WAVE-GUIDE.

AB is the line, the R.F. field of which energises the wave-guide through the slot C with which it is closely associated. Tuning arrangements not shown but may be carried out by piston or "screw-in" adjustment to alter the volume of the guide or cavity.

vide impedance transformations similar to those obtained with $\frac{\lambda}{4}$ and $\frac{\lambda}{2}$ lines. "Tuning out" may also be employed as in the case of transmission-line feeders.

In particular, it is possible to design joints between different portions of a wave-guide system where no mechanical connection is possible (*e.g.*, between a stationary and rotating section of guide), in which the mechanically open section appears as an electrical short-circuit to the wave, owing to the resonant properties of the region of space between the sections of guide (chokes).

The wave impedance of a guide of non-uniform cross-section contains, in general, both resistive and reactive elements. Thus we may regard constructions or obstacles placed in the path of waves travelling in a guide, as introducing reactive terms to the impedance presented to the travelling waves. The nature of these reactances will depend upon the geometrical configuration of the obstacles. The interposition of any type of obstacle within a guide will, in general, necessitate the introduction of some tuning adjustment calculated to annul the reactive component presented by the obstacle.

In the design of practical wave-guide feeder systems, it is necessary to introduce bends in the guide. The effect of a bend, or discontinuity, is to cause reflection, due to the corner not looking like Z_g. By rounding off the corner suitably, the impedance can be made to look like Z_g and the wave passes round the corner without appreciable loss.

In the case of H_1 waves propagated in rectangular guides, such "sharp" bends are made in the plane of the longer side of the rectangle, and do not alter the direction of the E vector (plane of polarisation). If a bend must be designed to rotate the plane of polarisation, the curvature must be comparatively small but the radius of the bend should be of the order of one or two wavelengths at least.

As stated in the chapter on Aerials, on reaching the open end of a wave-guide, waves are radiated into free space in the form of a directional beam containing a main lobe centred on the axis of the feeder, and a number of auxiliary side lobes. The angular width of the main beam, in planes parallel and perpendicular to the emergent E vector, will depend upon the dimensions of the aperture of the guide. This is dealt with under Aerials.

Methods of Matching

When guides are intended to radiate energy the output end is left open. In order to reduce standing waves due to that portion of the energy which is not radiated but reflected, the guide must be terminated properly so that it may deliver maximum power. This may be done by introducing a flare at the end of the guide

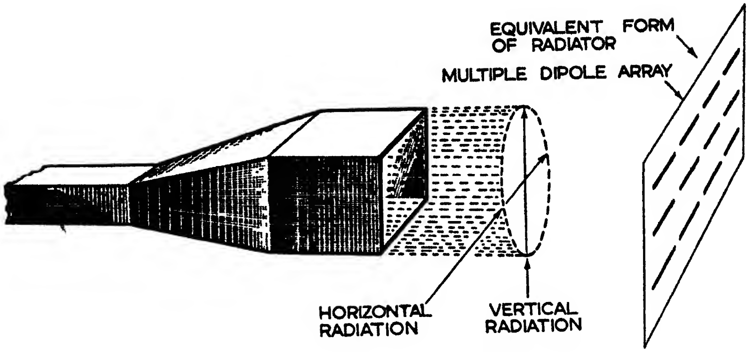


Fig. 18 (a).—PRINCIPLE OF MATCHING TO FREE SPACE BY A HORN, FLARE OR TRANSFORMER.

The vertical and horizontal widths of the beam are determined by the geometry of the flare and its dimensions in the vertical and horizontal directions.

shaped like a horn. The horn may be regarded as an energy transformer which matches the impedance of the guide to free space. This or any alternative method of terminating a guide in its characteristic impedance causes energy to be absorbed, thereby reducing the S.W.R. This match is, however, only a numerical match, and does not counteract any reactive elements that may

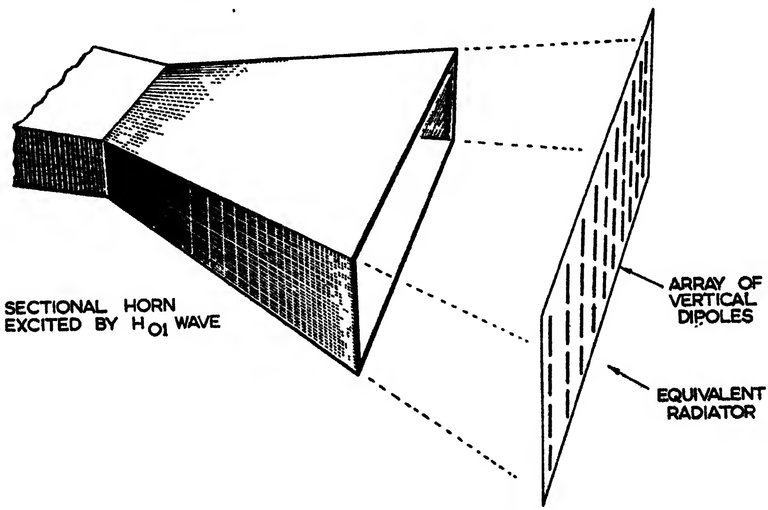


Fig. 18 (b).—ALTERNATIVE ARRANGEMENT.

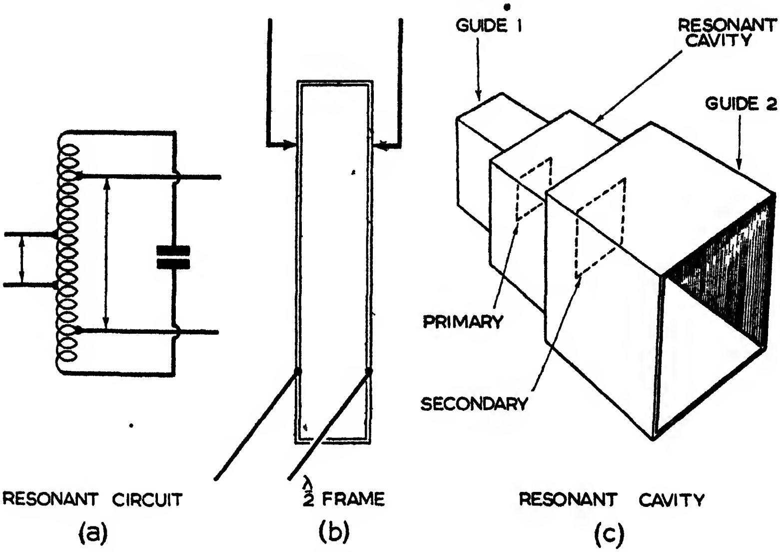


Fig. 19.—COMPARISON OF RESONANT CAVITY WITH OTHER MATCHING DEVICES.

be present. There is, therefore, in general, a S.W.R. due to these reactive elements (Fig. 18 (a)).

The characteristic impedance of a wave-guide may be thought of as the voltage/current ratio where both are moving in the same direction. The characteristic impedance of a given guide varies for every mode of operation and is directly proportional to the narrow dimension, x. If the other dimension, y, and the exciting frequency are fixed the impedance may vary between very wide limits for the various modes that may be excited in the guide by alternate methods of launching the R.F. energy into it.

The lowest characteristic impedance for a circular guide is about 350 ohms. For a rectangular guide it may vary between the values of 0 and 465 ohms, depending upon the dimensions and the mode employed.

When a wave-guide terminated in a horn is matched to space, the effect is similar to that produced by an array of dipoles—a beam of electromagnetic energy is transmitted. The ratio length of horn to its opening is important in determining the vertical and horizontal widths of the beam emitted.

Any abrupt change in the size or shape of a wave-guide causes reflections. In order, therefore, to pass energy efficiently from a guide of one impedance to a guide of a different impedance, some matching device must be provided.

The horn can also be employed as a device to match guides of different sizes since it permits the fields to expand or contract smoothly to fit different sizes of guides. The rate of change may therefore be regulated and abrupt changes can be avoided.

Another method of matching is by means of a cavity resonator (Fig. 19). Thus at comparatively low radio frequency coils and condensers are used as in (a). For higher frequencies the $\frac{\lambda}{2}$ frame achieves the same results. At micro-wave frequencies an equivalent matching device can be used as shown in Fig. 19 (c).

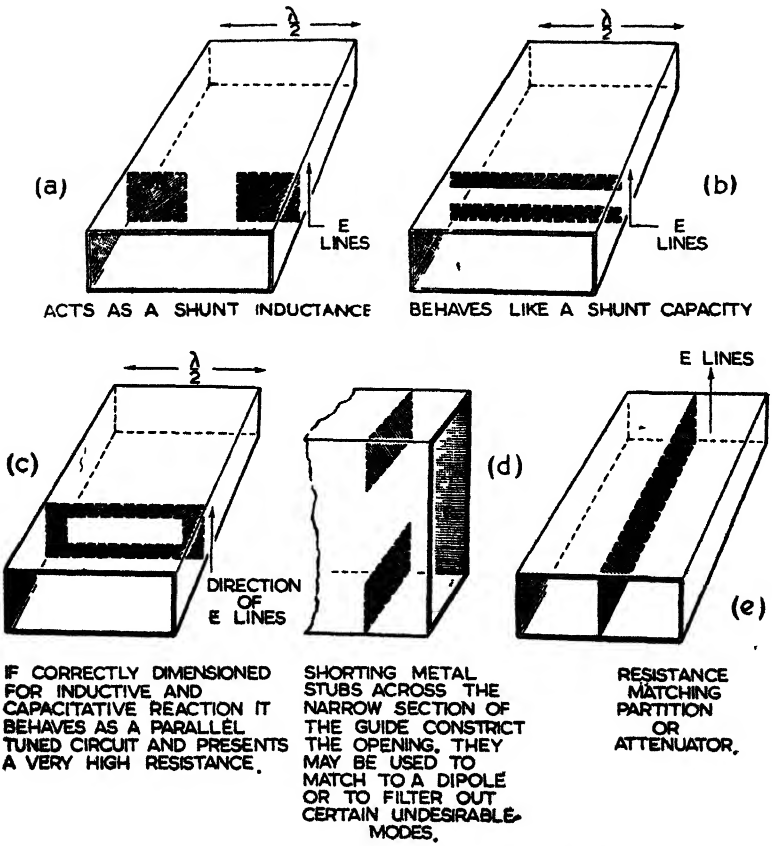


Fig. 20.—OBSTACLES IN WAVE-GUIDE TO CHANGE ITS IMPEDANCE AND TO PRESENT INDUCTIVE, CAPACITATIVE OR RESISTIVE REACTANCES.

Obstacles may be used as fixed tuning devices, or to attenuate the energy travelling down the guide.

Obstacles and Partitions in Wave-guides

Impedance characteristics of wave-guides may be changed by the use of metal partitions and obstacles of various types. Such obstacles or partitions may differ in shape and dimensions, also they may be placed so as to short the guide across the narrow side as in Fig. 20 (d), or they may be positioned merely to restrict free passage in the direction of the axis of the guide.

Fig. 20 (a). Edges of partition parallel to E lines. Current flows vertically in the partition. The magnetic field resulting causes the partition to behave as a shunt inductance.

Fig. 20 (b). Partitions turned crosswise. Electrons accumulate first on one side and then on the other. The resulting field is largely electric and the arrangement behaves like a shunt condenser.

Fig. 20 (c). Combination of Fig. 20 (a) and Fig. 20 (b). Behaviour

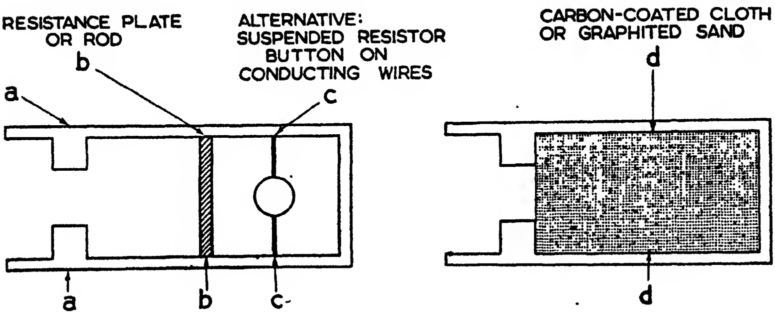


Fig. 21.—VARIOUS MEANS OF TERMINATING A WAVE-GUIDE IN Z_g (NON-RADIATING) FOR MEASUREMENT PURPOSES.

depends on dimensions. It may be inductive or capacitative. If dimensions are such that the capacity reactance resonates with the inductive reactance, the arrangement acts as a high shunt resistance.

The effect of partitions on modes having reversals in direction of lines of force from one side of the guide to the other is much more complicated; such may act as filters, rejecting all but certain bands of frequencies (see Fig. 20 (d)).

Fig. 20 (e). A lengthwise partition of a resistive material (graphite acts as an attenuator). This partition is dropped in through a slot, so that it cuts the electric field at maximum value.

In addition to a radiating horn, flared section and resonant cavity, other means as shown in Fig. 21 (a), (b), (c) and (d) may be used to terminate a wave-guide in Z_g, generally for purposes of measurement, since little or no radiation takes place under these conditions.

(1) Energy absorbed by flat plates of resistive material crosswise in the guide (Fig. 21 (a)).

(2) Energy dissipated by small rod of resistive material from top to bottom across the guide at high voltage points (Fig. 21 (b)).

(3) Energy dissipated by small particle of resistance material suspended by conducting wires across high voltage points (Fig. 21 (c)).

(4) Energy dissipated in a mass of carbon-coated cloth or graphited sand (Fig. 21 (d)).

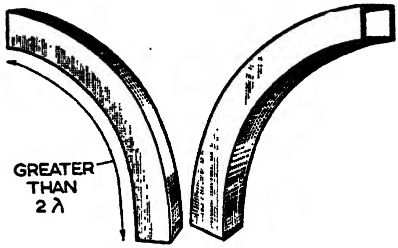


Fig. 22.—WAVE-GUIDE ELBOWS.

If the material has a low impedance at the frequency range of the guide and there are standing waves in the guide, the resistance material should be placed at a point where the ratio of E lines/H lines is low, conversely a high resistance should be placed where E lines are at a maximum. This condition is similar to that involved in choosing the proper $\frac{E}{I}$ point in a transmission line for matching purposes.

Connection Arrangements

Examples of connecting sections for turning corners (elbows) are shown in Fig. 22, while a section for rotating the field pattern is shown in Fig. 23.

When two wave-guides must be connected to allow for expansion, a choke joint is used (see Fig. 24).

The choke flange has cut in it a circular slot $\frac{\lambda}{4}$ deep, the middle of the broad face of the guide is $\frac{\lambda}{4}$ from the edge of the slot—point M in Fig. 24 (c).

Energy leak is prevented by standing waves set up at these $\frac{\lambda}{4}$ sections. The short circuit at A reflects to B as a high impedance, Fig. 24 (a). This high impedance is transformed to a low impedance across the $\frac{\lambda}{4}$ section B to C, thus the

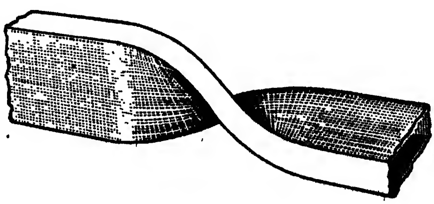


Fig. 23.—RECTANGULAR WAVE-GUIDE SECTION FOR ROTATING FIELD PATTERN THROUGH 90°.

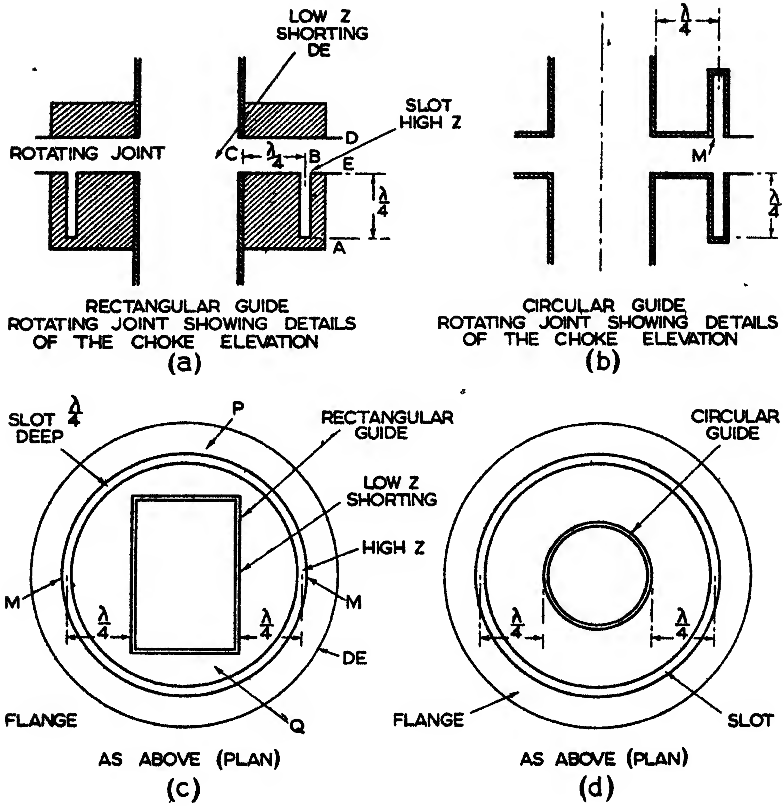


Fig. 24 (a)–(d).—JOINT BETWEEN MOVING SECTIONS OF WAVE-GUIDE, SHOWING DETAILS OF THE R.F. CHOKE EMPLOYED TO AVOID LEAKAGE.

choke joint effectively short circuits points D and E. The choke joint as applied to rectangular guides is relatively ineffective at P and Q, Fig. 24 (c) (*i.e.*, when the long side of one guide is rotated 90° relative to the long side of the other guide), but the R.F. voltage here is very low and there is little tendency to leakage. The same arrangement can be used with circular wave-guides to allow one to turn or move relative to the other. When used with circular guides it is much more effective than with rectangular guides.

Rotating Joint and Field Pattern Change

In a number of applications of radar it is necessary to feed the energy from a fixed transmitter to an aerial system which is continuously rotating on a vertical axis. In this case a rotating joint has to be inserted between the fixed and moving wave-guides. The mechanical details of this joint and the R.F. choke or chokes involved have already been described. The immediate purpose is to examine the problem in so far as it affects the field pattern of the electromagnetic waves in the wave-guide system.

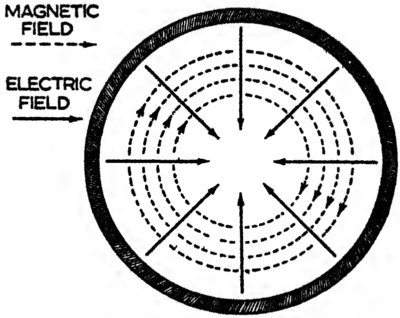


Fig. 24 (e) —THE E_0 MODE IS SYMMETRICAL ABOUT A COMMON AXIS FOR FIXED AND ROTATING CIRCULAR WAVE-GUIDE SECTIONS

Since the electric field has a definite direction in the case of the H_{01} wave in the rectangular guide, it is clear that the fixed wave-guide could not be connected to a portion of wave-guide (in prolongation of the fixed part) which could rotate about their common axis. The R.F. output would vary violently as the guide was turned. If on the other hand, a vertical circular guide using an H-type wave were used, the waves on being reflected by a 45° mirror on to the main aerial mirror, would render it impossible to keep the waves horizontally polarised, this being the condition required, *i.e.*, the H vector in the direction of propagation.

The solution lies in the use of a circular section of wave-guide with waves propagated in it in the E_0 mode ; that is the waves have only electric components along the axis of the wave-guide (see diagram, Fig. 24 (e)).

If the wave-guide is vertical an observer looking down into it (as if looking down a well) would see no change in the pattern of the electric components of the wave as he looked in from different directions. In other words, the pattern has lost any horizontal direction and there is radial symmetry.

The vertical circular guide with E_0 waves is used as follows (in the case of an aerial spinning on a vertical axis) :—

The wave comes along a fixed horizontal rectangular wave-guide in which larger side is a hole, leading into a fixed vertical circular guide. The H_{01} waves in the rectangular guide excite E_0 waves in the circular guide.

The E_0 waves enter a rotating circular guide, the joint between the fixed and moving parts being designed so as to reduce the loss of energy at the joint (see paragraphs on R.F. choke). The E_0 waves in the rotating guide now excite H_{01} waves in a *horizontal rectangular* wave-guide which can turn so as to point in any direction.

From this guide the energy can be fed into a rotating mirror.

Certain refinements must be added :—

It is important that no H waves are propagated along the

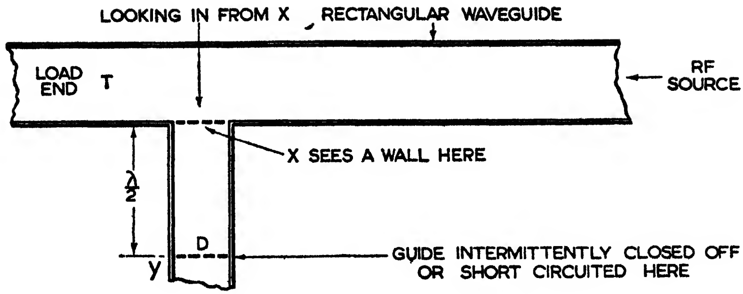


Fig. 25.—$\frac{\lambda}{4}$ SECTION OF WAVE-GUIDE INTERMITTENTLY SHORTED BY DISCHARGE ACROSS A SPARK GAP IN A CAVITY RESONATOR.

circular guide, because the efficiency of the system as a whole would change as the rotating parts moved. This would cause to be sent back, down the guide, varying amounts of energy which would upset the operation of the magnetron oscillator.

Filter rings or irises are therefore inserted in the circular guide which prevent the passage of the H waves. These rings have also some effect on the E_0 waves. This effect is reduced by careful design and spacing of the rings.

In some radar installations the aerial is given a limited action in elevation. In this case a horizontal circular wave-guide is also used as a rotating joint.

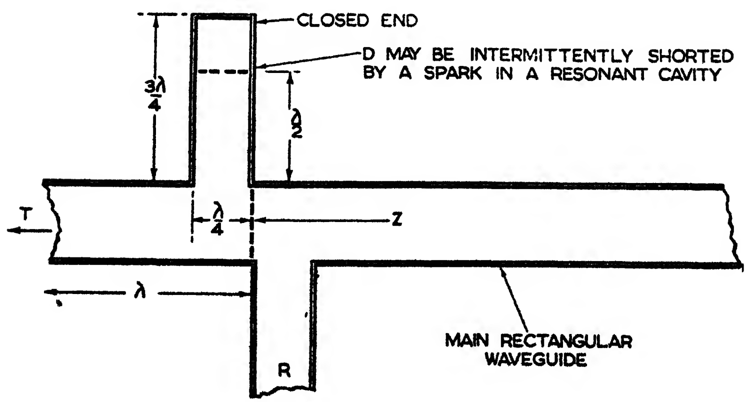


Fig. 26.—$\frac{\lambda}{2}$ AND $\frac{\lambda}{4}$ WAVE-GUIDE SECTIONS ASSOCIATED WITH SPARK GAPS IN CAVITY RESONATORS EMPLOYED AS T.R. AND T.B. SWITCHING DEVICES.

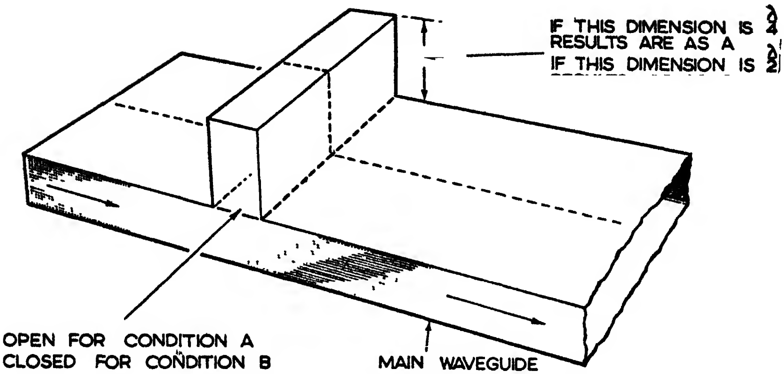


Fig. 27.—$\frac{\lambda}{2}$ AND $\frac{\lambda}{4}$ SECTIONS OF WAVE-GUIDES BEHAVE LIKE $\frac{\lambda}{2}$ AND $\frac{\lambda}{4}$ SECTIONS OF TRANSMISSION LINE.

A, a $\frac{\lambda}{4}$ shorted wave-guide section acts as a filter and passes all frequencies except one band. Its action is also similar to that of a quarter-wave shorted section of transmission line.

B is similar in effect to a $\frac{\lambda}{2}$ shorted section of a two-wire transmission line. It acts as though both terminations were short circuited. Thus a shorted $\frac{\lambda}{2}$ section transmits the short circuit across the far end to the junction without change, whereas the $\frac{\lambda}{4}$ section reflects the short across its far terminals as an open circuit to the junction, *i.e.*, no wall and free access. This short circuit may be regarded as a wall across the mouth of the opening.

Some Wave-guide Properties in Brief

(*a*) An electrical $\frac{\lambda}{4}$ or odd multiple $\frac{\lambda}{4}$ wave-guide *closed* at one end presents an *opening* where it joins another guide at the far end but a $\frac{\lambda}{2}$ closed section presents a solid wall (Fig. 27).

(*b*) Impedance matching may be accomplished by the arrangement in Fig. 19 (c).

(*c*) Chokes are used when wave-guide sections are used for expansion or rotating joints (Fig. 24).

(*d*) A wave-guide may be matched to air by increasing the internal cross-section of the guide or it may be flared in the form of a horn.

(*e*) A wave-guide attenuator may consist of a movable lengthwise partition as in Fig. 20 (e), or alternatively some other type of fixed or movable obstacle having an impedance equal to Z_r. It should act as a pure resistance at the excitation frequency.

(*f*) A section of wave-guide may be used as a tuned circuit or transformer (see Fig. 19 (a) and (b)).

(*g*) R.F. energy may be introduced in various ways, Figs. 14, 15 and 16 indicate some of the methods employed to excite the mode which it is desired to launch, assuming of course that the guide dimensions are such that the selected mode can exist.

(*h*) Sections of open and closed wave-guides may be used in switching circuits, *e.g.*, T.R. and T.B. switches.

In Fig. 25 an intermittent short circuit (such as the discharge across a spark gap) in a resonant cavity at D reflects a short and effectively a solid wall is seen " looking in " from X.

In Fig. 26, when D is shorted a solid wall results at the junction of the closed stub to the main wave-guide. Paths T and R may now receive energy.

When D is not shorted, a solid wall results looking in at Z which is λ from the closed end. No energy can, therefore, reach T, but the path to R is open.

The effect of $\frac{\lambda}{4}$ and $\frac{\lambda}{2}$ stubs at the main guide is shown in Fig. 27.

Rectangular wave-guides possess certain advantages over circular guides.

(*a*) Smaller amount of metal for a given frequency.

(*b*) Energy is propagated with a more constant type of distribution, a fact which is important when it is necessary to radiate the energy from the open end of the wave-guide with a definite polarisation.

Dielectrics as Wave-guides

Electromagnetic waves may be confined largely to solid rods of insulating materials. The effect is similar to the transmission of light through a Perspex rod.

The light rays travelling inside the rod are reflected back to the inside when they strike the surface of the rod at an oblique angle. In order to leave the material with appreciable strength the rays must strike a perpendicular wall as they do when they arrive at the end of the rod.

If the dielectric rod has the proper size and cross-section, electromagnetic waves are reflected at the surface where air and dielectric meet and most of the field is confined within the insulating material.

Dielectric wave-guides lose much more energy in radiation than do metal wave-guides, consequently attenuation of the wave is much greater.

The methods employed for checking standing waves in coaxial lines and wave-guides are described in Appendix IV.

APPENDIX III

THEORY OF CAVITY RESONATORS

PARALLEL-TUNED circuits, by virtue of the electrical properties which they possess, are required to perform functions in ultra-high frequency circuits for reasons similar to those which govern the use of parallel resonant circuits at lower frequencies.

The equivalent function is performed in ultra-high frequency circuits by cavity resonators.

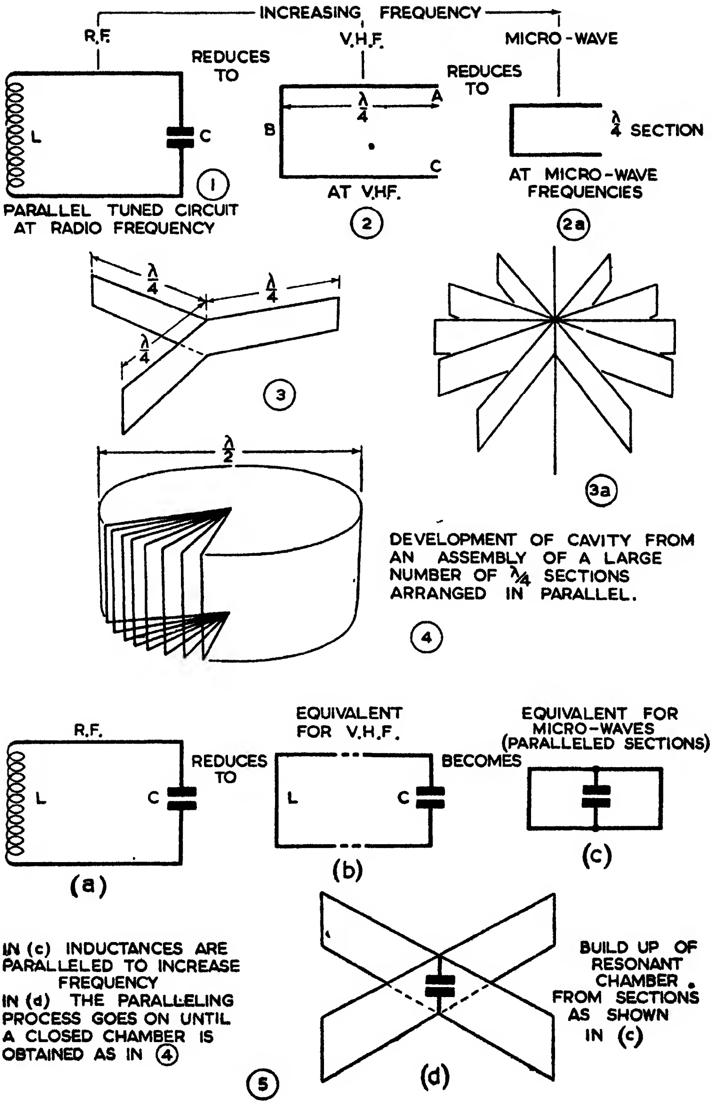
A cavity resonator may be thought of as a hollow container formed by arranging a large number of shorted quarter-wave sections of transmission line in parallel, sections being added until they ultimately merge to form a totally enclosed chamber.

Cavity resonators may be compared with acoustic resonators. For example, sound waves introduced into a chamber having smooth hard surfaces are reflected from wall to wall. If the frequency relative to the dimensions of the chamber is such that a resonance is produced, standing waves appear and the sound is reinforced at the anti nodes. Thus the resonant wavelength is proportional to the size of the resonator. The cavity cannot resonate if it is too small for the wavelength which it is desired to excite, but if R.F. energy is supplied at the resonant frequency for the cavity high amplitude micro-waves will propagate across and from top to bottom of the cavity and R.F. energy can be extracted.

The above conception assumes that each component quarter-wave section is an inductance with shunt capacity, reduced in length to some fraction of a turn of wire. Its length and shape must be such that it becomes a $\frac{\lambda}{4}$ shorted section of transmission line at the required frequency.

Figs. 2 and 2 (a) are comparative, and are intended to illustrate the manner in which the physical dimensions of conventional types of resonators shrink as frequency is increased. As the required frequency is increased (1) shrinks to (2) and eventually (2) itself shrinks towards 2 (a). All, however, retain the electrical properties of a parallel resonant circuit.

When the frequency becomes so high that ABC, Fig. 2, must shrink physically to inconvenient proportions, cavity resonators become practicable. However, it must not be thought that a cavity resonator is anything in the nature of a makeshift. On the contrary, it is a highly desirable device because of its high efficiency. It is not used at the lower frequencies because its bulk becomes pro-



Figs. 1–5.—ALTERNATIVE CONCEPTS TO ASSIST IN ANALYSING A SIMPLE CAVITY RESONATOR AS AN ASSEMBLY OF $\frac{\lambda}{4}$ ELEMENTS IN PARALLEL, AND AS AN ASSEMBLY OF MINUTE L/C ELEMENTS.

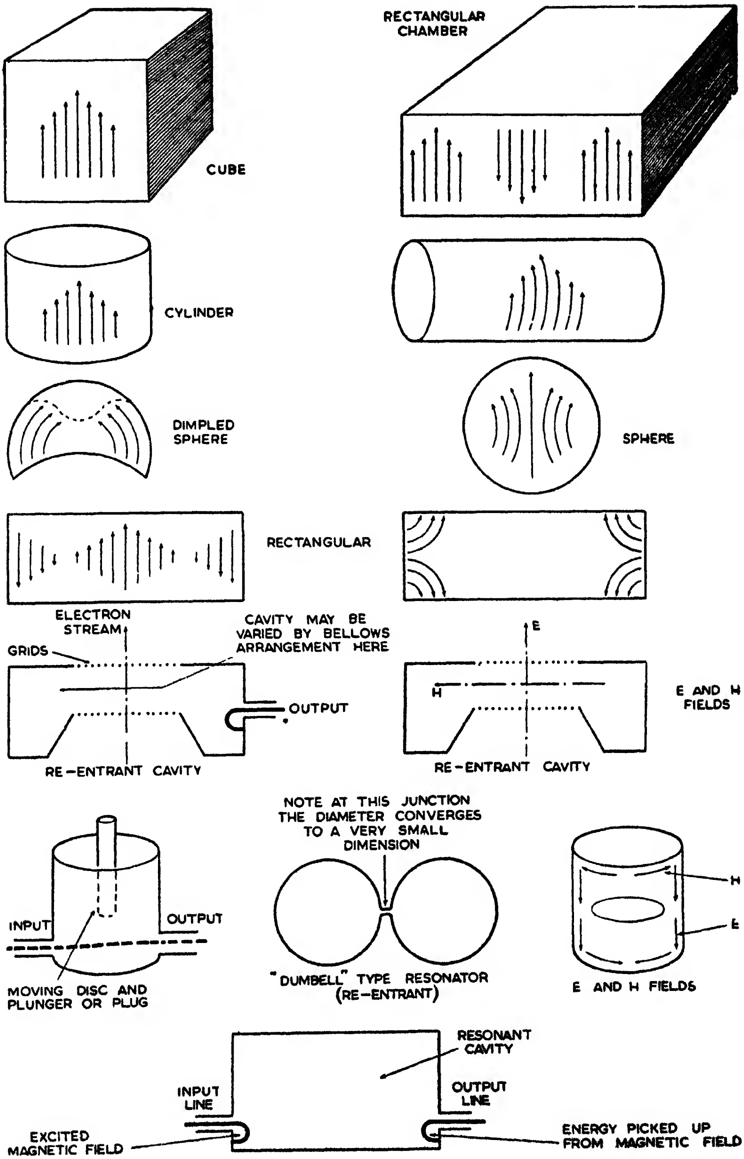


Fig. 6.—VARIOUS FORMS OF NON-RESONANT AND RE-ENTRANT CAVITIES.

hibitive. For example, at 1 megacycle a cavity equivalent to a conventional resonant parallel circuit containing a coil and a condenser would require to be about 100 ft. long in one dimension to exhibit similar properties.

Figs. 3 and 3 (a) show how a cavity resonator can be developed, theoretically, by paralleling a number of short-circuited quarter-wave sections of transmission line until a totally enclosed chamber is formed, as in Figs. 3 (a) and 4. As sections are added the total inductive reactance decreases, and the total capacity increases, but since the addition of any number of sections does not materially alter the resonant frequency of the assembly, as compared with individual quarter-wave sections of which it is composed, the parallel arrangement of quarter-wave sections shown in Figs. 3 and 3 (a) culminates in the totally enclosed chamber of Fig. 4, having properties similar to its $\frac{\lambda}{4}$ components.

Figs. 5 (a) to (d) is an analysis of a cavity build up using L and C concepts as an alternative to shorted $\frac{\lambda}{4}$ sections.

In Appendix I it was shown that the Q of a shorted section of transmission line is much higher than can be obtained with a conventional coil and condenser assembly, consequently the parallel arrangement of sections leading to the development shown in Fig. 4 must raise the value of Q for the device as a whole. In general, the Q of cavity resonators is very high and they are accordingly very selective.

Cavity resonators may take various shapes, such, for example, as cubes, spheres, cylinders, etc. They may also be dimpled or asymmetrical.

Classification of Cavity Resonators

Cavity resonators may be classified as non-resonant and re-entrant. Non-resonant cavities include such shapes as the cylinder, prism, sphere, ellipsoid, etc. The Q of these cavities is very high, but the oscillations which take place in them are generally of a forced character. Re-entrant cavities include such forms as the "dumbell," etc. In this class of cavity, the elements representing inductance and capacity are more easily distinguished, and may be regarded as "lumped," as in a conventional parallel tuned circuit. The Q in many cases may not be as high as in the non-resonant type. However, resonance effects are more in evidence and, since the Q is still very high, a greater degree of re-enforcement is generally obtained with the resonant types.

Standing waves are produced by reflection in both classes of cavity resonator, but true resonance effects are much more marked in the re-entrant type. This is to be expected since, in coil and condenser concepts, the ratio L/C is greater.

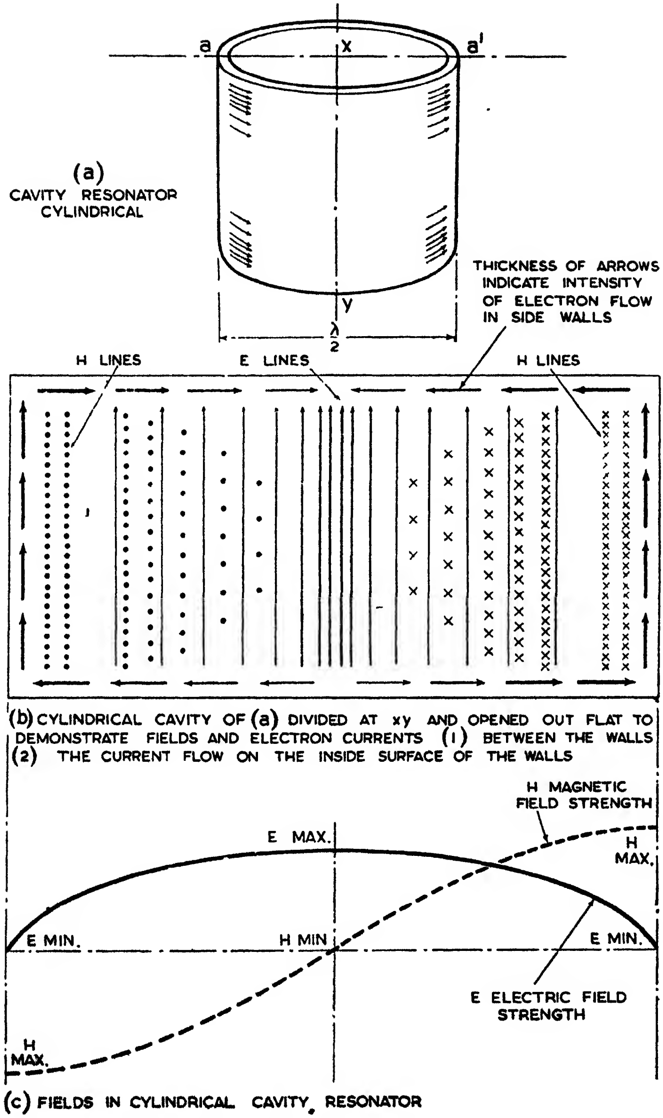


Fig. 7.—EXAMINATION OF THE ELECTRIC FIELD AND MAGNETIC FIELD IN THE CYLINDRICAL CAVITY (a), ALSO THE DIRECTION OF ELECTRON CURRENTS ON THE INNER SKIN OF THE CYLINDRICAL CAVITY.

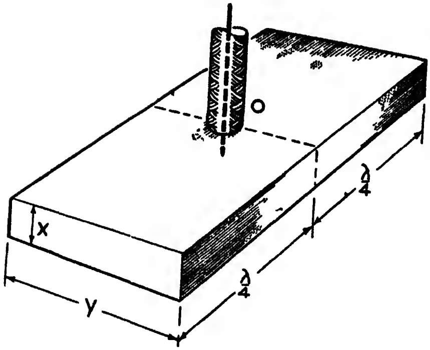


Fig. 8.—SECTION OF WAVE-GUIDE CLOSED AT BOTH ENDS AND EXCITED BY PROBE AT O IS, IN EFFECT, A RESONANT CAVITY IN WHICH MORE THAN ONE MODE MAY EXIST.

For a given frequency the number of possible modes is determined by the method of excitation and the X and Y dimensions of the cavity which also determine the reflection patterns from the various surfaces.

For very simple types of cavity resonators, symmetrical in shape, the resonant frequency can be calculated if the shape of the cavity is known. Also the rate of loss of energy establishes the Q for the cavity. The main factor determining Q is the volume to surface ratio of the device. Q may vary from approximately 28,000 for a cube, 31,000 for a cylinder and 26,000 for a sphere, when not loaded.

Fundamental Principles

The fundamental principles which control the functioning of cavity resonators are the same as for corresponding short lengths of wave-guides closed at both ends. In fact, there is no essential difference between a section of wave-guide short circuited at both ends and a cavity resonator of the same shape and dimensions.

The following fundamental principles apply, therefore, equally to cavity resonators as to wave-guides :—

(*a*) An E line can terminate only on an electric charge, wherefore it follows that the electric field, which is equivalent to a voltage, is short circuited when it exists across a perfect conductor.

(*b*) A magnetic field can cease suddenly only on a surface carrying an electric current. If the conductor is perfect, the magnetic field created by the current is exactly equal to the magnetic field which excites the current, but of opposite direction, so that the resultant field at the surface of the conductor is zero. In other words, action and reaction are equal and opposite.

The principles laid down in (*a*) and (*b*) establish the boundaries for the electric and magnetic fields inside the cavity, and it is clear, therefore, that these fields must exist wholly within the cavity itself, and are bounded by the walls of the resonant chamber.

Since the electrical conditions prevailing in the resonant chamber are similar to those associated with a similar section of wave-guide closed at both ends, the electric and magnetic fields are at right angles to one another. Also R.F. energy applied from the source oscillates alternatively, between the electrostatic and magnetic

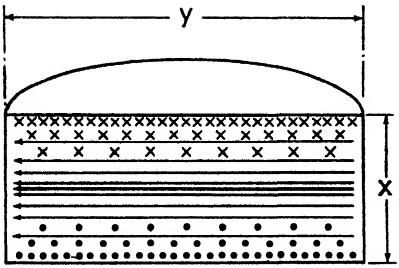


Fig. 9.—E LINES ARE NOW TRANSVERSE ACROSS THE xy AXIS OF FIG. 7 (a).

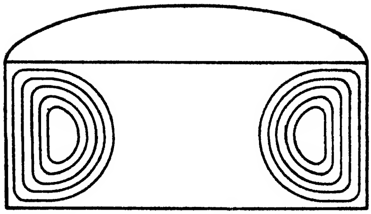


Fig. 10.—H LINES FORM CONCENTRIC CIRCLES VERTICAL TO THE xy AXIS OF FIG. 7 (a).

states, and standing electromagnetic waves of high amplitude are propagated across, and from top to bottom of the cavity, if it has in one direction the minimum dimension required by the frequency of the input (wavelength).

Having regard to condition (*a*) above, and to the fact that the E lines must have a maximum intensity at the centre of the chamber, and minimum intensity at the sides, it can be seen that these are the conditions for a short-circuited half-wave section, but it is essential that one dimension of the cavity should be an electrical half wave or a multiple of a half wave (see Fig. 7 (c)) in order to avoid the short circuit that would otherwise occur across the electric field, at the sides, for any smaller cavity dimension relative to wavelength employed.

Mode of Operation

For any cavity of given size and shape, various frequencies higher than the cut-off frequency (different modes) are possible, because wave energy may be reflected from various surfaces within the cavity. There is also the possibility of oscillation at some frequency that is a harmonic of the basic wave.

The mode of operation is largely dependent upon :—

(*a*) How oscillations are forced.
(*b*) How energy is removed.

Energy may be introduced or removed from a cavity inductively, capacitatively or by radiation.

Modes in a cavity resonator are designated by a three-number system. The third subscript signifies the number of half patterns crossed perpendicular to the transverse field. The mode in Fig. 7 (c) is classed as TM_{011} or E_0, when reduced to a half wavelength.

In Fig. 8 the R.F. energy is introduced at point O into a section of wave-guide closed at both ends and equivalent to a cavity resonator.

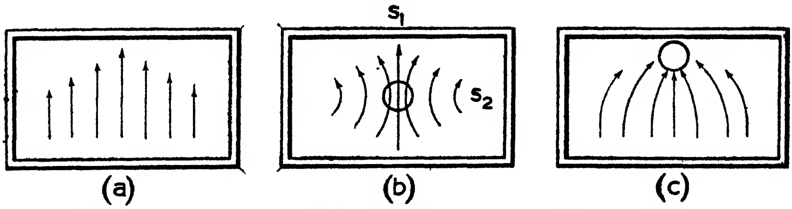


Fig. 11.—TUNING A RESONANT CAVITY BY THE INTRODUCTION OF A CONDUCTING BODY.

(a) E lines before tuning, *i.e.*, as seen in Fig. 7 (b).

(b) A metallic body (a sphere in this case) introduced at the point of maximum E line intensity shortens the E lines and increases the capacity of the chamber.

(c) A similar element introduced at the side of the cavity, where H lines have maximum intensity as seen in Fig. 7 (a) looking in from X to Y, shortens the H lines and produces an effect on frequency similar to that of decreasing inductance.

If the dimension XY is correct, the R.F. voltage at O will be reinforced by the in-phase waves reflected from X and Y.

Fig. 7 (b) shows the field pattern for one mode in a cylindrical resonant cavity like that shown in Fig. 7 (a). The graduated arrows show the direction and intensity of the electron flow on the inside surface at an instant when the field intensity is at a maximum. For the other half of the cycle all arrows are reversed in direction.

The guide patterns shown in Figs. 9 and 10 have the additional subscript 3 when one and a half wavelengths long, and 1. when reduced to half a wavelength.

Since the same cavity can operate in several different modes, depending upon the manner in which it is energised, etc., the cavity may operate at several fundamental frequencies as well as at harmonics of the fundamental. Figs. 9 and 10 are for an alternative mode for the cavities of Fig. 7 (a).

Electron flow is limited to a thin layer of metal on the inside of the cavity, consequently one essential feature is that the I^2R losses should be reduced to the absolute minimum, in order that the Q may be as high as possible. The construction is therefore generally from sheets of metal, usually copper, which are silver or gold plated, internally, the latter being employed when there is any danger of oxidisation taking place.

Cavity resonators for accurate frequency measurement are often constructed from a solid block of metal in order to secure greater rigidity.

A resonant cavity may be tuned between narrow limits by a suitable conducting body introduced at (*a*) the point of maximum E line intensity, or (*b*) at the point of maximum H line intensity, see Figs. 11 (a), (b) and (c). Since the point of H line intensity is a minimum for E line intensity, and *vice versa*, either the E or H field is affected, but not both. In either case, however, the effect is to shorten the E or the H lines according to position. Shortening the E lines decreases the frequency just as though the capacity of the cavity had been increased. On the other hand, shortening the H lines increases frequency just as though an inductance had been reduced.

Uses of Cavity Resonators

Cavity resonators in one form or another are used for a variety of purposes, (1) for matching wave-guides of different characteristic impedances, (2) for creating variable impedance points, (3) for matching to a given load impedance, (4) as link circuits for ultra-high-frequency oscillators, (5) as wavelength measuring devices (see Appendix IV), (6) as matching or filtering sections. Cavities performing all these functions, although not always designated as such, are described in Appendix II and the chapter on aerials. A cavity used for wavelength measurement is shown in Appendix IV.

The cavity resonators employed in resonant magnetrons are of the "dumbell" type. Each cavity is half a "dumbell"; the anode space common to all the cavities may be considered as forming the other half of the "dumbell," thus the cavity system may, in some cases, be classed as re-entrant.

APPENDIX IV

SOME NOTES ON RADAR TEST EQUIPMENT

SINCE the frequencies used in radar range from about 100 megacycles or less to 10,000 megacycles or more, and because the range for which each test equipment can be designed to function efficiently is limited, a large number of test equipments exist to cover the total range.

Also a change in technique occurs in the region of 400 megacycles and upwards, which necessitates a corresponding change in the test equipment for these higher frequencies.

Finally, there is test gear which is only suitable for use in the laboratory or for use under suitable working conditions.

In these circumstances it is not possible here to deal in detail with all test equipment for radar installations, since this, by its very complexity and numerous variations, is a subject for a separate manual.

This appendix, therefore, deals only with some of the general principles and is confined to such equipment as can be conveniently set up on site.

It is assumed that the reader is familiar with the test equipment normally used in the general communication frequency range, *i.e.*, at the lower frequencies round about 100 megacycles and up to about 400 megacycles, but a short note on the slotted line and its use in connection with transmission lines has been included.

Fault-tracing Procedure

Fault-tracing procedure follows the same general principles, whether it is applied in connection with high frequency or ultra-high frequency circuits, *i.e.*, classification of observed facts, analysis and deduction. Consequently fault-tracing procedure follows the same general lines whatever the frequency of the apparatus involved. The objectives are therefore similar, but the technique employed may be widely different, according to the frequencies involved.

In deducing the location of faults, the radio engineer relies largely upon information obtained by comparing actual voltage and current readings with a set of standard readings, usually supplied with the equipment, or recorded at the time of installation. In addition to this check, the radar engineer can also compare waveforms as observed with a cathode ray tube with a sketch or photograph of the standard waveform. Any serious dissimilarity in shape must be due to one or other of the three major causes of distortion; therefore, in the case of a normal failure, breakdown or disconnection, it should generally be possible to fix suspicion with a fair degree of certainty upon the faulty connection or component.

The test equipment usually supplied with radar sets operating at micro-wave frequencies consists of the following:—

Monitor

A monitor which may consist of an indicator using an "A" display and calibrated time base. The purpose of the monitor, as its name implies, is to check performance and it can also be employed to reproduce waveforms at test points, as and when required, for comparison with standard waveforms, fault-tracing purposes or during the installation of the equipment.

Field Strength Meter

A field strength meter (fluxmeter). This is supplied for the purpose of checking the field strength of the magnetic system of the magnetron.

The instrument consists of a search coil and probe movement, which is inserted between the pole pieces of the magnet. A meter gives a direct reading of the field strength when multiplied by a given factor. It is of the utmost importance for efficient operation of the magnetron that the magnetic field strength should remain constant, every care must be taken to avoid loss of field strength. The magnetic system must not come into contact with screwdrivers or other tools, and a "keeper" should *not* be fitted to permanent magnets when the magnetron is not in use.

Test Oscillator and Wavemeter

Another useful piece of test gear is *a test oscillator* in which is incorporated a magnetron. The object of the test oscillator is to excite oscillations of a known frequency in the transmitter or receiver. This piece of test gear is not always supplied, since it is, in effect, a facsimile transmitter of approximately fixed frequency but with a smaller output.

A wavemeter, which may incorporate arrangements for measuring peak power, standing wave ratio and observing the R.F. envelope, with particular reference to the spectrum width is obviously most essential.

Wavelength Measurement

A wave-guide section used as a resonator for wavelength measurement is shown in Fig. 1. Energy is allowed to enter the small opening in the end plate from a source of radiated fields. If the plunger is not set at the proper position for the frequency of the

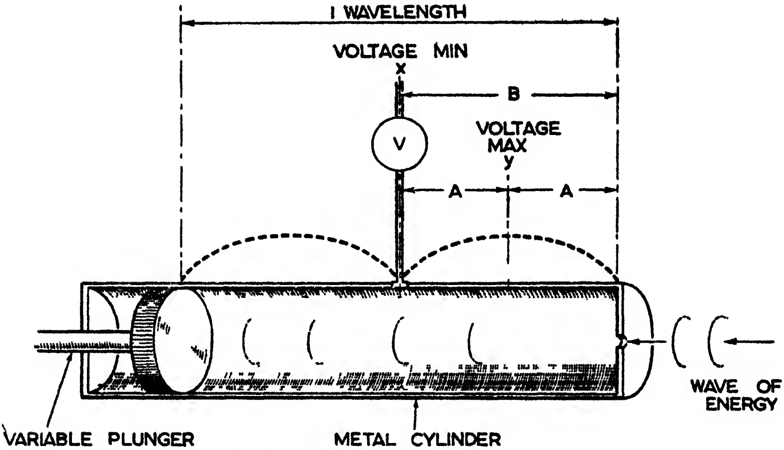


Fig. 1.—WAVE-GUIDE SECTION USED AS A RESONATOR FOR WAVELENGTH MEASUREMENT.

entering waves the cavity does not resonate, and hence there is no reading on the R.F. meter.

If the plunger is moved back and forth a point may be found where resonance takes place and the energy wave reflected from the piston reinforces the incoming wave.

A standing wave of electric field is then set up, and may be detected by the R.F. meter. If the cavity is long enough, a voltage maximum and minimum may be found by moving the R.F. probe in the lengthwise slot. Thus if x is found to be a voltage minimum or node and y is the adjacent voltage antinode, the wavelength may be determined. In Fig. 1 distance A or B determines the wavelength in the guide 2B or 4A. This value is always greater than in free space for the same frequency (see Appendix II).

Other methods may be used to tune the cavity to resonance. One method is to introduce a variable capacity at the point of highest voltage, as in Fig. 2 (a) and (b) in which the mode is assumed to be such that maximum voltage lies along HG.

Another method is to introduce a metal slug into the cavity, as in Fig. 3 (b) and (c). If the metal mass is at the position of maximum E lines, these lines are shortened and the capacity is increased, and the resonant frequency is lowered.

If the metal mass is at the position of maximum H lines, these lines are shortened, the inductance is decreased and therefore the resonant frequency is raised.

The Bolometer

In test equipment for use with wave-guides, metering is usually performed by a bolometer. This device consists of a piece of fine

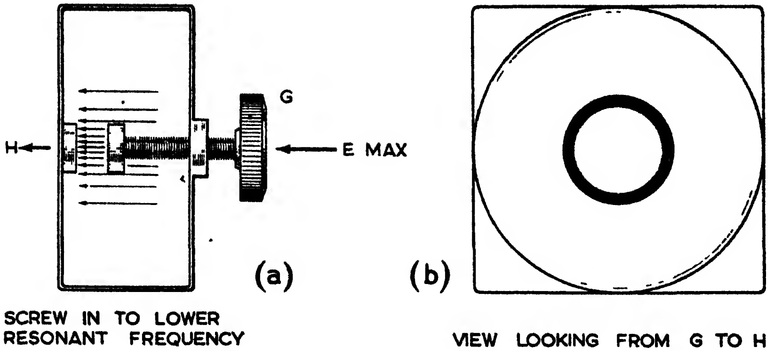


Fig. 2 —TUNING RESONANT CAVITY.

(a) Variable capacity introduced at the point of highest voltage.
(b) The mode is assumed to be such that the maximum voltage lies along H.G.

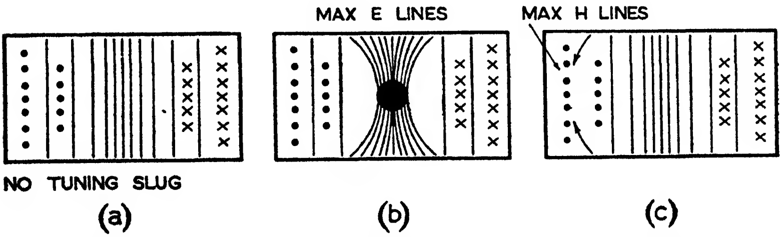


Fig. 3.—ALTERNATIVE METHODS OF TUNING RESONANT CAVITY.

metal wire sealed into a glass tube. The resistance of this wire changes as current is passed through it or R.F. power falls upon it.

If the bolometer is first calibrated against a known source of power, and its resistance-change curve plotted against current change, comparison can then be made directly between the mean R.F. powers which result in a given change of bolometer resistance.

The bolometer can be used for measuring power and also as a wavemeter indicator.

In the latter case, D.C. is passed through the bolometer to obtain a reading which acts as a standard of comparison. The test oscillator or the transmitter is then switched on (according to whichever is to be checked) and an increased reading will then be observed, over and above the standing current, *i.e.*, D.C. plus that due to R.F. power.

As the wavemeter micrometer is tuned to bring the cavity into resonance, the bolometer reading will suddenly dip. This is because

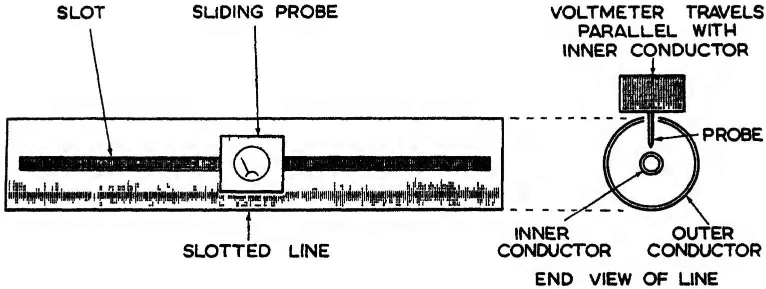


Fig. 4.—SLOTTED LINE

R.F. power is being absorbed in the wavemeter as it is tuned to resonance.

To Test the Efficiency of the T.R. and T.B. System

(*a*) The bolometer reading with the transmitter switched off is noted.

(*b*) Compare reading (*a*) with the bolometer reading in the receiver branch of the main guide when the transmitter is switched on.

(*c*) Repeat with the bolometer switched into the transmitter section of the main wave-guide instead of in the receiver branch. If the standing wave ratios obtained in this manner are a and b respectively then the ratio of ratios $\frac{b}{a} \times k$ (where k is some known constant that takes care of additional attenuation introduced by the test set connections), is a measure of efficiency of the T.B., T.R. system.

The transmitter peak power for a given set is a measure of the efficiency of the R.F. generator, all other factors being equal. This is measured by measuring the increase in resistance when the bolometer is read in the transmitter branch of the main wave-guide with the transmitter switched on. The mean power is then read off from a calibration curve in milliwatts. The R.F. peak power is milliwatts $\times$ 2,000 (approximately).

Use of Neon Tube

Standing waves of voltage on open wire transmission lines can be detected by a neon tube. If the line is correctly terminated the bulb will glow with constant brilliance as it is moved along the line because *no standing waves exist.* Better results may be obtained by using a valve rectifier capacity coupled to the line. If standing waves do exist the positions of nodes and antinodes will be indicated by variations in brilliance of the neon tube, alternatively by the values of the valve voltmeter readings.

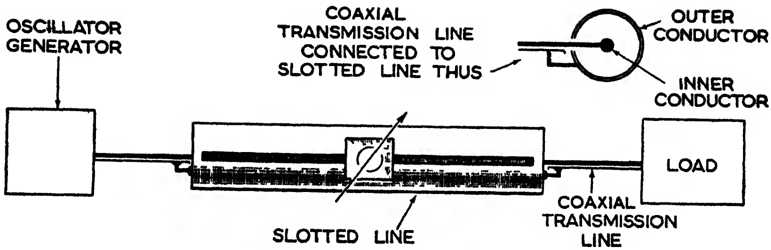


Fig. 5.—CONNECTIONS OF SLOTTED LINE TO LOAD FOR TEST PURPOSES.

The Slotted Line

A slotted line (Fig. 4) is a section of coaxial line with a slot along the outer tube to permit coupling an R.F. probe loosely to the inner conducting surface. It is joined in series with the line under test.

A slotted line may be used for the following purposes :—

(*a*) To determine the ratio of voltages at maximum and minimum points of standing waves along the line.

(*b*) To determine the position of these points with respect to a reference point.

From this information it is possible to deduce the resistive and reactive nature of a load at a specified frequency.

In practice a slotted line may be used to adjust the matching of aerial arrays to transmission lines, to determine the resistive and reactive components of a load (hence to ascertain the phase angle), and to adjust the input systems of receivers, dummy loads, etc., for correct match to a line.

Requirements for Accurate Measurement

The requirements for accurate measurement are uniform impedance throughout the length of the measuring device, and uniform spacing of the probe in its travel along the inner conductor. Other considerations are good earthing of the probe box to the outer conductor and rigidity of the coaxial assembly with minimum slip in travel of the voltmeter probe (see Fig. 5).

The characteristic impedance of the slotted line should be equal to the characteristic impedance of the associated coaxial line. Alternative inner conductors with different diameters are sometimes supplied, in order that the impedance of the slotted line may be changed to suit the characteristic impedance of the coaxial line which is to be measured.

The R.F. voltmeter as usually used with the slotted line is a diode or crystal detector with a current meter and tuned input capacitatively coupled to the inner conductor (Fig. 6). A high output oscillator is required to excite the line at the required frequency, since the diode or crystal sensitivity is comparatively low.

A feeder is usually adjusted for matching to an aerial (dipole) by

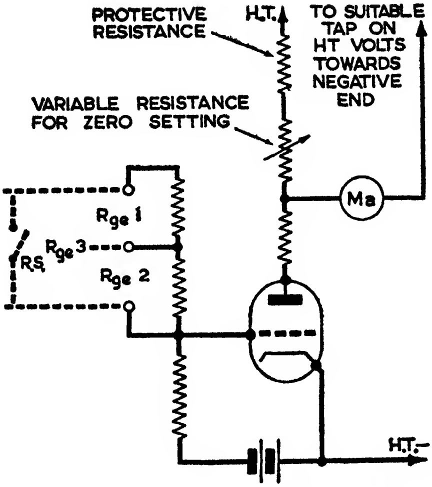


Fig. 6.—Circuit diagram for valve voltmeter for d.c. measurements.

changing the position or dimensions of the matching stub for minimum observed standing wave.

In some cases, however, it is necessary to determine the resistive and reactive components and the phase angle of the load. To accomplish this the standing wave ratio must be accurately measured, also the distance of some minimum (or maximum) voltage point with respect to some reference point must be determined. Firstly, for total reflection, *i.e.*, when the line is shorted at one end and maximum S.W.R. is present and, secondly, with the short circuit removed and the load substituted.

Additional considerations upon which accuracy of measurement depends are :—

(*a*) The input to the line should be a sine wave. If the output from the oscillator is not a pure sine wave harmonics will cause subsidiary nodes and antinodes to appear at points along the line which will cause inaccurate readings.

(*b*) If the characteristics of the rectifier associated with the voltmeter are not linear, readings plotted against distance along the line will not be a sine curve, even if the input is free from harmonics. The resulting voltage curve for the standing wave will be a distorted sine curve, as a result of which the voltage minima are not sharply defined.

When it is necessary to determine the standing wave ratio accurately the wave is plotted, as measured by the particular detector under working conditions and with the generator adjusted exactly for half-scale deflection at the maximum voltage point. A sine wave is then constructed geometrically and superimposed on the measured and plotted standing wave, with zero and maximum points coinciding. The sine wave indicates the results that would be obtained if the detector were linear. Thus by comparing the observed values of $\dfrac{\text{max. voltage}}{\text{min. voltage}}$ with the constructed sine curve, a suitable correction can be made to compensate for the non-linearity of the detector.

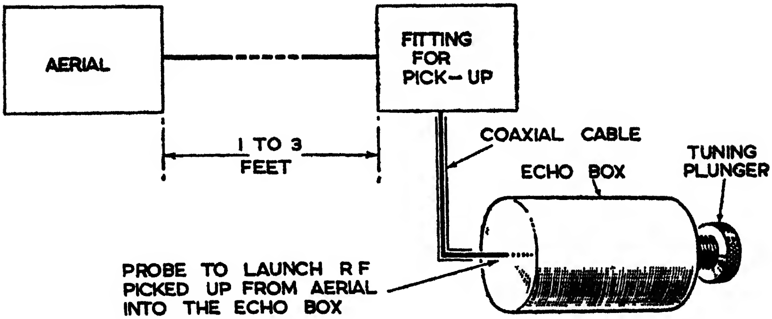


Fig. 7.—SCHEMATIC ARRANGEMENT FOR TEST, USING ECHO BOX.

Receiver and Transmitter Check Up

Echo boxes are resonant cavities designed to have a very high Q—around 35,000. They are used with other test equipment to provide an artificial target which may be used for tuning purposes when no real target is available. An echo box may also be used to provide a rough check of the transmitter output.

The principle on which the echo box operates is simple. Energy is fed in from the transmitted pulse, as a result of which it rings or resonates for some few microseconds after the transmitted pulse has ceased.

In this state it radiates energy back to the aerial and produces an artificial echo signal on the screen of the C.R.T.

Assuming that the cavity is tuned to resonate to the transmitter frequency and that its Q is normal, *i.e.*, has not deteriorated, the ringing time depends upon the peak power of the aerial output and the pulse shape of the R.F. envelope.

If the receiver is functioning normally this is a rough check on the transmitter output, provided the receiver is tuned to the artificial echo.

Conversely, although not so conclusive, all other things being equal, the echo from the artificial target may be used to tune the receiver to the transmitter average frequency, but it can scarcely be classed as a satisfactory check on the sensitivity of the receiver itself. It is only a help.

Echo boxes may take the form of a cylindrical cavity one and a half wavelengths long and oscillating in the H_{011} mode.

Fig. 7 shows schematically the arrangement of the echo box for test purposes.

When an echo box is permanently installed an increased width of range or echo pulse indicates an improvement in performance of the *system*.

During normal operation of the set the echo box is either detuned or disconnected in order to avoid confusion between artificial echoes and echoes from real targets.

Valve Voltmeters

Referring back to Fig. 6, the circuit diagram of a simple valve voltmeter for measuring D.C. voltages, the meter reading in the anode is proportional to the voltage applied to the grid.

R.S. is a range switch which enables several ranges of voltages to be measured. The anode volts and grid volts are constant.

The variable resistance is adjusted so that the meter in the anode circuit reads zero when no voltage is applied to the grid. For various settings of the range switch, the current flowing through the meter is proportional to the D.C. voltage applied to the grid. Direct readings of these values can be obtained from a calibrated scale or chart supplied with the instrument.

Determination of Impedance and Phase Angle of a Load

In order to determine the impedance and phase angle of any load, using a slotted line, it is necessary to employ in conjunction with it a special scale and with it to locate a reference point. The scale has the general appearance of Fig. 8 (a). This reference point is the position of a suitable voltage minimum when total reflection is present, *i.e.*, when the line is short circuited.

The test circuit is set up as in Fig. 5, together with the special scale. The load end of the transmission line is short-circuited and, with a known constant frequency input from the generator, the probe is adjusted to a voltage max. (reading given by the valve voltmeter). The next step (2) is to find a suitable point along the slotted line for min. voltage. This is done by moving the probe and comparing voltmeter readings. A voltage min. point or node should be selected near the centre of the slotted line.

The scale is then moved along the slotted line, so that 90° on the scale coincides with the reference point which has been located by step 2. This is step 3. Step 4 is to remove the short-circuit and with the load connected move the probe along the slotted line to locate a voltage minimum under the new conditions (*i.e.*, with a reflected wave less than total reflection and as determined by the load). The scale must not be moved since it marks the reference point previously found in step 2, but note the point on the scale where this new minimum is found.

Having noted the position and value of the voltage reading at the new minimum point, the next step is to determine the ratio $\dfrac{\text{voltage max.}}{\text{voltage min.}}$. Voltage min. has just been found, but the next voltage max. along the slotted line must now be located and measured, by moving the probe once more and noting the voltmeter

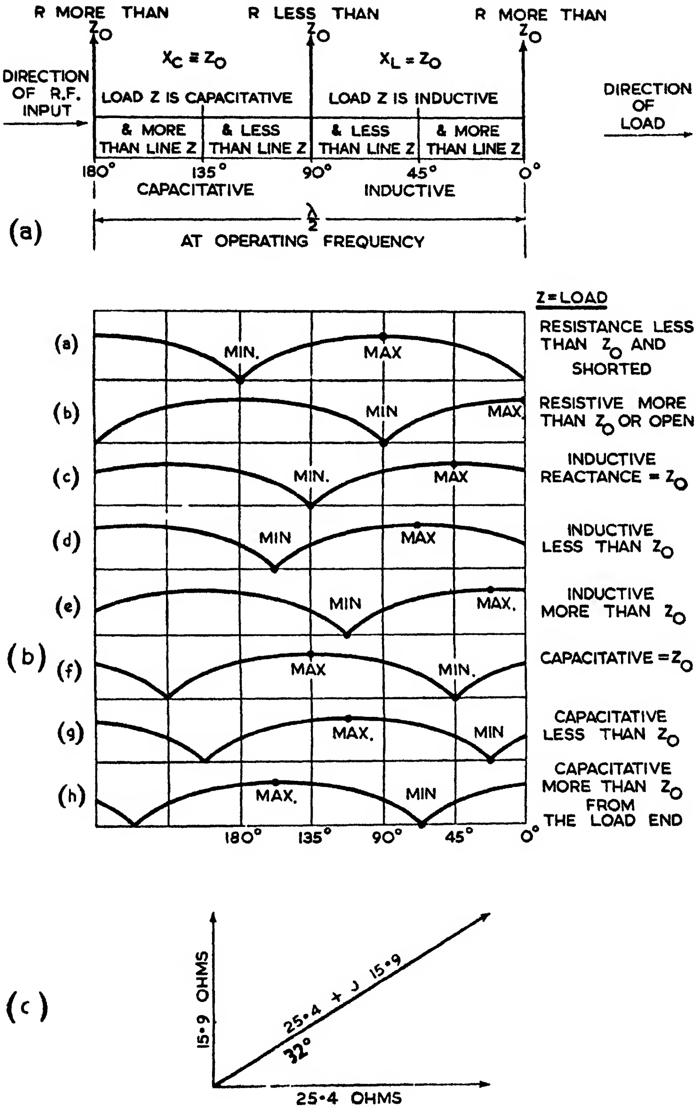


Fig. 8.—DETERMINATION OF LOAD IMPEDANCE AND PHASE ANGLE.

(a) Appearance of special scale. (b) Position of first min. and first max. from load end for resistive and reactive loads, equal to, or greater or less than Z_0. (c) Vector representation of load.

readings. The S.W.R. found in this manner is designated R. This is the S.W.R. of the load and it bears a definite and known relationship to the short-circuited condition where the phase angle is known to be 90°.

From Fig. 8 (b) it will be seen that at the outset, when shorted, the conditions at the line termination must be as (a) since the line is short circuited then. When the short is removed, the first min. from the load end moves to (b), (c), (d), (e), (f), (g) or (h), according to the nature of the load which terminates the slotted line after removal of the short.

This movement of the minimum and also the maximum imposed on the line by the nature and value of the load termination is repeated at intervals of 90° or $\frac{\lambda}{4}$ all along the line. Consequently, after removal of the short, when the probe is moved to find the position of the new min., its location when found, reflects the changed conditions at all points along the line up to the generator terminals. The direction of deviation under loaded conditions, of maxima and minima from their position, as found for complete antiphase under short-circuited conditions, indicates the reactive nature of the load and its magnitude relative to Z_0 of the line. This taken in conjunction with the ratio of the S.W.R.s makes it possible to determine the ratio $\frac{\text{load impedance,}}{\text{line impedance}}$ the nature of the reactance, whether capacitative or inductive, and the phase angle. Calculation is facilitated by special charts or curves provided for the purpose.

Thus, if R the standing wave ratio $= \frac{\text{Max. voltage}}{\text{Min. voltage}} = 2$ and if the probe has moved from the reference point forward to 60°. This means that the first min. must have moved 30° nearer the end of the line (*i.e.*, the load) after removal of the short. From Fig. 8 (b) the load is inductive and less than Z_0. From charts the ratio $\frac{\text{load impedance}}{\text{line impedance}}$ is shown to be 0·75 for this condition. If the characteristic impedance of the cable is taken as $Z_0 = 40$ ohms, the load impedance must be $40 \times 0{\cdot}75 = 30$ ohms inductive, and the phase angle given by the chart is 32°. The load impedance can, therefore, be expressed as $30 \cos 32° + J\,30 \sin 32°$

$$= 25{\cdot}4 + J\,15{\cdot}9 \text{ ohms,}$$

or vectorially as in Fig. 8 (c).

The power loss in an unmatched line to that lost in a matched line as a function of S.W.R.

$$\frac{\text{Power loss in unmatched line}}{\text{Power loss in matched line}} = \frac{1 + R^2}{2R}.$$

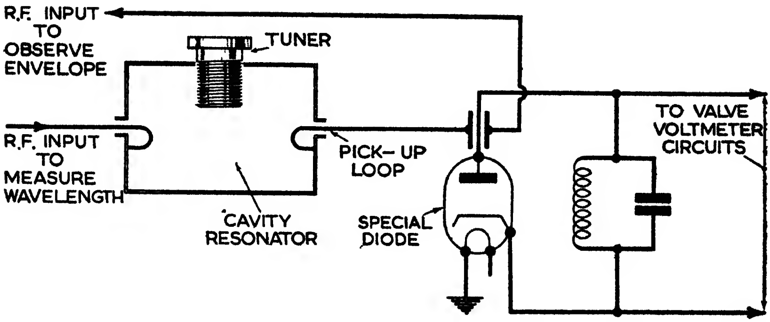


Fig. 9.—TYPICAL WAVEMETER FOR USE WITH MICRO-WAVE FREQUENCIES.

Measurement of impedance and phase angle with wave-guides is based on transmission line practice.

It is often more convenient to state the S.W.R. in terms of that part of the forward wave which is reflected.

Since the voltage and current at the measuring point must be the *vector* sum of the forward and reflected wave (voltage or current) at this point, the *vector* ratio of the reflected wave to the forward wave is termed "r," *i.e.*, the coefficient of reflection.

Hence the expression takes the form of

$$\text{S.W.R.} = \frac{1 + r}{1 - r}$$

where r is that portion of the forward wave which is reflected, *i.e.*, the reflection coefficient.

$$r \text{ (reflection coefficient)} = \frac{\frac{Z_L}{Z_0} - 1}{\frac{Z_L}{Z_0} + 1}$$

"r" is the *vector* ratio of voltage of the reflected wave to that of the voltage of the forward wave and this depends upon the *vector* ratio.

$$\frac{Z_L \text{ (line impedance)}}{Z_0 \text{ (characteristic impedance)}}$$

The *vector* ratio of voltage of the forward wave to the voltage at point of measurement is :—

$$\frac{\text{Voltage of forward wave}}{\text{Voltage of received wave}} = \frac{1 + \frac{Z_0}{Z_L}}{2}$$

Wavemeters for Measurement at Micro-wave Frequencies

Certain classes of wavemeters for use at micro-wave frequencies provide means for measuring the mean frequency and frequency spread (spectrum examination) of the transmitter pulse. It thus serves to indicate when the transmitter is tuned so as to radiate as much power as possible *within a frequency band equal to the acceptance band of the receiver*, the centre of the band remaining fixed from one pulse to another. In other words, the wavemeter enables the operator to tune the transmitter so as to obtain optimum spectrum width and stable working.

When this has been done, the wavemeter can be used to set the frequency of appropriate oscillators to the transmitter frequency, resulting from the above tuning procedure, and the oscillator can then be used to line up the receiver and check its performance.

The components are (1) the cavity resonator (or ringing box) tuned by a plunger controlled by a large calibrated dial, (2) a tuned diode line and a diode, (3) a valve voltmeter (see Fig. 9).

A small fraction of the transmitter output is collected by a probe, projecting not more than $\frac{1}{8}$ in. into the wave-guide run, and this is injected into the tunable cavity. The cavity has a Q of about 6,000 and only accepts frequencies over a narrow band. The output is rectified by a special diode connected across the diode line, and the rectified output can be measured on the valve voltmeter.

The deflection of the meter is proportional to the power picked up within a narrow band of frequencies, and the meter deflection gives an indication of the *spectrum* of transmitter frequencies.

The operation of measuring wavelength is carried out, by adjusting the screw plunger in the resonator, until a state of resonance is obtained. This is indicated by a dip of the meter, which is brought about when the probe begins to absorb energy from the cavity, and the corresponding wavelength can be read off from the calibrated dial of the screw adjustment to the resonator.

INDEX