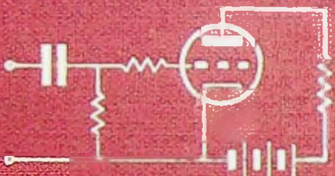
ULTRA-HIGH-FREQUENCY RADIO ENGINEERING



Ultra-High-Frequency Radio Engineering

THE MACMILLAN COMPANY
NEW YORK · BOSTON · CHICAGO · DALLAS
ATLANTA · SAN FRANCISCO

MACMILLAN AND CO., LIMITED
LONDON · BOMBAY · CALCUTTA · MADRAS
MELBOURNE

THE MACMILLAN COMPANY
OF CANADA, LIMITED
TORONTO

ULTRA-HIGH-FREQUENCY RADIO ENGINEERING

W. L. EMERY, M.S., FORMER INSTRUCTOR OF ELECTRICAL ENGINEERING, IOWA STATE COLLEGE

NEW YORK 1944

THE MACMILLAN COMPANY

PRINTED IN THE UNITED STATES OF AMERICA

PREFACE

This book is an outgrowth of the Electronics Teachers' Conference which was held at M.I.T. in the fall of 1941. As a result of the Conference, the author organized an ESMDT (later ESMWT) course in "Ultra-high-frequency Techniques" at Iowa State College. Since no text on the subject was then available, he prepared a set of lecture notes for the use of his students. Revisions and additions which were suggested as a result of teaching the course over a one and one-half year period were incorporated in the original notes and the present text was developed.

An effort has been made in the plan of the book to help the student correlate theory with practice even though the scope has been confined to basic principles. The problems which are included are of the type which involve the design of circuits or the prediction of their performance. The results of these problems are then checked in experiments which are associated with them. Usually temporary circuits assembled with clip leads are satisfactory for the purpose of making these tests. By a direct comparison between the theoretical and experimental results, the student is given an appreciation of variations which may be expected in such things as tube parameters, for example, as well as an appreciation of the errors which are sometimes involved in measuring techniques.

The book has been written for senior electrical engineering students, and assumes that they have a general background in elementary communication and electronics. Emphasis has been placed on the discussion of the component parts of ultra-high-frequency systems rather than on the systems themselves because the latter are readily broken down into their components. The mathematics is used as a tool, and the physical implications of the equations are emphasized. The derivations have been made as simple and direct as possible. Variations in particular pieces of apparatus are treated from a common viewpoint; for example, all oscillators are analyzed on the basis of "negative resistance." For the most part, the discussion has been concerned with ordinary circuits, but a few unusual circuits which illustrate the logical development of an idea have been discussed. A list of references is given at the end of each chapter which will help the student enlarge his acquaintance with the literature.

The author is indebted to many people for the contributions which they have made toward the preparation of the book. The series of lectures which were presented at the M.I.T. Conference in 1941 prompted the writing and provided one of the major sources of reference material. To the authorities who gave those lectures and especially to Dr. W. L. Barrow, who organized and directed the Conference, the author extends his sincere thanks. The Electrical Engineering Staff of Iowa State College have been most generous with their help and suggestions. In particular, Professors W. L. Cassell and J. D. Ryder, and Messrs. W. R. Abbott and E. M. Lonsdale have contributed ideas on experiments and presentation of theory which have been incorporated in the text. Professor Ryder and Mr. Lonsdale assisted in checking the manuscript in its final form. Suggestions which were made by students who used the original notes have helped materially in clarifying some of the more difficult parts of the theory. The illustrations are the commendable work of Mr. Dale Scott. Finally, there has been the continuous help and encouragement of the author's wife, Edith R. Emery. The cooperation of all these people is deeply appreciated.

W. L. Emery

Washington, D.C.

CONTENTS

EXPERIMENTS

Ultra-High-Frequency Radio Engineering

CHAPTER I

INTRODUCTION

1–1. Historical. While the extensive application of microwaves is a relatively recent development, the production of these waves is as old as radio itself, which began with the classic experiments of Hertz in 1887. These experiments demonstrated the optical properties of electric waves and thus confirmed Maxwell's Electromagnetic Theory of Light. It was inevitable that Hertz use very short waves because their properties are most nearly like those of light. The relationship between the long electric waves and light waves is less apparent and far more difficult to demonstrate.

Electric waves showed their merit as a means of communication in 1896 when Marconi successfully conducted tests over a distance of more than a mile. Each year thereafter he extended the range of his operation until, in 1901, he produced the first signals which were heard across the Atlantic Ocean. The vast commercial importance of these developments stimulated research in the art of radio communication, and study of some of the other uses of radio waves was temporarily neglected. Since it was soon recognized that the longer wavelengths propagated most successfully over the greater distances, attention was naturally directed toward them. Consequently, the use of longer wavelengths and the development of more powerful transmitters constituted the major progress made during the period which followed. A milestone in this progress was the invention of the vacuum-tube amplifier by DeForest in 1906.

The ideas concerning radio propagation were revolutionized shortly after the First World War. It was shown that phenomenally good reception of short waves was possible over very long distances. The quality of this reception, however, was dependent on the season and the time of day as well as the weather conditions. An explanation of these phenomena was formulated by Kennelly and Heaviside. They showed that an ionized layer of air far above the earth's surface could act like a reflector to radio waves and help them to travel very great distances. Experiments were set up which not only proved the existence of this conducting layer or ionosphere but also showed that it was made up of several parts. Numerous data were compiled on the ionosphere and the selection of a frequency to be used for a given communication problem became more scientific.

The growing interest in short waves brought forth very many unique applications of radio. Artificial-fever machines for the medical profession, radio beacons and navigation instruments for the aviation industry, and many other unusual devices worked more effectively when operated in the microwave bands; hence, studies were made on the production and transmission of these waves. Southworth began his investigation of wave guides. Hull had invented the magnetron, and Okabe together with others developed it as an efficient source of ultra-high-frequency power. The triode was adapted for high-frequency operation, and a radically new oscillator — the Klystron — was introduced by the Varian brothers. Since these efficient sources of high-frequency power have become available, the progress in the field of microwave radio has known no bounds.

The most universally useful measuring instrument in radio science as well as the heart of the modern television system is the cathode-ray tube. It originated in 1859 when Plücker observed a fluorescent glow on the glass of his two-electrode evacuated discharge tube. The fluorescence was obviously caused by something coming from the cathode. At first, it was thought that this something was a form of radiation; therefore, the name "cathode rays" was coined. The researches of J. J. Thomson in 1897, however, showed that "cathode rays" were in reality moving, charged particles which were later named "electrons."

During that same year, the first attempt to apply cathode rays in a measuring instrument was made by Braun. The Braun oscillograph was like the modern oscillograph in all its essential details. A stream of electrons was produced at one end of the tube and directed toward a fluorescent screen at the other end. Between the beam source and the screen were two magnetic deflecting systems which operated at right angles to each other. The current under observation by the instrument was used to produce the vertical deflection of the beam, and a periodic time-base current was used to produce the horizontal deflection. The figure traced on the screen was a temporal plot of the observed current. The images in these early tubes, however, were only partially satisfactory. Many improvements had to be made before present-day quality was obtained. One of the most significant of these improvements was introduced by Wehnelt in 1905. He replaced the inefficient cold-cathode source of electrons in the Braun tube by a thermionic source and thus obtained a more intense electron beam. The development of electron optics made it possible to narrow this beam into a very fine pencil and produce the sharp images which we know today. Modern inventors such as Farnsworth, vonArdenne, and Zworykin have extended the principle of the cathode-ray oscillograph to other apparatus and have firmly entrenched cathode-ray auxiliaries as essential components of high-frequency-radio systems.

1–2. Properties of Microwaves. Microwaves are very much like infrared light or heat waves. For this reason they are sometimes called quasi-optical waves. They are propagated very well through fog and mist, rain and snow, and in every kind of weather. Some authorities believe that very large hailstones may disperse them, but this has not been confirmed. The waves travel in straight lines and are not reflected by the ionosphere; therefore, the maximum range of their transmission is limited to line-of-sight distances. In order to extend this range, antennas are mounted as high above the surrounding territory as possible. The antennas for microwave installations are notably small. Indeed, the radiation process is so efficient at some of these frequencies that special precautions must be taken to prevent radiation from the circuits themselves. Highly efficient portable apparatus can therefore be constructed. Since the waves can be readily concentrated into very narrow beams, point-to-point microwave transmission is almost as effective as wired transmission.

Station crowding is not nearly the problem at high frequencies that it is at low frequencies. A very small change in wavelength is accompanied by a large change in frequency; therefore, an extremely wide band of frequencies is included in the ultra-high-frequency spectrum. It is especially well adapted for accommodating the wide-band services.

Nearly every solid and liquid surface is a good reflector for microwaves. The metallic surfaces are naturally the best and they are used in the form of parabolic mirrors to concentrate the waves for point-to-point communication. While the reflecting properties of the earth and sea are not nearly so good, they are good enough to permit the effective operation of absolute altimeters. Indeed, one of the serious problems of microwave transmission is the spurious reflections from the buildings in a metropolitan area.

1–3. Applications to Communication. A very wide band of frequencies is required for the transmission of both television and frequency-modulation signals. Consequently, the ultra-high frequencies are ideal for television and frequency-modulation services. The distribution of these signals for chain broadcasting takes up a large number of ordinary channels on a wire circuit; however, the signals can be effectively distributed by radio over a chain of point-to-point microwave repeaters spaced at appropriate intervals between the broadcasting centers. Of course, point-to-point beam transmission of this sort can also be used for an endless number of other communication services.

Portable combination transmitter and receivers operating in the microwave bands are taking an increasingly important place in the world of communication. The first units of this kind were used by broadcasters to relay play-by-play descriptions of some sports events. Nowadays, such units also assist in correlating the activities of the police or military forces.

1–4. Applications to Navigation. The absolute altimeter or terrain-clearance meter is one of the most fascinating developments of radio. Previously, airplane pilots had to depend on barometric altimeters whose readings only indicated elevation with respect to sea level and were seriously affected by changes in the weather. In mountainous districts during storms, the pilot had to judge the elevation of the terrain beneath him before he could determine which flying altitude would give him ample clearance with the ground. Naturally the only satisfactory altimeter for use under such circumstances is a terrain-clearance meter.

Several instruments of this sort have been devised. One type has a transmitting antenna mounted on one wing tip and a receiving antenna mounted on the other. An ultra-high-frequency signal which has been frequency modulated by a sawtooth wave is radiated from the transmitting antenna to the ground where it is reflected back to the receiving antenna. A portion of the signal also goes directly to the receiver along the underside of the wing. The time required for the arrival of the direct wave is constant; whereas, the time required for the reflected wave depends on the absolute altitude of the plane. A phase shift which is proportional to this altitude is thus introduced between the sawtooth modulations of the direct wave and the reflected wave. For a given altitude, there is a certain fixed frequency difference (except during a very small portion of the modulation cycle) between the direct signal and the reflected signal picked up by the receiver. This difference frequency is measured on a frequency meter whose scale has been calibrated to give terrain clearance in feet. A switch is provided which enables the operator to select either a high-accuracy, narrow-range calibration or a low-accuracy, high-range calibration.

Another fundamental navigation appliance is the microwave instrument-landing system for airplanes. This device enables a pilot to make a blind landing on an airport during adverse weather conditions. The basic idea for one such system was proposed by I. R. Metcalf of the CAA, who suggested that a fog light be placed on the ground near the start of the runway. Ahead of it and on either side of the runway center line would be two others at such a height that a surface passing through the three lights would make an angle with the ground equal to the best glide angle for the landing airplane. If the pilot came into the airport so that the ground light appeared midway between the other two and on a straight line connecting them, he would be directly above the runway center line and would be coming in at the proper angle for a landing. Since no visible lights will penetrate fog sufficiently well to be used directly, this system was adapted for microwaves.

The ground light is replaced by a set of highly directional horns which radiate ultra-high-frequency signals. One horn is directed at an angle slightly higher than the glide angle, while another is adjusted to a slightly lower angle. The signal from each of these horns is modulated with a characteristic audio-frequency for identification. When the audio tone from the upper horn comes through the receiver of the plane, the spot on the instrument-panel cathode-ray oscillograph is deflected upward from its center position. When the tone from the lower horn comes through, the spot is deflected downward. If both signals are received with equal intensity, the spot remains at the center of the screen. Another pair of horns are arranged so that their signal will indicate lateral departures of the plane from its normal path by moving the spot on the cathode-ray screen to the right or to the left of the center position. Thus the position of the spot on the screen shows at a glance if the plane is off the normal glide path, how much it is off, and in what direction. This indication in itself is sufficient to land the plane provided the pilot also simultaneously observes his other flying instruments. The necessity of watching several instruments can be avoided if their information is brought to the cathode-ray screen in the form of an additional pair of spots. The tilt of the line connecting these two additional spots is controlled by the artificial horizon gyroscope, and the gyroscopic compass moves them bodily to the right or left of the screen as the plane changes the direction of its flight. The three landing lights of the Metcalf system are thus simulated by these three spots on the cathode-ray screen. The illusion is made more perfect if the spacing between the pair of spots controlled by the flight instruments is made proportional to the strength of the glide-beam signals being received. In landing, the pilot flies into the position where the three spots lie in a straight, horizontal line with the center spot midway between the other two and at the exact center of the screen. Then as the flying field is approached, the two outer spots move toward the edges of the screen in exactly the same manner as the landing lights would appear to do in the Metcalf arrangement. A commutating switch enables an ordinary single-beam cathode-ray tube to produce the three spots required.

References

"Aircraft Radio 1939," *Electronics*, January 1939, pp. 10–14, 43.

Bowles, E. L., *et al.*, "The CAA–MIT Microwave Instrument Landing System," *A.I.E.E. Trans.*, 1940, pp. 859–865.

Encyclopaedia Britannica, 14th ed. London: Encyclopaedia Britannica Co., 1929.

Encyclopedia Americana. New York: Americana Corporation, 1934.

MacGregor-Morris, J. T., and J. A. Henley, *Cathode-ray Oscillography*. London: Chapman & Hall, Ltd., 1936.

Rice, C. W., "Transmission and Reception of Centimeter Radio Waves," *Gen. Elec. Review*, August 1936, pp. 363–369.

"Wireless Altimeter," *Wireless World*, February 2, 1939, pp. 101–102.

CHAPTER II

VOLTAGE-REGULATED POWER SUPPLIES

2–1. General. The power supplied by electric utilities is almost universally alternating current and must be converted to direct current or rectified before it can be used to operate radio equipment. Even when direct current is already available, it is often desirable to change it to alternating current first so that it can be readily transformed to a new voltage level and then reconverted to direct current. A unidirectional circuit element is used for the rectification process. Ordinarily, a diode vacuum tube is employed; however, gas-tube rectifiers are used when particularly high currents are necessary. The term "power supply" is defined as that portion of a circuit which includes the filament transformers and the rectifier together with the associated transformer, filter, and regulator.

Voltage regulators are usually incorporated in power supplies for microwave apparatus. They are necessary for the proper operation of some of the circuit elements used. The Klystron oscillator, for example, changes the frequency of its oscillation with changes in plate voltage. The wave form produced by some trigger circuits is also directly dependent on the voltages used.

2–2. Half-wave Rectifier. Figure 1 is an illustration of a simple half-wave rectifier. Current flows through the tube only when its plate is positive with respect to its cathode. The current in a pure resistance load, therefore, has the form of a series of half-cycle sine-wave pulses. The instantaneous current is equal to the instantaneous voltage applied by the transformer divided by the sum of the resistance of the tube r_p and the resistance of the load R_L. The d-c output current is equal to the average value of the load-current wave. The other components of the load-current wave are alternating current and indicate a lack of what would be perfect rectification. These other components collectively constitute the ripple of the power supply.

The output current of a half-wave rectifier is sinusoidal in form for the first half cycle and is zero for the second half cycle, and

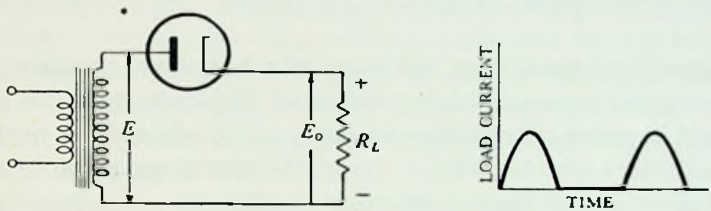


FIG. 1. Half-wave rectifier and output current with a resistance load.

since the average value of a sine wave taken over a half cycle interval is $\frac{2}{\pi}$ times the maximum value, the average of the output current is one-half this value or

$$I_o = \frac{1}{\pi} I_m$$

A half sine wave such as that being supplied by the transformer represents half of the power-producing possibilities of a full sine wave. Since the power dissipated in a circuit is proportional to the effective current squared, the effective value of a half sine-wave current squared is equal to one-half of the effective value of a full sine-wave current squared. The effective value of the current being supplied by the transformer is therefore equal to

$$I = \sqrt{\frac{1}{2}\left(\frac{I_m}{\sqrt{2}}\right)^2} = \frac{I_m}{2}$$

The efficiency of rectification is defined as the ratio between the d-c power output and the a-c power input. The d-c power output is equal to the d-c component of the load current squared and multiplied by the load resistance

$$P_o = I_o^2 R_L$$

The a-c power input is equal to the effective current squared and multiplied by the total resistance of the circuit

$$P_i = I^2(R_L + r_p)$$

Consequently, the efficiency of rectification is

$$\eta = \frac{P_o}{P_i} = \frac{\frac{I_m^2}{\pi^2} R_L}{\frac{I_m^2}{4}(R_L + r_p)} = \frac{4}{\pi^2}\frac{R_L}{R_L + r_p} = 40.6\% \frac{R_L}{R_L + r_p} \tag{1}$$

The theoretical maximum efficiency of a half-wave rectifier is thus 40.6 per cent; however, higher values of efficiency than this can be obtained in practical rectifiers by using tubes which have negligible loss and filters which reshape the load-current wave into a nearly square pulse rather than a sinusoidal pulse.

The d-c output voltage of a half-wave rectifier is equal to the d-c output current multiplied by the load resistance.

$$E_o = I_o R_L = \frac{I_m}{\pi} R_L \tag{2}$$

Since the maximum value of the voltage applied by the transformer is equal to the maximum value of load current multiplied by the total resistance of the circuit, the d-c output voltage can also be expressed as follows:

$$E_o = \frac{E_m}{\pi} - I_o r_p \tag{3}$$

The last term in the above equation is the drop through the rectifier tube. If it is neglected, the d-c output voltage in terms of the effective voltage E supplied by the transformer is $E_o = 0.45\,E$. Some types of filter circuits enable the half-wave rectifier to develop considerably more output voltage than this equation would indicate. If the load is a very high resistance which is shunted by a condenser of large capacity, the d-c output voltage will be nearly equal to $\sqrt{2}\,E$, the maximum value of the transformer voltage. The condenser charges up to the peak voltage and since the discharge through the bleeder is slow, the voltage across the condenser falls only a very small amount before another charging pulse arrives.

While the above analysis is intended for vacuum-tube rectifiers, the results can be extended to include gas-tube rectifiers. The main difference between the gas tube and the vacuum tube is that the former has a constant internal drop rather than an internal resistance, and it does not begin to conduct until a voltage higher than the breakdown value is applied. Consequently, the load current pulse using a gas tube is not the same shape as the load current pulse using a vacuum tube. It is in the form of a half sine wave with the ends clipped off. Integration must be used to find its average and effective values; however, these calculations are not necessary in practical cases. The tube drop is usually negligible in comparison with the applied voltage, and the equations which have been developed can be applied directly provided the plate resistance r_p is equated to zero.

2–3. Full-wave Rectifier. The circuit diagram for a full-wave rectifier is shown in Fig. 2. Each tube conducts for half the cycle so that the output wave is of the form shown in the diagram. The

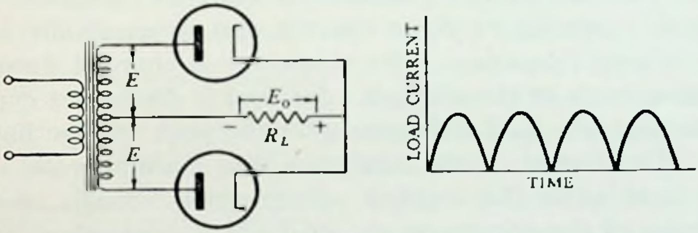


FIG. 2. Full-wave rectifier and output current with a resistance load.

output d-c current has a magnitude equal to the average value of each of the half-cycle sine waves or

$$I_o = \frac{2}{\pi} I_m$$

The effective value of the current wave delivered by the transformer is

$$I = \frac{I_m}{\sqrt{2}}$$

The efficiency can be calculated in the same manner as before. It is

$$\eta = \frac{\frac{4}{\pi^2} I_m^2}{\frac{1}{2} I_m^2} \frac{R_L}{R_L + r_p} = \frac{8}{\pi^2} \frac{R_L}{R_L + r_p} = 81.2\% \frac{R_L}{R_L + r_p} \tag{4}$$

Here the theoretical maximum efficiency is 81.2 per cent, but again this figure can be exceeded in practice by using a filter circuit which tends to square up the pulses of current being supplied to the load. The d-c output voltage for the half-wave rectifier is

$$E_o = I_o R_L = \frac{2\,E_m}{\pi} - I_o r_p \tag{5}$$

and if tube drop is neglected

$$E_o = \frac{2\sqrt{2}}{\pi}\,E = 0.9\,E \tag{6}$$

The effective voltage in these equations is measured between the center tap and one end of the power transformer.

2–4. Power Supply Filters. A low-pass filter circuit is inserted between the rectifier tube and the load for the purpose of smoothing out and suppressing the ripple voltages which would otherwise be present. The simplest low-pass filter is a shunt condenser. It has an infinite reactance to direct current and is practically a short circuit to high frequency. The condenser is charged during the conduction cycle of the rectifier tube, and it discharges exponentially through the load resistance after the peak voltage has been passed. The charge on the condenser thus maintains the voltage on the load while the rectifier voltage goes through zero. An oscillogram of the voltage on the load has a sawtoothed appearance. The performance of this filter is computed by applying the equation for the exponential discharge of a condenser through a resistance

$$e = E_m \epsilon^{-\frac{t}{RC}} \tag{7}$$

e is the instantaneous voltage across the load, E_m is the crest voltage of the transformer and the voltage to which the condenser is charged during the conduction cycle, C is the capacity of the condenser in farads, R is the resistance of the load in ohms, t is the time in seconds, and ϵ is the base of natural logarithms. The voltage across the load follows this curve until the condenser is again charged by the next cycle of rectification. A rectifier using this simple filter has very low efficiency because the current supplied by the transformer flows only for an extremely short portion of the cycle. The regulation or change in d-c output voltage with changes in load is also poor. In spite of these shortcomings, the circuit is useful in power supplies which are intended for low-current high-voltage devices such as cathode-ray tubes.

A series inductance must be added to the filter circuit if it is to square up the sine-wave pulses of current from the rectifier and thus improve the efficiency of rectification. Even more complicated circuits consisting of several series inductances and shunt capacities are often used. The computations on them are generally made with the assistance of design charts.

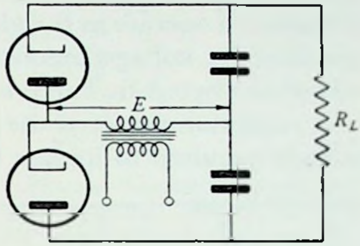


Fig. 3. Voltage-doubler circuit.

2–5. Voltage Doubler. The voltage doubler circuit is employed to obtain a higher d-c output voltage from an a-c source than would otherwise be possible. A typical circuit is shown in Fig. 3. The two condensers are charged during alternate portions of the cycle through the two rectifier tubes. The condensers are connected in series so that their voltages add and give a load voltage which approaches twice the crest voltage of the source. Other circuits permit the addition of more condensers in cascade and thus still greater multiplications of voltage are possible. All voltage doublers exhibit poor regulation and their current output is very limited. Their use is practically confined to power supplies for high-voltage cathode-ray tubes.

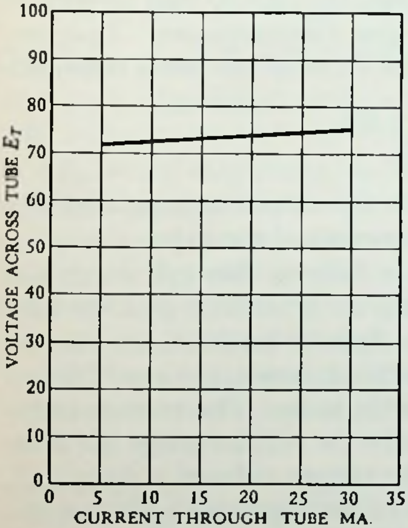


Fig. 4. Characteristic curve of the VR–75 regulator tube. Starting voltage 105 volts; operating current 5 to 30 ma.

2–6. Gas-tube Regulators. A two-electrode tube containing a small amount of gas has some very useful electrical properties. As the current through the tube is increased, a point is reached, at the breakdown voltage, where the drop across the tube falls to a lower level. Further increases in current are accompanied by practically no change in voltage. When this effect begins, the cathode starts to glow and as the current increases, the area of glow widens until it covers the entire cathode. Then the application of additional current causes a rise in tube voltage which will produce a spark between the electrodes if it is carried too far. Tubes which are designed to operate in the glow-discharge or constant-voltage range are used for voltage regulators. A typical example is the VR–75 whose characteristic curve is given in Fig. 4.

A regulator circuit of the gas-tube type is shown in Fig. 5. The voltage required to initiate the glow discharge is higher than that required to maintain it; therefore, the power supply must deliver a voltage equal to or greater than the starting voltage of the tube. A series resistor is connected between the input and the tube so that its current rating will not be exceeded under any circumstances. This prevents the occurrence of sparking. The value of the series resistance is given by

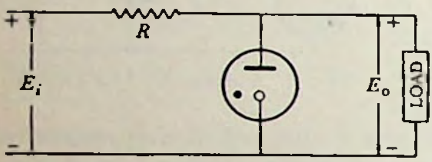


FIG. 5. Gas-tube voltage regulator.

$$R = \frac{E_i - E_o}{I_m} \tag{8}$$

where E_i is the supply voltage, E_o is the output or regulated voltage, and I_m is the maximum rated current of the tube.

The operation of the circuit is as follows: Enough current is taken by the tube so that the IR drop in the resistor plus the tube drop is equal to the supply voltage. As load is added, less current is taken by the tube and more by the load; hence, the total current and, therefore, the IR drop remains the same. The voltage across the regulator tube, which happens to be the voltage across the load, remains practically constant. If the supply voltage is increased, enough additional current will flow in the regulator tube to make the IR drop through the resistor compensate for the difference.

Figure 6 is a circuit which can be used to obtain an exactly constant voltage from a gas tube such as the VR–75 whose typical voltage vs. current characteristic is shown in Fig. 4. This characteristic is the same as that for a constant voltage E(71 volts in the case of the VR–75) in series with a resistance R_T whose magnitude is equal to the slope of the curve (133 ohms for the VR–75); hence,

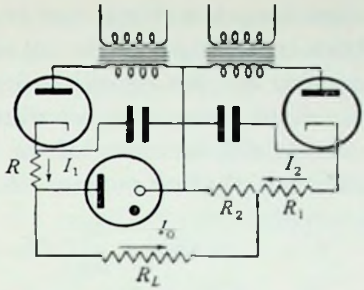


Fig. 6. Constant-voltage gas-tube regulator. The primaries of the supply transformers are connected either in series or in parallel.

the equation of a gas tube's characteristic can be written as

$$E_T = E + IR_T$$

where E_T is the voltage across the regulator tube. The object of the circuit shown in Fig. 6 is to introduce enough bucking voltage to eliminate the effect of the IR_T drop in the above equation. If the primaries of the transformers are connected in series, then a fixed relationship K exists between the currents in their secondaries or $I_1 = KI_2$. Writing Kirchhoff's laws for the loop including the gas tube, R_2 and R_L, we have:

$$E + IR_T = E + (I_1 - I_o)R_T = I_oR_L + (I_o + I_2)R_2$$

which is $$E + I_1R_T - I_2R_2 = I_o(R_L + R_2 + R_T)$$

or $$E_o = I_oR_L = \frac{R_L}{R_L + R_2 + R_T}[E + I_2(KR_T - R_2)] \quad (9)$$

If $KR_T = R_2$, everything on the right side of this equation is a constant or, in other words, the voltage across the load becomes independent of voltage changes in the supply circuit. In a similar manner, it can also be shown that a constant output voltage can be secured by paralleling the primaries of the supply transformers so that the two rectified voltages will be related by a constant factor.

2-7. Degenerative-amplifier Regulators. The action of this regulator in the control of output voltage is very similar to that of a thermostat in controlling the temperature of a room. A change in the output voltage E_o of the circuit shown in Fig. 7 causes the internal resistance of the tube to change in such a direction as to restore the output voltage to its original value. The output voltage must change before the regulatory action can begin; hence, this regulator will not produce a perfectly con-

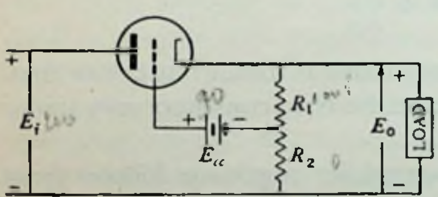


Fig. 7. Degenerative-amplifier regulator.

stant output voltage, but it will compensate for variations caused both by changes in the input voltage E_i and changes in the output current I_o. An equation for the effectiveness of regulation can be obtained from the calculus expression for the differential grid voltage dE_c in terms of the differential plate voltage dE_b and the differential plate current dI_b:

$$dE_c = \frac{\partial E_c}{\partial E_b} dE_b + \frac{\partial E_c}{\partial I_b} dI_b = -\frac{dE_b}{\mu} + \frac{dI_b}{g_m} \tag{10}$$

μ is the tube's amplification factor and g_m its mutual conductance. If $-K$ is the factor by which the change in output voltage must be multiplied to find the change in grid voltage, a ΔE_o change in output voltage will produce a $-K\,\Delta E_o$ change in the grid voltage of the tube. In the circuit shown, K is equal to the ratio

$$K = \frac{R_1}{R_1 + R_2}$$

When an amplifier is used to drive the grid, K is equal to the gain of the amplifier multiplied by the fraction of the change in output voltage used to drive its grid. The load resistor of the amplifier is connected either between the grid and plate or the grid and cathode of the regulator tube. Since the plate current of the tube I_b is equal to the output current of the regulator I_o, and the plate voltage E_b is equal to the difference between the input voltage and the output voltage $(E_i - E_o)$, the substitution of these values in equation 10 yields

$$\Delta E_o = \frac{\dfrac{\Delta E_i}{K\mu} - \dfrac{\Delta I_o}{K g_m}}{1 + \dfrac{1}{K\mu}} \tag{11}$$

Actually the performance of the regulator is somewhat better than this equation would indicate because the tube constants vary under extreme operating conditions.

The design of a degenerative-amplifier regulator follows from the analysis of its circuit. The values of the tube's average plate voltage and average plate current will determine the proper bias to use on its grid so that the operation will be on a suitable portion of the characteristic curve. The voltage rating of the battery is thus automatically determined since the bias voltage is equal to the battery voltage minus the drop across R_1. Higher values of K, hence, better regulation, can be obtained when a higher voltage battery is employed and R_1 is made much larger than R_2.

Example: A 2A3 tube is to be used as a degenerative-amplifier regulator. The load takes 45 ma at 100 volts.

(*a*) Select an appropriate value for the input voltage if the sum of R_1 and R_2 is 20,000 ohms.

(*b*) Find the magnitudes of R_1 and R_2 if E_{cc} is 90 volts.

The current through the resistance R_1 and R_2 is

$$I = \frac{100}{20{,}000} = 0.005 \text{ amp}$$

therefore, the total current through the tube is 50 ma. From the curves in the Appendix, we see that a plate voltage of 110 volts would provide a suitable operating point, i.e., where the tube currents and voltages could vary over a reasonable range. The input voltage must be 100 + 110 = 210 volts. The grid bias at the operating point is − 10 volts; hence, the drop across R_1 must be 90 − (− 10) = 100 volts.

$$R_1 = 20{,}000 \text{ ohms and } R_2 = 0$$

If E_{cc} had been 67 volts, then the drop across R_1 would have been

$$67 - (-10) = 77 \text{ volts and } R_1 = \tfrac{77}{100}\, 20{,}000 = 7700 \text{ ohms}$$

It is advisable to insert a grid-current-limiting resistor in the circuit also because the grid will become highly positive with respect to cathode if the input voltage is disconnected.

2-8. Bridge Regulators. The operation of bridge regulators is very much like that of a thermostat which is placed outdoors and used to control the temperature of a room. Such a thermostat would measure the outside temperature and supply just enough heat to compensate for the loss through the walls. If its mechanism were properly adjusted, it would keep the room temperature absolutely uniform because it does not depend on a change of room temperature for its operation. For the same reason, however, it would not recognize a change in heat loss such as that occasioned by opening a door and would not make any compensation for it. The bridge regulators measure the input voltage, and when it changes they make the compensations necessary to maintain the output voltage absolutely constant. They will not recognize changes in output voltage due to changes of output current; consequently, they will not regulate for these effects. They are used where the load is sensibly constant.

There are two forms of the bridge regulator — the mutual-conductance bridge and the amplification-factor bridge. The mutual-conductance bridge operates along the vertical line shown in the

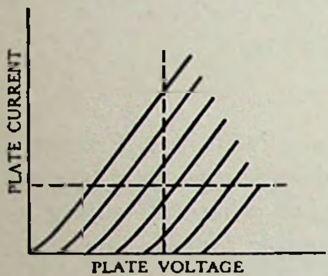


FIG. 8. Typical plate characteristics of a bridge-regulator tube. The mutual-conductance regulator operates along the vertical line and the amplification-factor regulator operates along the horizontal line.

plate characteristic curves of Fig. 8, and the amplification-factor bridge operates along the horizontal line. The regulation of the output voltage produced depends on the invariability of the tube "constant" involved. This is indicated by the uniformity of spacing between the crossings of the characteristic curves on the line of operation.

Mutual conductance is defined as the ratio of the change in plate current to the change in grid voltage when the plate voltage remains constant. A bridge to measure mutual conductance must fulfill these conditions. In other words, the bridge must be arranged so that changes in the grid voltage are not accompanied

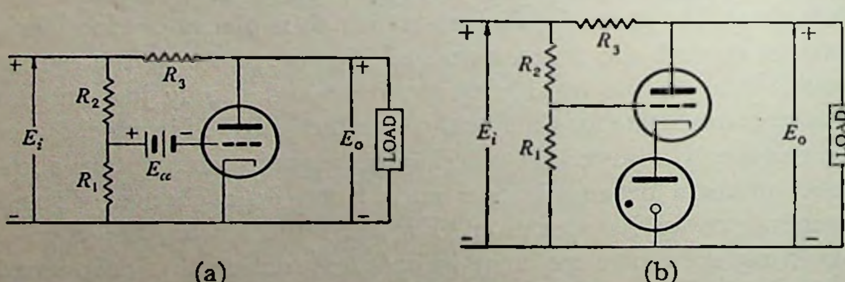


FIG. 9. Mutual-conductance regulator. The battery in circuit (*a*) is replaced by the gas tube in (*b*).

by any changes in the plate voltage. A circuit for doing this is shown in Fig. 9*a*. The change in grid voltage ΔE_c is equal to

$$\Delta E_c = \frac{R_1}{R_1 + R_2} \Delta E_i$$

If the accompanying change in plate current ΔI_b produces a voltage drop in R_3 which is exactly equal to the change in input voltage ΔE_i, the output voltage will remain the same. The ratio of the change in plate current to the change in grid voltage is the mutual conductance of the tube and if it is equal to

$$g_m = \frac{R_1 + R_2}{R_1 R_3} \tag{12}$$

the regulator will be in adjustment. The particular values of resistance used are not critical; however, they should not be so high that they interfere with the operation of the tube.

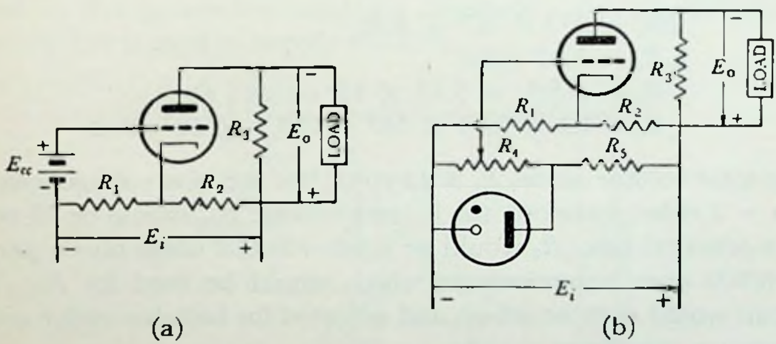


FIG. 10. Amplification-factor regulator. The bias battery in circuit (*a*) is replaced by the gas tube in (*b*).

Amplification factor is defined as the negative ratio of the change in plate voltage to the change in grid voltage when the plate current is kept constant. The constant plate current produced by the bridge circuit in Fig. 10*a* is passed through the resistance R_3 and the load, and the resultant drop constitutes the regulated voltage output. The bridge is balanced when

$$\mu = \frac{R_2}{R_1} \tag{13}$$

The bias battery can be eliminated by using the modified circuit shown in Fig. 10*b*.

Example: Find the size of the resistances and battery to be used in an amplification-factor bridge which employs a 6J5 tube, has an input voltage of 200 volts, an output voltage of 100 volts, and an output current of 3 ma.

If R_3 is considered to be the load resistance, it carries the output current of 3 ma. The tube's plate current is therefore 3 ma. From the characteristic curves in the Appendix, a suitable operating point for an $I_b = 3$ ma is found to be $E_b = 75$ volts, $E_c = -2$ volts. The voltage developed across R_2 must be equal to the sum of the output voltage and the plate voltage $100 + 75 = 175$ volts. For balance, $\dfrac{R_2}{R_1} = 20$, the amplification factor of the tube; hence, the drop across R_1 due to the current flowing through R_2 is $\frac{175}{20} = 8.7$ volts. In addition to this, R_1 also carries the plate current of the tube and so

$$\begin{aligned} 200 &= 175 + 8.7 + I_bR_1 \\ I_bR_1 &= 16.3 \text{ volts} \\ R_1 &= \tfrac{16.3}{0.003} = 5.43 \times 10^3 \text{ ohms} \\ R_2 &= 20 \times 5.43 \times 10^3 = 108.6 \times 10^3 \text{ ohms} \end{aligned}$$

The total voltage across R_1 is 25 volts, but the bias voltage selected was -2 volts; therefore, the battery voltage E_{cc} should be 23 volts. In a practical case, R_2 would be made 100,000 ohms plus a part of a 10,000 ohm potentiometer which would be used for R_1. The circuit would then be set up and adjusted for balance under actual operating conditions.

The choice between the two bridge circuits will depend on which side of the output is to be grounded. The negative side would be grounded in the mutual-conductance bridge and the positive side would be grounded in the amplification-factor bridge.

Erratic performance of vacuum-tube regulators can often be traced to changes in heater current which introduce spurious d-c voltages in the grid-cathode circuit. The heater current can be stabilized, and this source of trouble alleviated by the use of a current regulator or ballast lamp in the circuit. Physically, one of these lamps consists of an iron wire filament in a gas-filled bulb. Any tendency for the current to increase through such a tube is accompanied by an increase in the resistance of its filament due to the heating effect of the additional current.

2-9. Combinations. Many combinations and variations are possible with the four basic types of regulators. Combinations of gas-tube regulators with vacuum-tube regulators have already been mentioned. Two types of vacuum-tube regulators plus a gas-tube regulator appear in the single circuit of Fig. 11. A mutual-conductance bridge biased by a gas tube is used in place of the bias battery of a degenerative-amplifier regulator. In Fig. 12, a stage of amplification has been added to a degenerative-amplifier regulator. An amplification-factor bridge is used to supply the bias.

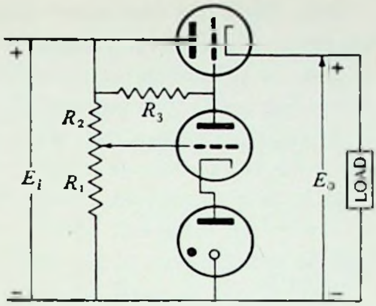


FIG. 11. Combination mutual-conductance and degenerative-amplifier regulator with a gas tube to provide the required constant bias voltage.

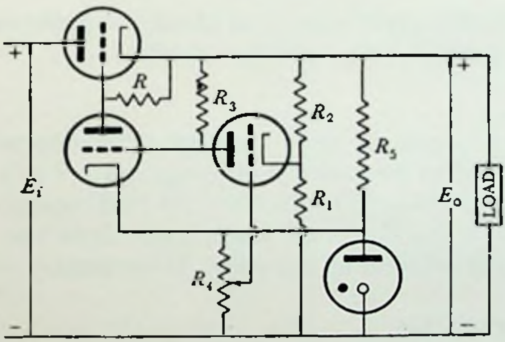


FIG. 12. Combination amplification-factor and degenerative-amplifier regulator with one stage of amplification and a gas-tube bias.

Problems

2-1. Show that the theoretical maximum efficiency of a half-wave rectifier is 50 per cent if the load-current wave is in the form of a square pulse rather than a half sine wave.

2-2. The load on a half-wave rectifier is a simple filter consisting of a 1 μf condenser and a bleeder resistance. The power transformer has a rating of 60 cps and 1000 volts rms. Find the minimum value which the bleeder resistance can have if the instantaneous voltage across it is never to fall below 1000 volts.

2–3. Find the resistance which should be used with a VR–75 voltage regulator if the input voltage is 125 volts.

2–4. A degenerative-amplifier regulator such as the one in Fig. 7 is delivering an output voltage of 100 volts to a resistance load of 3000 ohms. If R_1 is 50,000 ohms and R_2 is 5000 ohms, find the new output voltage when the input voltage is increased 10 volts. The tube used is a 2A3.

2–5. Solve problem 2–4 for the case where the regulator tube's grid is driven by an amplifier which has a net gain of $K = 12$. Assume that R_1 and R_2 are still part of the total load.

2–6. A 6J5 is used in a mutual-conductance regulator (Fig. 9*a*) to supply a current of 5 ma at 100 volts from a 150-volt supply.

(*a*) Select a suitable operating point for the tube.

(*b*) Calculate g_m graphically at the selected operating point and find the resistance of R_2 and R_3 if R_1 is 100,000 ohms.

(*c*) Calculate the voltage of the battery E_{cc}.

Experiment 1

Object:

The object of this experiment is to check the performance of the different kinds of vacuum-tube-regulator circuits.

Preliminary:

1. Compute and plot the output voltage of the degenerative-amplifier regulator described in the example of paragraph 2–7 as a function of the input voltage from 150 to 250 volts when the load current is held constant at 45 ma, also as a function of the load current from 0 to 60 ma with the input voltage held constant at 210 volts. Use a battery voltage E_{cc} equal to 90 volts.
2. Solve problem 2–6.

Performance:

1. Assemble the circuit described in paragraph 2–7. Use another 2A3 with variable grid bias as a load for the circuit and obtain the curves computed in the preliminary by adjusting the circuit to fulfill the conditions for which they were computed as nearly as possible. (*Note:* A separate filament transformer will have to be used for the load tube.) Plot experimental curves in a contrasting color on the same coordinate axes as the theoretical curves.
2. Assemble the circuit of problem 2–6. Adjust it for best operation and obtain data to plot a curve of its output voltage as a function of the input voltage.

References

AULMANN, A., "A Compensating Circuit for the Production of an Exactly Constant Voltage by Means of a Voltage Regulator Tube," *Zeits. f. Techn. Phys.*, 1938, pp. 320–322. (In German.)

HUNT, V. F., and R. W. HICKMAN, "On Electronic Voltage Stabilizers," *Review Scien. Instr.*, January 1939, pp. 6–21.

MILLMAN, J., and S. SEELY, *Electronics*. New York: McGraw-Hill Book Company, Inc., 1941, pp. 362–418.

NEHER, H. V., and W. H. PICKERING, "Two Voltage Regulators," *Review of Scien. Instr.*, February 1939, pp. 53–56.

REICH, H. J., *Theory and Applications of Electron Tubes*. New York: McGraw-Hill Book Company, Inc., 1939, pp. 365–368, 404–405.

SMITH, F. L., *The Radiotron Designer's Handbook*, 3rd ed. Sydney, Australia: The Wireless Press, 1941, pp. 185–198, 221–223.

CHAPTER III

ELECTRONIC SWITCHING AND SYNCHRONIZATION

3–1. General. Synchronization of the component parts of a complex high-frequency system is essential for its proper operation. In television, for example, it is necessary not only to keep the picture-forming circuits in step with each other, but also to keep them in step with the power-supply line voltage. If this were not done, an annoying band of shadow would continually drift across the picture. A phase discriminator, which detects changes in the phase angle between a locally generated control voltage and the power-supply line voltage, keeps these two voltages in step with each other. In turn, the low-frequency control voltage must synchronize with the high-frequency waves of the sweep circuits. This odd problem of synchronization is solved by means of an electronic switch of a particular type. Electronic switches are also employed wherever high-speed switching must be done.

In this discussion, the term "switching" is used in its most general sense. A d-c voltage is said to be "switched" if it is transferred from one position in a circuit to another, or if it is made to either appear or disappear across a particular circuit element.

3–2. Keying Circuits. A keying circuit is simply an amplifier which will operate when a switched voltage is applied to it and will cease operating when that voltage is removed. Keyed amplifiers are used in synchronizing systems to help form and control the synchronizing signal; however, they also have another very important application. A group of them feeding a common load impedance can be used in conjunction with a cathode-ray oscillograph for the purpose of observing several different wave forms simultaneously. The operation of the system is as follows: The input of each amplifier is connected to one of several signals which are to be observed. When one of the amplifiers is made operative or "keyed," its input signal will appear across the common load impedance and thus be applied to the cathode-ray oscillograph. When the control voltage is switched to the next amplifier in the line and keys it, the first amplifier becomes inoperative and its signal vanishes from the common load impedance. The signal from the second amplifier takes its place. This sequence is repeated over and over in synchronism with the oscillograph's sweep circuit; therefore, the several signals being applied to the amplifiers appear to be projected on the cathode-ray screen simultaneously.

There are two methods of keying an amplifier. In one, a triode amplifier whose tube has been biased far below cutoff is used. When a control voltage is switched in series with this bias voltage, the net bias on the tube is reduced to its normal value and the amplifier becomes operative. The other method of keying is applied to an amplifier employing a pentode tube. The screen-grid voltage is kept at zero with respect to the cathode; consequently, the amplifier is normally inoperative. It is keyed by applying rated voltage to the screen grid.

3–3. Gas-tube Switching Circuits. A gas triode can be used to close a d-c circuit and thus produce a switched voltage. This tube will not permit the flow of current until its grid bias is made less negative than the critical value. At that point, it begins to conduct rated current which will flow through the load resistor and produce the "switched" voltage. In order to stop the flow of current, however, the plate voltage must be cut off and kept off until the ions have had time to diffuse and the arc to go out. The grid then regains its control and can again initiate another discharge. To cut the plate voltage off, the circuit shown in Fig. 13 can be used.

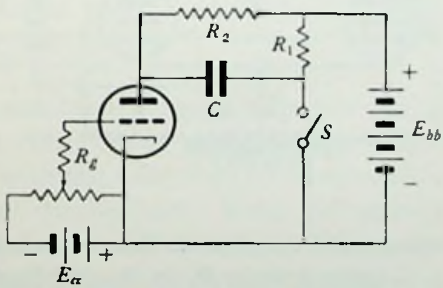


Fig. 13. Gas-tube switching circuit with manual switch for stopping conduction.

Assume that the grid potential has been driven to the point where the tube begins to conduct. A current will then flow from plus to minus through resistance R_2. The voltage appearing across R_2 will be equal to the applied voltage E_{bb} less the drop in the tube which is about 15 volts. This potential difference will charge the condenser C with a negative polarity on the left end through resistance R_1. With the condenser charged, the closure of switch S applies this negative voltage to the plate of the tube. If the condenser holds the plate negative long enough for the tube to become deionized, the current will stop flowing through it and the circuit will have been opened. The length of time the plate remains negative will depend on the rate at which the condenser discharges through R_1 and R_2. The time of discharge can be changed by adjustment of the circuit parameters and thus made long enough to ensure the complete deionization of the tube. On the other hand, the discharge time should not be made unnecessarily long because the waveshape produced across the load resistor is then badly distorted by the discharge current. A 0.005 μf condenser is typical for an 884 tube having a load resistance R_2 of 5000 ohms and an applied voltage E_{bb} of 100 volts.

If another gas triode is substituted for the switch S, as in Fig. 14, a basic switching circuit is obtained. The application of an a-c voltage to the grids of the tubes through a center-tapped transformer

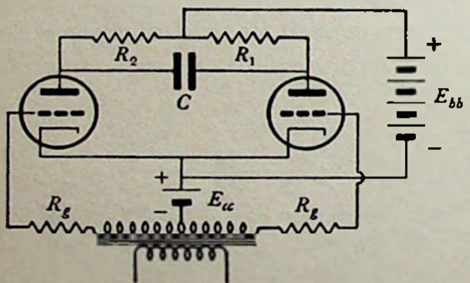


FIG. 14. Gas-tube switching circuit with another tube to replace the switch S in Fig. 13. This circuit is a typical inverter and can be used to convert the d-c power from E_{bb} into a-c power in the load resistors R_1 and R_2.

causes the current to flow alternately through them. The current is switched from R_1 to R_2 and back again in synchronism with the applied alternating current. The voltage wave appearing across either R_1 or R_2 or between the cathode and plate of either tube consists of a series of square-topped pulses. These pulses are switched d-c voltages.

If the conducting tube is arranged so that it shuts off the tube preceding it and primes the tube to follow it, the circuit can be made to switch the voltage to any number of independent positions. "Priming" means the reduction of the negative grid bias on the tube which is to be fired below the bias on any of the others, but still not low enough for it to conduct. A suitable impulse which is now applied to all of the grids will just fire the primed tube. If it too shuts off the preceding tube and primes the tube which is to follow, a group of tubes can be made to fire in sequence and repeat. A circuit for doing this is shown in Fig. 15.

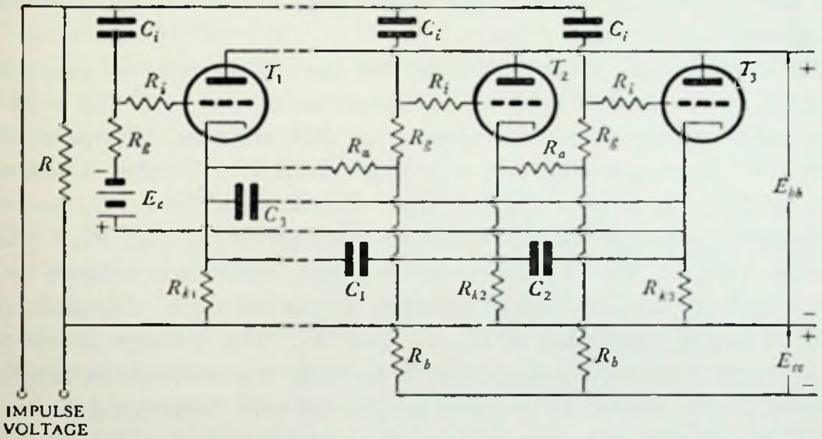


FIG. 15. Multiple gas-tube switching circuit which may be extended to include any number of tubes. The bias battery E_c on tube T_1 has been replaced by voltage-divider (R_a, R_b, and E_{cc}) bias supplies on the other two tubes. The resistance R or its equivalent is used to provide a discharge path for the condensers C_i between the applications of the triggering impulse voltage.

When none of the tubes are conducting, each tube has a bias equal to E_c which is obtained either from a series of separate batteries for each tube as shown with the tube on the left side of the diagram, or it is obtained more conveniently from the voltage divider arrangement shown with the tubes on the right. The latter system is the most advantageous because it permits the use of a single grid-bias supply. The drop across the resistance R_a takes the place of the individual battery which would otherwise be required. An additional resistance R_b is added to the divider for the purpose of isolating the bias voltages of each tube from one another as required by the circuit.

If the first tube starts to conduct, a current will flow through it and R_{k1}. The voltage across R_{k1} will be equal to E_{bb} minus the tube drop of about 15 volts. This voltage is the switched voltage E_s. The bias on tube T_2 is no longer E_c but rather $E_c - E_s$. Let E_i be an impulse voltage such that $E_c - E_s - E_i$ is less than the critical voltage of the tube and $E_c - E_i$ is more than it. If E_i is now applied to all the grids simultaneously, tube T_2 alone will fire. The resistance R_i in the grid circuit is necessary to prevent the impulse voltage from being lost through the low-impedance path of the conducting tube. The other grid resistance R_g limits the grid current to a safe value.

A capacitor is used to shut off the preceding tube as was done before. The shutoff will be independent of the grid bias and so the grid circuits can be neglected in this analysis. Suppose that tube T_2 is conducting. If the tube drop is 15 volts, a voltage $E_s = E_{bb} - 15$ appears across R_{k2}. Condensers C_1 and C_2 become charged up to this voltage through resistances R_{k1} and R_{k3} respectively. When tube T_3 begins to conduct, there is a voltage drop E_s across R_{k3} and a drop of 15 volts across the tube. Consider the closed circuit consisting of C_2, T_2, and T_3. The voltage across the condenser is in such a direction as to make the cathode of tube T_2 positive with respect to its plate and so the tube becomes deionized.

The cathode resistors used in the circuit must be selected with care so that the current rating of the newly conducting tube will not be exceeded. At the instant when the new tube first begins to conduct, its equivalent load is the parallel connection of the cathode resistors of all the other tubes in the circuit except the one which is being shut off. This is because the other shutoff condensers are effectively short circuits until they become charged. Consequently, the initial current in the new tube is equal to one less than the number of tubes n in the circuit multiplied by the steady state current

$$I = \frac{(n-1)E_s}{R_k} \tag{14}$$

This current must be within the rating of the tube.

To start a circuit of this kind, it is necessary to use a delay switch. The first position of the switch applies a temporary overpriming to the first tube. The second position applies the plate voltage and this tube begins to conduct. The third position removes the overpriming and applies the trigger voltage; consequently, the sequence starts.

3–4. Vacuum-tube Switching Circuits. When high-speed switching is necessary, vacuum-tube circuits must be used since gas tubes require an appreciable time to deionize. The Eccles-Jordan or Kipp relay shown in Fig. 16 is the basic circuit in this category. Its action depends on the fact that current flow in one of the tubes tends to stop the current flow in the other. If, for example, current is assumed to be flowing in the tube T_2, an IR drop will appear across R_2 making its right end more negative with respect to the cathode than it was before the current began to flow. Since this point is connected to the grid of the other tube, the latter tends to carry less current than it did before; hence, the drop in R_1 becomes less and the grid of the tube on the right becomes more positive.

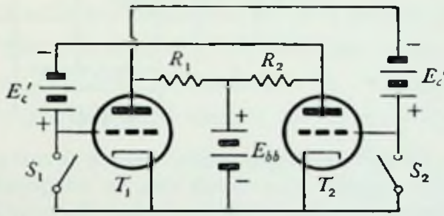


FIG. 16. Eccles-Jordan trigger circuit or Kipp relay.

From this it is seen that the action is cumulative and that the tube T_2 will conduct more and more current until the tube T_1 is completely cut off and conducts no current. If a positive pulse of voltage is now applied to the grid of the nonconducting tube T_1, more and more current is conducted by tube T_1 and the conduction of T_2 is reduced to zero. The changeover of conduction from one tube to the other is practically instantaneous; consequently, this circuit can be used as the basis for a high-speed electronic switch. The switched voltage appears across the resistance R_1 or R_2 or between the plate and the cathode of either tube. A negative pulse applied to the grid of the conducting tube or the closure of the switch (S_1 or S_2) on the nonconducting tube will also initiate the switching action.

The individual grid-bias batteries in the circuit of Fig. 16 can be replaced by a pair of voltage dividers R_a and R_b and a single bias supply E_{cc} as was done with the gas-tube switching circuit described in the preceding section. This revision is shown in Fig. 17. The resistances used for the voltage dividers should be high, whereas the resistances used in the plate circuit are of medium value. Almost any specific value of resistance will work provided the circuit remains symmetrical. The condensers C_i are for the purpose of applying the triggering impulse to the grids of the tubes from a simple battery and key circuit. The resistances R provide

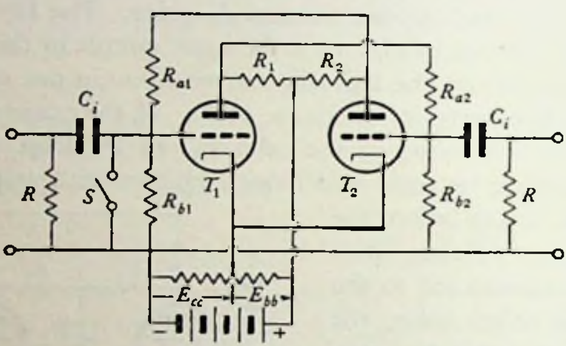


FIG. 17. Vacuum-tube switching circuit of Fig. 16 equipped with voltage-divider bias supply. The circuit can be triggered by applying an impulse to the grid through condenser C_i, by closing the switch S, or by any other means which will change the grid potential of one of the tubes.

a discharge path for the condensers when the impulse voltage is removed. These resistances and condensers can be eliminated if other triggering methods are applied.

The performance of the Eccles-Jordan circuit can be determined from the characteristic curves of the tubes used. A load line is drawn on these curves as shown in Fig. 18. The load line intersects the ordinate at the current which would be taken by the plate-load resistor if there were no voltage drop across the tube, and it intersects the abscissa at the plate supply voltage E_{bb}. The voltage drop across the plate-load resistor or the switched voltage E_s is determined from the intersection of the load line with the characteristic curve corresponding to the grid bias used. For the proper operation of the circuit, the switched voltage must be sufficient to drive the grid of the other tube below cutoff when it is added to the voltage across the divider circuit.

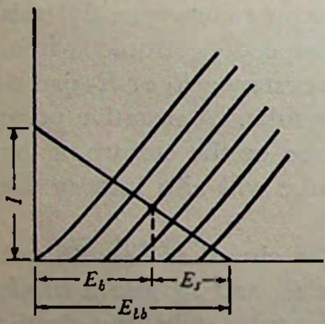


FIG. 18. Plate characteristic for a tube showing how to determine the switched voltage in a vacuum-tube switching circuit.

Example: A pair of 6J5 tubes are used in a vacuum-tube electronic switch or trigger circuit of the type shown in Fig. 17. The resistances R_1 and R_2 are each 10,000 ohms and the voltage divider resistances R_{a1}, R_{a2}, R_{b1}, and R_{b2} are each 1 megohm. Find the bias supply voltage E_{cc} which must be used to produce a switched voltage of 25 volts if the plate voltage E_{bb} is 100 volts.

The intersection of the load line on the ordinate is at

$$I = \frac{100}{10{,}000} = 0.01 \text{ amp or } 10 \text{ ma}$$

A switched voltage E_s of 25 volts will be produced with a grid bias of $E_c = -2$ volts. Neglecting the effect of the 10,000-ohm plate-load resistor in series with the divider resistances, the voltage $E_{bb} + E_{cc}$ will be divided equally between them; consequently, the grid bias $E_c = \dfrac{E_{bb} + E_{cc}}{2} - E_{cc}$ or $-2 = 50 - \dfrac{E_{cc}}{2}$ and $E_{cc} = 104$ volts.

An additional tube can be used to make the circuit sensitive to both positive and negative impulses. The impulse is applied simultaneously to the grids of the conducting tube and this auxiliary tube. The latter, which is biased to cutoff, is connected in parallel with the nonconducting tube (plate to plate and cathode to cathode). If the pulse is negative, the current in the conducting tube is decreased and the changeover started. If it is positive, the auxiliary tube conducts current through the resistance in the plate circuit of the nonconducting tube and thus initiates the switching operation.

If pentodes are substituted for triodes, a negative pulse applied to both control grids simultaneously will switch the current from whichever tube is conducting to the other one. When the function of the suppressor and the control grids are interchanged, the circuit will trigger with a pulse of either polarity.

If R_{a2} is replaced by a condenser and the grid bias of T_2 is changed as shown in Fig. 19, T_2 would normally conduct and the condenser C_2 would be charged to the normal plate voltage of the tubes. T_1 would be prevented from conducting due to the i_bR_2 drop. If a negative pulse is now applied through the condenser C_i to the grid of tube T_2, it stops conducting and removes the bias from T_1 which immediately begins to conduct. The resultant i_bR_1 drop is in the same direction as the voltage across the condenser C_2. The condenser, therefore, begins to discharge, and the discharge current flowing through R_{b2} maintains the grid of T_2 negatively biased even though the initial negative pulse has been removed. After a certain length of time which depends on the magnitude of R_{b2} and C_2, the discharge current falls so low that the

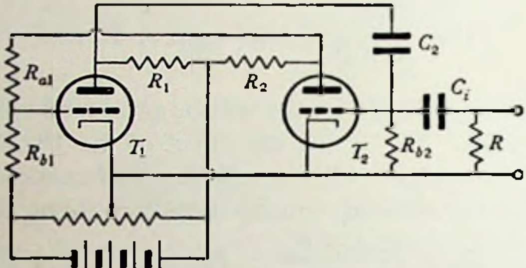


FIG. 19. Circuit for producing a single square-topped voltage pulse. The circuit can be triggered by applying an impulse voltage to C_i.

negative bias on T_2 is reduced to the point where it begins to conduct. The current flowing in R_2 increases the bias on T_1, which reduces the current in R_1, and so further reduces the bias on T_2. Consequently, tube T_1 suddenly stops conducting and tube T_2 resumes conducting as in the beginning. As a net result of this sequence of events, a square-topped voltage pulse, whose duration depends on the $R_{b2}C_2$ constant, is formed across the resistance R_1 or between the plate and the cathode of T_2.

Example: Find the duration of the voltage pulse formed across R_1 in the circuit of the preceding example if a condenser of 0.005 μf is substituted for the resistance R_{a2} and the circuit rewired to conform with Fig. 19 and triggered.

The condenser C_2 charges up to the supply voltage $E_{bb} = 100$ volts prior to the time of triggering. When the circuit is triggered by driving the grid of tube T_2 negative for an instant, T_1 begins to conduct and develops a voltage of 25 volts across its load resistor. This constitutes an excess voltage in the circuit C_2, R_1, E_{bb}, and R_{b2}; therefore, the condenser begins to discharge and develops a counteracting voltage of $-$ 25 volts across R_{b2}. As the condenser continues to discharge, the voltage across R_{b2} falls off exponentially and when it reaches the cutoff value of $-6\frac{1}{2}$ volts, tube T_2 begins to conduct again and the cycle is ended. The time required for this discharge can be computed from the following equation:

$$e_c = E_s \epsilon^{-\frac{t}{RC}}$$

$$-6\tfrac{1}{2} = -25\,\epsilon^{-\frac{t}{1\times 10^6 \times 0.005\times 10^{-6}}}$$

$$t = 1.35(0.005) = 0.00675 \text{ sec}$$

The effect of the discharge current flowing through R_1 is neglected because the voltage drop produced by it is negligibly small.

3-5. Multivibrator. The above circuit can be further modified so that the conduction will be automatically shifted from one tube to the other without the necessity of triggering. This is done by replacing the resistance R_{a1} by a condenser and changing the normal bias on tube T_1 as shown in Fig. 20. Assume that T_1 has just begun to conduct and that T_2 has just stopped. Condenser C_1 will start to charge up to the supply voltage. Meanwhile, as condenser C_2 discharges, the bias on T_2 is reduced to the point where it begins to conduct. The resultant i_bR_2 drop increases the bias on T_1 which decreases the i_bR_1 drop and, therefore, reduces the bias on T_2 causing it to conduct still more. The effect is cumulative so that T_2 is suddenly conducting its maximum current and T_1 is completely stopped. Due to the excess voltage i_bR_2, the condenser C_1 now begins to discharge and when its discharge current has dropped to the critical value as above, tube T_1 suddenly conducts its maximum current and T_2 stops conducting. Thus the circuit automatically switches from one tube to the other at a rate which depends on the magnitude of the grid resistors and the plate capacitors. The period of oscillation is approximately

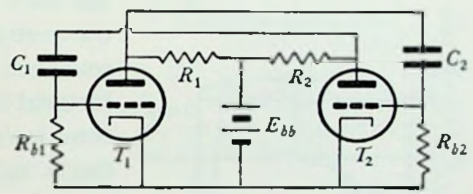


FIG. 20. Multivibrator circuit.

$$T = (R_{b1}C_1 + R_{b2}C_2) \text{ sec} \tag{15}$$

The voltage wave form produced across either of the plate-load resistors or between the plate and cathode of either tube consists of a series of square-topped pulses. The circuit is called a multivibrator.

A multivibrator has the useful property of falling into synchronism with a control voltage which is near its fundamental frequency or a multiple thereof. Consequently, the multivibrator can be used to maintain a synchronous relationship between the oscillations of two circuits whose frequencies are not the same but which are related by a common factor. The control voltage can be applied to one of the grids by means of a transformer inserted between the cathode and grid resistor or by means of a coupling condenser at the grid.

An oscillogram of the normal grid voltage on one of the tubes, T_2 for example, is shown in Fig. 21. When T_1 begins to conduct, this grid is driven rapidly negative and as the condenser discharges, the bias decreases exponentially until the cutoff voltage E_{co} of the tube is reached. Then T_2 begins to conduct; its grid is driven rapidly positive and remains there until the other tube conducts again and drives it negative. If a series sinusoidal control voltage is inserted in the circuit, the grid voltage will take the form shown dotted in the figure. The point at which the cycle is ended is now fixed by the control voltage. Each cycle of the multivibrator contains the same fixed number of control voltage cycles; hence, the multivibrator oscillates in synchronism with them. If the magnitude of the control voltage is increased, the phase angle between it and the output is shifted, but the frequency ratio between them remains the same. When the control voltage is increased still further, a point is reached where the multivibrator period is ended one control cycle earlier than before. The ratio between the two frequencies is changed. Consequently, there is an optimum value of control voltage for a given frequency ratio. This optimum is taken as one-half of the voltage at which the ratio between the frequencies changes.

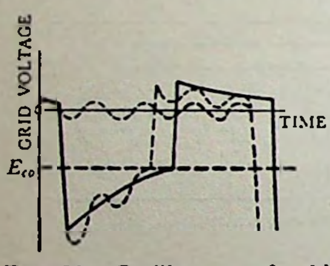


FIG. 21. Oscillogram of grid voltage on a multivibrator tube.

The reason for the above choice of the optimum voltage can be understood from Fig. 22. Here the reciprocal of the multivibrator's natural frequency is plotted as a function of control voltage. The natural frequency of the multivibrator is defined as the frequency at which it normally oscillates when no control voltage is present, and of course that depends on the values of resistance and capacitance used in the circuit. In order to obtain the curves of Fig. 22, the grid resistances of a multivibrator are changed while oscillations are synchronized by a certain amount of control voltage. When the frequency ratio changes, the control voltage is removed and the natural frequency of the multivibrator is observed. The reciprocal of this frequency plotted as a function of control voltage gives points on the curves bounding the various areas. The optimum control voltage for a given frequency ratio is seen to be near the center of its control area because the use of that voltage will permit a maximum variation of the circuit constants (due to heating, etc.) without the loss of control. Also the control voltage itself can vary within wide limits without any danger of the appearance of an unwanted frequency ratio or a drift into one of the uncontrolled regions.

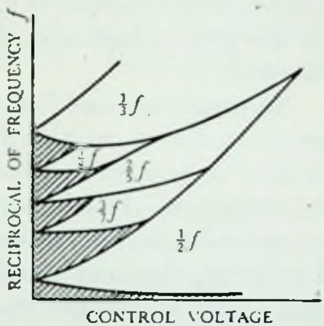


FIG. 22. Control area curves for a multivibrator. There is no control when operation is on the shaded areas.

Some connections for inserting control voltage favor synchronization at certain frequency ratios. If the control voltage is arranged to drive both grids in the same phase, the symmetrical multivibrator favors synchronization with control voltages which are even multiples of its frequency. Since the control voltage will shift its phase relationship with the multivibrator alternations until an equilibrium position is reached where each of the successive alternations are terminated on the same positive portion of the control-voltage cycle, each multivibrator alternation contains an integral number of these cycles. The two multivibrator alternations per cycle were assumed to be symmetrical (i.e., of the same duration); therefore, there is an even number (an integer multiplied by 2) of control cycles in each multivibrator cycle. On the other hand, if the grids are driven 180° out of phase, synchronization with odd multiples is favored. In this case, one alternation is terminated on a given portion of a positive half cycle of control voltage, and the other is terminated on the same portion of the corresponding negative half cycle; hence, each alternation contains an integral number of control cycles plus one half cycle. There is, therefore, an odd number of control cycles in each complete multivibrator cycle. Naturally, the above conclusions do not apply if the multivibrator output is not symmetrical. When the control frequency is applied to the grid of just one of the tubes, both odd and even ratios are favored.

An isolating amplifier must be used to supply the control voltage to the ordinary multivibrator; otherwise, the voltages produced by it may be transmitted back to the control generator and influence its operation. In order to avoid this complication, Potter devised a multivibrator which does not require an isolating amplifier. In his circuit, the condenser C_1 is completely eliminated. The cathodes and minus end of the power supply are connected together on one end of a single cathode resistor whose other end is connected to the resistances R_{b1} and R_{b2}. With the exception of these changes, the circuit is the same as that of Fig. 20. Since there is no multivibrator current flowing in the resistance R_{b1}, a control voltage connected across it will not receive any influencing voltage from the multivibrator.

3–6. Counter Circuit. The circuit of Fig. 23 is also popularly used to synchronize two signals having different frequencies. The

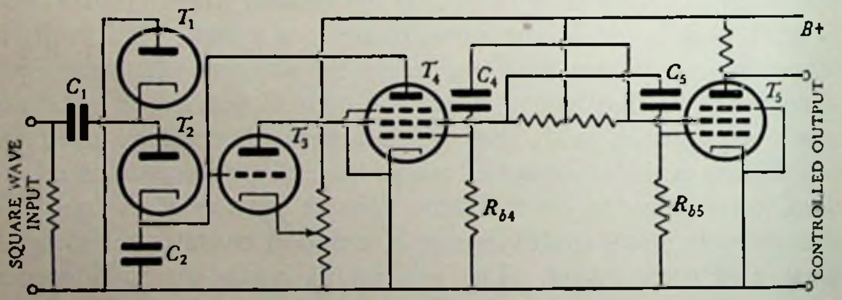


FIG. 23. Counter circuit.

input signal is converted to a square wave before it is applied to the counter. After a certain number of these input square-wave cycles have been passed by the counter, the multivibrator portion of the circuit starts its cycle. As the input wave rises to a positive value of E, the condensers C_1 and C_2 are charged in series. Since C_2 is ten to twenty times larger than C_1, it will charge up to a small fraction

$$\frac{C_1E}{C_1 + C_2} \tag{16}$$

of the applied voltage. When the input wave goes to zero, C_1 discharges through tube T_1. The next positive rise of voltage adds another increment of charge to the condenser C_2 or gives it an additional voltage equal to

$$\frac{C_1}{C_1 + C_2}\left(E - \frac{C_1E}{C_1 + C_2}\right) \tag{17}$$

Consequently, the voltage across C_2 rises in a staircase fashion with an additional step added with each positive pulse of the input wave. The grid of tube T_3 is biased so that it begins to conduct after a certain number of voltage steps have been built up on C_2. The current of T_3 flowing through the screen resistor of T_4 shuts off the conduction in tube T_5 and starts it in T_4. This multivibrator has the screen grids perform the same function as the plates of the tubes in the ordinary multivibrator. The condenser C_2 is discharged through T_4 in preparation for another cycle. The multivibrator of the circuit is made unsymmetrical so that the conduction period through tube T_4 is very much shorter than that through tube T_5. This is necessary because the condenser C_2 is effectively short-circuited as long as T_4 is conducting, and if that were too long the charge due to the first pulse of the new cycle would be lost. Consequently, the grid resistance R_{b4} is made much larger than R_{b5} if the condensers of both (C_4 and C_6) have the same capacity.

3–7. Shaping Circuits. When the output of electronic switches is used for synchronizing operations, it is most effective if it is perfectly square in form so that the beginning and end of the impulse are distinctly marked. The output of the multivibrator and other electronic-switch derivatives usually just approaches this ideal; hence, circuits have been devised for squaring them. These circuits together with others such as those which are used to lengthen or shorten the duration of pulses are known as shaping circuits.

A wave can be squared by either clipping or leveling. As an illustration of these processes, consider the problem of converting a sine wave into a square wave. It is apparent that a square wave can be formed from a sine wave by "clipping" off the top and bottom portions as shown in Fig. 24. The bottom is clipped off by passing the wave through the first stage of a resistance-coupled

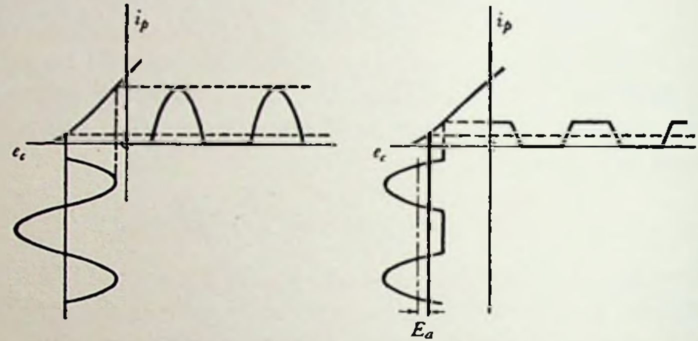


Fig. 24. Dynamic curves of a tube and an illustration of the clipping process. The dynamic curves are a plot of the instantaneous plate current as a function of the instantaneous grid voltage in a particular circuit.

amplifier, such as that in Fig. 25, whose bias is adjusted so that the undesired negative portions of the wave drive the tube below cut-off. The output of this stage will be reversed in phase; hence, the positive half of the wave is now on the underside and it too can be clipped off by passing the wave through another such stage. This process can be repeated until any desired degree of squareness is obtained. The stages of a clipper are often direct coupled so that the distortion of the output wave due to the condenser C is eliminated.

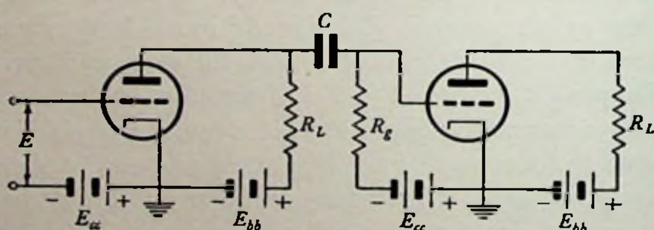


Fig. 25. Typical clipping circuit. Better results are obtained with this circuit if the condenser C is eliminated and direct coupling used instead.

Example: Two 6J5 tubes are used in the circuit of Fig. 25 for clipping a 60-cps 6.3-volt sine wave into a square wave. The plate load resistors R_L are 20,000 ohms and the plate supply voltage is 200 volts. Calculate the bias voltages which should be used on each of the two tubes if the wave front or time required for the output wave to rise from its minimum to its maximum value is $\frac{1}{5000}$ sec. Calculate the amplitude of the output wave.

The input voltage should be clipped at a point $\frac{1}{10,000}$ sec from the point where it crosses the zero axis. Converting this time to degrees we have:

$$\frac{\theta}{360} = \frac{60}{10,000} \qquad \theta = 2.16°$$

The voltage at that point is

$$6.3\sqrt{2} \sin 2.16° = 0.33 \text{ volt}$$

Since cutoff occurs at a grid voltage of -12 volts, the bias on the first tube should be adjusted to $-12 + 0.33 = 11.67$ volts. The load line intersects the ordinate on the tube characteristic curve at

$$I = \frac{200}{20,000} = 0.010 \text{ amp}$$

The grid voltage swings along this line from -12 to

$$-12 + 0.33 + 6.3\sqrt{2} = -2.77 \text{ volts}$$

The maximum voltage which appears across the load resistance is 90 volts. If we assume that the coupling condenser C and the grid resistance R_g of the second tube are both very high, they can be neglected and the grid swing of the second tube will also be 90 volts. For purposes of calculation we will assume that the wave shape at the second grid has not been distorted except for the clipping. The maximum value of the remaining half sine cycles is then

$$E = \frac{\sqrt{2}\,6.3}{\sqrt{2}\,6.3 + 0.33} \times 90 = 87 \text{ volts}$$

On the clipped side it rises 3 volts above the axis. The average value of this resultant wave must now be found in order to locate the position of its axis after the direct current has been eliminated by the coupling condenser. The area under the clipped portion is practically $3\,\pi = 9.4$ and the area under the sine portion is

$$\frac{2}{\pi} \times \pi \times 87 = 174$$

The average value of the new wave or the position of the new axis from the old is

$$\frac{174 - 9.4}{2\pi} = 26.2 \text{ volts}$$

This is the voltage E_a in Fig. 24. In order to clip the other portion of the wave off properly, the bias on the second tube should be

$$-12 + 3 - 26.2 = -35.2 \text{ volts}$$

The total grid swing is $3 \times 2 = 6$ volts, and from the load line the voltage developed at the output would be 55 volts.

The circuit shown in Fig. 26 is used to square a wave by leveling. The grid of the tube will conduct a current through R_1 when it is driven positive by the input wave. The grid current flowing in R_1 will produce an IR drop which will subtract from the impressed positive voltage and make the grid less positive with respect to the cathode. Due to this action, the positive side of the wave is "leveled" off. When a large driving voltage is used, the negative side of the wave is simultaneously clipped off so that a square wave is produced in this single stage.

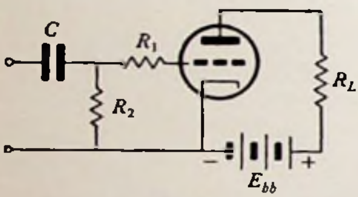


Fig. 26. Leveling circuit.

An RC series circuit can be used to shorten the duration of a square pulse. The application of a square-wave voltage to such a combination will produce an exponentially falling voltage across the resistance element. The use of low values of R and C will make this pulse very sharp and if its top and bottom are clipped off, a very narrow square pulse will be produced.

On the other hand, a wider pulse which has been delayed with respect to the applied pulse will result if the voltage across the condenser is used to drive the clipper stage. In this case, it is necessary to connect a high resistance between the grid of the first clipper tube and its cathode so there will be a circuit through which the charge accumulating on the grid can be removed and the grid prevented from floating. The action of the pulse narrowing and widening circuits can best be understood by an examination of Fig. 27.

Occasionally a sharp triangular pulse is wanted. Such a wave can be obtained from the output of the first stage in a pulse-narrowing circuit.

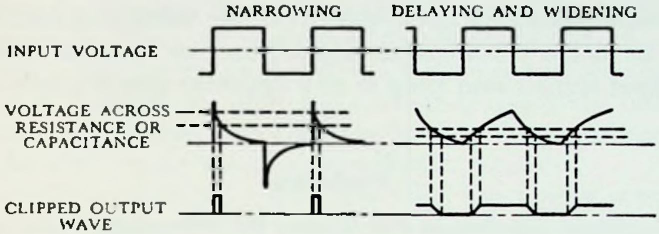


FIG. 27. Oscillograms illustrating the wave-shaping processes. The clipping levels are indicated by the dotted lines.

3-8. Phase Discriminator. The circuit diagram for a phase discriminator is shown in Fig. 28. The net voltage appearing across the two load resistors is a function of the phase angle between the two input voltages. The upper half-wave rectifier, for example, produces a d-c output voltage which is proportional to the vector sum of the two alternating input voltages while the lower rectifier

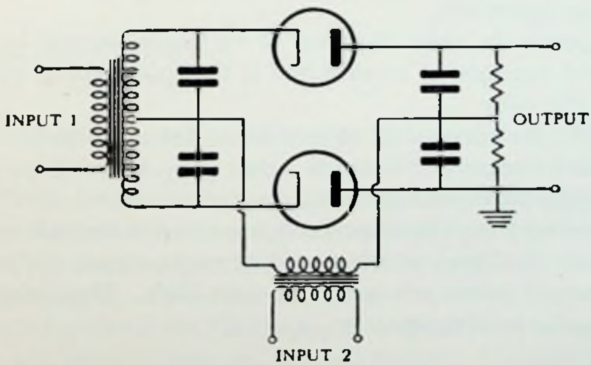


FIG. 28. Phase discriminator.

produces a d-c output voltage which is proportional to their vector difference. If the two input voltages are of the same magnitude and are in phase, it is apparent that double d-c output voltage will be produced across the load of the upper rectifier and that zero voltage will be developed across the load of the lower rectifier. A net positive double voltage will then be developed across the two resistors. When the input waves are 90° out of phase with each other, the same voltage will be developed across each load resistance and the net voltage across both of them would be zero. On the other hand, a 180° phase relationship between the input waves would produce a net negative double voltage across the total load. The net d-c voltage developed can be used to control one of the input signals and keep it at a constant phase angle with the other.

Problems

3–1. A circuit like that in Fig. 15 has the following circuit constants: $R_k = 5000$ ohms, $R_a = R_b = 100{,}000$ ohms, $E_{bb} = 100$ volts, and the tubes used are 884's (characteristic curve in Appendix). Find the bias supply voltage which must be used if the bias on the primed tube is to be 20 volts from the tube's critical value. What would be an appropriate value for the triggering impulse voltage to have?

3–2. Find the maximum switched voltage which could be obtained from the circuit in the first example of section 3–4 and the bias supply voltage E_{cc} corresponding to it. Also find the minimum switched voltage together with its bias supply which could be used with the circuit and still ensure proper operation.

3–3. Compute the time duration of the pulse formed by the circuit in the second example of section 3–4 if the grid supply voltage E_{cc} is changed to 100 volts.

3–4. A 60 cps symmetrical square wave has an over-all height of 25 volts. It is to be narrowed down to pulses $\frac{1}{1000}$ sec long spaced at intervals of $\frac{1}{60}$ sec by the use of a 200,000-ohm resistance and a 0.01-μf capacity in conjunction with the clipping circuit described in the example of section 3–7. Calculate the biases which should be used on each of the tubes if the narrowed output pulses are to be 50 volts high. How long would the crest of the pulse be if the base is $\frac{1}{1000}$ sec?

Experiment 2

Object:

The object of this experiment is to study the operating properties of gas-tube-switching circuits.

Preliminary:

1. Solve problem 3–1.
2. Compute the power rating of R_2 in Fig. 13 if its resistance is 5000 ohms and E_{bb} is 100 volts.

Performance:

1. Set up a circuit such as that in Fig. 13. Demonstrate that conduction is started in the tube by reducing the grid bias and that it is stopped by restoring the bias and closing the switch *S*.

2. Convert the circuit of part one to the circuit of Fig. 14. Apply a 60-cycle sinusoidal voltage to the input transformer and observe the output produced between the plate and cathode of one of the tubes with a cathode-ray oscillograph. Vary the size of the shutoff condenser *C* and observe the effect on the wave shape produced. Explain.

3. Assemble a circuit with three gas tubes like that of problem 3–1. Apply a positive pulse to all of the grids by means of a telegraph key and battery circuit and thus cause the conduction to transfer from one tube to the next. Observe this transfer of conduction by a series of milliammeters in the cathode circuits. Measure the voltages developed in various parts of the circuit and compare them with theoretical values.

Experiment 3

Object:

The object of this experiment is to study the operating characteristics of vacuum-tube-switching circuits.

Preliminary:

Solve problems 3–2 and 3–3.

Performance:

1. Assemble the circuit of problem 3–2. Trigger the circuit by touching one of the grids with a lead from the cathode terminal or with your finger. Detect the transfer of conduction from one tube to the other with a high-resistance voltmeter connected across one of the load resistors or by means of milliammeters in the plate circuits of the tubes. Vary the grid-bias supply voltage E_{cc} and check the results of your calculations.

2. Convert the above circuit to the one of problem 3–3. Trigger it by means of a pulse from a telegraph key and battery. Measure the duration of the square-topped pulse produced by the circuit by connecting the output in series with a standard oscillator and observing this combined wave on the cathode-ray oscillograph. The number of cycles picked up by the incoming square-topped pulse will be a measure of the pulse duration. Verify the results of problem 3–3. Replace the condenser C_2 by a much larger one and produce a pulse which is so long that it can be detected by a plate milliammeter inserted in the circuit of tube T_1.

3. Convert the above circuit to a multivibrator and observe the wave form produced by it on a cathode-ray oscillograph. Replace the con-

densers by ones which are large enough to give a long period of oscillation which can be observed by the deflection of milliammeters placed in the plate circuit.

Experiment 4

Object:

The object of this experiment is to study the operating characteristics of wave-shaping circuits.

Preliminary:

Solve problem 3–4.

Performance:

1. Assemble the clipping circuit described in the example of section 3–7. Produce a square-wave output from a sine-wave input and verify the results of the example. Determine the duration of the wave front by making measurements on the cathode-ray oscillograph image of the wave. Observe the effect of making the coupling condenser too small. Explain discrepancies.

2. Use a square-wave generator and the above circuit to obtain the narrowed pulse specified in problem 3–4. Check results against calculations and explain discrepancies.

3. Interchange the positions of the condenser and resistor in part 2 and obtain a pulse which has been lengthened and delayed. Calculate the time of delay and compare with the value measured by connecting the input and output of the circuit in series and observing the resultant wave form on the cathode-ray oscillograph.

4. Set up a circuit like that of Fig. 26 with the condenser 0.1 μf, $R_2 =$ 200,000 ohms variable, and $R_1 = 1$ megohm variable. Use a 6J5 tube with a 50,000-ohm load resistor and a supply voltage of 200 volts. Apply 50 volts at 60 cps to the input terminals and adjust R_1 and R_2 until a square wave is obtained at the output between the plate and cathode of the tube. Identify which portion of the output wave has been clipped and which has been leveled, and compare the relative effectiveness of the two processes. Sketch the wave shapes which are observed across the resistances R_1 and R_2.

References

Andrew, V. J., "The Adjustment of the Multivibrator for Frequency Division," *Proc. I.R.E.*, November 1931, pp. 1911–1917.

Bedford, A. V., and J. P. Smith, "Precision Television Synchronizing Signal Generator," *R.C.A. Review*, July 1940, pp. 51–68.

Fink, D. G., *Principles of Television Engineering*. New York: McGraw-Hill Book Company, Inc., 1940, pp. 402–414.

Hull, L. M., and J. K. Clapp, "A Convenient Method for Referring Secondary Frequency Standards to a Standard Time Interval," *Proc. I.R.E.*, February 1929, pp. 252–271.

MacGregor-Morris, J. T., and J. A. Henley, *Cathode-ray Oscillography*. London: Chapman & Hall, Ltd., 1936, pp. 179–183.

Potter, J. L., "Sweep Circuit," *Proc. I.R.E.*, June 1938, pp. 713–719.

Reich, H. J., "Trigger Circuits," *Electronics*, August 1939, pp. 14–17.

Shumard, C. C., "Some Electronic Switching Circuits," *Elec. Eng.*, May 1938, pp. 209–220.

CHAPTER IV

CATHODE-RAY TUBES AND SWEEP CIRCUITS

4–1. General. A cathode-ray tube consists of: a source of electrons which is usually thermionic, a grid for controlling the flow of these electrons, an accelerating system for giving them a high velocity, a lens system for focusing them into a narrow beam, a deflecting system for moving the beam up and down and from side to side, and finally a fluorescent screen which converts the energy of the moving electrons into light and thus marks the beam's position on the screen. The tubes fall into two main categories. There are the electrostatic-focusing electrostatic-deflecting or EE type and magnetic-focusing magnetic-deflecting or MM type. Combinations of the two types of focusing and deflecting systems occur in practice, but they are relatively few in number. The EE type is most nearly like the conventional cathode-ray tube, and the MM type is closely parallel to the kinescope, which is used in television. The EE type is most suitable for following high-speed phenomena. The MM type has the advantage of being considerably shorter for the same screen size than the EE type because of its greater deflection sensitivity. Also the difficult constructional alignment of the deflecting plates is avoided with it as are the difficulties with defocusing when grid modulation of the beam is used. The electrode leads are generally brought out at the base of all tubes for convenience. The spacing between the pins on the base of tubes intended for high-altitude service is somewhat wider than that on conventional tubes so as to prevent the formation of arcs between the wires at the reduced air pressure. Consequently, the base of these tubes is somewhat larger than usual.

4–2. Electron Lenses. The first lens following the electron emitter is practically always electrostatic and serves to accelerate the electrons as well as converge them into a beam. The narrowest portion of the beam leaving this lens is called the crossover and corresponds to the exit pupil of the lens system. The second lens focuses the image of this crossover area on the fluorescent screen of the tube. The grid, which is associated with the first lens system, controls the current and therefore the intensity of the beam by means of a space-charge barrier which restricts the area over which emission takes place. Changing the potential on the grid as in grid modulation may change the position of the crossover and thus introduce some defocusing.

Electron lenses are analogous to thick optical lenses. They can be considered to have two principal planes and two focal points as in Fig. 29. One focal point and one principal plane is determined

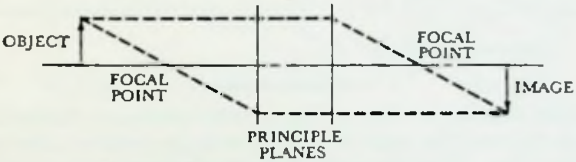


FIG. 29. Diagram illustrating the formation of an image by an electron lens.

by considering the path of an electron traveling parallel to the axis of the lens. The point at which the trajectory intersects the axis is the focal point. The intersection between a line tangent to the trajectory at the focal point and the initial line of travel is a point on the principal plane. The other principal plane and focal point can be obtained in the same way by considering the trajectory of an electron moving parallel to the axis in the opposite direction. When these four cardinal points have been determined, electron images can be handled by geometrical optics.

The two types of electrostatic lenses generally found in cathode-ray tubes are the aperture lens and the cylindrical lens. In the case of the aperture lens, the electrostatic-field intensity on each side of the aperture is different. This results in a convergence of the electrostatic lines of force passing from one side to the other. A paraxial electron moving through this field is deflected toward the axis due to the vertical component of electrostatic-field intensity at the aperture. The amount of deflection is proportional to the electron's distance from the axis; hence, the aperture acts like the ordinary convex optical lens.

The cylindrical lens consists of two successive coaxial cylinders which are operated at different potentials as in Fig. 30. When an electron passes through the plane between them, it is subjected to a component of electrostatic-field intensity directed toward the axis of the cylinders, and therefore the beam is bent like a ray of light passing through an optical lens. The exact mechanism of this process can be better understood from the gravitational analog of the lens. The equipotential lines of the figure correspond to levels of equal elevation or contour lines in the gravitational field. If we

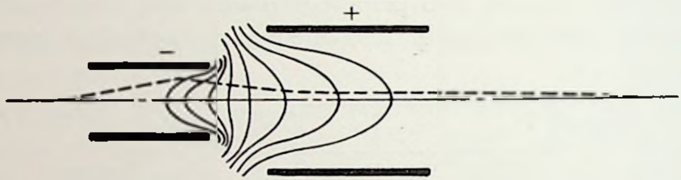


Fig. 30. Cylindrical lens. The section of the cylindrical electrodes is indicated by the heavy lines, and the path of the electron in passing through the lens is shown as a dotted line. The light solid lines lie along equipotential surfaces.

have the surface which is represented by these contour lines, the path followed by a small ball rolling on that surface is exactly the same as the path of an electron accelerating through the given electrostatic field. If the electrostatic lines of Fig. 30 are considered to be contour lines on a map, the more negative lines would correspond to the higher elevations since the electrons are accelerated from negative to positive regions. The analogous gravitational surface near the negative electrode is in the form of a cup or ravine which will cause the ball in our analogy to roll toward the center line of the field. When the ball begins to roll on the surface near the positive electrode, however, there is a tendency for it to roll away from the axis, but this tendency is vastly less pronounced than the former one. From this analogy, it is apparent that the path of the electron through the lens is as shown by the dotted line in the figure.

Gravitational models such as the above are frequently used to study the action of proposed electron lenses before they are incorporated in new cathode-ray tube designs. The distribution of electrostatic potential is found from an electrolytic model of the lens and that data is used to construct the gravitational model.

The same surface can also be obtained by stretching a thin sheet of rubber over some projecting edges at heights corresponding to the voltage on the particular electrodes of the lens which they represent. Small balls are then rolled over these surfaces to determine the trajectories of the electrons in the proposed lens and thus to predict its performance.

Two types of magnetic lenses are also used on cathode-ray tubes. They are the long focusing coil and the short focusing coil. The long focusing coil is shown in Fig. 31. It extends over a range

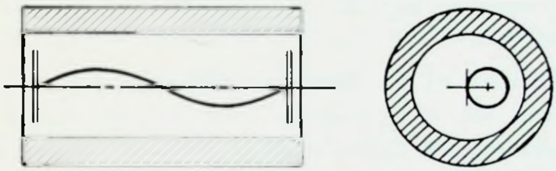


Fig. 31. Long focusing coil. The electron object lies on one of the planes at the end of the coil, and the image is formed on the other. The solid spiral line shows the path of the electron through the lens.

which includes both the object and the image. Consequently, the lines of magnetic force are parallel to the axis of the coil throughout this range. Electrons which have no radial component of velocity are not affected by the magnetic field and, therefore, occupy the same position on the plane of the object as they did on the plane of the image. Those electrons which have a radial velocity travel in a spiral path. The time required for any electron to complete the circle of its spiral flight depends only on the strength of the magnetic field. Hence, any electron from the object completes this circle in the same time, regardless of its initial velocity, and so lands on a point of the image plane exactly opposite to the point it left on the object plane. The long-coil magnetic lens has the distinctive property of producing an upright image. The other electron lenses are more like the optical analogs and produce inverted images.

The short focusing coil, as the name implies, occupies a short portion of the distance between the object and image. It is illustrated by Fig. 32. The electrons moving along the axis of the coil are unaffected by the magnetic field. Those which lie off the axis are subjected to a quick spiral twist as they pass through the magnetic field. If the circle is nearly completed, the electron is given a component of motion toward the axis of the field which persists after it passes through the lens. A focusing action such as that indicated in the figure is produced by the coil.

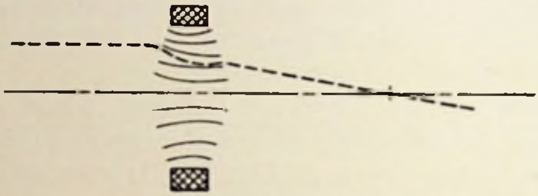


FIG. 32. Short focusing coil. The dotted line shows the path of the electron, and the solid lines indicate the magnetic flux.

4–3. Deflection of the Electron Beam. The action of the deflecting system for a cathode-ray tube is illustrated in Fig. 33. The dotted area of the figure represents the region occupied by the deflecting field. The electrons enter it along the horizontal axis after having been accelerated by the anode voltage E_a. The kinetic energy of each electron as it enters the field is equal to the

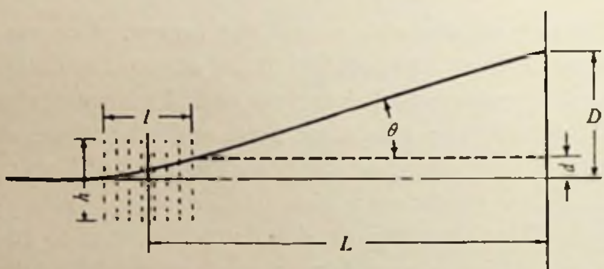


FIG. 33. Cathode-ray deflecting system. The deflecting field occupies the dotted area.

product of the accelerating anode voltage and the charge e on the electron. From Newton's law of motion, this kinetic energy is equal to one-half the mass m of the electron multiplied by its velocity v_x squared. Hence,

$$eE_a\, 10^7 = \tfrac{1}{2}\, mv_x^2 \text{ ergs}$$

or solving for v_x

$$v_x = \sqrt{\frac{2\, eE_a\, 10^7}{m}} \text{ cm/sec} \tag{18}$$

The factor 10^7 is necessary to convert the units of charge to coulombs and the voltage to practical volts.

If the area shown in the figure is assumed to be filled with a uniform vertical electrostatic field of intensity $\mathcal{E}_d$ volts per centimeter, the electrons will experience a vertical force of

$$F_y = \mathcal{E}_d e \, 10^7 \text{ dynes} \tag{19}$$

as they pass through it. This force is also equal to

$$F_y = ma \text{ dynes} \tag{20}$$

where a is the vertical acceleration of the electrons. The substitution of equation 19 into equation 20 and the solution for acceleration yields

$$a = \frac{\mathcal{E}_d e \, 10^7}{m} \text{ cm/sec/sec} \tag{21}$$

If t is the period of time during which the electrons remain under the influence of the deflecting field, the vertical velocity with which they leave will be

$$v_y = at$$

and since t is equal to the length of the field l divided by the horizontal velocity v_x,

$$v_y = \frac{\mathcal{E}_d e \, 10^7 \, l}{m v_x}$$

The angle θ at which the electrons leave the deflecting field is

$$\theta = \tan^{-1} \frac{v_y}{v_x} = \tan^{-1} \frac{\mathcal{E}_d l}{2 \, E_a} \tag{22}$$

where the ratio $\frac{v_y}{v_x}$ is obtained by the substitution of equation 18 into the equation for v_y.

The vertical distance d which the electrons travel before they leave the field is obtained from equations 18 and 21

$$d = \tfrac{1}{2} a t^2 = \tfrac{1}{2} a \frac{l^2}{v_x^2} = \tfrac{1}{2} \frac{\mathcal{E}_d l^2}{2 \, E_a}$$

and from comparison with equation 22

$$\frac{d}{\tfrac{1}{2} l} = \frac{\mathcal{E}_d l}{2 \, E_a} = \tan \theta$$

Consequently, the extension of the path back into the area of the field intersects the initial line of travel at the center of this area. The total deflection D produced on the screen is

$$D = L \tan \theta = \frac{L\mathcal{E}_d l}{2\ E_a} \text{ cm} \tag{23}$$

A nearly uniform vertical electrostatic field of intensity $\mathcal{E}_d$ is produced by applying a voltage

$$E_d = \mathcal{E}_d h \tag{24}$$

to the plates of a condenser which are h centimeters apart and which have a length slightly under the assumed length of the field. Fringing of the electrostatic lines at the edges of the plates makes them act as if they were longer than they actually are. The approximate physical length l_p of plates which will in effect produce the uniform field assumed is

$$l_p = l - h \tag{25}$$

The deflection sensitivity S_e is defined as the ratio of the deflection D to the voltage applied to the plates E_d.

$$S_e = \frac{D}{E_d} \tag{26}$$

A uniform magnetic field of flux density B lines per square centimeter occupying the dotted area shown in the figure will also produce a deflection of the cathode-ray beam. An electron moving with the velocity v_x constitutes a current element $i = ev_x$. The force on this current element due to the magnetic field is

$$F = \frac{Bev_x}{10} \text{ dynes} \tag{27}$$

This force is always perpendicular to the line of motion; hence, the electron moves in a circular path. The radius of the path R adjusts itself to a value such that the centrifugal force of the motion is equal to the deflecting force or

$$F = \frac{mv_x^2}{R} = \frac{Bev_x}{10} \text{ dynes} \tag{28}$$

The sine of the angle subtended by the arc of the electron's path while it is under the influence of the magnetic field is

$$\sin \theta = \frac{l}{R} \approx \tan \theta \text{ (for small angles)} \tag{29}$$

and it is the same angle as that with which the electron leaves the field. If this line of travel is extended backwards, it intersects the initial line of travel slightly beyond the center of the field. For practical purposes, however, it can be assumed that this intersection is at the center. From equations 28, 29, and 18, the deflection D which is produced then becomes

$$D = L \tan \theta \approx \frac{LlBe}{10\,mv_x} = \frac{LlB}{\sqrt{E_a}} \sqrt{\frac{e}{2\,m\,10^9}} \text{ cm} \tag{30}$$

or since the charge on the electron is $e = 1.6 \times 10^{-19}$ coulombs and the mass at ordinary velocities is $m = 9.1 \times 10^{-28}$ grams, the substitution of these values reduces it to

$$D = 0.297 \frac{LlB}{\sqrt{E_a}} \text{ cm} \tag{31}$$

A magnetic field of the type assumed can be easily obtained from a pair of coils mounted opposite each other alongside the path of the electron beam and carrying a current I_d. Occasionally an iron-cored yoke is used to concentrate the field more exactly, but usually an air-cored coil is considered sufficiently good. Relatively weak magnetic fields are required for deflecting purposes; hence, the iron-cored coil offers no particular advantage in this regard. The magnetic deflection sensitivity S_m is defined as the ratio of the deflection D to the current I_d.

$$S_m = \frac{D}{I_d} \tag{32}$$

The deflection produced by an electrostatic field varies inversely as the accelerating voltage, but the deflection produced by a magnetic field varies inversely as the square root of this factor. The reduction in deflection due to the use of higher accelerating voltages is far more pronounced in the electrostatic deflecting systems; therefore, magnetic deflection is used in high-voltage tubes whenever it is possible. The tube length can be made much shorter and the auxiliaries much less complex than for the equivalent electrostatic system. Electrostatic deflection must be used, however, if high-speed phenomena are to be observed. The inductance in the deflecting coils prevents the magnetic system from following rapid changes in current. Fortunately in both systems there is a linear relationship between the deflection and the factor which produces it. The electrostatic deflection is directly proportional to the field intensity between the plates or to the voltage which is being applied to them. Similarly, the magnetic deflection is directly proportional to the magnetic-field intensity between the coils or to the current which is flowing in them. Electrostatic deflection is used to measure voltage, and magnetic deflection is used to measure current. Before a voltage can be measured with a magnetic deflecting system, it must be converted to a current and vice versa.

4–4. Sinusoidal Sweep. A locally generated signal which is used to give the cathode-ray beam a particular motion with respect to time is known as a sweep. The simplest sweep is in the form of a sine wave. When it is applied to the horizontal deflecting system,

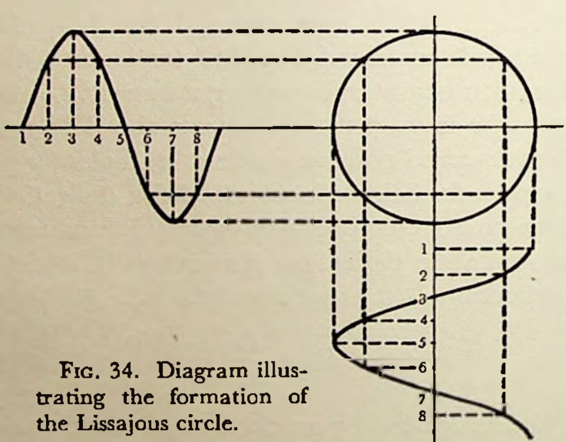


Fig. 34. Diagram illustrating the formation of the Lissajous circle.

the cathode-ray beam traces a single horizontal line on the screen. If in addition, however, a vertical sinusoidal deflection which is 90° out of phase with the horizontal sweep and of the same magnitude and frequency is applied to the tube, a circular image is produced as shown by the projection in Fig. 34. This image is known as a Lissajous figure. Other Lissajous figures are formed on the screen whenever there is an integral ratio between the frequency of a sine wave being applied to the vertical deflecting system and that of the horizontal sinusoidal sweep. The configuration of the resultant image depends on the phase relationship between the two waves as well as their frequency ratio. Several examples are given in Fig. 35. The phase angle between two waves having different frequencies has no physical significance unless it is defined for a particular instant of time. In Fig. 35, the phase angle between

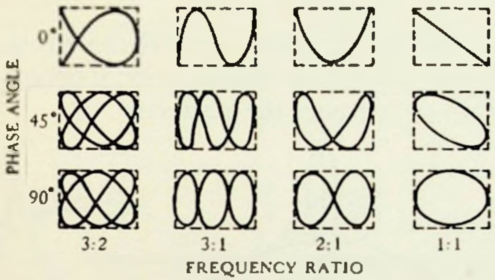


Fig. 35. Lissajous figures.

the two waves is taken as 90° minus the phase angle of the higher frequency wave at the instant when the lower frequency wave is passing through a positive maximum. The examples of Fig. 35 can be verified by the same point-by-point method of plotting as that used to find Fig. 34. It should be noted that Lissajous figures fit in a rectangular space and that the ratio between the frequency of the applied wave and that being used on the horizontal sweep is equal to the ratio between the number of tangencies on the top or bottom of the rectangle to the number of tangencies on the side. Since Lissajous figures are so easily identified with their corresponding frequency ratio, the major application of a standardized sinusoidal sweep is to determine the frequency of a test wave.

When the frequency ratio becomes large, the interpretation of a Lissajous figure becomes increasingly difficult. This is particularly true when the frequency of the oscillator being tested drifts slightly. A more satisfactory method then is to use a modification of the simple sinusoidal sweep and produce a pattern called a roulette. A roulette is formed by simultaneously applying a circular standard-frequency deflection and a circular test-frequency deflection to the electron beam. The test frequency causes the electron beam to trace a circle in one direction while the standard frequency causes it to trace a circle in the opposite direction. A construction illustrating the vector sum of these two simultaneous deflections or the form of the roulette for a 2 : 1 frequency ratio is given in Fig. 36. Patterns of other ratios are also shown in the figure. The ratio of the two frequencies is given by the equation

$$\frac{f_1}{f_2} = \frac{N_1 - N_2}{N_2} \tag{33}$$

where N_1 is the number of points around the circumference of the figure and N_2 is one more than the number of points skipped by the cathode-ray beam as it traces the pattern.

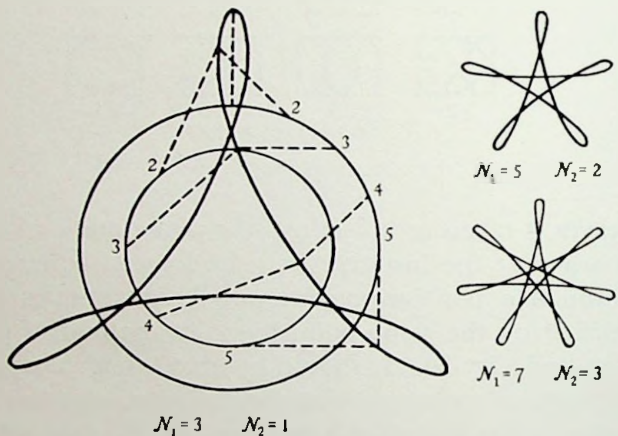


FIG. 36. Roulettes. The construction to the left shows how the figure is formed. The circumference of each circle is marked off into equal increments of time. Then, since the net position of the beam is equal to the vector sum of the positions it would have had if the forces had acted independently, the radii of the two circles which pass through corresponding points on their circumference are added together vectorially to give a point on the roulette. The frequency ratios shown are 2 : 1, 3 : 2, and 4 : 3.

With a tube having electrostatic deflection, a resistance capacity combination is used to obtain the 90° phase relationship between the two voltages which form one circle, and another resistance and capacity pair is used to obtain the other two voltages. A complete wiring diagram is shown in Fig. 37. These connections are such that circles are traced in opposite directions if each of the two frequencies act separately. This is exactly the condition specified in the beginning of the discussion. When the transformer in the wiring diagram has one of its coils reversed, however, the circles of the two frequencies acting separately would be traced in the same direction and the figure formed by them together is not a roulette. If queer figures are obtained, wiring should be checked.

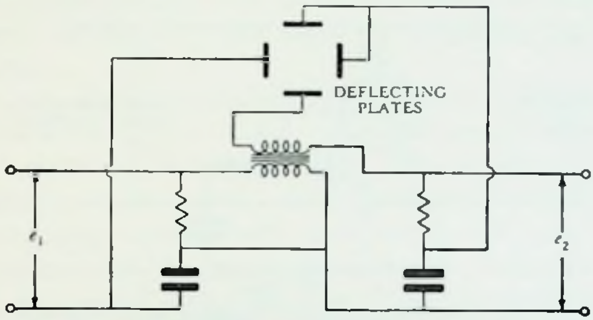


FIG. 37. Wiring diagram for obtaining roulettes.

The simple sinusoidal sweep is very useful in other ways. If a portion of the horizontal sweep voltage is applied to the plate of a vacuum tube and a voltage which is proportional to the plate current of the tube is taken from a small series resistance then amplified and applied to the vertical deflecting plates of a cathode-ray tube, the image formed on the screen will be a plot of the plate characteristic of the vacuum tube. Characteristic curves of many kinds can be obtained in this same manner, but care must be exercised in making these applications to avoid introducing phase shifts in the circuit which will distort the results.

4–5. Linear Sweep. Whenever the cathode-ray tube is used to determine the exact wave form of the voltage or current producing the vertical deflection, the horizontal sweep used must be linear; that is, the beam must be moved at a uniform rate across the tube until the end of the travel is reached and then it must instantly return to the beginning of the path and again repeat the process. This is a sawtoothed deflection. If the frequency of the sweep is the same as that of the applied wave, the image produced on the screen will be an exact plot of a cycle of the latter as a function of time since the image of each successive wave is superimposed on the preceding one.

A sawtoothed voltage wave for the production of a linear electrostatic sweep can be obtained from a motor-driven potentiometer, but this method is only suitable for relatively slow sweeps such as 1 to 20 cps. Higher speeds require the use of a circuit such as that shown in Fig. 38. Since the voltage across a condenser being

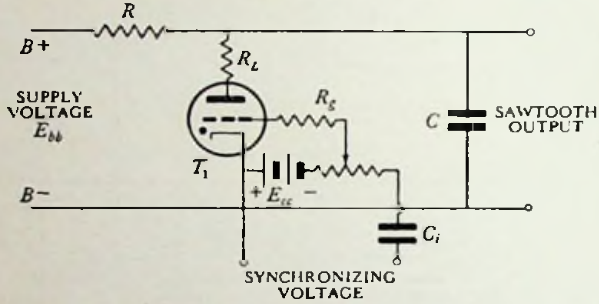


Fig. 38. Resistor-condenser charging circuit for the production of a cathode-ray sweep.

charged through a resistance varies practically linearly with time for the initial portion of the charging period, this combination is a suitable basis for a sweep circuit. The condenser C is charged through the resistance R to a fraction of the applied voltage E_{bb}, then suddenly discharged through a gas triode T_1. The triode forms a virtual short circuit across the condenser when the voltage across it exceeds the breakdown voltage of the tube. A sawtoothed

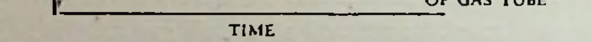


Fig. 39. Exponential charging curve illustrating how a sawtooth voltage is produced.

voltage is thus developed across the condenser as indicated in Fig. 39. The gas tube stops conducting as soon as the condenser voltage falls below the tube drop and another cycle is begun.

The oscillations can be synchronized with a wave applied to the grid of the gas tube through the condenser C_i if the positive peak of the applied wave arrives just before the normal cycle would be ended. The application of this positive pulse of voltage to the grid of the gas tube makes it start to conduct at a lower condenser voltage and thus ends the cycle earlier. Synchronization is not possible, however, if the normal cycle has ended before the pulse is applied. Consequently, it is essential that the normal frequency of the sweep circuit be lower than the frequency to which it is synchronized.

Example: A circuit like that of Fig. 38 is to produce a 500 cps sawtooth wave for a 2-in. sweep on a cathode-ray tube having a deflection sensitivity of 0.5 mm/volt. The supply voltage E_{bb} is 250 volts and the condenser C is 0.01 μf. The tube T_1 is an 884 and the synchronizing voltage will never drive the grid more than 10 volts positive with respect to the cathode. Compute the resistance of R, R_L, and R_g, also the bias E_c which should be used on the grid of T_1. The voltage swing of the sawtooth generated must be

$$\frac{2 \times 25.4}{0.5} = 101.6 \text{ volts}$$

The voltage across the condenser C will oscillate between the tube drop of 15 volts and $101.6 + 15 = 116.6$ volts. The equation for the instantaneous voltage across a condenser being charged through a resistance is $e = E\left(1 - \epsilon^{-\frac{t}{RC}}\right)$. After a steady state has been reached, the net voltage applied to the circuit at the start of a cycle is $E_{bb} - 15 = 235$. The time to complete a cycle is $\frac{1}{500}$ sec, therefore,

$$101.6 = 235\left(1 - \epsilon^{\frac{-10^6}{R \times 500 \times 0.01}}\right)$$

$$0.433 - 1.0 = -0.567 = -\epsilon^{\frac{-2 \times 10^5}{R}}$$

$$\frac{2 \times 10^5}{R} = 0.565$$

$$R = 354{,}000 \text{ ohms}$$

The maximum current rating of the 884 is 300 ma; hence, R_L should be at least

$$R_L = \frac{101.6}{0.3} = 338 \text{ ohms}$$

From the data on the tube, the grid resistor should be 1000 ohms for each volt that the grid is driven positive,

$$R_g = 1000 \times 10 = 10{,}000 \text{ ohms}$$

From the curve in the Appendix, the grid bias on the 884 to have it begin conducting at 116.5 volts is $E_c = -12$ volts.

Many combinations of R and C can be used to obtain a given sweep voltage, but R must be made large enough to prevent the current through it from the supply voltage E_{bb} from being sufficient to maintain conduction through the gas tube. If the resistance R is too small, the oscillator will "block" because the gas tube conducts continuously and a new cycle cannot be started.

The exponential charging circuit described above gives satisfactory linearity only when the sweep voltage peak is a small fraction of the supply voltage E_{bb}. The error in linearity can be conveniently defined by the following equation:

$$\text{percentage error} = \frac{E_L - E}{E} \times 100 \tag{34}$$

where E is the peak voltage which the actual sawtooth reaches, and E_L is the peak voltage it would have reached in the same time if the voltage had continued to rise at the same rate as it had at the beginning of the sawtooth. If the error in linearity is limited to 2 per cent, the peak of the sawtooth produced by an exponential charging circuit can only be 4 per cent of E_{bb}. Either a very high supply voltage must be used for the sweep, or, if a moderate supply voltage is used, the sawtooth output must be amplified before it can be used as a sweep. The latter procedure is the one generally followed.

A better utilization of the supply voltage is possible if the sweep is produced by the linear portion of the oscillatory voltage which appears across a condenser when it is charged through an inductance. The circuit is the same as Fig. 38 except that the resistance R is replaced by an inductance. In this case, the peak of the sawtooth can be 35 per cent of the supply voltage with a 2 per cent error in linearity.

If a condenser were charged with a constant current, the voltage across it would rise at a constant rate. This suggests that the substitution of a constant current device for the charging resistance would give perfect linearity. A pentode or saturated diode draws nearly a constant current for all values of applied voltage; consequently, one of these tubes is frequently substituted for the resistance R of Fig. 38. A circuit of this kind employing a pentode is shown in Fig. 40. The control grid bias is obtained by the drop across the cathode resistor R_k. As the condenser C charges and the current through the tube tends to decrease, the drop in this resistance also decreases and tends to restore the current to its original value. In this manner, an even more constant charging current is possible than the plate characteristic curves of the tube would

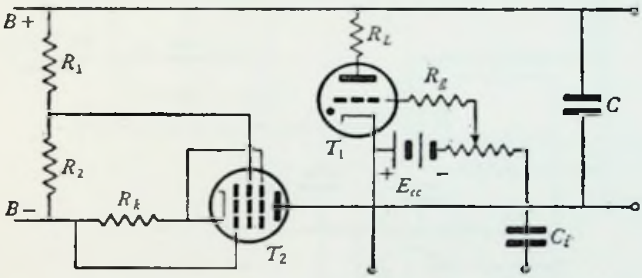


FIG. 40. Pentode-charging circuit for producing a sawtooth voltage.

indicate. Most of the improvement is lost, however, unless the screen-grid voltage is kept constant with respect to the cathode. This is done either by using a separate supply for the screen voltage or by using a large by-pass condenser between the screen grid and the cathode. The pentode charging circuit permits about a 75 per cent utilization of the supply voltage with an error in linearity of 1 or 2 per cent. It should be noted that separate heater transformers are necessary for the two tubes in Fig. 40. There is a considerable difference of potential between the cathodes which would break down the insulation between one of the heaters and its cathode if the heaters were tied together.

Rather than charging the condenser through the pentode and discharging it through the gas triode, these functions can be reversed by interchanging the position of the two tubes in the circuit. In this case, the sawtooth voltage produced across the condenser rises abruptly to its peak value and then decreases linearly to its minimum value.

The addition of a third tube to the pentode charging circuit makes it possible to develop an extremely linear electrostatic sweep. The circuit diagram is shown in Fig. 41. This circuit not only charges the condenser C with a constant current but also supplies the current taken by the resistance R_C, which is connected across the deflecting plates of the cathode-ray tube for the purpose of providing a path to ground for the stray charges which are picked up by the plates and which would defocus the beam if allowed to remain. The tube T_3 changes the characteristic of T_2 so that it takes

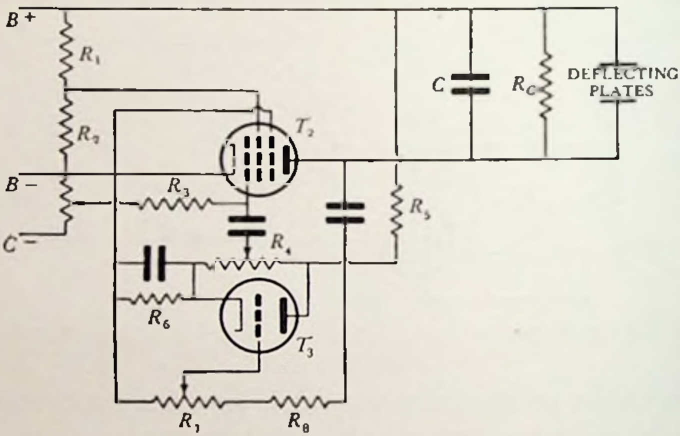


Fig. 41. Extremely linear sweep circuit. The gas tube for discharging the condenser C has been omitted from this diagram for simplicity. It is connected to this circuit in exactly the same manner as that indicated in Figs. 38 and 40.

more current as the voltage across it decreases. In other words, the tube is given a negative resistance characteristic which exactly cancels out the effect of the resistance R_C. The action of the circuit is as follows: A decrease in voltage across T_2 makes the grid of T_3 more negative; therefore, the tube T_3 draws less current through R_5 and the potential across R_4 rises. This rise in potential drives the control grid of tube T_2 more positive and causes it to take more current than before. The adjustment of R_4 and R_7 changes the resultant negative resistance of the tube and gives the sawtooth voltage developed across C almost any desired degree of linearity.

The frequency generated by any sweep circuit employing gas tubes is limited by the deionization time of the tubes. Extremely high-frequency sweeps require some kind of vacuum-tube switching circuit for their operation. A typical example of such a circuit is shown in Fig. 42. The circuit is the same as that in Fig. 38 except that a vacuum-tube trigger circuit has replaced the gas tube. There are two paths by which current can flow from the source E_{bb} through the tube. It can either flow through resistance R to the plate and cathode or it can flow through R_1 to the screen grid and cathode. Suppose that the condenser C_1 has an excess charge with a positive polarity on the plate connected to R_1. It will discharge through R_2 and thus bias the suppressor grid so negatively that the anode will be prevented from carrying current. During this period,

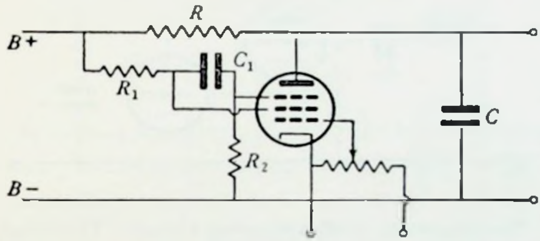


FIG. 42. High-speed sweep circuit using a vacuum-tube electronic switch to discharge the condenser C in place of the gas tubes of Figs. 38 and 40.

the condenser C is charged through the resistance R as in the gas-tube circuit. After a certain length of time, the discharge current from C_1 falls to such a low value that the bias it produces on the suppressor grid allows the plate to draw current. In so doing, the plate robs the screen grid of some of its current and the drop across R_1 is reduced. Therefore, C_1 begins to charge. The charging current flowing in R_2 is in such a direction that it makes the suppressor grid positive. The plate takes more current and the screen grid takes less until an equilibrium is reached. The conduction of current by the plate discharges the condenser C just the same as the gas tube did in the earlier example. When the plate voltage falls too low, the screen grid begins to draw more current which produces an excess voltage across R_1 and starts the discharge of C_1 as was assumed in the beginning; hence, another cycle is started. The circuit is synchronized by the application of a voltage pulse to the control grid.

Another modification of the basic sawtooth generator which uses a vacuum tube to discharge the condenser rather than a gas tube is shown in Fig. 43. The resistance R_p has been added to adapt the circuit for magnetic as well as electrostatic deflection. Its exact function will be discussed later. For the present, consider that it is zero. As in the circuit of Fig. 38, the condenser C of Fig. 43 is charged through the resistance R and after a certain interval of time it is quickly discharged through tube T_2. The pulses of voltage which bias tube T_2 positively during these conduction intervals are obtained from the blocking-tube oscillator which

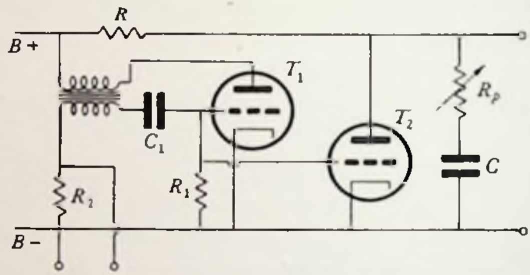


Fig. 43. Blocking-tube oscillator sweep circuit. The resistance R_p has been added to peak the output for magnetic deflection.

consists of tube T_1 and its associated circuit. The transformer is arranged so that an increase of plate current through tube T_1 will induce a positive voltage on the grid of the tube. The application of this additional positive voltage to the grid causes the plate to conduct more current and thus induces more voltage on the grid which in turn charges condenser C_1 to a higher voltage. Eventually the saturation current of the tube is reached and since the plate current then has a steady value it will produce no change in the magnetic flux of the transformer and will induce no voltage on the grid of the tube. The condenser C_1 is left with a large charge, and, since there is no other source of voltage in the circuit, it begins to discharge through the resistances R_1 and R_2. The discharge current flowing through R_1 biases the grid strongly negative and cuts off the plate current abruptly. The tube remains "blocked" until the discharge current through the resistor R_1 gives a smaller bias on the tube than the cutoff value. The tube then begins to conduct, the change in plate current drives the grid positive, and the cycle starts anew. During the short intervals of time that tube

T_1 is conducting, tube T_2 also conducts and discharges the condenser C. The frequency of oscillation depends on the rate at which C_1 discharges (i.e., the product R_1C_1). Synchronization is accomplished by applying a positive voltage across R_2 which starts conduction in tube T_1 prior to the instant when the conduction would normally begin.

A sawtooth voltage wave cannot be utilized directly as a source of the sawtooth current wave which is required for magnetic deflection. The reason for this is that the deflecting coils have inductance and the current flowing in an inductance does not have the same wave shape as the applied voltage. The fundamental equation relating the current to the voltage is

$$e = L\frac{di}{dt} \tag{35}$$

The integration of this equation shows that the current in an inductance L will increase linearly with time t only when the applied voltage e is constant

$$i = \frac{e}{L}t \tag{36}$$

If a sawtooth of current like Fig. 44*a* is to flow in a pure inductance, the applied voltage must be of the form shown in Fig. 44*b*. The small positive pulse of voltage increases the current linearly at a slow rate of speed and the large negative pulse of voltage decreases the current linearly at a high rate of speed. The magnitude of the voltage required for a given rate of increase or decrease of current can be computed directly from equation 36.

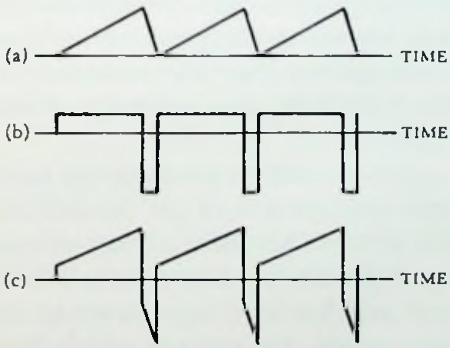


FIG. 44. Oscillograms of voltage across a coil containing (*a*) a pure resistance, (*b*) a pure inductance, and (*c*) a combination of resistance and inductance when it carries a sawtooth current.

In addition to its inductance, a practical deflecting coil also has some resistance. The sawtooth of current flowing in this resistance will give a sawtooth voltage drop which must be added to the constant voltage drop of the inductance to get the total voltage drop across the coil. The applied voltage must have the same shape as this drop and should look like Fig. 44c.

The required voltage wave shape for magnetic deflection is obtained by the addition of a "peaking" resistor R_p to the ordinary charging circuit as in Fig. 43. When the condenser C begins to charge, the high initial charging current flowing through R_p gives a sudden rise of voltage across that resistance and provides the sharp rise of voltage required in Fig. 44c. As the condenser continues to charge, the voltage across R_p falls slightly, but the rate at which it falls is only a fraction $\frac{R_p}{R_p + R}$ of the rate at which the voltage across the condenser rises. Consequently, there is a net linear increase in voltage as required. When the tube T_2 begins to conduct, the initial discharge current of the condenser produces a high negative voltage drop across the peaking resistor R_p so that the net voltage across the combination falls abruptly. The continued discharge of the condenser produces the required linear decrease in voltage. When the condenser begins to charge again, the initial charging current produces an abrupt rise in voltage across R_p and thus the combination. The function of R_p is to add the abrupt peaks which are necessary for magnetic deflection; hence, it is known as the peaking resistance. The peaking resistance is much smaller than the resistance R so that its effect in decreasing the height of the sawtooth developed across the condenser is negligible.

Another way to produce the sawtooth current required for magnetic deflection is to use the deflecting coils as the load impedance in a pentode amplifier circuit and apply any sawtooth voltage to the input terminals. Since the plate resistance of the pentode is very high and the load impedance of the deflecting coils is very low in comparison, the current which flows in the load will have exactly the same wave form as the voltage which is being applied to the grid. This statement is easily verified by looking at the equivalent circuit of a vacuum tube (see section 5–1, Fig. 47). Some of the more exactly linear sweep circuits can be used for magnetic deflection when the pentode amplifier is used to convert their output from a sawtooth voltage to a sawtooth current.

4–6. Polar Sweep. The polar sweep in effect plots a test wave as a function of time in polar coordinates rather than the rectangular coordinates of the linear sweep. It has the following advantages over the linear sweep: The time base is traced at a perfectly uniform velocity. The sweep voltage is easy to obtain from any simple sinusoidal oscillator. In a given tube, the length of the sweep is more than three times that which would be possible using a linear sweep. No retrace is necessary and so the beam is constantly on the screen; hence, no portion of a transient input wave is lost as may happen if such a wave were to occur during the retrace period of a linear sweep.

The base for the polar sweep is the Lissajous circle described in section 4–4. The beam moves with a uniform velocity in tracing this figure; hence, it makes an ideal time base for a cathode-ray tube. The radial deflections are produced in a number of ways. One of the earliest schemes added the test voltage to the accelerating voltage varying the sensitivity of the cathode-ray tube and thus the diameter of the circle formed on its screen. This system had the disadvantages of defocusing and the lack of a linear relationship between the applied voltage and the radial deflection produced.

A Heising balanced modulator is capable of overcoming these difficulties. Essentially it modulates the amplitude of the voltages producing the Lissajous circle on the screen in accord with the test voltage being applied. A circuit of this sort is shown in Fig. 45.

If the tubes are operated on the parabolic portion of their curves, their plate current is

$$i_p = A + Be_g + De_g^2 \tag{37}$$

where A, B, and D are constants which depend on the tubes used. The voltage applied to the grid of the first tube (T_1 or T_3) is

$$E \cos \omega t + f(t)$$

and that applied to the grid of the second tube (T_2 or T_4) is

$$-E \cos \omega t + f(t)$$

Since the voltage which appears across the resistance in the plate circuit of the first tube

$$R_L i_p = R_L[A + B\{E \cos \omega t + f(t)\} + D\{E \cos \omega t + f(t)\}^2]$$

is subtracted from the voltage across the plate circuit resistance of the second tube, the resultant voltage is found to be

$$E_r = 2\,ER_L[B + 2\,D\{f(t)\}]\cos\omega t \tag{38}$$

This is a sine wave whose amplitude is modulated in accordance with the applied voltage wave.

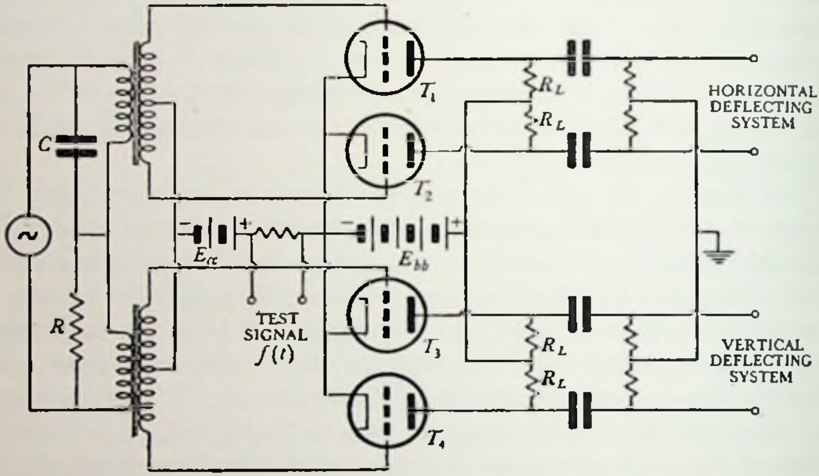
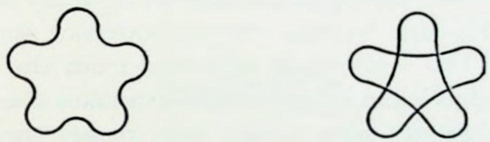


FIG. 45. Modulating circuit for obtaining a polar-coordinate deflection on a cathode-ray oscillograph.

Tetrodes or pentodes can be substituted for the triodes shown in Fig. 45. The test signal is applied to the screen grids and the sweep voltages are applied to the control grids. The change in screen voltage changes the internal resistance of the tube and its amplification is changed in proportion. Therefore, the diameter of the Lissajous circle formed on the fluorescent screen is varied in accordance with the test signal.

Special cathode-ray tubes have been devised for polar coordinate sweeps. Conical deflecting plates are used to obtain radial deflections of a cathode-ray beam which is swung in a circular path. The two deflecting systems producing the circular trace have a common center at the point where the conical deflecting plates would reach their apex if they were extended. This necessitates the use of a combination electrostatic and magnetic deflecting system. If a magnetic system were used alone, there would be difficulty in maintaining the proper phase relationship between the deflecting currents in the coils and if an electrostatic system were used alone, there would be unavoidable interactions between the vertical and the horizontal plates. The combination system avoids these difficulties and also automatically maintains the proper phase relationship between the sweep current and voltage if the inductance and capacity of the system are made part of the oscillator circuit supplying the sweep.

Polar sweeps are particularly useful in comparing an unknown frequency to a standard. If the test frequency is a multiple of the sweep frequency, a gear-shaped pattern will be formed on the screen. The number of teeth divided by the number of times the electron beam completes its sweep before the figure is closed is equal to the frequency ratio. Some examples are given in Fig. 46.



5:1 RATIO 5:2 RATIO

FIG. 46. Images obtained when a sinusoidal test signal having an integral relationship with the sweep frequency is applied to a polar-coordinate cathode-ray oscillograph.

The spiral sweep is a derivative of the polar sweep. It is produced by applying a synchronized slow speed sawtooth to the radial deflecting system of a polar sweep. The test signal is inserted in series with the sawtooth and appears as radial deflections of the spiral formed on the screen. An important advantage of a spiral sweep is its extreme length.

4–7. Fluorescent Screens. When an electron beam strikes a phosphor, it raises the orbital electrons of the phosphor into an unstable higher energy state. The electrons remain in this state for a period of time which depends on the type of phosphor. Then they fall back into their original positions and in so doing emit light. The amount of light emitted depends on the watts (power) put into the phosphor. With a given accelerating voltage, it varies nearly directly with the anode current. The amount of light also depends on the particle size and the thickness of the screen. The color of light depends on the type of material used.

The most important phosphor is zinc orthosilicate. It radiates a green light. This phosphor is prepared synthetically to assure uniformity of the resultant product. The pure material must be contaminated with a small amount of manganese to get the most radiation. Other common phosphors are calcium tungstate and zinc sulphide. The former radiates blue light and is particularly useful in photography. The latter has a long persistence characteristic.

Green screens are most desirable from a physiological point of view since they seem brighter and are more restful to the eye. Materials which produce green light have a relatively short persistence and so are not suitable for applications requiring a long persistence screen. The range of colors produced by long persistence screens is very limited and allows little or no choice.

It is often desirable to view the cathode-ray screen in a well-lighted room. The white light reflected from the surface of the phosphor cuts down the contrast and so makes the image appear dim. When a gelatin filter which only passes the color of light given off by the phosphor is placed in front of the screen, the contrast and therefore relative brightness is increased because most of the extraneous light reflected by the screen is eliminated.

Problems

4-1. The image formed on a cathode-ray oscillograph screen is in the form of an equilateral triangle when the horizontal sweep is a pure sine wave. Plot the magnitude of the vertical deflection being applied as a function of time.

4-2. Construct the roulette formed by a pair of equal voltages which have a frequency ratio of 3 : 2.

4-3. Calculate the deflection sensitivity of a cathode-ray tube having deflecting plates which are 1.5 cm long and 0.5 cm apart if the distance from the center of the plates to the screen is 15 cm and the accelerating voltage is 1000 volts.

4-4. A cathode-ray sweep of 100 cps is produced by charging a condenser through a resistor. If the maximum voltage appearing across the condenser is one-half the applied voltage E_{bb} and the sweep on the cathode-ray oscillograph is 10 cm long, find:

(*a*) The length on the screen of the first half cycle of a 1000-cps sine wave viewed with this sweep.

(*b*) Repeat for the last half cycle of the 1000-cps wave.

(*c*) Compute the percentage error in linearity as defined by equation 34.

4–5. A 6J7 charges a 0.02-μf condenser to produce a 180-volt (peak) 500-cps sawtooth from a supply voltage of 250 volts. Compute:

(*a*) The resistance of the cathode resistor R_k in Fig. 40.

(*b*) The bias to be used on the 884 discharge tube for the circuit.

(*c*) The magnitude of a single dropping resistor which could be connected between the positive side of the supply and the screen grid for the purpose of obtaining the screen-grid voltage of 100 volts. (*Note:* See caption under Fig. 51 for theory behind this calculation.)

4–6. A certain magnetic deflecting coil has an inductance of 100 mh and a resistance of 75 ohms. Find the voltage wave which must be applied to it to produce a sawtooth current which rises from 0 to 30 ma in $\frac{1}{500}$ sec and falls back to 0 ma in $\frac{1}{7000}$ sec.

Experiment 5

Object:

The object of this experiment is to study the two major methods of frequency comparison.

Preliminary:

Solve problems 4–1 and 4–2.

Performance:

1. Calibrate an oscillator against a standard-frequency source by the use of Lissajous figures. Sketch some of the figures obtained and label them with the frequency ratio they represent.

2. Assemble a circuit for producing roulette patterns and repeat the calibration using these figures. Sketch some of the patterns used and label them with their frequency ratio.

3. Plot the calibration curve of the oscillator tested.

Experiment 6

Object:

The object of this experiment is to study electrostatic deflection and sweep circuits.

Preliminary:

Solve problem 4–5.

Performance:

1. Connect a variable d-c voltage to the vertical plates of a cathode-ray oscillograph. Start the internal sweep of the oscillograph and measure the vertical deflection on the screen for a series of applied voltages. Plot the deflection as a function of the d-c voltage. Compute the deflection sensitivity.

2. Assemble the sweep circuit of the example in section 4–5. Observe the wave form produced on a cathode-ray oscillograph and measure its frequency. Check the linearity by using the output as a sweep to observe a high-frequency sine wave applied to the other set of plates. Compare the length of the first half cycle on the screen with the last half cycle.

3. Assemble the sweep circuit of problem 4–5 and compare its performance with specifications. (*Note:* Use separate filament transformers for the two tubes in the circuit to avoid breakdown between the cathode and filament.) Observe the wave form produced on a cathode-ray oscillograph.

4. Rearrange the circuit of part 3 so that the condenser is charged through the 884 tube and discharged through the 6J7. Observe the output wave produced and compare it with part 3.

Experiment 7

Object:

The object of this experiment is to study magnetic deflection and sweep circuits.

Preliminary:

Solve problem 4–6.

Performance:

1. Connect a small battery, variable resistance, and milliammeter in series with the vertical deflecting coil of an oscillograph. Start the internal sweep and measure the vertical deflections on the cathode-ray screen for a series of currents. Plot a deflection curve and compute the deflection sensitivity.

2. Set up the blocking-tube oscillator of Fig. 43. Obtain a sawtooth voltage across C when R_p is cut out. Vary R_p and note its effect on the voltage wave produced. Observe the voltage wave shape across R_1 and across the oscillation transformer. Sketch wave forms observed.

3. Add a simple triode amplifier connected to the deflecting coil load. Adjust the peaking to produce a sawtooth of current in the coil. Compare the amount of peaking required with a calculated value.

4. Replace the triode amplifier with a pentode amplifier connected to the deflecting coil load. Observe that the current in the coil has the correct form.

Note: If an oscillation transformer is not available, fair results can be obtained by using an ordinary interstage audio transformer instead. One of the coils is connected in reverse from the markings indicated. Any inductance can also be substituted for the deflecting coil. The wave shape of the current is checked by observing the voltage drop across a small resistance in series with the inductance.

Experiment 8

Object:

The object of this experiment is to study the polar sweep.

Preliminary:

1. If the tubes in Fig. 45 are 6J5's, the load resistors R_L are 50,000 ohms, and the supply voltage E_{bb} is 250 volts, select a suitable value for the bias voltage E_{cc} and explain why you made your selection.

2. If the sweep frequency is 500 cps and R is 200 ohms, find the value of C in Fig. 45.

Performance:

1. Assemble the Heising-modulator circuit of Fig. 45. Vary the bias on the modulator tubes and note the change in diameter of the circle formed on the screen.

2. Put another oscillator in series with the bias voltage on the modulating tubes. Compare the frequencies of the oscillators by means of the gear-shaped images produced.

3. Replace the sinusoidal test signal of part 2 with a low-frequency sawtooth from the oscillograph being used and adjust the circuit to produce an acceptable spiral sweep.

References

Fink, D. G., *Principles of Television Engineering*. New York: McGraw-Hill Book Company, Inc., 1940, pp. 142–158.

Fleming-Williams, B. C., "Single-valve Time-base Circuit," *Wireless Engineer*, April 1940, pp. 161–163.

Goubau, G., "A Type of Radial Deflection for the Cathode-ray Oscillograph," *Hochf-tech. u. Elek-akus.*, July 1932, pp. 1–3. (In German.)

Myers, L. M., *Electron Optics Theoretical and Practical*. London: Chapman & Hall, Ltd., 1939.

PIEPLOW, H., "The Production of Voltages Which Vary Linearly with Time," *Arch. f. Elek-tech.*, December 1938, pp. 815–821. (In German.)

The Staff of the Radio Research Station, "A Circular Time-base Giving Radial Deflections for Use with the Cathode-ray Oscillograph," *Jour. Inst. Elec. Eng.*, June 1932, pp. 82–83.

REICH, H. J., *Theory and Applications of Electron Tubes*. New York: McGraw-Hill Book Company, Inc., 1939, pp. 542–544, 592–596, 601–605.

VON ARDENNE, M., *Cathode-ray Tubes*. London: Sir Isaac Pitman & Sons, Ltd., 1939, pp. 86–91.

VON ARDENNE, M., "A New Polar Co-ordinate Cathode-ray Oscillograph with Extremely Linear Time Scale," *Wireless Engineer*, January 1937, pp. 5–12.

WATT, R. A. W., *et al.*, *Applications of the Cathode-ray Oscillograph in Radio Research*. London: His Majesty's Stationery Office, 1933.

ZWORYKIN, V. K., and G. A. MORTON, *Television, The Electronics of Image Transmission*. New York: John Wiley & Sons, Inc., 1940, pp. 69–126, 336–393, 455–462.

CHAPTER V

AMPLIFIERS

5–1. General. Amplifiers are classified according to the purpose for which they are intended and according to their operating conditions. A voltage amplifier, for example, is designed to produce a high-voltage gain whereas a power amplifier is designed for a high-power gain. The operating grid bias on the amplifier tube is indicated by its designation as class A, B, or C. A tube operating in class A conducts plate current throughout the entire cycle of the applied alternating grid voltage. Class B operation indicates that the bias on the tube is at or near cutoff; hence, the plate current flows only during approximately half of each grid-voltage cycle. In class C operation, the tube is biased considerably below cutoff. The plate current flows only when there is a signal applied to the grid and then for less than one-half of each grid-voltage cycle.

A perfect amplifier which would faithfully reproduce all signals applied to its input has to amplify each component frequency of the signal an equal amount and also delay the components by the same period of time so that they will occupy relatively the same phase position with respect to each other in the output as they did in the input. A constant delay is obtained if the phase angle of each component is shifted a multiple of 180° (i.e., completely reversed) on passing through the amplifier or if the phase angle of each component is shifted an amount which is directly proportional to its frequency.

Practical amplifiers operate over a limited range of frequencies. If they are to reproduce every signal whose components lie within their operating band, they must fulfill the criteria for a perfect amplifier throughout that band. In some cases, such as with video or television-picture amplifiers, the operating band of frequencies must be very wide, but in other cases the amplifier owes its usefulness to the fact that it selectively amplifies only a narrow band of frequencies or even a single frequency.

Class A amplifiers are the only ones which inherently reproduce the complete signal applied to the grid. It may seem, therefore, that they would be the only amplifiers which would give faithful reproduction of the input wave. This is not true. Class B amplifiers can be associated with circuits which enable them to give faithful reproduction also. The class B amplifier has a higher efficiency than the class A; consequently, it is generally used as a power amplifier. Voltage amplifiers, on the other hand, are usually operated class A. The class C amplifier is not well suited for the exact reproduction of the input wave; instead, it is particularly well adapted for other uses.

The first part of the chapter will be devoted exclusively to a discussion of class A voltage amplifiers. Class B and C amplifiers will be discussed at the end of the chapter.

Since the tube constants remain practically unchanged throughout the operating range of a class A amplifier, these amplifiers are conveniently analyzed by the use of an alternating-current equivalent circuit for the vacuum tube. Either of two possible equivalent circuits can be used. One is based on Thévenin's theorem and the other is based on Norton's theorem. Thévenin's theorem states that a complex network can be replaced by a single impedance in series with a single emf if the emf is made equal to the output terminal voltage of the network when the output current is zero, and the series impedance is made equal to the impedance of the network as seen from the output terminals when the voltages in the network are short circuited. The alternating voltage in the output of a vacuum tube is the alternating plate voltage E_p. When the alternating current in the output is equal to zero, the alternating plate voltage of the tube is equal to the alternating grid voltage E_g multiplied by the amplification factor μ. A minus sign is also introduced to take account of the 180° phase relationship between the grid and plate voltages. The impedance of a vacuum tube as seen from the output is given by the ratio of the change in plate voltage to the change in plate current when the internal emf's (i.e., the alternating grid voltage) is made zero. This is the definition of plate resistance; therefore, the equivalent circuit of a vacuum tube is a generator with an emf of $-\mu E_g$ in series with the resistance r_p as shown in Fig. 47*a*. Norton's theorem is the same as Thévenin's theorem except that the voltage generator is replaced by a constant-current generator delivering the same current as that from the network when the terminal voltage is zero, and this constant-current generator is connected in parallel with the same single

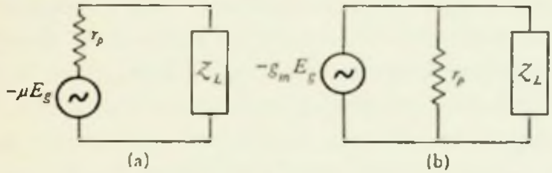


FIG. 47. (*a*) Thévenin's theorem equivalent circuit and (*b*) Norton's theorem equivalent circuit of a vacuum tube when it is operated in class A.

impedance as above. The ratio of the change in plate current to the change in grid voltage is defined as the mutual conductance g_m; hence, the Norton's theorem or constant-current equivalent circuit has a generator delivering a current $-g_m E_g$ to the parallel combination of plate resistance and load as shown in Fig. 47*b*. Either form of the equivalent circuit can be used for any tube, but the constant-current form is generally most convenient for pentodes.

5-2. Resistance-coupled Amplifiers. Resistance-coupled amplifiers are designed for operation in the audio or video range of

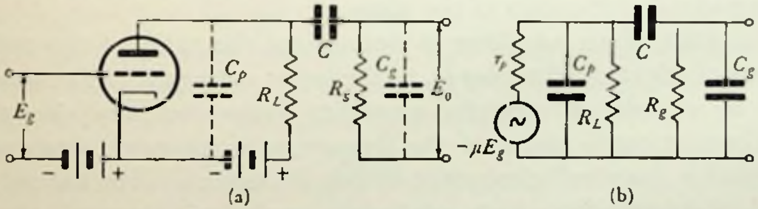


FIG. 48. (*a*) Typical resistance-coupled amplifier and (*b*) its equivalent circuit.

frequencies. A typical example together with its equivalent circuit is shown in Fig. 48. The output voltage of the tube is developed across the resistance R_L and is coupled to the next stage through the condenser C. The coupling condenser is necessary to isolate the grid of the succeeding tube from the high plate voltage. The dotted capacity C_p represents the output capacity C_{pk} of the tube together with the wiring capacity of the circuit, and the dotted capacity C_g represents the input capacity of the next stage of the amplifier. Essentially this input capacitance consists of the grid-to-cathode capacity of the tube C_{gk} plus the capacity added by the shunting effect of the grid-to-plate capacity C_{gp}. The shunting effect of C_{gp} is proportionally greater than that of C_{gk} due to the fact that the a-c voltage developed across it is greater than that developed across C_{gk}. The a-c voltage which appears across C_{gp} is equal to the sum of the a-c voltage applied to the grid and the a-c voltage developed across the load. Equations which will be developed in subsequent paragraphs show that, over the normal frequency range, the voltage developed across the load is equal to the voltage applied to the grid multiplied by the mid-frequency gain A of the stage. Consequently, the actual shunting effect of the grid-to-plate capacity of the tube is $(1 + A)$ times greater than the effect which would be produced if this same amount of capacity were connected across the input terminals of the tube. The input capacity of the tube becomes

$$C_g = C_{gk} + (1 + A)\, C_{gp}$$

For triodes it is a function of the gain of the stage. In the case of pentodes, however, the presence of the screen grid reduces the grid-to-plate capacitance to negligible proportions, and the input capacitance of the tube becomes a constant which is practically independent of the gain of the stage.

The gain of an amplifier is defined as the ratio of the output voltage E_o developed across R_g to the input voltage E_g. The magnitude of a resistance-coupled-amplifier gain gradually rises to a certain maximum value as the frequency is increased from zero. It remains practically constant at this maximum value for quite a range of frequencies which are known as the mid-frequency band. As the frequency is increased further, however, the magnitude of gain gradually drops to zero. Accompanying the change in magnitude, there is also a change in phase angle between the input and output voltage. At zero frequency, the phase angle is 270°. As frequency is increased, the phase angle gradually decreases until the mid-frequency band is reached. Throughout that band, it remains practically constant at 180°. At higher frequencies, the phase angle again begins to decrease and approach a value of 90°.

From this discussion, it is apparent that the resistance-coupled amplifier satisfies the criteria for a perfect amplifier throughout the mid-frequency band. Both the amplitude and the phase angle of gain are constant in this range. The solution of the equivalent circuit of the amplifier verifies these statements.

At medium frequencies the small capacitances C_p and C_g have such a high reactance that they can be neglected. Also the capacity C which is much larger in comparison can be neglected because its reactance is relatively small at the medium frequencies. The impedance Z_L of Fig. 47 reduces to the parallel combination of R_L and R_g

$$Z_L = \frac{R_L R_g}{R_L + R_g}$$

Since the output voltage E_o which appears across R_g is the same as the voltage across Z_L of the equivalent circuit,

$$E_o = \frac{-\mu E_g Z_L}{Z_L + r_p} \tag{39}$$

The gain A at medium frequencies is

$$A = \frac{E_o}{E_g} = \frac{-\mu Z_L}{Z_L + r_p} = \frac{-\mu R_L R_g}{R_L R_g + R_L r_p + R_g r_p} \tag{40}$$

The phase angle of the gain represents the phase shift in passing through the amplifier. In this case, it is 180° as indicated by the minus sign in front of the equation.

Most pentodes have such a very high plate resistance that the products $R_L R_g$ and $R_L r_p$ of equation 40 can be neglected in comparison with $R_g r_p$. This reduces the equation to

$$A = \frac{-\mu R_L R_g}{R_g r_p} = -g_m R_L \tag{41}$$

For this reason, the value of plate resistance is often omitted from the data on these tubes and the value of mutual conductance given instead. The equations which follow can also be adapted for them by making the same simplifications as are indicated here and by substituting $\frac{\mu}{g_m}$ for r_p. It should also be noted that equation 41 can be derived directly from the Norton's theorem equivalent circuit by making the same approximations as above.

At low frequencies, the effect of the coupling condenser C cannot be neglected. The effect of the tube and wiring capacitances, however, is even less than at the medium frequencies. The load impedance reduces to the series combination of C and R_g in parallel with R_L. The solution of this circuit for the gain at low frequencies is

$$A = \frac{E_o}{E_g} = \frac{-\mu R_L R_g}{(R_L R_g + R_L r_p + R_g r_p) - j\dfrac{(R_L + r_p)}{\omega C}} \qquad (42)$$

where $\omega = 2\pi f$ and f is the frequency of the applied sinusoidal voltage E_g. The magnitude of the above vector equation indicates the relative amplitude of the input and output sine wave and the angle represents the phase shift. As the frequency approaches zero, the phase shift approaches 270°.

The lower frequency limit f_l of the amplifier response is defined as that frequency at which the magnitude of the gain is $\frac{1}{\sqrt{2}}$ times the mid-frequency value (equation 40). This occurs at the point where the real and imaginary parts in the denominator of equation 42 are equal.

$$(R_L R_g + R_L r_p + R_g r_p) = \frac{(R_L + r_p)}{2\pi f_l C}$$

and

$$f_l = \frac{(R_L + r_p)}{2\pi C(R_L R_g + R_L r_p + R_g r_p)} \qquad (43)$$

The phase angle at this point is 225°.

Example: An amplifier uses a 6J5 tube and has the following circuit constants: $R_L = 50{,}000$ ohms, $R_g = 500{,}000$ ohms, $C = 0.01\ \mu f$. Calculate the gain at 100 cps.

This is in the low-frequency range and so equation 42 applies. The μ of a 6J5 is 20 and the r_p is 7700 ohms.

$$\omega C = 2\pi \times 100 \times 0.01 \times 10^{-6} = 6.28 \times 10^{-6} \text{ ohms}$$

$$A = \frac{-20(5 \times 10^4 \times 5 \times 10^5)}{(25 \times 10^9 + 3.85 \times 10^8 + 3.85 \times 10^9) - j\dfrac{5.77 \times 10^4}{6.28 \times 10^{-6}}}$$

$$A = \frac{-20 \times 25}{29.2 - j\,9.2} = \frac{500\underline{/180^\circ}}{30.6\underline{/-17.5^\circ}}$$

$$A = 16.4\underline{/197.5^\circ}$$

The gain is 16.4 and the phase shift is 197.5°.

At the high frequencies, the coupling condenser acts like a short circuit and the shunting effect of the tube and wiring capacitances becomes very pronounced. The equivalent load impedance becomes, in effect, the parallel combination of R_L, R_g, C_p, and C_g. The solution of this circuit for the gain gives

$$A = \frac{E_o}{E_g} = \frac{-\mu R_L R_g}{(R_L R_g + R_L r_p + R_g r_p) + j\omega(C_p + C_g)(R_L R_g r_p)} \qquad (44)$$

The magnitude of the gain drops off and the phase-shift angle gradually decreases from 180° toward 90° as higher and higher frequencies are approached.

The upper frequency limit f_u is defined as the point where the gain falls back to $\frac{1}{\sqrt{2}}$ times the mid-frequency value. It occurs at the point where the real and imaginary parts of the denominator of equation 44 are equal.

$$f_u = \frac{(R_L R_g + R_L r_p + R_g r_p)}{2\pi(C_p + C_g)(R_L R_g r_p)} \qquad (45)$$

The phase angle at this point is 135°.

The high-frequency response of an amplifier can be improved by reduction of the wiring capacity through the use of short and direct leads and by the proper selection of low-capacity tubes. For example, pentodes are used in wide-band amplifiers because they have much lower interelectrode capacities in comparison with their gain than do triodes. The low-frequency response is improved by making the coupling condenser C larger. If the condenser is made too large, however, the high-frequency response will be impaired by the capacitance between it and ground which is added to the wiring. Also the amplifier will tend to "block" or become inoperative following the application of a positive voltage at the input. The reason for this is that the voltage causes a charge to build up on the condenser which produces a negative bias on the tube as it slowly leaks off through R_g. The bias may hold the tube below cutoff for an appreciable length of time and thus prevent it from operating. In order to extend the width of the mid-frequency band beyond the limits which are imposed by these factors, negative feedback or special compensating circuits are used.

5–3. Feedback. A feedback amplifier is one which has a portion of its output reapplied at the input. The feedback can be positive and add to the input signal or negative and subtract from it. Positive feedback produces a higher gain in a given amplifier, but it often leads to instability of operation and distortion. Negative feedback, on the other hand, lowers the gain, but it gives a more uniform response characteristic and it also tends to limit distortion.

If β is the factor (which may be a complex number) by which the output voltage E_o must be multiplied to obtain the feedback voltage, the net voltage which is applied to the grid of the tube is

$$E_g = E_i + \beta E_o \tag{46}$$

where E_i is the alternating input voltage. It should be particularly noted that in this equation and the ones which follow, β carries the sign associated with the type of feedback which is used (i.e., with negative feedback, the sign of β is also negative). The gain of the amplifier without feedback is

$$A = \frac{E_o}{E_g} \tag{47}$$

and the gain with feedback is

$$A_f = \frac{E_o}{E_i} \tag{48}$$

The substitution of equations 47 and 48 into equation 46 yields

$$\frac{1}{A} = \frac{1}{A_f} + \beta \quad \text{or} \quad \frac{1}{A_f} = \frac{1}{A} - \beta$$

and

$$A_f = \frac{A}{1 - \beta A} \tag{49}$$

when the factor βA is large in comparison with 1, the above equation reduces to

$$A_f = -\frac{1}{\beta} \text{ approximately} \tag{50}$$

This shows that the gain with feedback is made independent of the tube and its circuit; instead, the nature of the feedback constant determines the performance of the amplifier. For this reason, negative-feedback amplifiers are particularly well adapted for use as component parts of measuring instruments. Almost any characteristic can be given to an amplifier by the proper selection of the feedback circuit.

There are two general types of negative feedback. Current feedback inserts a voltage at the input which is proportional to the current flowing in the load, and voltage feedback inserts a voltage at the input which is proportional to the voltage across the load.

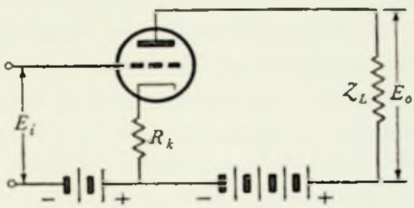


FIG. 49. Typical amplifier with current feedback. The voltage drop across the resistance R_k helps to bias the tube negatively in addition to supplying the feedback.

A circuit diagram for obtaining current feedback is shown in Fig. 49. The same current flows through both R_k and Z_L; hence, the ratio of the feedback voltage to the output voltage or the factor β is equal to the ratio of these two impedances

$$\beta = -\frac{R_k}{Z_L} \tag{51}$$

Voltage feedback can be obtained from the circuit in Fig. 50. If the condenser C is large enough that it can be neglected, the total output voltage appears across the series combination of R_1 and R_2. The feedback factor or the fraction of the output voltage which is reintroduced at the grid is

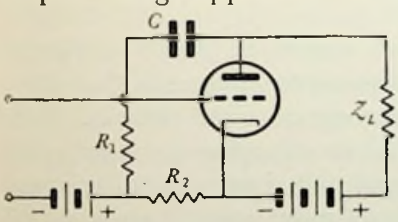


FIG. 50. Amplifier with a typical voltage-feedback circuit.

$$\beta = \frac{-R_2}{R_1 + R_2} \tag{52}$$

In computing the gain of this circuit, the resistances R_1 and R_2 should also be considered as part of the load impedance Z_L. If so desired, a combination of both current and voltage feedback can be employed in a single circuit.

A cathode resistor such as the one shown in Fig. 49 is also often used to obtain the grid-bias voltage E_{cc}. The size of the resistance necessary in this case can be calculated by Ohm's law from the cathode current and the bias voltage required. If no feedback is intended, the alternating current must be shunted around the resistor by means of a by-pass condenser so that practically no a-c voltage drop will be introduced in the grid circuit. The reactance of the by-pass condenser should be less than one-tenth the cathode resistance for the lowest frequency concerned. The circuit of Fig. 51 shows a typical cathode-bias arrangement.

The cathode-follower stage has the same wiring diagram as Fig. 49 except that the impedance Z_L is almost invariably omitted and the plate is connected directly to the supply voltage. The output voltage is taken from across the cathode resistor R_k. By the use of this connection, a low-impedance output which is at ground potential is obtained. Since the output voltage is equal to the feedback voltage, $\beta = -1$ and the equation for the mid-frequency gain of the stage (equation 49) becomes

$$A = \frac{\dfrac{\mu R_k}{r_p + R_k}}{1 + \dfrac{\mu R_k}{r_p + R_k}} = \frac{\mu R_k}{r_p + (\mu + 1)R_k} \tag{53}$$

It is apparent from the above equation that the gain of a cathode-follower stage is always less than unity. The function of such a stage is, therefore, to match a low-impedance load rather than to obtain a voltage gain.

A unique application of feedback occurs in the Wien-bridge oscillator. This oscillator generates sine-wave voltages in the audio- and video-frequency range without the use of tuned circuits. Only resistance and capacitance elements are employed; hence, the oscillator can be made small in size and with a very high frequency stability. The Wien bridge consists of four arms. One arm (*a*) consisting of a series combination of resistance and capacity lies opposite a resistance arm (*b*); the third arm (*c*) consisting of a parallel combination of resistance and capacity lies opposite another resistance arm (*d*). The input voltage to the bridge is applied between the corner formed by the junction of arms (*b*) and (*c*) and the corner formed by the junction of arms (*a*) and (*d*). The output from the bridge is taken from the corner formed by the junction of arms (*a*) and (*c*) and the corner formed by the junction of arms (*b*) and (*d*). When properly adjusted, the bridge is in balance at a single a-c frequency. For that frequency there is no output voltage. If one of the resistance arms is changed a small amount, however, there is a slight output voltage which is in phase with the applied input voltage. Small changes in frequency give a slightly greater output voltage, but this voltage is thrown considerably out of phase with the input. In the Wien-bridge oscillator, a slightly unbalanced bridge of this sort is used as a voltage-feedback network between the output of a 2-stage amplifier (which has a net phase shift of zero) and its input. If a small amount of voltage at the proper frequency — from the noise and other normal voltage variations of the amplifier — is applied to the input, it is amplified and reintroduced by the bridge network in the proper phase relationship to add to the initial voltage. Consequently, there is a tendency for this particular frequency of oscillation to build up indefinitely. By the use of a ballast lamp as one of the resistance arms of the bridge, this tendency is checked before the oscillations become so great that they drive the amplifier beyond class A operation and cause distortion of the wave form. As the oscillations become excessively large, the ballast lamp heats up and increases in resistance. This action brings the bridge more nearly into balance and reduces the amount of voltage reintroduced at the input of the amplifier. As a result, an equilibrium condition is obtained whereby the output voltage is large enough to be useful yet small enough to avoid distortion.

5-4. Resonance Compensation. The compensating circuits which are effective at high frequencies produce no change in the low-frequency response and vice versa. Therefore, the two problems can be considered separately.

Equation 44 indicates that the high-frequency response can be improved if the effective impedance of the tube and wiring capacity is increased at the higher frequencies. The first thing that comes to mind for increasing the reactance of a capacitive circuit is to put it into parallel resonance. Naturally this procedure will produce the greatest increase in impedance at the resonant frequency, but if the Q (ratio of reactance to resistance) of the circuit is lowered, the resonance peak is flattened out and spread over a wider band. When such a resonant peak is made to occur in the region where the normal gain of the amplifier falls off rapidly, a considerable extension of the mid-frequency band, with the gain equal to the mid-frequency value, is made possible.

The resonating inductance or shunt peaking coil, as it is called, is placed in series with the load resistance R_L as shown in Fig. 51. The Norton's theorem equivalent circuit of the amplifier is also shown in the figure. Since r_p and R_g are very large and the reactance of C is very small at the high frequencies, these elements can be neglected and the equivalent circuit takes the form shown in Fig. 52. C_T is the total tube and wiring capacitance, i.e., the sum of C_p and C_g.

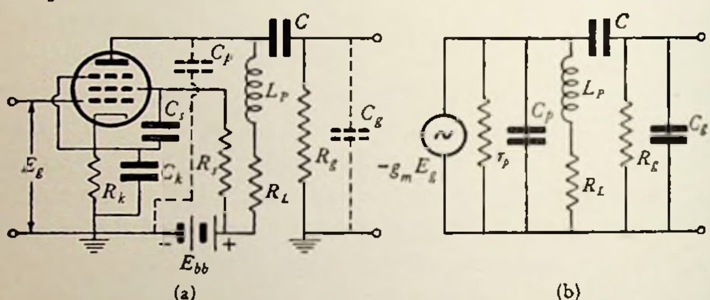


Fig. 51. (a) An amplifier compensated with a shunt-peaking coil together with (b) its equivalent circuit. The grid bias is obtained by the voltage drop through the cathode resistor R_k as explained in section 5–3. The screen current of a pentode is practically constant for all values of plate voltage (see Fig. 134 in the Appendix); therefore, this current is passed through a resistance R_s to produce a voltage drop which will subtract the proper amount from the supply voltage to give the required screen voltage with respect to cathode. The small alternating voltage which is produced across R_s with changes in plate voltage is shunted out through the capacitor C_s which should have a negligible reactance at the lowest frequency concerned with respect to the dropping resistor R_s.

A very flat gain curve which has an extremely small variation in phase angle is obtained if the inductive reactance of the coil is made 0.3 times the reactance of the capacity C_T at the upper frequency limit f_u of the amplifier.

$$2\,\pi f_u L_P = \frac{0.3}{2\,\pi f_u C_T} \tag{54}$$

The upper frequency limit of a compensated amplifier is defined as the highest frequency with a gain practically equal to the mid-frequency value. The reason for the choice of the same symbol, but different definition for this term, will be explained in the next paragraph. The proper value of Q for the resonant circuit is obtained by making the load resistance equal to 0.83 times the capacitive reactance at this frequency.

$$R_L = \frac{0.83}{2\,\pi f_u C_T} \tag{55}$$

A less constant but much higher gain is obtained if the load resistance is made equal to the capacitive reactance.

$$R_L = \frac{1}{2\,\pi f_u C_T} \tag{56}$$

and the inductive reactance is made half as large as the capacitive reactance.

$$2\,\pi f_u L_P = \frac{1}{4\,\pi f_u C_T} \quad \text{or} \quad L_P = \frac{R_L}{4\,\pi f_u} \tag{57}$$

at the upper frequency limit. Equation 56 is the definition of the upper frequency limit for an uncompensated amplifier (equation 45); therefore, the same symbol is used for it here as was used there. The gain of the compensated amplifier at this frequency, however, is the same as the mid-frequency gain.

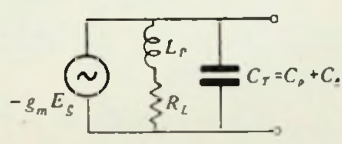


FIG. 52. The equivalent circuit of the amplifier in Fig. 51 at high frequencies.

The first step in the design of a peaking coil is to measure the tube and wiring capacity of the amplifier. This is done by measuring the frequency at which the gain of the uncompensated amplifier drops to $\frac{1}{\sqrt{2}}$ times its mid-frequency value or the upper frequency limit for the load resistance used. In order to make an accurate measurement, the input capacity of the measuring vacuum-tube voltmeter is isolated from the stage being measured by the tube which is fed by the voltage across R_g. The load resistor of this latter tube is made very small (say 100 ohms) so that its gain remains at the mid-frequency value far beyond the upper frequency limit of the stage being tested. The vacuum-tube voltmeter is connected across the load resistor of the second tube and thus is effectively isolated from the stage. The magnitude of the shunting capacity C_T is calculated by substituting the measured value of f_u in equation 56. Equations 56 and 57 are then used to calculate the values of R_L and L_P which will make the amplifier meet the design specifications.

Example: An amplifier circuit such as Fig. 51 uses a 6AC7/1852 tube. When R_L is 2000 ohms, the measured upper frequency limit of the uncompensated amplifier is 2.6 mc. Calculate the size peaking coil and load resistor to give the amplifier an upper frequency limit of 4.0 mc. Also calculate the mid-frequency gain of the compensated amplifier.

From equation 56 the total wiring and tube capacity is

$$C_T = \frac{1}{2\,\pi f_u R_L} = \frac{1}{2\,\pi \times 2.6 \times 10^6 \times 2000} = 30.6\ \mu\mu\text{f}$$

Again from equation 56

$$R_L = \frac{1}{2\,\pi \times 4 \times 10^6 \times 30.6 \times 10^{-12}} = 1300 \text{ ohms}$$

From equation 57

$$L_P = \frac{R_L}{4\,\pi f_u} = \frac{1300}{4\,\pi \times 4 \times 10^6} = 25.8\ \mu\text{h}$$

Since the g_m for the 1852 is 9000, the mid-frequency gain from equation 41 is

$$A = -\,g_m R_L = -\,9000 \times 10^{-6} \times 1300 = 11.7\underline{/180^\circ}$$

Another compensating circuit which is known as "series peaking" is shown in Fig. 53. This circuit has the advantage of separating the tube capacitances from each other and thus giving a higher gain with less phase shift than the shunt peaking circuit. The idea

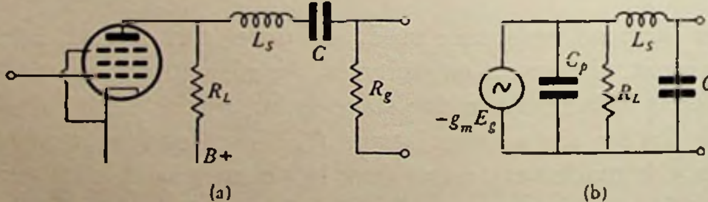


FIG. 53. (*a*) Connection for series peaking and (*b*) the corresponding high-frequency equivalent circuit. In this diagram, the plate load alone has been drawn for simplicity. All the other connections are the same as those shown in Fig. 51.

here is to take the decreased voltage developed across the parallel combination of C_p and R_L and apply it to the series resonant circuit L_S and C_g. The capacity C_p is determined by measuring the upper frequency limit of the tube and the load alone with a vacuum-tube voltmeter having a known input capacitance. C_g can then be calculated from the measurement of C_T. It should be twice as large as C_p. If this is not the case, the coupling condenser C can be put on the other side of the peaking coil to assist in the adjustment of the ratio. The inductance and capacitance are resonated at $\sqrt{2}$ times the upper frequency limit; hence,

$$L_S = \frac{1}{8\,\pi^2 f_u^2 C_p} \tag{58}$$

The load resistance for most satisfactory operation with series peaking is

$$R_L = \frac{3}{4\,\pi f_u C_T} \tag{59}$$

If the ratio of C_g to C_p is reversed, the same formulas still apply. In effect the network is just turned around and will work equally well because of network reciprocity.

A combination of series and shunt peaking is shown in Fig. 54. This circuit allows a much higher gain than shunt peaking. The design equations are:

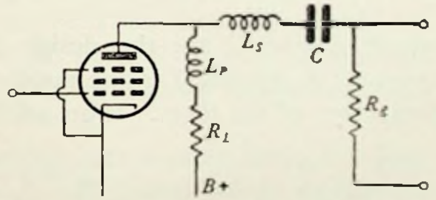


FIG. 54. Connection of the plate-load circuit for a combination of series and shunt peaking. The other connections are the same as those shown in Fig. 51.

$$\frac{C_g}{C_p} = 2 \qquad L_P = 0.12\, C_T R_L^2$$
$$R_L = \frac{1.8}{2\,\pi f_u C_T} \qquad L_S = 0.52\, C_T R_L^2 \tag{60}$$

The load resistance for this circuit is 80 per cent higher than for a simple shunt peaking coil; therefore, the gain is 80 per cent higher than the latter.

5-5. Filter Compensation. Wheeler has approached the problem of compensation from another viewpoint. He has shown that a low-pass filter section can also be used as a coupling network between stages. The tube and wiring capacitance are used for the terminal capacitor of a filter having an input impedance which is designed to be constant over the pass band. One of the two terminals of the filter coupling network is connected to the plate of the tube and through the coupling condenser C to the resistor R_g. The other terminal is connected to the B supply. The effective load on the amplifier is then the constant input impedance of the filter.

A typical filter coupling consists of a constant K half section to provide the input terminals with a shunting capacity and this is followed by an m-derived half section with $m = 0.6$ for the purpose of matching the filter impedance to the terminating load resistor R_L. The input impedance of this combination is the characteristic impedance of a pi or mid-shunt filter and is

$$Z_0 = \frac{\sqrt{\frac{L}{C}}}{\sqrt{1 - \frac{f^2}{f_c^2}}} \tag{61}$$

where L and C are the design inductance and capacity of the constant K section, f is the applied frequency, and f_c is the cutoff frequency of the filter. If an additional capacity C_a

$$C_a = \frac{n}{2\pi f_c \sqrt{\frac{L}{C}}} \tag{62}$$

is placed across the input terminal, the input impedance becomes

$$Z_i = \frac{1}{\frac{1}{Z_0} + j2\pi f C_a} = \frac{\sqrt{\frac{L}{C}}}{\sqrt{1 - \frac{f^2}{f_c^2}} + j\frac{nf}{f_c}} \tag{63}$$

When n is made equal to 1, the magnitude of this input impedance becomes a constant.

$$|Z_i| = \sqrt{\frac{L}{C}} = R = R_L \tag{64}$$

The corresponding phase shift curve varies from zero at $f = 0$ to 90° at $f = f_c$ in nearly a straight line relationship. If n is made slightly greater then 1, the phase shift curve has an even more uniform slope and the condition of having the angle of phase shift directly proportional to frequency is attained as in a perfect amplifier. The capacity of the wiring and tubes is used for the terminal capacity of the filter $\frac{1}{2}C$ plus the additional capacity C_a.

$$C_T = C_a + \tfrac{1}{2}C \tag{65}$$

If n is equal to 1 in equation 62, that equation becomes one-half the regular design equation for the capacity C of a low-pass filter having a cutoff frequency f_c since $\sqrt{\frac{L}{C}} = R$. The substitution of equations 64 and 65 in equation 62 then gives

$$C_T = \frac{1}{\pi f_c R_L} \tag{66}$$

The cutoff frequency f_c of the filter section is equal to the upper frequency limit f_u of the amplifier because the load impedance is only constant up to this point. Equation 66 is used to determine the magnitude of load resistance for a filter coupling circuit. The other circuit constants are designed with the usual filter formulas. It should be noted that equation 66 gives a slightly higher load resistance and therefore a higher gain for a given upper frequency limit than do equations 60.

A somewhat higher gain is possible over a given band if a four-terminal section is used rather than the two-terminal section described above. A four-terminal filter separates the capacity C_p from C_g by adding a section to the constant K filter. C_g then becomes the shunt capacity of the second section and is separated from C_p by an inductance. Equation 66 then gives the relationship between C_p and R_L; hence, a higher gain with a given upper frequency limit is possible. The gain of a four-terminal section is

$$A = -g_m Z_0 \sqrt{\frac{C_p}{C_g}} \tag{67}$$

The low-pass amplifiers designed on the basis of these equations can be adapted for band-pass service by using the band-pass equivalents of these circuits.

5–6. Figure of Merit. An important problem in the design of a wide-band amplifier is the evaluation of the relative merit of different tubes for this service. Since the four-terminal filter gives the highest gain per stage of any compensating network, it is used as a basis for the comparison of different tubes. The figure of merit of a tube is defined as the upper frequency limit of an amplifier employing the tube if the gain is unity and a four-terminal filter coupling is used. Since the input impedance is equal to R_L, and C_T reduces to C_p, equation 66 becomes

$$C_p = \frac{1}{\pi f_c Z_0} \tag{68}$$

The solution of 68 for Z_0 and the substitution into equation 67 gives

$$|A| = \frac{g_m}{\pi f_c C_p}\sqrt{\frac{C_p}{C_g}} \tag{69}$$

If the magnitude of the gain is unity, then

$$f_c = \frac{g_m}{\pi\sqrt{C_p C_g}} \tag{70}$$

which represents the upper frequency limit or band width of the tube and is called the figure of merit. Tubes which are best suited for wide-band service have a high figure of merit. For the 6SJ7, it is 125 mc while for the 6AC7/1852, it is 385 mc.

A radically different tube has been designed which gives a still higher figure of merit. It is the orbital-beam secondary-electron multiplier. A small tube structure consisting of a cathode, control grid, and screen grid is used to obtain a controlled beam of electrons which strike a secondary electron emitter and produce a much larger current for the plate. The use of secondary emission thus materially increases the effective mutual conductance of the tube. The figure of merit for an orbital-beam tube is 935 mc.

5–7. Low-frequency Compensation. A simple low-frequency compensating circuit is shown in Fig. 55. As the frequency is decreased, the reactance of C_F increases and in effect makes a portion of R_F act like an additional load resistance, thus tending to increase the gain of the stage as the coupling condenser tends to reduce it. The circuit also introduces a phase shift which cancels out the phase shift due to C and R_g. Compensation is effective as long as R_F is about ten times the reactance of C_F. For proper operation, the time constant R_LC_F of the load resistor and compensating condenser must be made equal to the time constant R_gC of the grid resistor and coupling condenser as can be shown by the solution of the equivalent circuit. R_F is made as large as a reasonable B supply voltage will permit. The design procedure is to select a value for C_F and calculate the time constant with the load. The grid time constant is then adjusted accordingly.

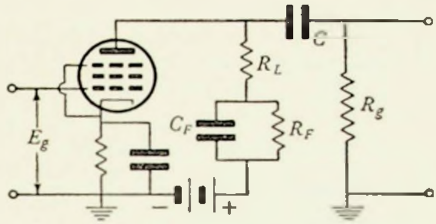


FIG. 55. Low-frequency compensating circuit. The screen supply voltage has been omitted from the diagram for simplicity. It is the same as that shown in Fig. 51.

5-8. Radio-frequency Amplifiers. Tuned radio-frequency amplifiers are used to select and amplify a particular radio frequency or band of radio frequencies. A typical example of such a circuit is shown in Fig. 56. The primary P of the radio-frequency transformer contains an inductance L_p and a resistance R_p. The secondary S contains an inductance L_s and a resistance R_s. The mutual inductance between the coils of the transformer is M. The equivalent circuit of the amplifier takes the form shown in the figure.

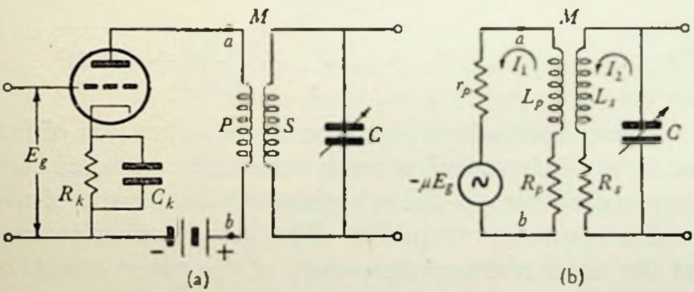


FIG. 56. (*a*) Singly tuned radio-frequency amplifier and (*b*) its equivalent circuit. The addition of another variable condenser between the points *a* and *b* would make this circuit a doubly tuned radio-frequency amplifier.

The gain of this amplifier is defined in the same manner as before and can be computed by the solution of the equivalent circuit. The voltage drops in the primary side are

$$[(r_p + R_p) + j\omega L_p]I_1 + j\omega MI_2 = \mu E_g \tag{71}$$

and since R_p is very small in comparison with r_p, it is neglected and the equation becomes

$$(r_p + j\omega L_p)I_1 + j\omega M I_2 = \mu E_g \tag{72}$$

Similarly, the voltage drops in the secondary are

$$j\omega M I_1 + \left[R_s + j\left(\omega L_s - \frac{1}{\omega C}\right)\right] I_2 = 0 \tag{73}$$

The solution of the simultaneous equations 71 and 72 for the secondary current I_2 gives

$$I_2 = \frac{-\mu E_g j\omega M}{[r_p + j\omega L_p]\left[R_s + j\left(\omega L_s - \frac{1}{\omega C}\right)\right] + (\omega M)^2} \tag{74}$$

The output voltage E_o is the product of the current I_2 and the reactance of the capacitance C

$$E_o = I_2\left(\frac{-j}{\omega C}\right) = \frac{-\mu E_g \dfrac{M}{C}}{[r_p + j\omega L_p]\left[R_s + j\left(\omega L_s - \frac{1}{\omega C}\right)\right] + (\omega M)^2} \tag{75}$$

The gain is equal to equation 75 divided by the input voltage E_g.

$$A = \frac{E_o}{E_g} = \frac{-\mu \dfrac{M}{C}}{[r_p + j\omega L_p]\left[R_s + j\left(\omega L_s - \frac{1}{\omega C}\right)\right] + (\omega M)^2} \tag{76}$$

A plot of the magnitude of equation 76 as a function of frequency is in the form of the familiar series-resonance curve of current vs. frequency and illustrates the selectivity characteristic of the singly tuned radio-frequency amplifier. The gain reaches its maximum value at the series resonant frequency of the tuned circuit or when $\omega L_s = \dfrac{1}{\omega C}$. The gain at resonance is

$$A_r = \frac{-\mu \dfrac{M}{C}}{(r_p + j\omega_r L_p)R_s + (\omega_r M)^2} \tag{77}$$

Since the term $j\omega_r L_p R_s$ is negligible in comparison with the other two terms of the denominator, equation 77 becomes

$$A_r = \frac{-\mu \dfrac{\omega_r M}{R_s \omega_r C}}{r_p + \dfrac{(\omega_r M)^2}{R_s}} = \frac{-\mu \omega_r M Q_s}{r_p + \dfrac{(\omega_r M)^2}{R_s}} \tag{78}$$

Where Q_s is the Q of the secondary coil of the radio-frequency transformer.

If equation 78 is maximized by differentiating with respect to $\omega_r M$ and setting the derivative equal to zero, the optimum value of $\omega_r M$ for maximum gain is found to be

$$\omega_r M = \sqrt{R_s r_p} \tag{79}$$

At resonance, the input impedance to the primary coil of the radio-frequency transformer is

$$Z_i = R_p + j\omega_r L_p + \frac{(\omega_r M)^2}{R_s} \tag{80}$$

When the optimum value of $\omega_r M$ is used, this input impedance becomes practically equal to the plate resistance of the tube

$$Z_i = R_p + j\omega_r L_p + \frac{r_p R_s}{R_s} \approx r_p \tag{81}$$

neglecting $R_p + j\omega_r L_p$ which is very small in comparison.

The addition of another tuning condenser between the points *a* and *b* on Fig. 56 makes the circuit a doubly tuned radio-frequency amplifier. In the same manner as the above, the solution of the equivalent circuit of this amplifier gives an equation for its gain. Considering that both the primary and secondary coil of the radio-frequency transformer have the same inductance L, the same resistance R, and are tuned with the same amount of capacity C, the equation for gain is

$$A = \frac{j \dfrac{g_m M}{\omega C^2}}{\left[R + j\omega L - j\dfrac{1}{\omega C}\right]^2 - (j\omega M)^2} \tag{82}$$

At the resonant frequency ω_r, the gain becomes

$$A_r = \frac{j \dfrac{g_m M}{\omega_r C^2}}{R^2 + (\omega_r M)^2} \tag{83}$$

The value of mutual inductance which gives the maximum gain at resonance can be found by differentiating equation 83 with respect to M and equating the derivative to zero. The result is

$$\omega_r M = R \tag{84}$$

The substitution of this value into equation 83 gives the maximum magnitude of the gain.

$$|A_{\max}| = \frac{g_m}{2R(\omega_r C)^2} \tag{85}$$

When the mutual inductance has the value indicated by equation 84, the radio-frequency transformer is said to be critically coupled. The gain magnitude vs. frequency curve has a single peak and that occurs at the resonant frequency. The magnitude of the peak is given by equation 85.

When the mutual inductance is made less than critical, the transformer is said to be undercoupled and the gain curve has a single peak like that of critical coupling, but the magnitude of that peak is considerably reduced.

On the other hand, a larger mutual inductance than critical causes the gain curve to have two peaks — one on either side of the resonant frequency. The gain at each of these peaks has the magnitude given by equation 85 and the gain at the resonant frequency is considerably less.

The reason for the two peaks can best be seen by a rearrangement of the terms of equation 82. If the denominator of this equation is expanded and the term $j\omega M\left(R - j\frac{1}{\omega C}\right)$ is added to and subtracted from the resulting expansion, it can be factored and the equation for gain becomes

$$A = \frac{j\dfrac{g_m M}{\omega C^2}}{\left[R + j\omega(L - M) - j\dfrac{1}{\omega C}\right]\left[R + j\omega(L + M) - j\dfrac{1}{\omega C}\right]} \tag{86}$$

When the reactance in either factor of the denominator reduces to zero, the gain will be close to a maximum. Then the upper frequency at which a peak occurs is

$$\omega_u = \frac{1}{\sqrt{(L - M)C}} \tag{87}$$

and the lower frequency peak is at

$$\omega_l = \frac{1}{\sqrt{(L + M)C}} \tag{88}$$

The ratio between the gain at the two peaks and that at the resonant frequency ω_r can be found by dividing equation 85 by the magnitude of equation 83.

$$\frac{A_{\max}}{A_r} = \frac{R^2 + (\omega_r M)^2}{2\,R\omega_r M} = \frac{1}{2}\left[\frac{R}{\omega_r M} + \frac{\omega_r M}{R}\right] \tag{89}$$

Since $M = KL$ where K is the coefficient of coupling

$$\frac{A_{\max}}{A_r} = \frac{1}{2}\left[\frac{R}{\omega_r LK} + \frac{\omega_r LK}{R}\right] = \frac{1}{2}\left[\frac{1}{QK} + QK\right] \tag{90}$$

By the proper choice of Q and K, it is possible to obtain a fairly uniform gain over a band of frequencies; hence, the doubly tuned radio-frequency amplifier is well adapted to band-pass service. Even better characteristics can be obtained if several doubly tuned amplifiers are cascaded and adjusted to compensate for the deficiencies of each other. That is, the peak of one section can be used to fill in the dip in the other. Extremely wide-band service, however, requires the use of filter coupling networks such as those described in section 5–5.

5–9. Power Calculations. The power output of an amplifier is equal to the product of the effective values of the load alternating current and voltage. When the load is a pure resistance, the data for the computation of power output can be found graphically by a load line on the characteristic curves of the tube (Fig. 57). The intersection of the load line with the grid-voltage curves corresponding to the peaks of the alternating grid-voltage cycle gives the maximum and minimum values of the instantaneous plate current and plate voltage. If the characteristic curves are assumed to be uniformly spaced straight lines, the difference between the maximum and minimum values of current and voltage gives a current or voltage which is two times the peak value of the alternating current or voltage on the load. Therefore, the effective value of the alternating current in the load is

$$I = \frac{(I_{\max} - I_{\min})}{2\sqrt{2}} \tag{91}$$

and the effective value of the voltage across the load is

$$E = \frac{(E_{\max} - E_{\min})}{2\sqrt{2}} \tag{92}$$

The product of equations 91 and 92 is the power dissipated in the load

$$P_o = EI = \frac{(E_{\max} - E_{\min})(I_{\max} - I_{\min})}{8} \tag{93}$$

The plate efficiency is a measure of the ability of an amplifier to convert the d-c power supplied to the plate into alternating power. It is defined as the ratio of the alternating-power output (equation 93) to the d-c power input and it is the opposite of rectification efficiency which was defined in section 2-2. The d-c power input is equal to the product of the average plate current I_b and the plate supply voltage E_{bb}.

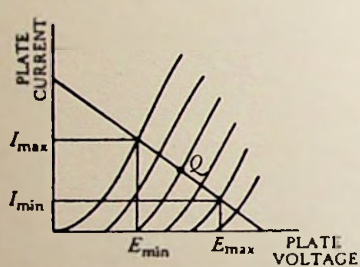


Fig. 57. Construction on the characteristic curves of a class A amplifier to calculate its efficiency. The quiescent point of operation is Q.

In class A operation, the maximum allowable instantaneous grid voltage would make $I_{\min} = 0$, and, since the grid could not be driven positive, $I_{\max}$ would be given by the intersection of the curve for zero grid bias with the load line. On the assumption that the characteristic curve is a straight line passing through the origin, the voltage corresponding to this intersection is

$$E_{\min} = I_{\max} r_p \tag{94}$$

Also the difference between the minimum and maximum voltage is equal to the maximum current multiplied by the load resistance

$$E_{\max} - E_{\min} = I_{\max} R_L \tag{95}$$

A resistance-coupled amplifier is series fed (i.e., the supply voltage is in series with the load); therefore, the supply voltage E_{bb} is equal to $E_{\max}$ or $E_{bb} = I_{\max} R_L + I_{\max} r_p$. The average supply current I_b is one-half the maximum current $I_{\max}$. The substitution of these values in equation 93 yields

$$\eta = \frac{P_o}{E_{bb}I_b} = \frac{I_{\max}R_L \times I_{\max}}{8(I_{\max}R_L + I_{\max}r_p)\left(\frac{I_{\max}}{2}\right)} = \frac{R_L}{4(R_L + r_p)} \tag{96}$$

This shows that the efficiency of a resistance-coupled class A amplifier cannot exceed 25 per cent. If the load resistance is made equal to the plate resistance as it should for maximum power transfer, then the theoretical maximum efficiency is only $12\frac{1}{2}$ per cent.

A better efficiency can be obtained with class A operation if impedance coupling is used. In this case, the plate is shunt fed (i.e., the supply voltage is effectively in parallel with the load). The supply voltage is then equal to the quiescent voltage E_b and the theoretical maximum efficiency becomes 50 per cent. Transformer-coupled amplifiers also have a maximum theoretical efficiency of 50 per cent. The output from both transformer and impedance-coupled amplifiers is a maximum when the apparent load in the plate circuit is two times as large as the plate resistance.

5-10. Class B and C Amplifiers. Efficiencies of higher than 50 per cent require the use of class B or C operation. For example, the theoretical maximum plate efficiency for a class B amplifier is $78\frac{1}{2}$ per cent and the efficiency of class C is even higher. These high efficiencies, however, are accompanied by certain limitations. Class B is the only one of the two which can be used to reproduce the complete input wave and it requires the use of two tubes when operated in this service. One of the tubes reproduces the upper half cycle and the other reproduces the lower half cycle. The tubes are connected together in a "push-pull" circuit which combines their individual outputs into the complete wave form. Class C operation is practically limited to radio-frequency applications.

When the distorted output current wave of either class B or C operation (Fig. 58) is applied to a tuned circuit as a load impedance (Fig. 59), it will produce a sinusoidal output voltage since the impedance of the circuit and therefore the IZ drop is large for the component of current which has a frequency near resonance and is negligibly small for the other frequency components of current present. These amplifiers are frequently used for high-efficiency radio-frequency amplifications because they do produce a single-frequency sinusoidal output. When the resonant circuit is tuned to the second harmonic rather than the fundamental component of the output current, the output voltage is twice the frequency of the input and the amplifier becomes a frequency doubler. Still higher frequency multiplication can be obtained if the resonant circuit is tuned to one of the higher harmonics. A frequency multiplier is very useful in obtaining a stable high-frequency output from a standard low-frequency crystal oscillator.

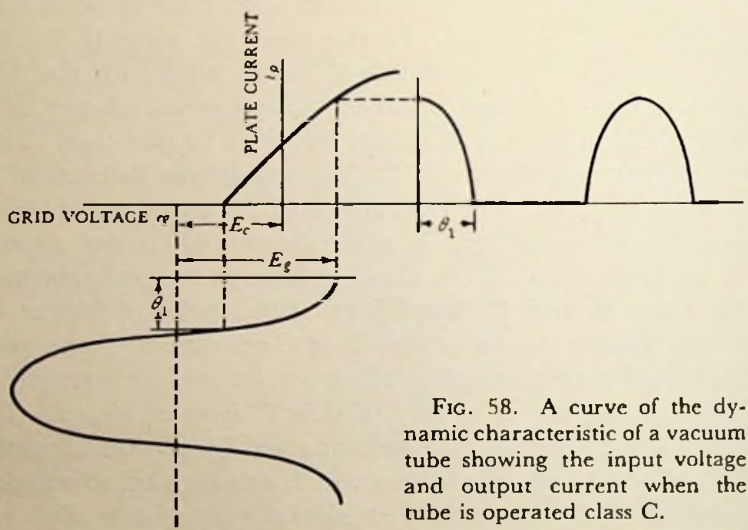


FIG. 58. A curve of the dynamic characteristic of a vacuum tube showing the input voltage and output current when the tube is operated class C.

Everitt has analyzed the performance of class B and C amplifiers on the basis of an approximate expression for the dynamic curve of a tube. The amount of plate current taken by a tube depends on both the instantaneous grid voltage e_g and the instantaneous plate voltage e_p. It is equal to

$$i_p = g_m e_g + g_p e_p \tag{97}$$

where g_m is the mutual conductance and g_p is the plate conductance or reciprocal of the plate resistance. If the resonant circuit is tuned to the frequency of the fundamental current I_1 and has an impedance in the form of a pure resistance R_L, the instantaneous voltage drop across the circuit is $I_1 R_L \cos\theta$ and the instantaneous plate voltage is

$$e_p = E_{bb} - I_1 R_L \cos\theta \tag{98}$$

The instantaneous grid voltage is

$$e_g = -E_c + E_g \cos\theta \tag{99}$$

and, therefore, the plate current becomes

$$\begin{aligned} i_p &= g_m(-E_c + E_g \cos\theta) + g_p(E_{bb} - I_1 R_L \cos\theta) \\ &= (g_m E_g - g_p I_1 R_L)\cos\theta - (g_m E_c - g_p E_{bb}) \\ &= E' \cos\theta - E \end{aligned} \tag{100}$$

where E' and E are defined by the above equation. The cutoff angle θ_1 at which the current stops flowing is equal to

$$\theta_1 = \text{arc} \cos \frac{E}{E'} \tag{101}$$

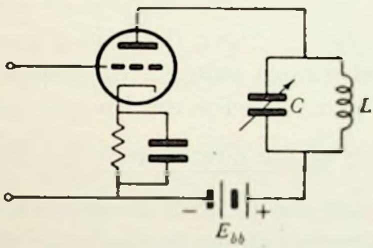


FIG. 59. Class B or C radio-frequency amplifier. The grid is generally driven positive during a portion of the cycle (see Fig. 58) so that a resistance in the grid circuit can be used for biasing, however, the cathode bias shown in the figure or a fixed bias is also generally included in the circuit to prevent the excessive flow of plate current when the input signal is removed.

The Fourier series (section 6–2) of equation 100 shows that the fundamental component of plate current has the value

$$I_1 = \frac{E'}{\pi}[\theta_1 - \sin\theta_1 \cos\theta_1] \tag{102}$$

The substitution for E' gives

$$I_1 = \frac{(g_m E_g - g_p I_1 R_L)}{\pi}(\theta_1 - \sin\theta_1 \cos\theta_1)$$

and the solution of this equation for I_1 gives

$$I_1 = \frac{\mu E_g}{R_L + \left(\dfrac{\pi}{\theta_1 - \sin\theta_1 \cos\theta_1}\right) r_p} \tag{103}$$

since $\frac{1}{g_p} = r_p$ and $\mu = \frac{g_m}{g_p}$. The output voltage appearing across the tuned circuit is equal to this fundamental component of current multiplied by the impedance R_L of the circuit at resonance; therefore,

$$E_o = I_1 R_L = \frac{\mu E_g R_L}{R_L + \left(\frac{\pi}{\theta_1 - \sin\theta_1 \cos\theta_1}\right) r_p} \tag{104}$$

The output voltage of a single-tube class B radio-frequency amplifier has a magnitude which is directly proportional to the amplitude of the voltage applied to the grid. Consequently, it can faithfully reproduce a single-frequency wave even when that wave has a varying amplitude. Since there is a linear relationship between the input and output voltage, this amplifier is also called a linear amplifier. The reason for the linearity can be seen by substituting $\frac{\pi}{2}$ (the value of the cutoff angle for class B operation) for θ_1 in equation 104. All the terms of the equation become constants except E_g and the output voltage is seen to be directly proportional to it.

On the other hand, the class C amplifier will only give faithful reproduction of constant-amplitude single-frequency waves. In spite of this limitation, however, the class C amplifier is nevertheless very important in other ways. When the magnitude of the input voltage is constant, for example, the magnitude of the alternating output voltage produced across the tuned circuit is directly proportional to the voltage supplied to the plate E_{bb}. Hence, the amplitude of the output can be modulated in a linear manner by simply changing the supply voltage (see Chap. IX). The linear relationship between the sinusoidal output voltage and the plate-supply voltage can be verified by substituting various assumed values of the conduction angle θ_1 in equation 104 and solving for the corresponding output voltage E_o and the supply voltage E_{bb}. The curve between these two voltages is practically a straight line as indicated in the above discussion.

Both class B and C amplifiers are useful as frequency multipliers.

Problems

5–1. Calculate the upper and lower frequency limits for the amplifier in the example of section 5–2 if the wiring capacity is 10 $\mu\mu$f and the next tube in line is another 6J5 which is connected to a pure resistance load Z_L = 50,000 ohms.

5–2. A cathode resistor is used to obtain a $-$ 4 volts bias on an amplifier using a 6J5 tube. The load resistance R_L is 20,000 ohms and the supply voltage E_{bb} is 200 volts.

(*a*) Calculate the resistance of the cathode resistor.

(*b*) Calculate the mid-frequency gain if the cathode resistor is suitably by-passed.

(*c*) Calculate the gain if the by-pass condenser is removed.

5–3. Compute the magnitude and phase angle of gain as a function of frequency for the amplifier in the example of section 5–4 both with shunt compensation and without, and plot the results on semilog paper. Assume that R_L = 1300 ohms in each case and that C = 0.01 μf and R_g = 250,000 ohms.

5–4. A circuit using a 6AC7 tube has C_p = 10 $\mu\mu$f and C_g = 20 $\mu\mu$f. Compute the series peaking coil, load resistor, and mid-frequency gain if the upper frequency limit is to be 4.0 mc.

5–5. Design a two-terminal filter-coupling network (a half section of constant K followed by an m-derived half section for the purpose of matching to the load resistance) for an amplifier having the specifications of problem 5–4. Compute the mid-frequency gain.

5–6. The amplifier of problem 5–3 is compensated for low-frequency response by the circuit shown in Fig. 55.

(*a*) Specify the magnitude of the components required to give the amplifier a satisfactory response down to 50 cps.

(*b*) Assume that the normal plate current drawn by the tube is 10 ma and calculate the additional amount of plate supply voltage which would be required when the compensating circuit is added.

(*c*) Set up an expression for the low-frequency gain of a compensated amplifier and show that it reduces to the expression for the mid-frequency gain when the conditions specified in section 5–7 are realized.

5–7. Compute and plot the gain vs. frequency curve of the radio-frequency amplifier in Fig. 56 if the tube used is a 6J5. The primary of the radio-frequency transformer has an inductance of 25 μh with a Q of 200, the secondary has an inductance of 100 μh with a Q of 150, and the mutual inductance is 30 μh. The capacity is adjusted to resonance at 1 mc.

5-8. Specify the self and mutual inductance of a doubly tuned radio-frequency transformer if it is to be operated at 465 kc and have a band width (between peaks) of 10 kc. The tuning condensers are each 100 $\mu\mu$f. (*Note:* Assume that each peak is 5 kc from resonance and average the results.) What is the Q of the coils if the ratio between the gain at the peaks and the gain at the resonant frequency is 1.1?

5-9. Compute and plot a curve of alternating-voltage output from the tuned circuit of a class C amplifier as a function of the supply voltage E_{bb}. Assume that the tube is a 6J5 operated at $-$ 45 volts bias and connected to a tuned circuit consisting of a 0.1 μf condenser in parallel with a 251 mh inductance which has an internal resistance of 150 ohms. The signal applied to the grid has a peak value of 45 volts and a frequency of 1000 cps (the audio frequency is assumed for convenience in checking experimentally).

Experiment 9

Object:

The object of this experiment is to study the effects of negative feedback on the performance of an amplifier.

Preliminary:

1. Calculate and plot a curve of mid-frequency gain as a function of the load resistance (up to 250,000 ohms) for an amplifier employing a 6J7 tube operated without feedback at its normal rating. On the same coordinate axis plot the gain curve of the amplifier with 10 per cent negative feedback.

2. Plot a curve of gain as a function of percentage feedback when the load resistance is 50,000 ohms.

Performance:

1. Assemble a circuit like Fig. 50 and make R_1 and the load resistor Z_L variable so that various percentages of feedback and load can be obtained. Use a 6J7 tube rather than the triode indicated in the figure.

2. Take data to verify the calculations of part 1 in the preliminary. Plot the experimental results in a contrasting color on the same coordinate axes as the computed curves.

3. Obtain an experimental curve to verify the calculations of part 2 in the preliminary. Plot the results in a contrasting color on the same coordinate axes as the calculated curve.

4. Obtain data to plot a curve of gain vs. supply voltage both without feedback and with 10 per cent negative feedback. Explain the results.

5. Insert the output from a 6.3-volt 60-cycle transformer in series with the load resistor of the amplifier to simulate "hum." Without feedback,

observe the magnitude of the alternating voltage across the load by means of a cathode-ray oscillograph. Insert the feedback circuit and adjust it until the hum has been reduced 50 per cent. Determine the loss in gain from the data of part 3 and compare with the loss of hum.

Experiment 10

Object:

The object of this experiment is to study the performance of a wide-band amplifier.

Preliminary:

Solve problems 5–3 and 5–6.

Performance:

1. Measure the gain as a function of frequency on an uncompensated amplifier similar to the one in the example of section 5–4. As suggested in the text, isolate the measuring vacuum-tube voltmeter from the stage by making the measurement across a small load resistor in the plate circuit of another tube which is fed by the amplifier. The gain of the second tube can be measured easily at mid-frequency and is valid for all frequencies; hence, the effect of the second amplifier can be accounted for in calculations of the net gain of the stage. Plot the net gain as a function of frequency on semilog paper.
2. Add a shunt peaking coil and repeat the above measurement.
3. Measure the approximate phase shift by means of a Lissajous figure on a cathode-ray oscillograph for a few frequencies.
4. Compare the experimental results with the calculations of problem 5–3.
5. Connect the low-frequency compensating circuit and measure the low-frequency gain as a function of frequency. Plot results on curve sheet of part 1.

Experiment 11

Object:

The object of this experiment is to measure the gain of a class A radio-frequency amplifier.

Preliminary:

Solve problem 5–7.

Performance:

1. Set up a circuit similar to the one of problem 5–7. Apply a 1-volt 1-mc output from a standard signal generator to the grid of the tube.

Connect a vacuum-tube voltmeter across the tuned circuit and adjust the condenser until a maximum deflection is obtained.

2. Keep the condenser fixed and measure the output voltage at 1.0, 1.005, 1.01, 0.995, and 0.99 mc. Plot the gain as a function of frequency and compare the results with those of the problem.

3. Remove the tuned circuit and substitute for it a variable coupling i-f transformer. Connect the output of the transformer to a diode detector. Apply a frequency-modulated signal to the grid of the tube and synchronize a cathode-ray oscillograph connected across the diode load with the rate of frequency modulation. Observe the wave form produced on the oscillograph (which is a plot of gain vs. frequency) for various values of transformer coupling. Sketch results.

Experiment 12

Object:

The object of this experiment is to study the operating characteristics of a class B and class C amplifier.

Preliminary:

Solve problem 5–9.

Performance:

1. Assemble a circuit similar to the one described in problem 5–9. Use a Brooks inductometer for the inductance of the tuned circuit and adjust it to resonate with the condenser at 1000 cps. Detect resonance by the maximum reading of a voltmeter across the tuned circuit or by a drop in the plate current of the circuit. Observe the wave form of the plate current by means of a cathode-ray oscillograph connected across a small resistance in series with the plate lead.

2. With a constant input voltage having a maximum value equal to the grid bias (− 45 volts), vary the supply voltage E_{bb} and measure the alternating output voltage across the tuned circuit with a vacuum-tube voltmeter. Plot a curve of output voltage vs. supply voltage in a contrasting color on the same coordinate axes as the solution to problem 5–9. Select a point at about the middle of the linear portion of the experimental curve and adjust the supply voltage to that value. Insert a 60-cycle voltage having a value which will keep the maximum swing on the linear portion of the curve in series with the supply and observe output voltage across the tuned circuit with a cathode-ray oscillograph.

3. Remove the alternating voltage from the plate and change the frequency of the input voltage to 500 cps. Determine the frequency of the output by Lissajous figures.

4. Adjust the grid bias for class B operation and with the supply voltage constant at 250 volts measure the alternating output voltage for various values of input voltage. Plot a curve of output voltage vs. applied voltage. Explain results.

References

EASTMAN, A. V., *Fundamentals of Vacuum Tubes*, 2nd ed. New York: McGraw-Hill Book Company, Inc., 1941, pp. 249–402.

EVERITT, W. L., *Communication Engineering*, 2nd ed. New York: McGraw-Hill Book Company, Inc., 1937, pp. 179–214, 526–596.

MILLMAN, J., and S. SEELY, *Electronics*. New York: McGraw-Hill Book Company, Inc., 1941, pp. 573–670.

SEELEY, S. W., and C. N. KIMBALL, "Analysis and Design of Video Amplifiers," *R.C.A. Review*, October 1937, pp. 171–183.

TERMAN, F. E., R. R. BUSS, W. R. HEWLETT, and F. C. CAHILL, "Some Applications of Negative Feedback with Particular Reference to Laboratory Equipment," *Proc. I.R.E.*, October 1939, pp. 649–655.

WAGNER, H. M., and W. R. FERRIS, "The Orbital-beam Secondary-electron Multiplier for Ultra-high-frequency Amplification," *Proc. I.R.E.*, November 1941, pp. 598–602.

WHEELER, H. A., "Wide-band Amplifiers for Television," *Proc. I.R.E.*, July 1939, pp. 429–438.

CHAPTER VI

SQUARE-WAVE TESTING AND TRANSIENT RESPONSE

6–1. General. The performance of an amplifier has thus far been considered in terms of the phase and amplitude response to steady-state sinusoidal waves of various frequencies, but the input to an amplifier is often of a transient nature and the performance under these conditions is not obvious from the steady-state characteristic. There is a definite relationship between transient and steady-state performance, however, and when one is known, the other can be calculated directly. For example, two observations of the transient response of an amplifier are often used to evaluate its steady-state performance because these two observations are usually sufficient to tell the whole story. In contrast, the measurement of the steady-state performance directly requires a large number of observations. For this reason, transient testing is much more rapid than sinusoidal testing and is, therefore, frequently preferred.

The fundamental transient wave is the Heaviside unit function. It has a magnitude of zero up to a certain time, then rises abruptly to a value of one and remains at that value thereafter. The wave is in the form of a step. It can be produced by suddenly closing a switch on a battery connected to a circuit. A square wave is a close approximation to the unit function if the period of the square wave is long compared with the transient effect under observation. Consequently, the square-wave testing technique is the primary method of determining transient response.

The effect of nonsinusoidal waves or transients can be studied in terms of pure sine waves and the sinusoidal response characteristic by means of Fourier analysis. The impressed wave or pulse is resolved into a series of sinusoidal waves. The response of the amplifier to each of these different frequency components is known; hence, the resultant magnitude as well as the phase delay (the angle between the vectors representing the impressed and the output wave) of each harmonic passing through it is determined. The superposition of these output components gives the resultant output wave. A comparison between it and the assumed input reveals the amount of distortion produced by the amplifier and provides a criterion for evaluating its performance. Such calculations also indicate modifications in the coupling circuits which will yield still better performance.

6–2. Fourier Series. If a wave is a function of time $F(t)$ and has a period

$$T = \frac{1}{f} = \frac{2\pi}{\omega_0} \tag{105}$$

it can be expressed as a series of pure sine waves thus:

$$F(t) = \frac{A_0}{2} + \sum_{n=1}^{\infty} (A_n \cos n\omega_0 t + B_n \sin n\omega_0 t) \tag{106}$$

where n is an integer which varies from one to infinity. The coefficients of the above equation are determined in the following manner:

$$A_n = \frac{\omega_0}{\pi} \int_{-\frac{\pi}{\omega_0}}^{\frac{\pi}{\omega_0}} F(t) \cos n\omega_0 t \, dt \tag{107}$$

$$B_n = \frac{\omega_0}{\pi} \int_{-\frac{\pi}{\omega_0}}^{\frac{\pi}{\omega_0}} F(t) \sin n\omega_0 t \, dt \tag{108}$$

The A coefficient can be verified by multiplying both sides of equation 106 by $\cos n\omega_0 t \, dt$ and integrating over a cycle. All the terms become zero except the ones which reduce to equation 107. The equation for the B coefficient can be verified in the same manner.

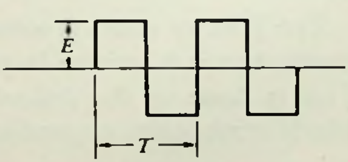


FIG. 60. Symmetrical square wave.

Example: Find the Fourier series for the symmetrical square wave of Fig. 60. The amplitude is E and the frequency is $f = \frac{\omega_0}{2\pi} = \frac{1}{T}$.

The magnitude of the wave $F(t)$ is $-E$ from $-\frac{1}{2f}$ to zero and E from zero to $\frac{1}{2f}$; hence, equation 107 becomes

$$A_n = 2f\int_{-\frac{1}{2f}}^{0}(-E)\cos n\,2\pi ft\,dt + 2f\int_{0}^{\frac{1}{2f}}(E)\cos n\,2\pi ft\,dt$$
$$= \frac{-2fE}{2\pi nf}\sin 2\pi nft\Big]_{-\frac{1}{2f}}^{0} + \frac{2fE}{2\pi nf}\sin 2\pi nft\Big]_{0}^{\frac{1}{2f}}$$
$$= \frac{E}{\pi n}\sin(-\pi n) + \frac{E}{\pi n}\sin(\pi n) = 0$$

for all values of n from one to infinity. Also from equation 107 when n is zero

$$A_0 = 2f\int_{-\frac{1}{2f}}^{0}(-E)\,dt + 2f\int_{0}^{\frac{1}{2f}}(E)\,dt = -E + E = 0$$

On the other hand, equation 108 becomes

$$B_n = 2f\int_{-\frac{1}{2f}}^{0}(-E)\sin 2\pi nf\,dt + 2f\int_{0}^{\frac{1}{2f}}(E)\sin 2\pi nft\,dt$$
$$= \frac{2fE}{2\pi nf}\cos 2\pi nft\Big]_{-\frac{1}{2f}}^{0} - \frac{2fE}{2\pi nf}\cos 2\pi nft\Big]_{0}^{\frac{1}{2f}}$$
$$= \frac{E}{n\pi}[1 \pm 1] - \frac{E}{n\pi}[\mp 1 - 1] = \begin{matrix} 0 \text{ for } n \text{ even} \\ \dfrac{4E}{n\pi} \text{ for } n \text{ odd} \end{matrix}$$

and the series is

$$F(t) = \frac{4E}{\pi}\left[\sin 2\pi ft + \frac{1}{3}\sin 2\pi(3f)t + \frac{1}{5}\sin 2\pi(5f)t + \cdots + \frac{1}{n}\sin 2\pi(nf)t\right]$$

The Fourier series is somewhat more convenient to use in some cases when it is written in polar or exponential form ($E\underline{/\theta} = E\epsilon^{j\theta}$). This is done in the following way: From a comparison of the Maclaurin's series expansions of $\epsilon^{j\theta}$, $\cos\theta$, and $\sin\theta$, we know that

$$\epsilon^{j\theta} = \cos\theta + j\sin\theta \qquad \epsilon^{-j\theta} = \cos\theta - j\sin\theta$$

hence

$$\cos\theta = \tfrac{1}{2}(\epsilon^{j\theta} + \epsilon^{-j\theta}) \qquad j\sin\theta = \tfrac{1}{2}(\epsilon^{j\theta} - \epsilon^{-j\theta}) \qquad (109)$$

Substituting these expressions in the equations for the coefficients of the Fourier series (equations 107 and 108), we have

$$A_n = \frac{\omega_0}{\pi}\int_{-\frac{\pi}{\omega_0}}^{\frac{\pi}{\omega_0}} F(t)\,\tfrac{1}{2}(\epsilon^{jn\omega_0 t} + \epsilon^{-jn\omega_0 t})\,dt$$

$$B_n = -j\frac{\omega_0}{\pi}\int_{-\frac{\pi}{\omega_0}}^{\frac{\pi}{\omega_0}} F(t)\,\tfrac{1}{2}(\epsilon^{jn\omega_0 t} - \epsilon^{-jn\omega_0 t})\,dt$$

The series (equation 106) becomes

$$F(t) = \frac{A_0}{2} + \sum_{n=1}^{\infty}\left[A_n\,\tfrac{1}{2}(\epsilon^{jn\omega_0 t} + \epsilon^{-jn\omega_0 t}) - jB_n\,\tfrac{1}{2}(\epsilon^{jn\omega_0 t} - \epsilon^{-jn\omega_0 t})\right]$$

or

$$F(t) = \frac{\omega_0}{2\pi}\int_{-\frac{\pi}{\omega_0}}^{\frac{\pi}{\omega_0}} F(t)\,dt$$

$$+ \sum_{n=1}^{\infty}\left[\frac{\omega_0}{\pi}\int_{-\frac{\pi}{\omega_0}}^{\frac{\pi}{\omega_0}} F(t)\,\tfrac{1}{2}(\epsilon^{jn\omega_0 t} + \epsilon^{-jn\omega_0 t})\,dt\,\tfrac{1}{2}(\epsilon^{jn\omega_0 t} + \epsilon^{-jn\omega_0 t})\right.$$

$$\left.- \frac{\omega_0}{\pi}\int_{-\frac{\pi}{\omega_0}}^{\frac{\pi}{\omega_0}} F(t)\,\tfrac{1}{2}(\epsilon^{jn\omega_0 t} - \epsilon^{-jn\omega_0 t})\,dt\,\tfrac{1}{2}(\epsilon^{jn\omega_0 t} - \epsilon^{-jn\omega_0 t})\right]$$

If the above equation is expanded and the terms are collected, the result is

$$F(t) = \frac{\omega_0}{2\pi}\int_{-\frac{\pi}{\omega_0}}^{\frac{\pi}{\omega_0}} F(t)\,dt + \sum_{n=1}^{\infty}\left[\frac{\omega_0}{2\pi}\int_{-\frac{\pi}{\omega_0}}^{\frac{\pi}{\omega_0}} F(t)\epsilon^{-jn\omega_0 t}\,dt\right]\epsilon^{jn\omega_0 t}$$

$$+ \sum_{n=1}^{\infty}\left[\frac{\omega_0}{2\pi}\int_{-\frac{\pi}{\omega_0}}^{\frac{\pi}{\omega_0}} F(t)\epsilon^{jn\omega_0 t}\,dt\right]\epsilon^{-jn\omega_0 t}$$

This expression simplifies to a very compact form of the Fourier series.

$$F(t) = \sum_{n=-\infty}^{\infty} a_n\epsilon^{jn\omega_0 t} \tag{110}$$

where the coefficients a_n are determined by the equation

$$a_n = \frac{\omega_0}{2\pi}\int_{-\frac{\pi}{\omega_0}}^{\frac{\pi}{\omega_0}} F(t)\epsilon^{-jn\omega_0 t}\,dt \tag{111}$$

The integer n in equation 110 has negative as well as positive values. The "negative frequencies" implied by the negative n's simply indicate an opposite rotational sense of the vector.

6–3. Square-wave Testing. The technique of square-wave testing is based on the fact that a square wave really consists of a multitude of sine waves of different frequencies which are related to one another in a specific manner (see the example of the preceding section). An observation of the output wave on a cathode-ray oscillograph when a square wave is applied to the input of a circuit is then a collective observation on the individual effects of the components making up the square wave. The distortion of the output gives data on the circuit response at each of the component frequencies and gives those data simultaneously. For this reason it is often possible to tell everything about the properties of a circuit from a single square-wave test.

When the low-frequency response of a circuit is studied, the square-wave cycle used must be long enough so that the low-frequency deficiencies will be manifest. A typical output wave

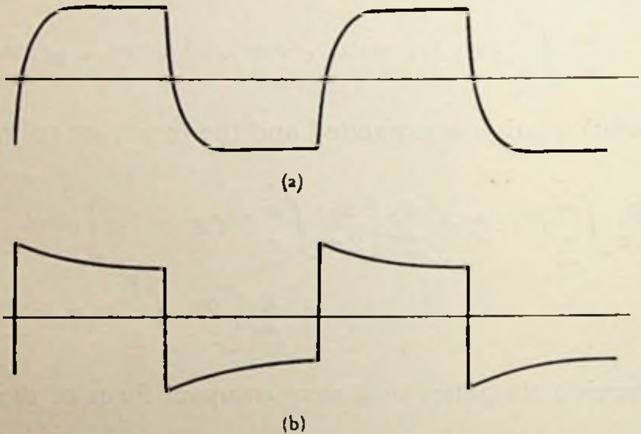


FIG. 61. Output wave form observed on an ordinary resistance-coupled amplifier when (*a*) a high-frequency square wave is applied to the input and (*b*) when a low-frequency square wave is applied to the input.

from an ordinary resistance-coupled amplifier when tested with a low-frequency square wave is shown in Fig. 61*b*. It takes the shape indicated due to the nature of the low-frequency equivalent circuit of the amplifier which is essentially the series combination of the coupling condenser C and the grid resistor R_g. When a square wave is applied to this combination, the accumulation of charge on the condenser gradually reduces the voltage which appears at the output across the grid resistor. The rate at which this voltage falls is a measure of the low-frequency response. Specifications of an amplifier are often written in terms of this square-wave test because the test is very simple and the results are easily interpreted in terms of sinusoidal response.

When the high-frequency performance is being studied, it is most convenient to use a high-frequency square wave for testing so that the front of the output wave is spread out over a greater distance on the cathode-ray screen, and the high-frequency effects are more readily discernible. An ordinary resistance-coupled amplifier has an output wave such as the one illustrated in Fig. 61*a* when it is tested in this manner. The time required for the wave to rise to its final value is proportional to the time required for the tube and wiring capacity of the circuit C_T to be charged through the load resistor R_L. This time constant is an index of upper frequency limit of the amplifier and can be evaluated directly from the square-wave observation.

If the output wave has a superimposed oscillation of the type shown in Fig. 62*a*, there is a peak in the response curve of the circuit at the frequency of oscillation, or, if the oscillation is less pronounced, the circuit has a rapid cutoff at the frequency of oscillation instead. A more complete quantitative interpretation of an output wave, however, requires a detailed analysis such as the one described below or considerable experience with square-wave testing.

Bedford and Fredendall have recently described a method for computing the steady-state sinusoidal response of an amplifier or a network from a cathode-ray oscillograph observation of the output wave when a square wave is applied to the input. A square-wave input is a close approximation of a unit function or step wave; hence, it contains sinusoidal components of almost every conceivable frequency (see section 6–4). Due to the effect of the amplifier or network, these component frequencies are so changed in magnitude and shifted in phase angle that they produce a distorted output wave shape which may take a form like that of Fig. 62*a*. The analysis of such an output wave simply consists of finding the change in magnitude and phase of each component which is necessary to produce the shape indicated.

For the purposes of analysis, the wave is approximated by a series of steps which are displaced from each other by equal periods of time (Fig. 62*b*). Since the input was essentially a step wave, the magnitude of a particular frequency component in one of the output steps is related to the magnitude of that same frequency component of the input by the ratio between the heights of the two step waves. The actual magnitude of the particular frequency component being considered is of no consequence. The first step of the output "staircase" is assumed to start at the same instant

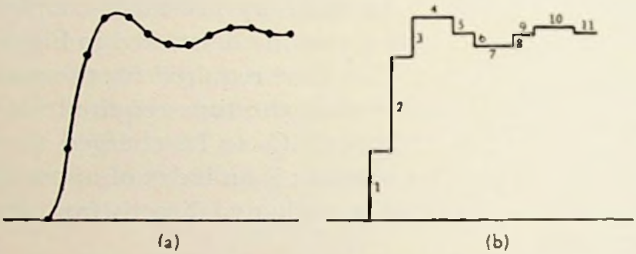


FIG. 62. (*a*) Typical output wave from a circuit with a sharp cutoff and (*b*) a stepped wave which has approximately the same characteristics.

as does the input step wave; therefore, the phase angle between the two components is the angle which corresponds to the time delay between the first step and the step under consideration.

Since the output wave consists of a series of these steps, the net magnitude and phase angle of a particular frequency component in the output relative to that same component in the input, i.e., the response to the amplifier at that particular frequency, is given by the vector sum of the component ratios under consideration from each of the steps which together make up the output. The sinusoidal response of the amplifier at another frequency is found by carrying through the same calculations using angles which correspond to the new frequency. Naturally, the accuracy of the method depends on the spacing between the steps which are assumed to make up the output. The method is best suited for calculating the high-frequency response.

For convenience in securing the necessary data, time intervals are marked on the cathode-ray trace of the output wave by modulating the cathode-ray beam with a standard frequency source. This is done by connecting an oscillator between the grid of the cathode-ray tube and ground. The oscillator is isolated from the high voltage on the grid by means of a coupling condenser. The timing "dots" which make up the cathode-ray trace are prevented from moving along the trace by synchronizing the square-wave input with the timing wave. Data for the calculation can then be readily taken from this image.

Example: When a square wave is applied to a certain network, the output as observed on a cathode-ray oscillograph is as shown in Fig. 62*a*. A 40,000-cps sine wave was used for timing and dotted the output wave at the points indicated. Find the amplitude and phase angle of the response of the network at 8000 cps assuming that the steady-state square-wave output has the same amplitude as the input.

The phase angle between the 8000-cps components of the successive steps in the equivalent output wave shown in Fig. 60*b* is

$$\theta = \frac{8000}{40{,}000} \times 360^\circ = 72^\circ$$

The amplitude of the steps is equal to the difference in amplitude between the successive points on the output wave as indicated in the figure. Tabulating calculations:

No. of step	*Relative amplitude of step*	*Phase angle component makes with same frequency component of first step*	*Complex form of component in the step*
1	0.38	0°	$0.38 + j\,0$
2	0.50	72°	$0.154 + j\,0.475$
3	0.21	144°	$-0.170 + j\,0.124$
4	0.00	216°	$0.00 + j\,0.00$
5	− 0.09	288°	$-0.028 + j\,0.085$
6	− 0.06	360°	$-0.06 + j\,0.00$
7	0.00	432°	$0.00 + j\,0.00$
8	0.05	504°	$-0.041 + j\,0.029$
9	0.04	576°	$-0.032 - j\,0.023$
10	0.00	648°	$0.00 + j\,0.00$
11	− 0.03	720°	$-0.03 + j\,0.00$
		Sum	$0.173 + j\,0.690$

The polar form of the sum is $0.71\underline{/76^\circ}$; therefore, the ratio of the output to input voltage at 8000 cps would be 0.71 and the phase angle would be 76°. In an analysis of this sort, the steps are taken up to the point where the wave reaches a steady-state value.

The calculations involved in determining the steady-state sinusoidal response from a square-wave test are lengthy, but they can be simplified by the use of polar charts for making the vector additions (see first reference, page 131). In spite of the length of the calculations, however, there are occasions when it is more convenient to determine the sinusoidal response from the square-wave test than to measure it directly. Furthermore, experience with square-wave testing enables an operator to get a good evaluation of an amplifier's performance without the necessity of making the calculations.

The transient response of an amplifier can be calculated from the sinusoidal response by an application of the above principles. Charts have also been prepared to facilitate making these calculations.

6–4. Heaviside Unit Function. The Fourier integral is used to represent a nonrecurrent wave in much the same manner that a Fourier series is used to represent a cyclic wave. As a matter of fact, the Fourier integral is derived from the Fourier series by making the period infinitely long. This spreads a single cycle over the entire time domain, and makes the analysis of transients possible.

Since the frequency of the harmonics of a Fourier series is $n\omega_0$, the difference in frequency $\Delta n\omega_0$ between the successive harmonics is equal to the frequency ω_0 of the fundamental. The series can, therefore, be written as

$$F(t) = \sum_{n=-\infty}^{\infty} \frac{a_n}{\omega_0} \epsilon^{jn\omega_0 t} \, \Delta n\omega_0 \tag{112}$$

by multiplying and dividing equation 110 by ω_0. As the period of a wave is made longer and longer, the frequency of its fundamental gets lower and lower. The difference in frequency between the adjacent harmonics approaches zero or, in other words, it becomes differential in magnitude. Hence, *a continuous spectrum of frequencies is required to represent a transient wave, whereas a series of discrete harmonics is sufficient for the representation of a cyclic wave.* In the limit, the Fourier series of equation 112 becomes the definition of an integral with the frequency $\omega = n\omega_0$ as the variable, and it can be written as

$$F(t) = \int_{-\infty}^{\infty} \frac{a_n}{\omega_0} \epsilon^{j\omega t}\, d\omega = \int_{-\infty}^{\infty} g(\omega)\epsilon^{j\omega t}\, d\omega \tag{113}$$

and the amplitude-distribution function $g(\omega)$ is

$$g(\omega) = \frac{a_n}{\omega_0} = \frac{1}{2\pi}\int_{-\frac{\pi}{\omega_0}}^{\frac{\pi}{\omega_0}} F(t)\epsilon^{-j\omega t}\, dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(t)\epsilon^{-j\omega t}\, dt \tag{114}$$

This analysis, as with the Fourier series, is only valid when the above integrals are finite. From equation 114, it is evident that the individual amplitudes of the transient's harmonics a_n must approach zero together with ω_0 so that $g(\omega)$ will remain finite. Hence, a transient resolves into an infinite spectrum of harmonics whose amplitudes are infinitesimally small. This result is to be expected since a transient has an infinite number of components in any finite portion of its frequency spectrum. The infinitesimal magnitude of these components, however, does not invalidate the usefulness of the Fourier integral concept. The function $g(\omega)$ gives the relative magnitude of the different frequency components thus indicating their relative importance. The band-pass requirements of an appropriate amplifier for a given transient wave are intimated by its amplitude-distribution function $g(\omega)$.

As an illustration of these concepts, consider the harmonic composition of the Heaviside unit function which is zero up to time $t = 0$ and unity thereafter. Unfortunately, this transient wave does not fulfill the integrability condition of Fourier analysis; that is $\int_{-\infty}^{\infty} | F(t) |\, dt$ is not finite, but this difficulty is overcome by using a new function $F(t) = \epsilon^{-at}$ and letting its exponent a approach zero. The new function is then essentially unit function for the region $0 < t < \infty$. Its amplitude distribution curve is given by equation 114

$$g(\omega) = \frac{1}{2\pi}\int_0^{\infty} \epsilon^{-at}\epsilon^{-j\omega t}\, dt = \frac{1}{2\pi}\left(\frac{-1}{a + j\omega}\right)\epsilon^{-at}\epsilon^{-j\omega t}\Big]_0^{\infty} = \frac{1}{2(a + j\omega)}$$

In the limit

$$\lim_{a \to 0} g(\omega) = \frac{1}{2\pi j\omega} \tag{115}$$

This is the curve of an equilateral hyperbola. In other words, a unit function consists of a spectrum of harmonics whose magnitudes vary inversely as their frequencies. The magnitudes of the components in a square wave vary in exactly this same manner (see the example of section 6–2); therefore, a square wave of sufficiently long period is a good approximation of the Heaviside unit function.

A square-topped pulse can be analyzed in a similar manner. Assume that it has unit height, a duration of t_1 seconds, and is symmetrical with respect to the y axis. The amplitude distribution is then given by equation 114 with $F(t) = 1$ for $-\frac{t_1}{2} < t < \frac{t_1}{2}$ and $F(t) = 0$ for $t < -\frac{t_1}{2}$ or $t > \frac{t_1}{2}$

$$g(\omega) = \frac{1}{2\pi}\int_{-\frac{t_1}{2}}^{\frac{t_1}{2}} \epsilon^{-j\omega t}\,dt = \frac{1}{2\pi}\left(\frac{1}{-j\omega}\right)\epsilon^{-j\omega t}\Big]_{-\frac{t_1}{2}}^{\frac{t_1}{2}}$$

$$= \frac{1}{2\pi j\omega}\left(\epsilon^{+j\omega\frac{t_1}{2}} - \epsilon^{-j\omega\frac{t_1}{2}}\right) = \frac{1}{\pi\omega}\sin\omega\frac{t_1}{2}$$

or

$$g(\omega) = \frac{t_1}{2\pi}\,\frac{\sin\omega\frac{t_1}{2}}{\omega\frac{t_1}{2}} \tag{116}$$

The curve of this amplitude-distribution function is shown in Fig. 63.

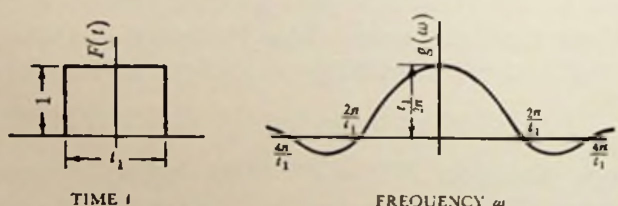


Fig. 63. Square-topped pulse of unit height and its corresponding amplitude distribution function.

A square-topped pulse of this sort can also be represented by a positive unit function followed t_1 seconds later by a negative unit function. The phase relationship between the components of these two unit functions is such that when they are added together, the result is also equation 116.

6-5. Unit Impulse. Although the unit function is the fundamental wave shape of transient analysis, a more useful wave for calculations is its derivative. This is called unit impulse because it is zero for all values of time except $t = 0$. There, the value is infinite. At first sight, it may seem that such a queer function would be hard to handle; however, it is simply a limiting case of the square-topped pulse described in the preceding section.

The current flowing into a condenser is equal to the capacity multiplied by the derivative of the applied voltage and if the condenser has a capacity of 1 farad, the current flowing into it is *equal* to the derivative of the applied voltage. From this analogy, the relationship between a unit impulse and a unit function can be deduced. If the voltage being applied to the condenser rises at

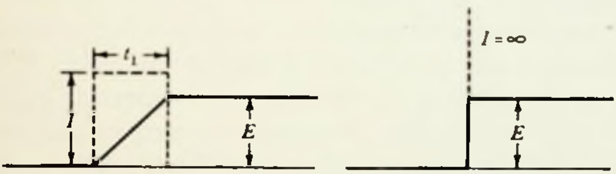


FIG. 64. Oscillograms illustrating the relationship between unit impulse and a square-topped wave.

a constant rate instead of abruptly and reaches its final value in t_1 seconds, the current flowing into the condenser rises abruptly to a value $I = C\dfrac{E}{t_1}$ and remains there for t_1 seconds; i.e., until the voltage reaches its final value (Fig. 64). The charge on the condenser would be the area under this square-topped current pulse. If the voltage rises more rapidly, the current is again a square pulse but of shorter duration and greater height so that the area under the pulse, i.e., the charge, will remain the same because it depends only on the end value of the voltage. In the limit this current pulse has an infinitesimal duration and an infinite height. It corresponds to the unit impulse.

The amplitude-distribution function of the unit impulse is equal to the amplitude-distribution function of a square-topped pulse when its duration t_1 approaches zero and its height approaches infinity. Multiplying equation 116 by I so that it represents a current pulse of this magnitude and taking the limit, we have

$$g(\omega) = \lim_{t_1 \to 0} \frac{It_1}{2\pi} \frac{\sin \omega \frac{t_1}{2}}{\omega \frac{t_1}{2}} = \frac{q}{2\pi}$$

Since a *unit* impulse of current flows when a unit function of voltage is applied to a 1-farad condenser, the amplitude-distribution function for a *unit impulse* is

$$g(\omega) = \frac{1}{2\pi} \tag{117}$$

because the charge q would be 1 coulomb. This particular amplitude-distribution function is a constant for all values of frequency; therefore, it is very useful in making calculations.

6–6. Response of a Network. An analysis of the response of an ideal low-pass filter to the application of a unit function voltage will serve to illustrate the method of transient analysis. If the cutoff frequency of the filter is ω_c, its frequency response curve can be drawn as shown in Fig. 65*a*. The negative frequencies are included to take care of the negative frequencies which occur in the Fourier integral (equation 113) and have the same physical meaning as they do with the Fourier series of equation 110. The amplitudes of the output harmonics will be equal to the input amplitudes multiplied by the response characteristic of the filter. In this case, all of the harmonics up to ω_c will be passed without attenuation and all those above ω_c will be zero in the output. Also, each harmonic would suffer a certain phase delay due to the time required for the wave to go through the filter, but we will neglect this factor here for simplicity. The output amplitude-distribution function when substituted in the Fourier integral (equation 113) gives the equation for the output wave. The mathematics are considerably simplified if the derivative of unit function, i.e., unit impulse, is used for the applied wave and later integrated for the answer. Multiplying the amplitude-distribution function of the input unit impulse by the filter characteristic, there results

$$
\begin{aligned}
g(\omega) &= \frac{1}{2\pi} && -\omega_c < \omega < \omega_c \\
g(\omega) &= 0 && \omega < -\omega_c \\
& && \omega > \omega_c
\end{aligned}
$$

which is the output-wave amplitude-distribution characteristic and has the same form as the filter characteristic (Fig. 65*a*). Substituting this amplitude-distribution function into the Fourier integral (equation 113), the output wave for an applied unit impulse is found to be

$$F(t) = \int_{-\omega_c}^{\omega_c} \frac{1}{2\pi} \epsilon^{j\omega t}\, d\omega = \frac{1}{2\pi j t} \epsilon^{j\omega t}\Big]_{-\omega_c}^{\omega_c}$$

$$F(t) = \frac{1}{\pi t} \sin \omega_c t = \frac{\omega_c}{\pi} \frac{\sin \omega_c t}{\omega_c t} \tag{118}$$

This wave is shown in Fig. 65*b*.

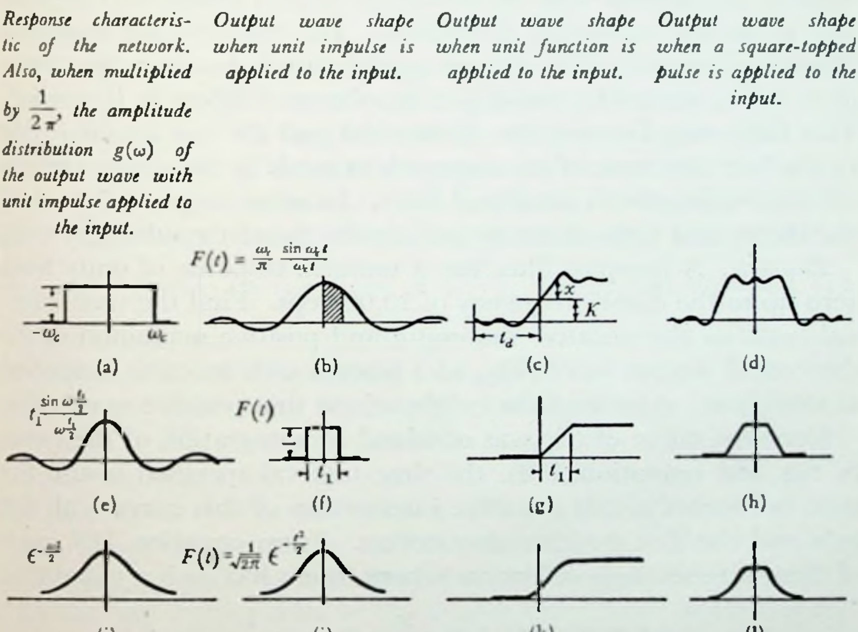


FIG. 65. Theoretical output wave forms for various network characteristics and applied input waves. The ordinate x on curve (c) is equal to the shaded area under curve (b) where the time from the origin is the same in each case. The anticipatory transient shown on curve (c) does not occur in an actual output (Fig. 62*a*).

When a unit function voltage is applied to the filter, the output voltage is the above equation integrated; i.e., an ordinate on the new curve is equal to the area under the above curve between zero and the given ordinate. The integration process is shown graphically in Fig. 65*b*. The constant of integration K moves the axis of the curve upward as shown in Fig. 65*c*. If the phase delay of each component frequency is t_d seconds, the whole curve is also moved to the right t_d seconds as shown. Since a square pulse consists of a unit function followed t seconds later by a negative unit function, the output of the filter has the form indicated in Fig. 65*d* when a square pulse is applied.

The theoretical output wave as shown is characterized by both an anticipation and a follow-up oscillatory transient. The follow-up oscillations are observed whenever a square-topped pulse is applied to a circuit with an abrupt cutoff and sometimes they give rise to serious operating difficulties. The anticipatory transient, however, does not occur in any actual circuit because the effects of a wave, naturally, could not be observed before it is applied. This difference between the theoretical and the test results is due to the fact that some of the assumptions made in the above analysis are not realizable in an actual filter. In other respects, however, the theoretical form of the output checks the test results very well.

Example: A low-pass filter has a uniform response of unity from zero up to the cutoff frequency of 10,000 cps. Find the time interval between the negative minimum and positive maximum of the theoretical output wave (Fig. 65*c*) when a unit function is applied at the input. Also find the height of the first positive maximum.

Since the curve of 65*c* was obtained by integration of the curve in Fig. 65*b* (equation 118), the time interval specified is also the time between the first negative intersection of this curve with the axis and the first positive intersection. From equation 118, each of these intersections will occur where time t has such a value that

$$\omega_c t = \pi \quad \text{or} \quad t = \frac{\pi}{2\pi f_c} = \frac{1}{2 \times 10{,}000}$$

The desired time interval is twice this value or

$$t_f = \frac{1}{10{,}000} \text{ sec}$$

This is also the period of the oscillation and confirms the general statement on square-wave response which was made in section 6–3.

The height of the first oscillation is obtained by integration of equation 118 between zero and the time where $\omega_c t = \theta = \pi$

$$\int_0^{\frac{\pi}{\omega_c}} \frac{\omega_c}{\pi} \frac{\sin \omega_c t}{\omega_c t}\, dt = \frac{1}{\pi} \int_0^{\pi} \frac{\sin \theta}{\theta}\, d\theta = 0.588$$

(*Note:* The integral of $\frac{\sin \theta}{\theta}$ is in the form of a series.) The constant K is

$$K = \frac{1}{\pi} \int_0^{\infty} \frac{\sin \theta}{\theta}\, d\theta = \frac{\pi}{2\pi} = 0.5$$

and the total height becomes $0.5 + 0.588 = 1.088$. In an actual filter, the anticipatory oscillation of Fig. 65*c* would be missing. Instead, the output would start at zero and rise directly to the overshoot value of 1.088.

Naturally, any practical amplifier must have a finite cutoff frequency. The tube capacitances themselves shunt out the very high frequencies. An amplifier with a uniform response up to its cutoff frequency is, therefore, subject to the above oscillation troubles. The oscillations can be eliminated, however, if the cutoff is made less abrupt. When a low-pass coupling network has a response function in the form of a $\frac{\sin \theta}{\theta}$ curve like equation 116 (Fig. 65*e*) and a unit impulse is applied, the output is in the form of a square-topped wave having a duration of t_1 seconds (Fig. 65*f*). Accordingly, the output from such a circuit when a unit function is applied is this square-topped-pulse integrated. The results are shown graphically in Fig. 65*g*. The oscillations are completely eliminated and the corners of the wave are quite square. The performance of this network ceases to be satisfactory, however, when a square-topped pulse of such short duration is applied to the input that the two sloping sides of the output wave meet to form a triangle instead of a trapezoid (Fig. 65*h*). If the network is to handle pulses of shorter duration than this, the frequency range has to be extended; i.e., the frequency of the first intersection of the $\frac{\sin \theta}{\theta}$ characteristic curve with the zero axis must be higher.

Other low-pass circuits also have properties which tend to eliminate oscillations. One such is a circuit which has a frequency characteristic in the form of an error function (Fig. 65*i*). When a unit impulse is applied, the output pulse is also in the form of an error function. If a unit function voltage is applied to such a circuit, therefore, the output voltage takes the form shown in Fig. 65*k*.

While the above results were derived for low-pass circuits such as those used in video amplifiers, the same theory nevertheless also applies to band-pass amplifiers of the type used in radio-frequency circuits. A radio-frequency wave modulated with a square-topped pulse or an abrupt transient is subject to the oscillation discussed above when it is passed through a band-pass network which has a sharp cutoff. The effect on the output wave is shown in Fig. 66.

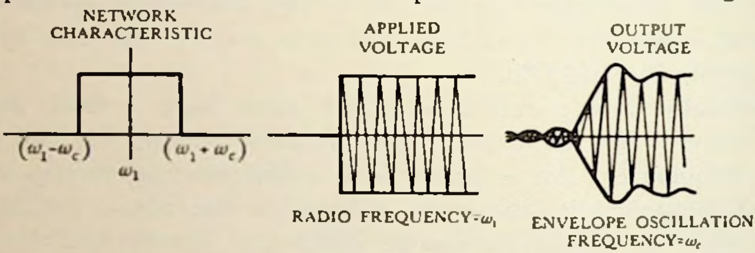


FIG. 66. Oscillogram of output voltage when an abruptly modulated wave is applied to a band-pass network having a very sharp cutoff.

The response characteristic of Fig. 65*a* or Fig. 66 is very similar to that of an overcoupled radio-frequency transformer. Consequently, a circuit of this kind gives bad performance when used with abruptly modulated waves. On the other hand, a critically coupled or undercoupled radio-frequency transformer has a characteristic similar to that of Fig. 65*i*; therefore, its performance is more satisfactory provided, of course, that the band width is large enough so that the time required for the output wave to build up or die down is not excessively long. In this connection, resonant circuits should be analyzed from the viewpoint of the damped oscillation decay rate. The resolution of an abruptly modulated wave will suffer if the decay time is too long.

The current in a resonant circuit after the driving voltage has been removed is given by the equation

$$i = I_0\epsilon^{-\frac{Rt}{2L}} \sin 2\pi ft \tag{119}$$

Therefore, the ratio of the maximum current in the $(n + m)$th cycle to that in the nth cycle is

$$\frac{I_{(n+m)}}{I_n} = \frac{\epsilon^{-\frac{R\left(t+\frac{m}{f}\right)}{2L}}}{\epsilon^{-\frac{Rt}{2L}}} = \epsilon^{-\frac{Rm}{2\pi f L}} = \epsilon^{-\frac{m\pi}{Q}} \tag{120}$$

where Q is defined as the ratio between the inductive reactance and the resistance. For a given application, the minimum allowable rate of decay (i.e., the current ratio after a certain number of cycles have elapsed) is specified and the maximum allowable Q which the resonant circuit can have is determined by equation 120.

6–7. Design of a Network. From Figs. 65*e*–65*h*, it is apparent that an ideal coupling network for the reproduction of square waves would have a response characteristic in the form of $\frac{\sin\theta}{\theta}$ curve. Unfortunately, this kind of response is extremely difficult, if not impossible, to obtain. A second choice would be something like the probability curve of Fig. 65*i*. It is more within the realm of possibility. The problem of finding a network which will fit an attenuation curve of arbitrary shape has not been satisfactorily solved. Consequently, there are no general formulas such as those found in filter theory. "Main strength and awkwardness" methods are necessary and often approximations have to be made. These difficulties arise from the fact that resistive elements are usually required in networks having an arbitrary response and the formulas involved are correspondingly more complex.

There is a useful network property, however, which helps to simplify the problem. If the resistance in a network is across the output terminals and all of the other elements are pure reactances, the voltage response ratio is a function of the resistance and the input conductance alone. The input admittance is

$$Y(\omega) = \frac{I_i}{E_i} = G(\omega) + jB(\omega)$$

where each of these factors is a function of the applied frequency. The power put into the circuit is all dissipated in the terminal resistance; therefore,

$$|E_i|^2\, G(\omega) = \frac{|E_o|^2}{R}$$

or

$$\left|\frac{E_o}{E_i}\right| = \sqrt{RG(\omega)} \tag{121}$$

which is a simple expression for the ratio between input and output voltage. In the same manner, a similar expression can also be obtained for the transfer impedance $\left(\frac{E_o}{I_i}\right)$ of the network.

A network consisting of a capacity C in parallel with the terminal resistance R and connected to the input through a series inductance L can be made with a response curve similar to the probability function of Fig. 65*i*. The input impedance of this combination is

$$Z(\omega) = j\omega L + \frac{1}{\frac{1}{R} + j\omega C}$$

and the admittance is

$$\begin{aligned} Y(\omega) &= \frac{1}{Z(\omega)} = \frac{1 + j\omega CR}{j\omega L - \omega L\omega CR + R} \\ &= \frac{R + j[-\omega L + \omega CR(R - \omega L\omega CR)]}{(\omega L)^2 + R^2 - 2\,\omega L\omega CR^2 + (\omega L)^2(\omega C)^2R^2} \end{aligned}$$

If
$$\omega L = 2\,\omega CR^2 \tag{122}$$
and
$$x = \sqrt{2}\,\omega CR \tag{123}$$

the equation for the input admittance reduces to the following form:

$$Y(\omega) = \frac{1}{R(1 + x^4)} - j\frac{x}{\sqrt{2}\,R}\frac{(1 + x^2)}{(1 + x^4)} \tag{124}$$

According to equation 121 the response of this circuit is then

$$\left|\frac{E_o}{E_i}\right| = \sqrt{RG(\omega)} = \frac{1}{\sqrt{1 + x^4}} \tag{125}$$

A plot of this equation shows that it has a shape similar to that of Fig. 65*i*. The frequencies at which the response falls to $\frac{1}{\sqrt{2}}$ times the value at zero frequency are defined as the limits of the pass band. They occur at the points where $x = \pm 1$ and are indicative of the rapidity with which the output wave rises to its final value when a unit function input is applied. A plot of the susceptance portion of equation 124, however, shows that it does not have the proper form to give the circuit a satisfactory input impedance for matching purposes. It can be improved by the addition of a shunt capacity C' at the input terminals of the network where

$$C' = C \tag{126}$$

Consequently, equations 122, 123, and 126 can be used to design an amplifier coupling network which will reproduce transient waves of the unit function type without introducing the troublesome oscillations which are often produced by the ordinary coupling networks.

If the input wave to the coupling network is in the form of a unit function current instead of a unit function voltage as it would be if the network were supplied by a pentode tube, the transfer impedance rather than the voltage ratio should have the form illustrated in Fig. 65*i* if satisfactory performance is to be obtained. A cut and try process similar to the above gives a different set of design equations for this condition. They are:

$$\omega L = 2\,\omega C R^2 \qquad x = 2\,\omega C R \qquad C' = 3\,C \tag{127}$$

The coupling circuit is connected as a resistance-terminated pi section in exactly the same manner as before and the same symbols apply to the individual circuit elements. The frequency limits of the pass band are again defined by $x = \pm 1$.

When a network is designed for low-pass service with either of the above sets of equations, a simple transformation adapts it for band pass. For a given application, the shape of the response curves for the two networks should be identical. Since the response characteristic of the band pass coupling is symmetrical about the mid-frequency of the band ω_m instead of about zero frequency as with the low-pass network, a means of adapting a low-pass design to a band-pass service is suggested. The inductive elements of the low-pass design are replaced by circuits which have practically the same characteristics around a frequency of ω_m as the given inductive elements have around zero frequency, and a similar substitution is made for the capacitive elements.

An inductive reactance is represented by a straight line which passes through the origin and has a slope equal to the magnitude of the inductance L_0. Its counterpart in the band-pass network must have a reactance curve which is a straight line passing through the frequency ω_m and having a slope equal to L_0. This curve is approximated by the reactance curve of a series inductance L_1 and capacity C_1 combination which is resonant at ω_m provided the width of the band is small compared with the resonant frequency ω_m. The slope of the reactance curve for such a combination is equal to 2 L_1 at the point of resonance; hence,

$$L_1 = \tfrac{1}{2} L_0 \tag{128}$$

Also

$$\omega_m = \frac{1}{\sqrt{L_1 C_1}}$$

Therefore

$$C_1 = \frac{2}{\omega_m^2 L_0} \tag{129}$$

In a similar manner, a capacitive susceptance ωC_0 is represented by a straight line through the origin with a slope equal to C_0. In a narrow range near the mid-frequency of the band ω_m, this is approximated by a parallel circuit of C_2 and L_2 which resonates at ω_m. The slope of the susceptance curve is 2 C_2 at ω_m; hence,

$$C_2 = \tfrac{1}{2} C_0 \tag{130}$$

and from

$$\omega_m = \frac{1}{\sqrt{L_2 C_2}}$$

$$L_2 = \frac{2}{\omega_m^2 C_0} \tag{131}$$

Problems

6–1. Find the Fourier series for a wave which is defined in the interval of a single cycle as follows: It is zero from $t = -\ 0.0025$ sec to $t = -\ 0.0005$ sec at which point it rises abruptly to a value of 100 and remains there until $t = 0.0005$ sec. It then drops to zero and remains there until the end of the cycle at $t = 0.0025$ sec. Plot the magnitude of the components as a function of frequency and compare the results with Fig. 63.

6–2. Consider that the square wave in the example of section 6–2 is applied to an amplifier which reproduces all the components up to and including the fifth harmonic (third term of the series) exactly and cuts off all the others.

(*a*) Find the shape of the output wave graphically by adding the component sine waves together point by point.

(*b*) Find the shape of the output wave if the gain of the amplifier at the different harmonics is as follows: gain at $f = 0.89$, gain at

$3f = 0.73$, gain at $5f = 0.35$, and the gain at all other harmonics is zero.

(*c*) Which amplifier more nearly reproduces the input wave?

6–3. Find the amplitude and phase angle for the response of the network described in the example of section 6–3 at 5000 cps.

6–4. (*a*) Find the amplitude distribution $g(\omega)$ for the oscillatory discharge wave $E\epsilon^{-\alpha t} \cos \omega_0 t$. The wave starts at $t = 0$. (The answer to this problem shows why a spark transmitter can be heard on all frequencies, but is heard best on the frequency of its fundamental. See section 8–1).

(*b*) Show that the answer to part (*a*) reduces to equation 115 when $\alpha = 0$ and $\omega_0 = 0$.

6–5. Following are data on the gain of a certain amplifier:

Frequency (cps)	*Gain*
0	13.0
5,000	12.5
10,000	11.2
15,000	9.4
20,000	7.3
25,000	5.9
30,000	4.3
35,000	3.4
40,000	2.9

On the assumption that the response of this amplifier is the same as that of an amplifier having the $\dfrac{\sin \theta}{\theta}$ characteristic which most nearly fits the above data, calculate the time required for the output wave to build up to its final value when the input is a unit function.

6–6. Find the maximum Q a resonant circuit can have if the current in it must decay at least 20 db in 500 cycles.

6–7. (*a*) Specify the components of a network which will give good reproduction of a unit function voltage applied at the input terminals. The capacitor at the output terminals is 0.005 μf and the frequency limit of the pass band is 7000 cps. Sketch the circuit diagram.

(*b*) Adapt the above design for band-pass use at 465 kc. Sketch the circuit diagram.

6–8. Solvè problem 6–7 for a unit function current input.

Experiment 13

Object:

The object of this experiment is to demonstrate the validity of the Fourier integral as a means of representing a transient wave.

Preliminary:

1. Plot the amplitude-distribution curves for unit function and for a square-topped pulse of $\frac{1}{1000}$-sec duration.
2. Solve problem 6–1.

Performance:

1. Connect a telegraph key, a small battery, and the input terminals of a wave analyzer in series. Produce a unit function by depressing the key and observe the maximum ballistic kick of the wave-analyzer meter. Since this kick is proportional to the amplitude of the harmonic indicated on the frequency dial, take a series of these readings for an experimental amplitude distribution curve.

Repeat the observation at each frequency setting several times and record the most usual reading obtained. This procedure will tend to eliminate errors which occur when the key "bounces" as it is depressed and produces a transient wave shape different from a true unit function. (Omit readings at or near zero frequency because they will be erroneous.)

2. Assemble a circuit like Fig. 19 for producing a square pulse with a duration of $\frac{1}{1000}$ sec when a voltage from a battery is applied to C_i by a telegraph key. Adjust the circuit until the measured pulse duration is $\frac{1}{1000}$ sec as observed on a cathode-ray oscillograph. Connect the square-topped pulse output to the wave analyzer and obtain an experimental amplitude-distribution curve for this wave by the same method as used in part 1.

3. Multiply the experimental data of part 1 by a constant factor which will make it match the theoretical curve of the preliminary part at 100 cps. Plot this experimental curve in a contrasting color on the same coordinate axes as the theoretical curve.

4. Multiply the experimental data of part 2 by a constant factor which will make it match the theoretical curve at 500 cps. Plot this resultant curve in a contrasting color on the same coordinate axes as the theoretical curve.

5. Discuss results.

Experiment 14

Object:

The object of this experiment is to demonstrate the connection between the response characteristic of a coupling network and its ability to reproduce a unit function wave, and to obtain a correlation between transient and sinusoidal response.

Preliminary:

1. Solve problem 6–7.

2. Plot the sinusoidal response $\frac{E_o}{E_i}$ of the circuit in problem 6–7 as a function of frequency, and from it estimate the probable duration of the output wave front (t_1 in Fig. 65*g*) when a unit function wave is applied to the input.

Performance:

1. Set up the circuit of problem 6–7, but make the output terminal resistor R variable (0–5000 ohms).

2. Apply a square-wave voltage to the input of the above and observe the output voltage with a cathode-ray oscillograph when the terminal resistor is set at 5000 ohms. Repeat the observation for several different frequencies of the applied wave.

3. Modulate the beam of the cathode-ray oscillograph with a 40-kc timing wave while observing the transient output of part 2. Obtain data for calculating the sinusoidal response of the network. Be sure to synchronize the square-wave input with the timing wave.

4. Adjust the terminal resistance to give the most nearly square output wave. Measure this resistance and compare with the calculated value from part 1 of the preliminary. Find the highest frequency the applied square wave can have and still produce a trapezoidal output. Compare this frequency with the predicted results of part 2 in the preliminary.

5. Adjust the terminal resistance to a lower value than the one used in part 3 and observe the output wave shape.

6. Change the terminal resistance back to 5000 ohms and apply a sinusoidal voltage to the network. Measure the input and output voltage for a series of frequencies with a vacuum-tube voltmeter and plot this response curve in a contrasting color on the same coordinate axes as a plot of the response curve calculated from the results of part 3 above. The latter calculations will be greatly facilitated by adapting the Bedford-Fredendall charts to the problem.

References

Bedford, A. V., and G. L. Fredendall, "Analysis, Synthesis, and Evaluation of the Transient Response of Television Apparatus," *Proc. I.R.E.*, October 1942, pp. 440–457.

Bush, V., *Operational Circuit Analysis*. New York: John Wiley & Sons, Inc., 1937, pp. 165–175.

Guillemin, E. A., *Communication Networks*. New York: John Wiley & Sons, Inc., 1935, vol. II, pp. 461–551.

CHAPTER VII

ULTRA-HIGH-FREQUENCY CIRCUIT ELEMENTS

7–1. General. The conventional ideas of inductance and capacitance are no longer valid at the ultra-high frequencies. For example, the capacitive reactance of a small condenser is virtually nonexistent to 10-cm, i.e., 3000-mc waves. A condenser which only consists of two plates having an area of 0.1 cm^2 and spaced 0.1 cm apart has a reactance of 6×10^{-10} ohms. The inductive reactance of the condenser leads alone would probably be a great deal more than this. The usual formulas for inductance indicate that a single turn of wire having a radius of about 0.1 cm and connected between the two plates would produce a circuit which would resonate at the 3000-mc frequency, but the equations on which these calculations are based neglect many factors which cannot be overlooked at the extremely high frequencies and solutions made by them are in error. Assuming that the calculations were approximately correct, however, it would be absurd to attempt the use of such microscopic elements in a piece of apparatus.

It is next to impossible to identify any portion of a circuit specifically as a capacity or an inductance since both of these factors are present to a significant degree in all high-frequency circuit elements. The distributed parameters of all elements must be taken into account. Even the wires connecting a circuit together must be treated as transmission lines because they are usually long in comparison with the wavelength. Since a transmission line is analyzed on the basis of its distributed parameters and its performance can be calculated with a reasonable degree of accuracy at the high frequencies, it was quite natural for early experimenters to investigate the possibilities of using this circuit element in other ways than the mere transfer of energy from one point to another. As a result of these investigations, it has been found that sections of transmission lines together with combinations of such sections can be used to simulate almost any electrical circuit element. Transmission lines can be used for resistance, inductance, and capacitive elements as well as resonant circuits, impedance-matching transformers, filters, and even insulators.

The open-wire line in its various forms (single wire, parallel wire, or multiple wire) is perhaps the most familiar of all types. A two-wire line made up of wires having a diameter d and a distance D between their centers has a distributed capacity of

$$C = \frac{27.8 \times 10^{-12}}{\cosh^{-1} \dfrac{D}{d}} \text{ farads/meter of length} \tag{132}$$

where d and D are each in the same units. When the distance between the centers of the wires is greater than ten times their diameter, the nonuniform distribution of the current on the conductor need not be considered and the following approximate formula can be used:

$$C = \frac{12.06 \times 10^{-12}}{\log_{10} \dfrac{2\,D}{d}} \text{ farads/meter of length} \tag{133}$$

The distributed inductance of a parallel wire line is

$$L = \left[10 + 92 \log_{10} \frac{2\,D}{d}\right] \times 10^{-8} \text{ henries/meter of length} \tag{134}$$

In spite of the fact that the resistance of a transmission line increases with frequency due to the skin effect, it is negligible in comparison with the inductive reactance of the line at the extremely high frequencies because the resistance increases as the square root of the frequency, whereas the latter increases directly with frequency. Similarly, the conductance of a line is negligible in comparison with the susceptance. Consequently, the characteristic impedance which is defined as

$$Z_0 = \sqrt{\frac{R + j\omega L}{G + j\omega C}} \tag{135}$$

(where C is the distributed capacitance, L is the distributed inductance, R the distributed resistance, G the distributed conductance, and ω is equal to the frequency f multiplied by 2π) becomes a pure resistance

$$Z_0 = \sqrt{\frac{L}{C}} \tag{136}$$

at these frequencies. If the constant portion of the inductance equation 134 is neglected, the substitution of equations 133 and 134 into equation 136 for characteristic impedance gives

$$Z_0 = \sqrt{\frac{92 \times 10^{-8} \log_{10}^2 \frac{2D}{d}}{12.06 \times 10^{-12}}} = 276 \log_{10} \frac{2D}{d} \text{ ohms} \tag{137}$$

When the spacing is less than ten times the diameter of the wires, the nonuniform distribution of current must be considered and the equation for characteristic impedance becomes

$$Z_0 = 120 \cosh^{-1} \frac{D}{d} \tag{138}$$

The construction of radio-frequency open-wire lines follows the usual power-line practice. Even the same materials are often used. The insulators introduce electrical irregularities in the line, but these effects are usually negligible. On the other hand, when the wires change their position with respect to each other, the characteristic impedance of the line also changes; therefore, precautions are taken to prevent this difficulty. The radiation loss, which is increasingly important at the higher frequencies, is reduced by using very close spacings. As an index to the magnitude of this loss, the following approximate formula for the radiation from a two-wire line terminated in its characteristic impedance is useful.

$$\text{Radiation loss} = 160\, I^2 \left[\frac{\pi D}{\lambda}\right]^2 \text{ watts approximately} \tag{139}$$

where I is the rms current being carried by the line and λ the wavelength of the transmitted energy expressed in the same units as the line spacing D. The above equation is based on the assumption that the line spacing is less than $\frac{1}{10}\lambda$ and that the line length is greater than 2λ. More exact formulas which are not limited by these restrictions are given in the Sterba and Feldman paper. In general, however, it can be said that the coefficient of equation

139 increases from 0 in almost direct proportion with the length of the line up to a line length of $\frac{\lambda}{2}$ at which time it has a magnitude of about 135. As the length of the line is increased beyond this value, the coefficient asymptotically approaches the magnitude of 160.

The coaxial transmission line is perhaps less familiar, but it is far more important in microwave practice than the open-wire line.

It consists of an inner conductor with an outside radius of a centimeters surrounded by an outer conductor having an internal radius of b centimeters. The important advantage of this line is that it is well shielded. If it is grounded at frequent intervals or buried in the ground, there is practically no effect from external electrical disturbances. Also, the coaxial line does not tend to radiate at ultra-high frequency as does the open-wire line. Consequently, it finds a major application here even though its construction is more expensive.

The capacity of a coaxial line is equal to

$$C = \frac{10^{-9}}{41.4 \log_{10} \frac{b}{a}} \text{ farads/meter of length} \tag{140}$$

and its distributed inductance is

$$L = 4.6 \times 10^{-7} \log_{10} \frac{b}{a} \text{ henries/meter of length} \tag{141}$$

Neglecting losses, the characteristic impedance becomes

$$Z_0 = 138 \log_{10} \frac{b}{a} \text{ ohms} \tag{142}$$

Occasionally, it is necessary to calculate the radio-frequency resistance of a coaxial line. For this purpose, the following formula can be used:

$$R = \sqrt{\frac{10\,\rho\mu f}{\mu_0}} \left(\frac{1}{a} + \frac{1}{b}\right) 10^{-3} \text{ ohms/meter of length} \tag{143}$$

where f is the operating frequency of the line in cycles per second. The resistivity ρ is in ohms per centimeter cube and a and b are in centimeters. μ is the permeability of the material from which the line is made, and μ_0 is the permeability of vacuum. For coaxial lines made of copper, the radio-frequency resistance per unit length becomes

$$R = 4.16 \times 10^{-6}\sqrt{f}\left(\frac{1}{a} + \frac{1}{b}\right) \text{ ohms/meter of length}$$

The thermal expansion of long coaxial lines presents a major constructional problem. One solution is to mount the lines on flexible supports which permit it to buckle and thus compensate for temperature changes. Otherwise, expansion joints of some sort must be provided. Glazed-porcelain spacing insulators are to be preferred over the porous type because dirt can be more readily removed from their surfaces. At microwave frequencies, the synthetic resins such as polystyrene "A" are commonly used; however, their insulating properties are impaired if plasticizers have been added.

When radiation is neglected, the general transmission-line equations give an exact solution for the sending-end current and voltage. Since a line has a distributed impedance per unit length equal to $Z = R + j\omega L$ and a distributed admittance per unit length of $Y = G + j\omega C$, the differential change in voltage and change in current along the line are

$$\frac{dE}{dx} = IZ \quad \text{and} \quad \frac{dI}{dx} = EY$$

where x is measured from the receiving end to the point in question. Differentiating the first of these equations with respect to x and substituting in the second,

$$\frac{d^2E}{dx^2} = \frac{dI}{dx}Z = EZY \tag{144}$$

This differential equation has the following solution for the line voltage E at a distance x from the receiving end

$$E = A\epsilon^{\sqrt{ZY}x} + B\epsilon^{-\sqrt{ZY}x}$$

or in terms of a new pair of constants

$$\begin{aligned} E &= \left(\frac{C + D}{2}\right)\epsilon^{\sqrt{ZY}x} + \left(\frac{C - D}{2}\right)\epsilon^{-\sqrt{ZY}x} \\ &= C \cosh \sqrt{ZY}x + D \sinh \sqrt{ZY}x \end{aligned} \tag{145}$$

The constants C and D are evaluated from the terminal conditions that the line voltage is equal to the receiving-end voltage or $E = E_r$ when $x = 0$ and the derivative of the line voltage is equal to the receiving-end current multiplied by the distributed impedance or $\frac{dE}{dx} = I_r Z$ when $x = 0$; therefore,

$$C = E_r \quad \text{and} \quad D = \frac{I_r Z}{\sqrt{ZY}} = I_r\sqrt{\frac{Z}{Y}}$$

Since $\sqrt{\frac{Z}{Y}}$ is defined as the characteristic impedance Z_0 and $\sqrt{ZY}$ is defined as the propagation constant γ, the equation for the line voltage at a distance x from the receiving end becomes

$$E = E_r \cosh \gamma x + I_r Z_0 \sinh \gamma x \tag{146}$$

Similarly, the equation for the line current at a distance x from the receiving end becomes

$$I = I_r \cosh \gamma x + \frac{E_r}{Z_0} \sinh \gamma x \tag{147}$$

The propagation constant consists of a real and an imaginary portion and is written as

$$\gamma = \alpha + j\beta = \sqrt{(R + j\omega L)(G + j\omega C)} \tag{148}$$

The name comes from the function of this factor in the transmission of an electric wave. It determines the magnitude and phase angle of the wave at all points along the line. From equations 146 and 147, for example, the ratio between the sending-end voltage and the receiving-end voltage or the sending-end current and the receiving-end current of a unit length of line terminated in its characteristic impedance Z_0 becomes

$$\frac{E_s}{E_r} = \frac{I_s}{I_r} = \epsilon^{\gamma} = \epsilon^{\alpha}\epsilon^{j\beta} \tag{149}$$

This equation also shows the significance of the real and complex portions of the propagation constant. It is recalled that $\epsilon^{j\beta}$ is simply another way of writing the angle of a complex quantity in polar form; therefore, β represents the phase angle between the sending- and receiving-end current and voltage. On the other hand, ϵ^{α} is a real number which gives the ratio of the magnitudes of the sending- and receiving-end current and voltage. Consequently, the real portion α is called the attenuation constant of the line and the imaginary portion β is called the phase constant.

If both the resistance R and the conductance G in equation 148 are neglected as they can be with ultra-high-frequency lines, the propagation constant reduces to

$$\gamma = \sqrt{(j\omega L)(j\omega C)} = j\omega\sqrt{LC}$$

But from equations 140 and 141 or 133 and 134, the factor $\frac{1}{\sqrt{LC}}$ is found to be equal to the velocity of light $c = 3 \times 10^8$ meters/sec; therefore, the propagation constant becomes

$$\gamma = j\frac{2\pi f}{c} = j\frac{2\pi}{\lambda} \tag{150}$$

where λ is the wavelength corresponding to the frequency. If the conductance alone is neglected and the resistance is assumed to be small in comparison with the inductive reactance, the expression for the propagation constant can be expanded by the approximate formula for the square root of a binomial in which one term is appreciably smaller than the other. The results are

$$\gamma = \alpha + j\beta = \sqrt{(R + j\omega L)(j\omega C)} = j\omega\sqrt{LC} + \frac{j\omega CR}{2j\omega\sqrt{LC}} \tag{151}$$

The real portions of the terms on each side of the above equation must be equal to each other and the imaginary terms on each side must also be equal to each other; hence,

$$\alpha = \frac{R}{2}\sqrt{\frac{C}{L}} = \frac{R}{2Z_0} \tag{152}$$

and

$$\beta = j\omega\sqrt{LC} = j\frac{2\pi}{\lambda} \tag{153}$$

Equations 152 and 153 are used when the attenuation must be taken into account and equation 150 is used for general purposes.

The coaxial line attenuation constant is computed by substituting equations 142 and 143 in equation 152.

$$\alpha = \frac{R}{2Z_0} = \frac{\sqrt{\frac{10\,\rho\mu f}{\mu_0}}\left(\frac{1}{a} + \frac{1}{b}\right) \times 10^{-3}}{276 \log_{10}\frac{b}{a}} \tag{154}$$

The attenuation constant has a minimum value when the ratio of $\frac{b}{a} = 3.6$. This can be shown by differentiating the above equation with respect to a and finding the minimum. The attenuation is low for quite a range of radii ratios near 3.6; however, it becomes very pronounced outside of that range.

The ratio of the conductor radii of a coaxial line has a slightly different value than the above if the maximum voltage gradient is to have a minimum value. The maximum voltage gradient $\mathcal{E}_m$ occurs at the surface of the inner conductor and is equal to

$$\mathcal{E}_m = \frac{E}{a \log_e \frac{b}{a}} \tag{155}$$

where E is the voltage which is applied between the inner and outer conductor. This equation has a minimum value when the ratio $\frac{b}{a} = 2.72$ as can be shown by differentiating the equation with respect to the ratio and equating the derivative to zero.

7-2. Transmission-line Impedance Elements. A transmission line which has been terminated in its characteristic impedance makes a useful standard resistance at high frequencies. More important, however, from an application point of view, are its reactive properties. Short lengths of short-circuited line are pure inductive reactance and open-circuited lines have pure capacitive reactance. Longer lengths of these elements have the opposite kind of reactance. When the losses of the line are neglected, the general transmission line equations (146 and 147) are simplified considerably because the propagation constant becomes a pure imaginary from equation 150, and the identities $\cosh j\theta = \cos\theta$ and $\sinh j\theta = j \sin\theta$ can be used to replace the hyperbolic functions of the general equations

$$E = E_r \cos\frac{2\pi x}{\lambda} + jI_r Z_0 \sin\frac{2\pi x}{\lambda} \tag{156}$$

$$I = I_r \cos\frac{2\pi x}{\lambda} + j\frac{E_r}{Z_0}\sin\frac{2\pi x}{\lambda} \tag{157}$$

A short-circuited line has a receiving-end voltage $E_r = 0$. Hence, its terminal impedance from equations 156 and 157 is equal to

$$Z_{sc} = \frac{E_s}{I_s} = \frac{jI_rZ_0 \sin \frac{2\pi x}{\lambda}}{I_r \cos \frac{2\pi x}{\lambda}} = jZ_0 \tan \frac{2\pi x}{\lambda} \tag{158}$$

The positive sign of the j indicates an inductive reactance, and since the tangent varies from zero to infinity, the magnitude can be of any value. Indeed, when x is equal to a quarter wavelength, the input impedance is infinite or, in other words, the transmission line is virtually an insulator. "Insulators" of this sort are often used at ultra-high frequency. A typical example of their use on a

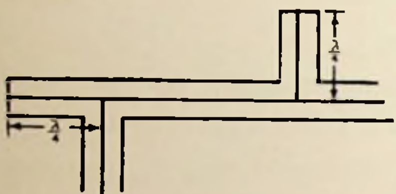


FIG. 67. Diagram illustrating the use of short-circuited quarter-wavelength sections of line as ultra-high-frequency insulators.

coaxial transmission line is shown in Fig. 67. Such "insulators" are advantageous in that they can be used to provide a d-c path between the conductors and at the same time effectively block the radio frequency.

In a similar manner, an open-circuited line has a receiving-end current $I_r = 0$. Its impedance is

$$Z_{oc} = \frac{E_s}{I_s} = \frac{E_r \cos \frac{2\pi x}{\lambda}}{j\frac{E_r}{Z_0} \sin \frac{2\pi x}{\lambda}} = -jZ_0 \cot \frac{2\pi x}{\lambda} \tag{159}$$

a capacitive reactance. In either case, the opposite kind of reactance is obtained when the length of the section of line lies between a quarter and a half wavelength because the sign of the trigonometric function changes.

If these lines are made a quarter wavelength long at a wavelength of λ_0, the factor $\frac{2\pi x}{\lambda}$ can be written as

$$\frac{2\pi \frac{1}{4} \lambda_0}{\lambda} = \frac{\pi f}{2 f_0} \tag{160}$$

where the f's are the frequencies corresponding to the wavelengths λ and λ_0. Then

$$\frac{2\pi x}{\lambda} = \frac{\pi}{2}\left(1 + \frac{f - f_0}{f_0}\right) = \frac{\pi}{2} + \delta$$

where
$$\delta = \frac{f - f_0}{f_0}\frac{\pi}{2} \tag{161}$$

In the region near $\frac{\pi}{2}$, the tangent is practically equal to the negative reciprocal of the angular difference from $\frac{\pi}{2}$ or $\tan\frac{2\pi x}{\lambda} = -\frac{1}{\delta}$. Also in the same neighborhood, $\cot\frac{2\pi x}{\lambda} = -\delta$. Hence, the input impedance of a short or open-circuited section of line becomes

$$Z_{sc} = -\frac{jZ_0}{\delta} \tag{162}$$

$$Z_{oc} = jZ_0\delta \tag{163}$$

When two such lines with different characteristic impedances are connected in parallel, their combined impedance is:

$$Z = \frac{Z_{sc}Z_{oc}}{Z_{sc} + Z_{oc}} = \frac{Z_0'Z_0''}{-\frac{jZ_0'}{\delta} + jZ_0''\delta} = -j\frac{Z_0'\delta}{\delta^2 - \frac{Z_0'}{Z_0''}} \tag{164}$$

an equation which is of the same form as the terms in Foster's reactance theorem. This theorem provides a design procedure for reactance networks in terms of the frequencies at which the reactance of the network is to be either zero or infinite. It states that the curve of reactance vs. frequency for a pure reactance network must always have a positive slope and that the points of resonance and antiresonance alternate. Since the factor δ is simply another expression for frequency, the curve of reactance vs. δ for a network takes the form shown in Fig. 68. If the theorem is restricted to a band of frequencies, the input impedance of a reactance network can be written

$$Z = jA\delta\left[\frac{B_0}{\delta^2 - \delta_0^2} + \frac{B_2}{\delta^2 - \delta_2^2} + \cdots + \frac{B_{2n}}{\delta^2 - \delta_{2n}^2}\right] \tag{165}$$

because a plot of this equation also has the form indicated in

Fig. 68. The coefficients B are given by the following equation in terms of the points where the zeros and infinities are to occur.

$$B_m = \frac{(\delta_m^2 - \delta_1^2)(\delta_m^2 - \delta_3^2) \cdots (\delta_m^2 - \delta_{2n-1}^2)}{(\delta_m^2 - \delta_0^2)(\delta_m^2 - \delta_2^2) \cdots (\delta_m^2 - \delta_{m-2}^2)(\delta_m^2 - \delta_{m+2}^2) \cdots (\delta_m^2 - \delta_{2n}^2)} \qquad (166)$$

A series connection of a group of parallel combinations such as those of equation 164 will have an impedance equation in the form of equation 165. Consequently, a general reactance designed on the

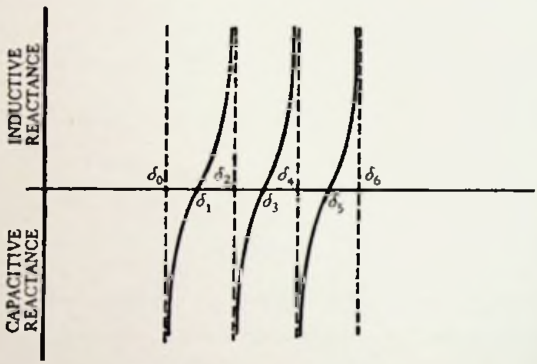


FIG. 68. A curve illustrating the variation of the reactance of a network as a function of frequency.

basis of equations 165 and 166 can be built by making parallel combinations of quarter wavelength short-circuited (or half wavelength open-circuited) and quarter wavelength open-circuited (or half wavelength short-circuited) transmission lines and connecting these combinations in series provided

$$AB_m = -Z_0' \qquad (167)$$

and

$$\frac{Z_0'}{Z_0''} = \delta_m^2 \quad \text{or}$$

$$Z_0'' = \frac{Z_0'}{\delta_m^2} \qquad (168)$$

The above equations give the same resultant network irrespective of the sign of the set of δ's used. This means that a complete plot of the network reactance as a function of frequency would have the design set of zeros and infinities in the range above the base frequency f_0 and an image set of zeros and infinities in the range below the base frequency. In addition, other image responses of this sort would occur around those frequencies where the length of the lines in the network becomes an odd multiple of a quarter wavelength.

7-3. Transmission-line Impedance Transformers. Whenever a transmission line is terminated in an impedance other than its characteristic impedance, standing waves are produced. If the terminal impedance is a pure resistance and is less than the characteristic impedance, the standing waves have the same general form as those on a short-circuited line; i.e., a current maximum occurs on the load end and a current minimum is a quarter of a wavelength back toward the generator. The magnitude of the standing waves becomes correspondingly less as the magnitude of the load impedance approaches that of the characteristic impedance. Higher impedances reflect traveling waves of the same sort as an open circuit; therefore, the current distribution under these circumstances is a minimum at the end of the line and a maximum a quarter of a wavelength back toward the generator. Since standing waves on a transmission line introduce high I^2R and E^2G losses in the line, they reduce the efficiency of transmission. They can be eliminated by terminating the line in its characteristic impedance. This is done by the use of a section of line alone or together with a stub as an impedance-matching transformer.

A line terminated in an impedance $Z_r = \dfrac{E_r}{I_r}$ has an input impedance of

$$Z_s = \frac{E_s}{I_s} = \frac{I_r Z_r \cos \dfrac{2\pi x}{\lambda} + jI_r Z_0 \sin \dfrac{2\pi x}{\lambda}}{I_r \cos \dfrac{2\pi x}{\lambda} + jI_r \dfrac{Z_r}{Z_0} \sin \dfrac{2\pi x}{\lambda}} \tag{169}$$

$$Z_s = Z_0 \frac{Z_r + jZ_0 \tan \dfrac{2\pi x}{\lambda}}{Z_0 + jZ_r \tan \dfrac{2\pi x}{\lambda}} \tag{170}$$

If x is some multiple of a half wavelength, the above equation reduces to $Z_s = Z_r$. In other words, a transmission line which is any multiple of a half wavelength long can be used to match any two identical impedances. On the other hand, if x is some odd multiple of a quarter wavelength, equation 169 becomes

$$Z_s = \frac{jI_r Z_0}{jI_r \frac{Z_r}{Z_0}} \quad \text{or} \quad Z_r Z_s = Z_0^2 \tag{171}$$

That is, a quarter-wave section of line will match two impedances Z_r and Z_s provided its characteristic impedance Z_0 is equal to the geometric mean of the two impedances ($Z_0 = \sqrt{Z_r Z_s}$). The best operation is secured with this section when the two terminating impedances are in the same order of magnitude; however, a series

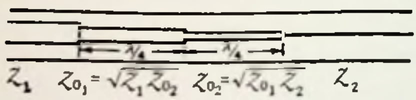


FIG. 69. A series of quarter-wavelength sections of line used to match two widely different impedances.

of quarter-wave sections can be used to obtain the match of the two terminal impedances if the two terminal impedances are too widely different. Such an application is shown in Fig. 69.

The input admittance of a line terminated in a given load is a function of the length of the line to the load. There is some length of line l for which the real portion of the input admittance is equal to the reciprocal of the line's characteristic impedance. If a susceptance which will cancel out the imaginary portion of the input admittance is shunted across the line at this distance l from the load, the line beyond that point works into its characteristic impedance and is therefore matched to the load. When the imaginary portion of the input admittance is positive or capacitive, an inductive length

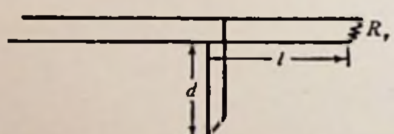


FIG. 70. A stub used as an impedance-matching transformer. The same connection can also be used on coaxial lines.

of short-circuited line can be used as the neutralizing susceptance. This short-circuited "stub" is mounted at right angles to the transmission line so that coupling between it and the line is reduced. An open-circuited "stub" is used in the same way if the imaginary portion of the input admittance is negative (Fig. 70).

When a line is terminated in a pure resistance R_r, the input impedance of the line from equation 170 is

$$Z = Z_0 \frac{R_r + jZ_0 \tan \dfrac{2\pi x}{\lambda}}{Z_0 + jR_r \tan \dfrac{2\pi x}{\lambda}}$$

Consequently, the input admittance is

$$Y = \frac{1}{Z} = \frac{1}{Z_0}\left[\frac{Z_0R_r + Z_0R_r \tan^2\left(\dfrac{2\pi x}{\lambda}\right)}{R_r^2 + Z_0^2 \tan^2\left(\dfrac{2\pi x}{\lambda}\right)} + j\,\frac{(R_r^2 - Z_0^2)\tan\left(\dfrac{2\pi x}{\lambda}\right)}{R_r^2 + Z_0^2 \tan^2\left(\dfrac{2\pi x}{\lambda}\right)}\right] \qquad (172)$$

$x = l$ when the real portion of this equation is equal to the reciprocal of the characteristic impedance

$$\frac{1}{Z_0} = \frac{1}{Z_0}\left[\frac{Z_0R_r + Z_0R_r \tan^2\left(\dfrac{2\pi l}{\lambda}\right)}{R_r^2 + Z_0^2 \tan^2\left(\dfrac{2\pi l}{\lambda}\right)}\right]$$

or

$$\tan^2 \frac{2\pi l}{\lambda}(Z_0^2 - Z_0R_r) = Z_0R_r - R_r^2$$

and

$$l = \frac{\lambda}{2\pi} \arctan \sqrt{\frac{R_r}{Z_0}} \qquad (173)$$

The substitution of equation 173 in the imaginary portion of the input admittance equation 172 yields

$$jB = j\,\frac{1}{Z_0}\,\frac{(R_r^2 - Z_0^2)\sqrt{\dfrac{R_r}{Z_0}}}{R_r^2 + Z_0^2\dfrac{R_r}{Z_0}} = j\,\frac{(R_r - Z_0)}{Z_0\sqrt{R_rZ_0}}$$

or

$$|X| = \frac{Z_0\sqrt{R_rZ_0}}{(R_r - Z_0)} \qquad (174)$$

The type of stub used will depend on which of the possible values of l is selected and the corresponding sign which the tangent has. An examination of equation 172 reveals that the susceptive portion of the input admittance is inductive for the first quarter of a wavelength from the load, capacitive for the next quarter, etc. when the load resistance R_r is less than the characteristic impedance. If the load resistance is greater than the characteristic impedance, the input reactance is capacitive for the first quarter wavelength as measured from the load and inductive for the second quarter. In either case, the input susceptance is inductive in the region of less than a quarter of a wavelength measured toward the load end from a current minimum and capacitive in the same region when the measurement is made toward the sending end. In other words, a short-circuited stub is appropriate when it is to be connected on the sending-end side within a quarter of a wavelength of a current minimum, and an open-circuited stub is appropriate when connected on the load side of the current minimum. These conclusions are shown graphically in Fig. 71. The length of stub required can

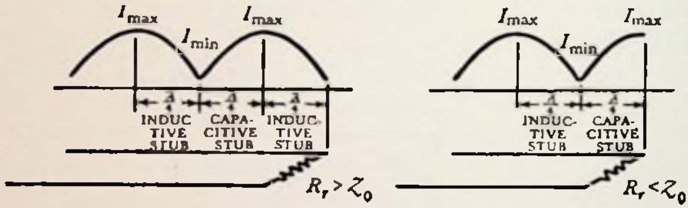


Fig. 71. Current distribution on lines with different loads together with an indication of the type of stub which is appropriate for various locations along the line.

be determined by the substitution of equation 174 into either 158 or 159. The signs are obtained automatically by selecting the type of stub to be used in the manner outlined above.

In practical cases, it is often easier to make calculations in terms of current rather than impedance ratios. The ratio of maximum to minimum current as measured along the line is equal to the ratio of the terminal impedance to the characteristic impedance. This can be demonstrated by a solution of the sending-end current equation (157). If $R_r > Z_0$, the distance l of equation 173 can also be measured from a current minimum or

$$l_{min} = \frac{\lambda}{2\pi} \arctan \sqrt{\frac{R_r}{Z_o}} = \frac{\lambda}{360^\circ} \arctan \sqrt{\frac{I_{max}}{I_{min}}} \tag{175}$$

When $R_r < Z_0$, the current minimum is a quarter of a wavelength or $\frac{\pi}{2}$ radians from the end of the line. In this case

$$l_{\min} = \frac{\lambda}{2\pi}\left(\frac{\pi}{2} - \arctan\sqrt{\frac{R_r}{Z_0}}\right) = \frac{\lambda}{360°}\arctan\sqrt{\frac{I_{\max}}{I_{\min}}}$$

Consequently, equation 175 can be used to calculate the position of the stub regardless of the magnitude of the terminal load. The distance $l_{\min}$ is measured either way from a current minimum. When the measurement is made toward the load end, an open-circuited stub is used. Its length is calculated from equations 159 and 174. Only absolute magnitudes are involved; therefore, the sign of equation 174 has no significance and can be reversed.

$$\frac{Z_0\sqrt{Z_0R_r}}{(R_r - Z_0)} = +Z_0\cot\frac{2\pi d}{\lambda} \quad \text{and} \quad \frac{2\pi d}{\lambda} = \operatorname{arccot}\frac{\sqrt{\dfrac{R_r}{Z_0}}}{\left(\dfrac{R_r}{Z_0} - 1\right)}$$

Therefore,

$$d = \frac{\lambda}{360°}\operatorname{arccot}\frac{\sqrt{\dfrac{I_{\max}}{I_{\min}}}}{\left(\dfrac{I_{\max}}{I_{\min}} - 1\right)} \tag{176}$$

If the measurement is made toward the sending end, a short-circuited stub is required whose length is

$$d = \frac{\lambda}{360°}\arctan\frac{\sqrt{\dfrac{I_{\max}}{I_{\min}}}}{\left(\dfrac{I_{\max}}{I_{\min}} - 1\right)} \tag{177}$$

Example: Find the length and position of a short-circuited stub to match a parallel-wire transmission line consisting of two number 18 wires spaced $1\frac{1}{4}$ in. apart to a load impedance of 73 ohms. The wavelength used is 1 meter.

From equation 137, the characteristic impedance of the line is

$$Z_0 = 276\log_{10}\frac{2D}{d} = 276\log_{10}\frac{2.5}{0.04} = 495 \text{ ohms}$$

The distance to the stub l is given by equation 173 and Fig. 71.

$$l = \frac{\lambda}{360°}\arctan\sqrt{\frac{73}{495}} = \frac{\lambda\, 159°}{360°} = 0.442 \quad \text{or} \quad 44.2 \text{ cm}$$

The length of the stub is then found to be

$$Z_0 \tan \frac{2\pi d}{\lambda} = \left| \frac{Z_0\sqrt{73 \times 495}}{(73 - 495)} \right| \quad \text{or} \quad \frac{2\pi d}{\lambda} = \arctan 0.45$$

and $d = 0.0672\,\lambda$ or 6.72 cm.

A short-circuited stub could also be placed at the point on the line which calls for an open-circuited stub if the short-circuited stub were made longer so that its susceptance would be capacitive rather than inductive. It should be noted that equations 175 and 177 could also have been used to solve this problem since the current ratio is equal to the impedance ratio.

The illustrations in this discussion imply the exclusive use of open-wire transmission lines for stub matching, but the theory of stub matching can also be applied to coaxial lines in exactly the same manner.

The choice between an open-circuited and short-circuited stub will depend on whether or not the two conductors are to be isolated from each other with respect to the direct current in the line.

A series rather than a shunt stub can also be used to match impedances. Its position is determined by finding the point at which the real portion of the input impedance of the line is equal to the characteristic impedance. A series reactance stub is introduced at that point to cancel out the effect of the input reactance. This connection is not generally used, however, because of the constructional difficulties which are involved.

The choice between a quarter-wavelength matching section and a stub depends on whether or not it is convenient to have lines of different characteristic impedance in the connection.

Sometimes it is inconvenient to adjust the distance l between the stub and the load when matching the load to the line. Under these circumstances, matching can be accomplished by the use of two stubs. One stub is connected across the load and the other is connected across the line a fixed distance from the load. The adjustment of the first stub has an effect similar to that caused by varying the distance l in the case of a single stub. It makes the conductance of the load as seen at the position of the second stub equal to the characteristic admittance of the line $\frac{1}{Z_0}\cdot$ The second stub is used to cancel out the susceptance which is simultaneously produced at this point. As a consequence, a match is obtained. The application of the double stub is limited by the fact that it cannot be used to match the line to all load impedances. The range over which a match is possible is determined by the spacing between the stubs.

Example: A pair of stubs are used to match a load conductance G_r to a line conductance G_s. Find the susceptances of the stubs in terms of these conductances, the characteristic admittance $\left(G_0 = \frac{1}{Z_0} \text{ which is a pure conductance}\right)$ of the line between, and its angular length $\theta = \frac{2\pi x}{\lambda}$. Also find a criterion to determine whether or not a match is possible.

From equations 156 and 157 the input admittance of a line having a length θ and terminated with an admittance $Y_r = \frac{I_r}{E_r}$ is

$$Y_s = \frac{I}{E} = \frac{I_r \cos\theta + j\dfrac{E_r}{Z_0}\sin\theta}{E_r\cos\theta + jI_rZ_0\sin\theta} = \frac{Y_r\cos\theta + jG_0\sin\theta}{\cos\theta + j\dfrac{Y_r}{G_0}\sin\theta}$$

The terminal admittance is equal to the load conductance G_r plus the terminal stub susceptance jB_r; hence, $Y_r = G_r + jB_r$. If the second stub has a susceptance jB_s then the input susceptance where it is connected to the line must be $-jB_s$ so that the net susceptance will reduce to zero. Consequently, the input admittance of the line must be $Y_s = G_s - jB_s$.

$$Y_s = G_s - jB_s = \frac{(G_r + jB_r)\cos\theta + jG_0\sin\theta}{\cos\theta + j\dfrac{(G_r + jB_r)}{G_0}\sin\theta}$$

and

$$(G_s - jB_s)(G_0\cos\theta - B_r\sin\theta + jG_r\sin\theta) = G_rG_0\cos\theta + jG_0(B_r\cos\theta + G_0\sin\theta)$$

Two simultaneous equations are obtained by expanding the above and equating the real and imaginary terms. From the real terms we have

$$G_sG_0\cos\theta - G_sB_r\sin\theta + B_sG_r\sin\theta = G_rG_0\cos\theta$$

which yields

$$B_s = \frac{G_r G_0 \cos\theta + G_s B_r \sin\theta - G_s G_0 \cos\theta}{G_r \sin\theta}$$

From the imaginary terms we have

$$B_s(-G_0 \cos\theta + B_r \sin\theta) + G_r G_s \sin\theta = G_0 B_r \cos\theta + G_0^2 \sin\theta$$

The substitution of the solution for B_s from the first equation into the above yields the following quadratic equation:

$$\begin{aligned}(G_s \sin^2\theta)B_r^2 + (-\,2\,G_s G_0 \sin\theta\cos\theta)B_r \\ - G_r G_0^2 + G_s G_0^2 \cos^2\theta + G_r^2 G_s \sin^2\theta = 0\end{aligned}$$

The solution for B_r is obtained by applying the quadratic formula

$$B_r = \frac{G_s G_0 \sin\theta\cos\theta \pm \sqrt{G_s G_r \sin^2\theta (G_0^2 - G_s G_r \sin^2\theta)}}{G_s \sin^2\theta}$$

This is one of the required results. The other susceptance is obtained by substituting the numerical value of B_r into the above equation for B_s. A match is impossible unless the factor B_r is a real number; therefore, a criterion for matching is given by

$$G_0^2 > G_s G_r \sin^2\theta$$

Since the characteristic admittance G_0 of the line between the stubs is usually equal to the conductance G_s to which the load is being matched, the above criterion reduces to

$$\sin^2\theta < \frac{G_0}{G_r}$$

The matching connections described in the above paragraphs are impedance transformers in which the ratio of transformation changes with a change in frequency. An extended analysis of the short-circuited stub connection illustrated in Fig. 70 shows that it acts like a fixed-ratio impedance transformer over a wide band of frequencies if the length d is made equal to l and the characteristic impedance Z_0' of the stub is different from the characteristic impedance Z_0'' of the line. These characteristic impedances are related to the sending- and receiving-end impedance by the following equation:

$$\frac{R_r}{R_s} = 1 + \frac{Z_0''}{Z_0'} \tag{178}$$

The length of the line and of the stub are related to the mid-frequency f_m of the constant-impedance-ratio band by the equation

$$l = d = \frac{c}{4f_m} \tag{179}$$

where c is the velocity of light. The ratio of the band width to the mid-frequency for relatively narrow bands is

$$\frac{f_u - f_l}{f_m} = \frac{4}{\pi}\sqrt{\frac{R_s}{R_r}} \text{ approximately} \tag{180}$$

where f_u and f_l are the upper and the lower frequency limits of the band over which a practically constant impedance match is obtained.

7-4. Transmission-line Filters. The general transmission-line equations can be applied to any four-terminal network whether it be a section of line, a combination of lumped circuit elements, or a combination of transmission-line elements. For convenience in handling, the equations are written in terms of the so-called general circuit parameters A, B, C, and D. Thus

$$E_s = E_r \cosh \gamma + I_r Z_0 \sinh \gamma = AE_r + BI_r \tag{181}$$

$$I_s = I_r \cosh \gamma + \frac{E_r}{Z_0} \sinh \gamma = DI_r + CE_r \tag{182}$$

When these equations are applied to the four-terminal network, Z_0 represents the iterative impedance which is the input impedance to a series connection of an infinite number of identical networks, or the input impedance to one such network when it is terminated in Z_0. The propagation constant γ has the same meaning as with the transmission line, but since the distance x has no significance in a four-terminal network, it does not appear in equations 181 and 182. In effect, it is included in the propagation constant.

It has been pointed out that γ is a complex number $\alpha + j\beta$. The real portion causes an attenuation or decrease in magnitude of the current or voltage passing through the network whereas the imaginary portion causes a phase angle to exist between the sending and receiving current or voltage. In order to determine the nature of the propagation constant for a given four-terminal network, it is necessary to write an equation for its sending-end voltage in terms of its receiving-end voltage and current. The coefficient A of E_r then corresponds to and is equal to $\cosh \gamma$. In nondissipative networks, i.e., circuits with negligible loss, this coefficient is real. If it lies between -1 and $+1$ for a given frequency range then γ must be a pure imaginary and the network will have no attenuation in that range. This can be made more clear by writing the equation for the angle whose hyperbolic cosine is A.

$$\cosh^{-1} A = \sinh^{-1}\sqrt{A^2 - 1} = \gamma$$

When $|A| < 1$, $\sinh \gamma$ is a pure imaginary number and that is possible only when γ is also a pure imaginary or when the attenuation α is zero. The network will have attenuation for all other values of A. Impedance elements can be connected together in such a way that the magnitude of the A coefficient of the network is less than one for a particular band of frequencies and greater than one for other bands. In other words, the network can be used as a filter having a pass band in the region where $|A| < 1$. Ordinary electric filters use combinations of coils and condensers to fulfill this criterion, but coils cannot be used at extremely high frequencies. Combinations of transmission lines or combinations of transmission lines and condensers are used for the elements of high-frequency filter networks instead.

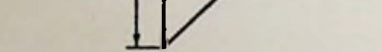


FIG. 72. A typical transmission-line filter.

For example, consider a network consisting of a shorted length of line shunted across another section of line at its center (Fig. 72). As a special case, consider that the length of the shorted section l is the same as each of the two segments on either side of it. If the characteristic impedance of the shorted line is Z_0', that of the other line is Z_0'', and θ is used for $\frac{2\pi l}{\lambda}$, the following equations can be written:

$$E_m = E_r \cos\theta + jI_r Z_0'' \sin\theta \tag{183}$$

$$I_3 = I_r \cos\theta + j\frac{E_r}{Z_0''} \sin\theta \tag{184}$$

$$I_2 = \frac{E_m}{jZ_0' \tan\theta} \tag{185}$$

$$I_1 = I_2 + I_3 = -j\frac{E_r \cos\theta}{Z_0' \tan\theta} + \frac{I_r Z_0''}{Z_0'}\cos\theta + I_r\cos\theta + j\frac{E_r}{Z_0''}\sin\theta$$

The substitution of the above equations into the equation for the sending-end voltage

$$E_s = E_m \cos\theta + jI_1 Z_0'' \sin\theta \tag{186}$$

gives

$$E_s = E_r\left(\cos^2\theta + \frac{Z_0''}{Z_0'}\cos^2\theta - \sin^2\theta\right) + jI_r Z_0''\left(2\sin\theta\cos\theta + \frac{Z_0''}{Z_0'}\sin\theta\cos\theta\right) \tag{187}$$

also

$$I_s = I_r\left(\cos^2\theta + \frac{Z_0''}{Z_0'}\cos^2\theta - \sin^2\theta\right) + j\frac{E_r}{Z_0''}\left(2\sin\theta\cos\theta - \frac{Z_0''\cos^2\theta}{Z_0'\tan\theta}\right) \tag{188}$$

The A, B, C, and D coefficients of the above network are determined by comparing equations 187 and 188 with equations 181 and 182. The iterative impedance is the square root of the B coefficient divided by the C coefficient. It is equal to

$$Z_0 = Z_0''\sqrt{\frac{2 + \dfrac{Z_0''}{Z_0'}}{2 - \left(\dfrac{Z_0''}{Z_0'}\right)\cot^2\theta}} \tag{189}$$

There is no attenuation when the magnitude of the A or D coefficient is less than unity or when

$$\cos^2\theta + \frac{Z_0''}{Z_0'}\cos^2\theta - \sin^2\theta < 1$$

hence

$$\frac{Z_0''}{Z_0'}\cos^2\theta < 2\sin^2\theta$$

and the pass band for the network occurs where

$$|\tan\theta| > \sqrt{\frac{Z_0''}{2Z_0'}} \tag{190}$$

The curve of attenuation vs. frequency has the form shown in Fig. 73. If it is necessary that the filter have just one pass band, two filters can be connected in series which have the same pass band, but whose higher pass bands occur at different frequencies. Other combinations of transmission line either alone or in combination with condensers are used to obtain attenuation and impedance curves of almost any desired character. Many different examples of these networks are given in the Mason and Sykes paper.

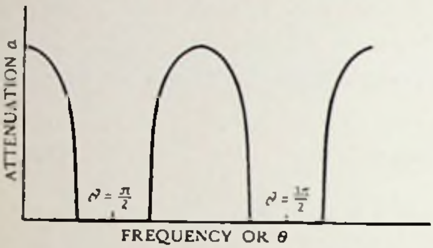


Fig. 73. Attenuation curve for the filter of Fig. 72.

7–5. Transmission-line Resonators. A transmission line can be tuned to a given frequency. The application of Lecher wires for the measurement of wavelength is a typical example. In this case, a parallel-wire transmission line is shorted on both ends. When the length of line between the shorted ends is exactly one-half wavelength long (or a multiple thereof), waves started down the line are successively reflected in exactly the right phase to cause a high build-up of current on each end of the line or a build-up of voltage a quarter wavelength from the shorting bars (i.e., in the middle). In a similar manner, a quarter-wavelength line (or odd multiple thereof) which is shorted on one end and open on the other also exhibits large standing waves when it is in tune. If we assume that the voltage is applied to the open end of this line, the response is very similar to that of a parallel resonant circuit. The input impedance of the line becomes increasingly high as resonance is approached. Also, at resonance the current flowing at the short-circuited end of the line is many times that flowing in at the open end. The line is, therefore, a current multiplier the same as a parallel circuit is a current multiplier.

The Q or figure of merit of a resonant transmission line is defined as the ratio of the reactive power in the line to the watts loss. The distribution of rms current along a tuned line can be assumed as sinusoidal varying from a maximum value of I at the shorted end to practically zero at the open end. If the distributed resistance of the line is R ohms per meter, the watts loss per quarter wavelength of line is equal to

$$P = \left(\frac{1}{\sqrt{2}} I\right)^2 R \frac{\lambda}{4} = \frac{I^2 R \lambda}{8} \tag{191}$$

The reactive power is equal to the reactive voltage multiplied by the current. For a quarter wavelength of line, this is

$$\text{vars} = 2\,\pi f L \left(\frac{1}{\sqrt{2}} I\right)^2 \frac{\lambda}{4} \tag{192}$$

where L is the distributed inductance in henries per meter. The Q of the line is therefore

$$Q = \frac{2\,\pi f L}{R} \tag{193}$$

The Q of a coaxial line is a maximum when the ratio of the conductor radii is $\frac{b}{a} = 3.6$ because the attenuation is then a minimum. The substitution of equations 141 and 143 together with the optimum ratio of radii into equation 193 shows that the maximum Q obtainable with a coaxial line made of copper is

$$Q = 1460 \frac{b \text{ (cm)}}{\sqrt{\lambda} \text{ (meters)}} \tag{194}$$

The fact that the Q of a line varies inversely with wavelength suggests that high-Q lines might well be used to control frequency at the shorter wavelengths where quartz crystals cannot be used. This has been done in a number of installations.

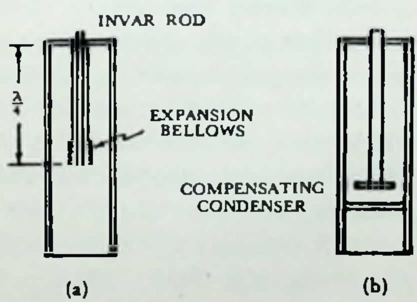


FIG. 74. Typical transmission-line resonators.

The temperature coefficient for frequency drift is closely related to the coefficient of linear expansion for the material of the line because most of the frequency change is due to a change in linear dimension. The easiest way to reduce thermal expansion is to provide the inner conductor of the transmission line with an expansion bellows and maintain the length of the conductor constant by means of an internal invar rod (Fig. 74*a*). The frequency can be quickly changed by changing the position of the invar rod. The expansion of the transmission-line elements can also be used to actuate a compensating condenser on the line as indicated in Fig. 72*b*. The materials used for the inner and outer conductor have a different coefficient of expansion. For example, the inner conductor could be made of copper and the outer conductor made of aluminum. In Fig. 72*b* the line used is less than a quarter of a wavelength long because the capacity on the end resonates with the line which acts like an inductive reactance.

The important points to observe in the design of a resonant transmission line are to make it large enough to dissipate the power it will carry and with a wide enough spacing between conductors to prevent voltage breakdown. Care must also be taken to secure positive contact between the conductors of the line at the short-circuited end, otherwise the oxide film on their surfaces will insulate them from one another. The voltage is a minimum at this end of the line; therefore, it will not break down some of the thinnest films. Usually a soldered connection or a tight clamp is used to secure this contact. When the length of line must be readily adjustable, the shorting plug is provided with spring collars which make contact on the inner and outer conductors at some distance away from the shorting plug and at a point where the voltage is sufficient to break down the film between the edge of the collar and the conductor.

7–6. Cavity Resonators. The action of cavity resonators is similar to that of transmission-line resonators. They differ in that the waves are propagated in the space of the cavity rather than along a conductor. Consider an electromagnetic wave directed toward a conducting metallic sheet. A standing wave would be produced, which has zero electric-field intensity at the surface of the sheet. Another null point would occur one-half wavelength in front of the sheet. A conducting surface could be placed at this point without disturbing the field. Finally, top, bottom, and sides could be added and we would have a typical cavity resonator. The performance of this rectangular box resonator can be calculated exactly by using Maxwell's equations. These results, together with those for a few other simple shapes which can be calculated, are given in Fig. 75.

The formula for Q which is given in the figure is based on the assumption that Q is equal to the ratio of the energy stored in the magnetic field at the peak of the cycle to that lost in the resistance of the resonator material. Roughly, the loss in this material will be directly proportional to the depth of penetration Δ of the "skin

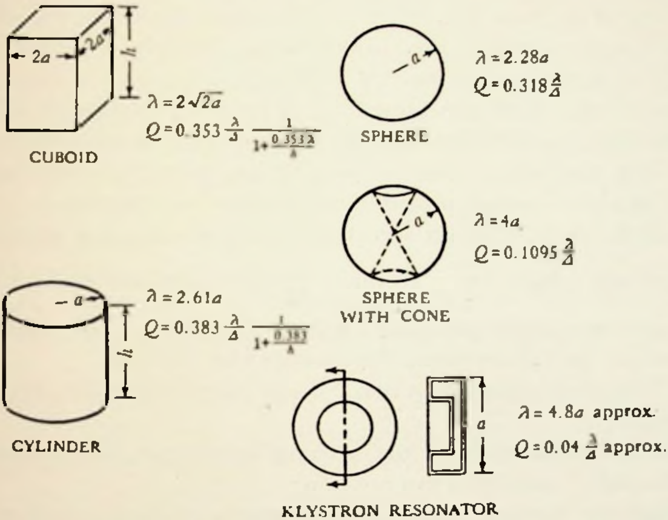


FIG. 75. Cavity resonators.

effect." Hence, Q is inversely proportional to this factor. The wavelength λ comes in as a scaling factor, hence, Q varies directly with it. The ratio $\frac{\lambda}{\Delta}$ is equal to 8.32×10^4 for copper at wavelengths of 10 cm. At this wavelength, the Q's for the resonators of Fig. 75 range from 3000 to 26,500 on the assumption that $h = a$.

The resonant wavelength for shapes which cannot be readily calculated is best determined by building a model, testing it, and then scaling the results down to the desired size.

Problems

7-1. The inside diameter of the outside conductor of a coaxial line is $\frac{1}{2}$ in. and the diameter of the internal conductor is $\frac{1}{8}$ in. The material used is copper.

(*a*) Find the characteristic impedance.
(*b*) Find the attenuation constant.

(*c*) Find the voltage on the transmission line 22 meters from a 73-ohm resistive load if the line delivers 500 watts. The wavelength used is 3 meters.

7–2. Design a transmission-line reactance network with zero reactance at 1090 and 1180 mc and an infinite reactance at 1000, 1125, and 1250 mc. Use 900 mc as the base frequency and 10 ohms for the lowest characteristic impedence of any line used in the circuit. Sketch the wiring diagram.

7–3. The outside conductor of a 77-ohm coaxial line has an inside diameter of $\frac{1}{2}$ in. Find the dimension of an inside conductor which will match the line to a 150-ohm load at 300 mc. Sketch connection.

7–4. Find the length and position of an open-circuited stub which could be used for a solution to the example given in section 7–3.

7–5. Find the lengths (in terms of wavelength λ) of a pair of short-circuited stubs which are a distance $\frac{\lambda}{12}$ apart and are used to match a 50-ohm line to a pure resistance load of 100 ohms. The characteristic impedance of the line between the stubs is also 50 ohms.

7–6. The outside conductor of a coaxial transmission line has an inside diameter of $\frac{1}{2}$ in.

(*a*) Find the diameter of the internal conductor which would give this line a minimum attenuation.

(*b*) Find the location, length, and diameter of the internal conductor of a short-circuited line which would match the above line to a 110-ohm load over a wide band of frequencies when the mid-frequency is 1000 mc. Assume that the same diameter outside conductor is used for both the stub and the line.

(*c*) Compute the approximate frequency band over which a successful match could be expected.

7–7. A transmission-line filter of the type described in section 7–4 is made up of lines having a length $l = 10$ cm. The characteristic impedance $Z_0'' = 77$ ohms, and $Z_0' = 60$ ohms.

(*a*) Find the cutoff frequencies of the first and second pass bands.

(*b*) Compute and plot the iterative impedance through both of these bands.

Experiment 15

Object:

The object of this experiment is to demonstrate resonance and standing waves on a transmission line and to match the impedance of the line to a load.

Preliminary:

Compute the length and position of a short-circuited stub to match a transmission line consisting of two number 18 wires spaced $1\frac{1}{4}$ in. apart[1] to a 100-ohm resistive load.

Performance:

1. Inductively couple a 1-meter oscillator to a short-circuited open-wire transmission line of the dimensions given in the preliminary. Tune the line into resonance by moving one of the shorting bars along the length of the line. Use the maximum deflection of a thermal galvanometer connected to a small pickup loop which is near a short-circuiting bar as the indication of resonance. Then move the loop along the line and plot the reading of the galvanometer, which is proportional to current, as a function of the position of the loop with respect to a short-circuiting bar.

2. Measure the voltage distribution along the resonant line of part 1 by means of a pentode voltmeter.[2] Plot these results on the same axes as those of part 1.

3. Tune the line when it is open circuited on one end and short circuited on the other. Measure the distribution of the current and voltage on the line as a function of the distance from the open end. Plot the results in the same manner as parts 1 and 2.

4. Terminate the line in a noninductive (i.e., carbon) resistance equal to the characteristic impedance. Measure the standing wave ratio (maximum current or voltage divided by minimum current or voltage) and compare with the expected value.

5. Terminate the line in a 100-ohm noninductive resistance. Measure the standing wave ratio and compare with the expected value. Attach

[1] The reader may prefer to use a line which is mechanically sturdier than this for these experiments. In that event, it is suggested that a line consisting of two $\frac{1}{4}$ in. diameter copper tubes or rods spaced 2 in. between centers be used instead of the wires.

[2] A detailed description of the pentode voltmeter is given by King, *loc. cit.* (See References.) Essentially it consists of two pentodes such as 6J7's arranged so that each wire of the transmission line is near one of the grid caps. A high resistance (several megohms) is connected between the grid and cathode of each tube. The plate and screen potentials are applied in the usual manner. Provision is made for reading the change in plate current taken by the two tubes when their grid bias is changed due to the presence of a radio-frequency voltage on the transmission line. The action of the tube is explained as follows: Once during each cycle the high-frequency voltage which is capacitively coupled to the grids drives them positive. During this time, the grids attract electrons from the space charge which is near by and the resultant current which flows through the grid resistor puts a negative bias on the latter. The bias voltage does not have time to leak off between cycles, but it is removed in a relatively short period of time when the strength of the radio-frequency field is changed. Consequently, the change in plate current taken by the tube which is proportional to the change of the grid bias is also proportional to the strength of the radio-frequency field present.

the stub computed in the preliminary and again measure the standing wave ratio. If it is different from unity, adjust the length and position of the stub until a match is obtained. Discuss results.

References

Guilleman, E. A., *Communication Networks*, Vol. 2. New York: John Wiley & Sons, Inc., 1935, pp. 184–221.

Hansell, C. W., and P. S. Carter, "Frequency Control by Low Power Factor Line Circuits," *Proc. I.R.E.*, April 1936, pp. 597–619.

Hansen, W. W., "A Type of Electrical Resonator," *Jour. Applied Phys.*, October 1938, pp. 654–663.

Hansen, W. W., and R. D. Richtmyer, "On Resonators Suitable for Klystron Oscillators," *Jour. Applied Phys.*, March 1939, pp. 189–199.

King, R., "Electrical Measurements at Ultra-high-frequencies," *Proc. I.R.E.*, August 1935, pp. 885–934.

Mason, W. P., and R. A. Sykes, "The Use of Coaxial and Balanced Transmission Lines in Filters and Wide-band Transformers for High Radio Frequencies," *Bell Sys. Tech. Jour.*, July 1937, pp. 275–302.

Ramo, S., "Synthesis of a High Frequency Reactance," *Jour. Applied Phys.*, February 1939, pp. 138–139.

Salzberg, B., "Graphs for Transmission Lines," *Electronics*, January 1942, pp. 47–48.

Slater, J. C., *Microwave Transmission*. New York: McGraw-Hill Book Company, Inc., 1942, pp. 7–78.

Sterba, E. J., and C. B. Feldman, "Transmission Lines for Short-wave Radio Systems," *Proc. I.R.E.*, July 1932, pp. 1163–1202.

CHAPTER VIII

OSCILLATORS

8–1. General. The function of an oscillator is to convert the energy of a power supply into alternating current of a particular frequency. This conversion process can be accomplished by a simple expedient which is suggested by a study of electrical transients. It is well known that an oscillatory current flows from a charged condenser when it is connected across an inductance which has a sufficiently low resistance. If E is the voltage to which the condenser is charged, the current which flows in such a circuit is

$$i = \frac{E}{\omega L} \epsilon^{-\frac{R}{2L}t} \sin \omega t \tag{195}$$

where ω is 2π times the frequency of oscillation f and is equal to

$$\omega = \sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}} \tag{196}$$

R, L, and C are the resistance inductance and capacity of the circuit.

It is apparent that an oscillator could be constructed by an arrangement such that a condenser is alternately connected to a d-c charging voltage and to a discharge inductance. The complication of a mechanical switch for doing this can be avoided by using the impedance properties of a spark gap. The condenser is left connected to the d-c supply and a spark gap is connected in series with the condenser and the attached discharge coil. As the condenser is charged, the voltage across it rises until the breakdown voltage of the gap is exceeded. The arc which is formed as a result suddenly reduces the impedance of the gap from infinity to practically zero; therefore, the coil becomes connected to the condenser. During the oscillatory discharge which follows, the condenser is prevented from receiving additional charge from the power supply by some chokes in the lines connecting these two elements. As the oscillatory transient dies out, the voltage across the spark gap is reduced to the point where it ceases to conduct and again it becomes an open circuit. The condenser receives another charge from the power supply, the gap breaks down, and the cycle is repeated.

Spark-gap oscillators of this sort were used to produce the first radio waves and at one time were the primary source of radio-frequency energy. They rapidly fell into disrepute, however, because the output from them contains a large number of frequencies besides the fundamental (see problem 6–4), and these spurious frequencies interfered with the transmission from other stations in near-by frequency bands. Today, the use of spark-gap oscillators as a source of radio-frequency power for communication is outlawed except for the emergency transmission of distress signals.

It is possible, however, to obtain a constant-amplitude single-frequency output from the above resonant circuit. An examination of equation 195 shows that this result is obtained if the circuit resistance is reduced to zero or neutralized. In this connection, the term "negative resistance" appears quite frequently. A negative resistance is used to neutralize the positive resistance of a resonant circuit. It is present whenever the voltage drop across a portion of the circuit is in the same direction as the current flowing through it or whenever the slope $\left(\frac{\Delta E}{\Delta I}\right)$ of the characteristic curve is a negative quantity. The simplest example of a negative resistance is a generator since the terminal voltage of a generator is in the same direction as the output current. If an a-c generator having a frequency f (equal to the resonant frequency of the circuit) and a generated voltage E were inserted in series with the above circuit, a constant-amplitude alternating current

$$i = \frac{E}{R} \sin 2\pi f t \qquad (197)$$

would flow because this generator would effectively neutralize the circuit resistance between the current limits $\frac{E}{R}$ and $-\frac{E}{R}$.

Any element which will neutralize the resistance of an RLC circuit will maintain the flow of alternating current in that circuit.

Consequently, when a resonant circuit and a resistance-neutralizing element which takes its power from a d-c supply are connected in series, they make up an oscillator. The frequency of the alternating current produced is that frequency at which the total reactance of the circuit is zero, and the amplitude of the current is such that the net resistance of the circuit is zero. In a similar manner, any element connected across a parallel resonant circuit which will neutralize the conductance of the latter, will sustain oscillations in the circuit at a frequency such that the total susceptance of the circuit is zero. All constant-amplitude oscillators depend on the process of resistance neutralization for their operation.

8-2. The Dynatron. The plate resistance or plate conductance of a tetrode vacuum tube is negative throughout a portion of the tube's operating range. An examination of the typical plate characteristic curve given in Fig. 76*a* shows that this is true. The

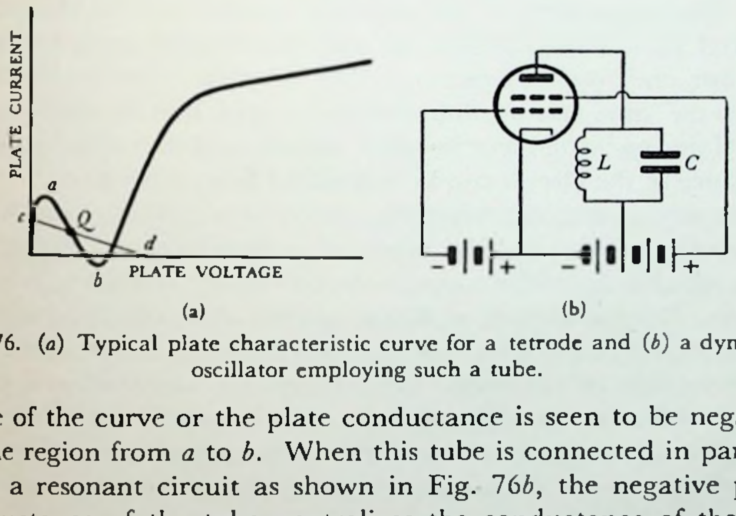


FIG. 76. (*a*) Typical plate characteristic curve for a tetrode and (*b*) a dynatron oscillator employing such a tube.

slope of the curve or the plate conductance is seen to be negative in the region from *a* to *b*. When this tube is connected in parallel with a resonant circuit as shown in Fig. 76*b*, the negative plate conductance of the tube neutralizes the conductance of the resonant circuit and oscillations are sustained in the latter. The combination of the tube and the resonant circuit is called the dynatron oscillator.

The load line *c-d* of the resonant circuit in Fig. 76*b* is a line which passes through the quiescent point *Q* of the tube and has a negative slope equal to the circuit conductance. Oscillations occur whenever the slope of this line is less in magnitude than the negative slope of the plate characteristic curve. When the circuit conductance is lower than the maximum negative conductance of the tube, as it is for the case illustrated in Fig. 76*a*, the magnitude of the oscillation produced increases to the point where the dynamic negative plate conductance is exactly equal to the load conductance. In other words, the plate current swings over such a wide range that the tube is operated on the positive conductance portion of its plate characteristic curve during a part of each cycle. Enough positive conductance is introduced in this manner to reduce the negative conductance of the tube to the magnitude of the resonant circuit conductance. The net conductance of the oscillating circuit is therefore zero. Oscillations stop if the circuit conductance exceeds the maximum negative conductance of the tube.

Because of these phenomena, the dynatron oscillator can be used to measure the radio-frequency resistance of a resonant circuit. The magnitude of the negative conductance of the tube is reduced by making the bias on the control grid more and more negative until oscillations stop. The negative plate conductance which the tube has at that particular grid bias is equal to the radio-frequency conductance of the circuit, and the radio-frequency resistance of the circuit can be computed from this value.

The screen-grid characteristic curve of a pentode also has a negative-resistance portion when a proportional relationship is maintained between the suppressor-grid voltage and the screen-grid voltage. The transitron oscillator utilizes this negative-resistance characteristic in the same manner as a dynatron utilizes the negative resistance of a tetrode. Oscillations are sustained in a tuned circuit which is connected between the screen grid and the supply voltage. The proportionality between the suppressor voltage and the screen voltage is maintained by supplying the suppressor bias through an isolating resistance and coupling the suppressor to the screen grid through a condenser.

8–3. The Tuned-plate Tuned-grid Oscillator. Under certain conditions, the input admittance of a triode has a negative conductance. This negative conductance can also be used to produce oscillations. Referring to Fig. 77, we see that the triode has three capacitances which occupy the positions indicated in the diagram. In the equivalent circuit, the tube itself is replaced by a series combination of a resistance r_p and a generator $-\mu E_g$ connected between the points a and b (also see Fig. 47a). The alternating voltage between these two points is the output voltage E_o. On the assumption that the tube is operated so that the grid does not take

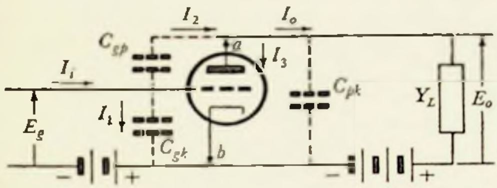


FIG. 77. Diagram of a triode amplifier and the notation used to determine its input admittance.

any current, the currents which would flow in the circuit are those indicated in the figure. The output admittance $Y_o = G_o + jB_o$ consists of the admittance of the load Y_L plus the susceptance of the capacity C_{pk}. The equations of current and voltage are

$$I_2 = I_3 + I_o \tag{198}$$

and
$$E_o = E_g - I_2\left(\frac{-j}{\omega C_{gp}}\right) = I_3 r_p - \mu E_g = \frac{I_o}{G_o + jB_o} \tag{199}$$

Solving for I_3 and I_o in equation 199 and substituting these results in equation 198, we have

$$I_2 = I_2\left(\frac{j}{\omega C_{gp}}\right)(G_o + jB_o) + E_g(G_o + jB_o) + \frac{I_2}{r_p}\left(\frac{j}{\omega C_{gp}}\right) + \frac{(\mu + 1)E_g}{r_p} \tag{200}$$

The admittance at the input terminal of C_{gp} is equal to I_2 divided by E_g; hence, from equation 200, it is

$$Y = \frac{I_2}{E_g} = \frac{G_o + jB_o + \dfrac{\mu + 1}{r_p}}{\left[1 - \dfrac{j}{\omega C_{gp}}(G_o + jB_o) - \dfrac{j}{r_p \omega C_{gp}}\right]} \tag{201}$$

This equation can be simplified by substituting the mutual conductance g_m for $\dfrac{\mu}{r_p}$ and substituting the plate conductance g_p for the reciprocal of plate resistance. Making these substitutions and rationalizing, we see that equation 201 becomes

$$Y = \frac{\omega C_{gp}[(G_o + g_m + g_p) + jB_o][(\omega C_{gp} + B_o) + j(G_o + g_p)]}{(\omega C_{gp} + B_o)^2 + (G_o + g_p)^2} \tag{202}$$

The input admittance of the triode is equal to this admittance plus the susceptance of the capacity C_{gk} or

$$Y_i = \frac{\omega C_{gp}[(G_o + g_m + g_p)\omega C_{gp} + g_m B_o]}{(\omega C_{gp} + B_o)^2 + (G_o + g_p)^2} + \frac{j\omega C_{gp}[B_o(\omega C_{gp} + B_o) + (G_o + g_p)(G_o + g_m + g_p)]}{(\omega C_{gp} + B_o)^2 + (G_o + g_p)^2} + j\omega C_{gk} \tag{203}$$

The conductance of this equation is negative if the output susceptance B_o is negative and

$$|g_m B_o| > |(G_o + g_m + g_p)\omega C_{gp}| \tag{204}$$

In other words, a triode has a negative input conductance when it is connected to a sufficiently large inductive load.

Oscillations are produced when this negative input conductance is used to neutralize the conductance of a tuned circuit which is connected across the input of the tube (Fig. 78). The necessary inductive load is supplied by another tuned circuit which is connected to the plate of the tube and is tuned to the inductive side of resonance. The connection is known as the "tuned-plate tuned-grid oscillator." The magnitude of the oscillations produced is fixed by an equilibrium condition with the tube constants. Larger oscillations change the tube constants in such a direction that the negative input conductance is reduced to the point where it is exactly equal to the conductance of the tuned circuit.

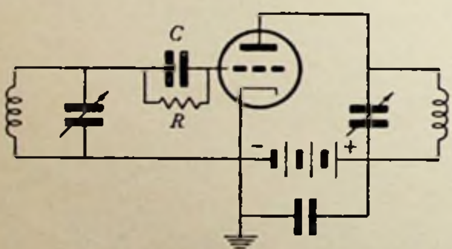


FIG. 78. Tuned-plate tuned-grid oscillator.

The best performance is secured when the tube is operated in class C, but no oscillations would start if a fixed bias this large were used. Consequently, an automatic bias such as the combination of the resistance R and the capacity C in Fig. 78 is generally employed. During a portion of each cycle, the grid is driven positive and conducts a current through the resistance R. The resultant voltage drop charges the condenser C, and, due to the slow discharge of the condenser, the bias is held on the grid for the remainder of the cycle. When the tube is first put into operation, the bias is zero and the circuit readily begins to oscillate. As soon as the oscillations finish building up to the equilibrium magnitude, the proper operating bias is supplied to the tube.

The instantaneous voltage developed across the resistance-capacity combination on the peak of the last build-up cycle is sufficient to stop the resistance-neutralization process if an appreciable portion of this voltage is not removed by leakage before the arrival of the next peak on the oscillation cycle. This happens when the time constant RC is too large. If, under these circumstances, the Q of the tuned circuit including its load is small, the oscillations will die out rapidly and will not start again until the bias produced by the resistance-capacity combination has fallen to a sufficiently low value. The oscillations then build up and stop again as before. Intermittent operation of this sort can be avoided by making the time constant of the bias combination smaller.

Another phenomenon known as "blocking" is occasionally observed in the operation of a self-biased oscillator. Oscillations stop suddenly and the grid current is reversed due to secondary emission from the grid. The reversal of current biases the grid positively and causes the plate to draw such an excessive current that the tube is burned up in the process. Blocking can be avoided by conservative design.

The connections for a tuned-plate tuned-grid oscillator are the same as those of a radio-frequency amplifier. Since self-sustained oscillations are undesirable in the latter, precautions must be taken to prevent their occurrence. An examination of equation 203 shows that the magnitude of the negative input conductance is directly proportional to ωC_{gp}. The negative input conductance causes the oscillations; therefore, they can be eliminated by reducing it to a value which is less than the conductance of the input tuned circuit. In other words, the grid-to-plate capacity of the tube must be reduced. The simplest way to do this is to put a screen grid in the tube. For that reason, screen-grid tubes are generally employed in radio-frequency amplifiers. The capacity can also be neutralized by the use of a circuit such as that shown in Fig. 79. The current flowing in the capacity C_n is 180° out of

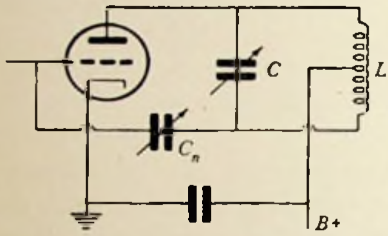


FIG. 79. A typical circuit for neutralizing the effect of the grid-to-plate capacity in a radio-frequency amplifier tube.

phase with the current flowing in grid-to-plate capacity C_{gp} of the tube. If the reactance of C_n is adjusted so that these two currents are equal, the net current flowing between the input and output side of the tube is reduced to zero. That is to say, the effect of the grid-to-plate capacity on the operation of the circuit is eliminated.

The adjustment of the neutralizing condenser is quite simple. A radio-frequency voltage is applied to the grid, and both the input and the output circuits are tuned to resonance. The B supply to the amplifier tube is disconnected and the tuning again adjusted for resonance. C_n is then adjusted to the point where no alternating current flows in the output resonant circuit and consequently no alternating voltage appears across it.

The usual crystal oscillator is an adaptation of the tuned-grid tuned-plate oscillator. A typical circuit diagram is shown in Fig. 80. A quartz crystal has an equivalent electrical circuit which

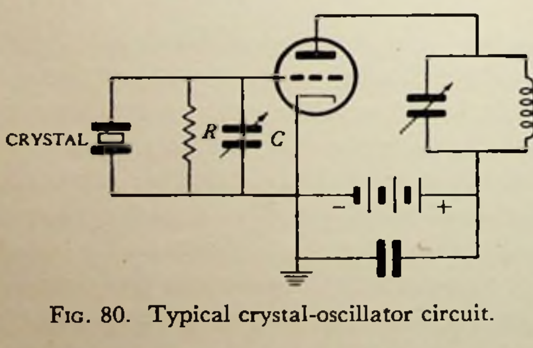


FIG. 80. Typical crystal-oscillator circuit.

corresponds to the ordinary inductance- and capacity-tuned circuit. As a matter of fact, the equivalent inductance, capacity, and resistance of a crystal can be computed from its dimensions. The operation of the circuit shown in Fig. 80 is, therefore, the same as the operation of the circuit shown in Fig. 78. The resistance R must be connected as shown to provide a conduction path between the grid and the cathode of the tube, and it is also used in conjunction with the capacity C to supply the grid bias in the same manner as before.

Quartz-crystal oscillators are very important due to their high-frequency stability. Since the frequency of the oscillations produced by a tuned-grid tuned-plate oscillator is determined by the resonant frequency of the circuit which is connected to the grid, a high-frequency stability is obtained when this circuit has a high Q and is negligibly affected by external influences. Both of these requirements are fulfilled by a quartz crystal. Ultra-high frequencies can be produced with crystal oscillators by operating them on one of the harmonics of the crystal.

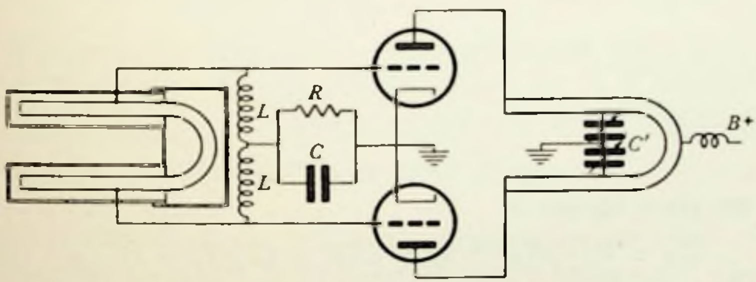


Fig. 81. Push-pull oscillator circuit which uses transmission lines for resonant elements and is intended for the generation of ultra-high-frequency.

Figure 81 shows an ultra-high-frequency version of the ordinary tuned-grid tuned-plate oscillator. Transmission lines are used for the resonant circuits. In this particular example, the usual single tube has been replaced by a pair of tubes which are operated in push-pull. High-frequency stability is obtained by using a high-Q line on the grid circuit. The shorted transmission line on the output side is tuned to the inductive side of resonance by the ganged condenser C'. A resistance capacity of combination provides the grid bias automatically the same as in the other examples.

8–4. The Colpitts Oscillator. Resistance neutralization can also be produced by other vacuum-tube connections. An example of one of these is the Colpitts circuit which is illustrated in Fig. 82. If we consider the circuit as it appears to the high-frequency currents and neglect the d-c power supply for the time being, the

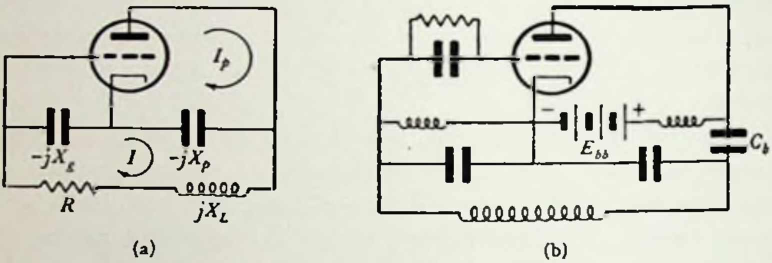


FIG. 82. (*a*) The high-frequency equivalent circuit of the Colpitts oscillator and (*b*) a typical diagram of the actual connections of such an oscillator.

connection takes the form shown in Fig. 82*a*. The alternating plate current I_p is given by equation 97 or

$$I_p = g_m E_g + g_p E_p$$

where E_g is the alternating voltage of the grid with respect to the cathode and E_p is the alternating voltage of the plate with respect to the cathode. The grid voltage is equal to

$$E_g = -I(-jX_g) = jIX_g \tag{205}$$

and the plate voltage is

$$E_p = I(-jX_p) - I_p(-jX_p) = jX_p(I_p - I) \tag{206}$$

The substitution of equations 205 and 206 into the equation for the plate current yields

$$I_p = jg_m IX_g + jg_p X_p(I_p - I) \tag{207}$$

Solving this equation for the plate current, we have

$$\begin{aligned} I_p &= \frac{jI(g_m X_g - g_p X_p)}{1 - jg_p X_p} \\ &= I\left[\frac{g_p X_p(g_p X_p - g_m X_g) + j(g_m X_g - g_p X_p)}{1 + (g_p X_p)^2}\right] \end{aligned} \tag{208}$$

The total voltage drop around the tuned circuit is E and is equal to

$$E = I[R + jX_L - j(X_p + X_g)] - I_p(-jX_p) \tag{209}$$

Substituting equation 208 into equation 209, we have

$$E = I[R + jX_L - j(X_p + X_g)] + jX_pI\left[\frac{g_pX_p(g_pX_p - g_mX_g) + j(g_mX_g - g_pX_p)}{1 + (g_pX_p)^2}\right] \tag{210}$$

But the net impedance Z of the tuned circuit is equal to the total voltage drop divided by the current or

$$Z = \frac{E}{I} = R + \frac{X_p(g_pX_p - g_mX_g)}{1 + (g_pX_p)^2} + j\left[X_L - (X_g + X_p) + \frac{X_p(g_pX_p)(g_pX_p - g_mX_g)}{1 + (g_pX_p)^2}\right] \tag{211}$$

The resistance of this circuit reduces to zero when

$$R + \frac{g_pX_p^2}{1 + (g_pX_p)^2} = \frac{g_mX_gX_p}{1 + (g_pX_p)^2} \tag{212}$$

and constant amplitude oscillations will be maintained in it.

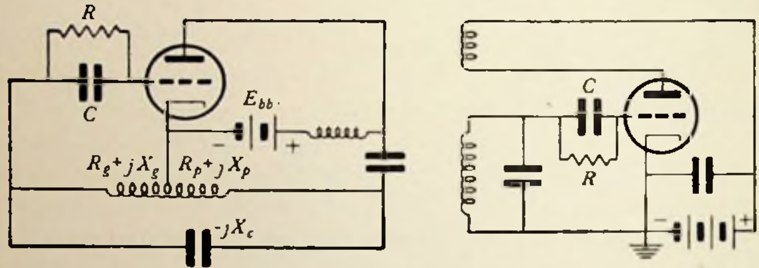
The frequency of these oscillations is the frequency at which the total reactance of the circuit reduces to zero. It is approximately equal to the frequency at which the capacitive reactance of the circuit $(X_g + X_p)$ is equal to the inductive reactance of the circuit X_L; however, equation 211 shows that the exact frequency is actually a function of the tube constants. The approximate expression is nevertheless used in making engineering calculations because a solution of equation 211 in terms of the generated frequency is not practical.

The amplitude of the oscillations increases to the point where the dynamic constants of the tube acquire a value such that the real term of equation 211 is reduced to zero.

As yet nothing has been said about the manner of supplying power to the tube. One method is to connect the plate supply voltage between the cathode and plate through a radio-frequency choke as indicated in Fig. 82*b*. This is called a shunt feed because the power is supplied in parallel with the tube and the plate load. The choke is necessary to prevent the radio-frequency current from being shorted out through the power supply. Another radio-frequency choke provides the necessary d-c path between the cathode and the grid, and the grid bias is supplied by the usual condenser and resistance combination. A blocking condenser C_b

must also be added to avoid the short circuit to the power supply which would otherwise exist through the inductance of the resonant circuit and the radio-frequency choke between the cathode and the grid.

8–5. The Hartley Oscillator and Others. A typical circuit diagram of a Hartley oscillator is shown in Fig. 83*a*. An analysis similar to the above shows that this connection is also capable of neutralizing the resistance of a tuned circuit. If R_g and X_g are the



(a) (b)

Fig. 83. (*a*) Hartley oscillator. (*b*) Tuned-grid oscillator.

resistance and reactance of the portion of the coil between the grid and the cathode, and if R_p and X_p are the resistance and reactance of the remainder of the coil, the net resistance R_T of the tuned circuit is found to be

$$R_T = R_g + R_p + \frac{[g_m R_g - g_p R_p][R_p + g_p(R_p^2 + X_p^2)] - X_p[g_m X_g - g_p X_p]}{(1 + g_p R_p)^2 + (g_p X_p)^2} \tag{213}$$

This equation can be reduced to zero; therefore, oscillations are maintained in the tuned circuit. The frequency of the oscillations which are produced is that frequency at which the net reactance of the circuit X_T becomes zero. This reactance is

$$X_T = X_g + X_p - X_c + \frac{[g_m X_g - g_p X_p][R_p + g_p(R_p^2 + X_p^2)] + X_p[g_m R_g - g_p R_p]}{(1 + g_p R_p)^2 + (g_p X_p)^2} \tag{214}$$

Again the frequency produced depends on the tube parameters, but it is approximately equal to the frequency at which the inductive reactance of the circuit $(X_g + X_p)$ is equal to the capacitive reactance of the circuit X_c.

There are many other connections of a vacuum tube which will produce resistance neutralization and consequent oscillations. An example of one of these is the tuned-grid oscillator shown in Fig. 83*b*. With nearly all of the connections, the frequency of the oscillations produced depends to some extent on the tube parameters. Oscillators have been designed, however, in which the frequency produced is independent of the tube characteristics. A discussion of them is given by Llewellyn in the *Proc. I.R.E.*, December 1931.

The efficiency of an oscillator is defined as the ratio of the alternating power output to the d-c power input. The calculations are made in the same way as are those for the efficiency of a radio-frequency amplifier. An efficiency of 80 per cent is not uncommon for a properly adjusted oscillator operating in the medium frequency range.

8-6. U.H.F. Negative-grid Oscillators. The negative-grid oscillators which have been considered thus far must be modified for operation at ultra-high frequencies. Transmission lines are used as the resonating elements rather than coils and condensers. One possible connection has already been considered in section 8-3. Usually, however, a tuned transmission line is simply connected between the plate and the grid of the triode. The active length of the line is adjusted by the position of a shorting condenser C such as that illustrated in Fig. 84. An additional shorting condenser is placed a quarter of a wavelength from C to provide a high-impedance path between the resonant section of line and the cathode so that the radio-frequency current will not be shorted out by the power supply and lost. Sometimes this high-impedance section of line is placed in the filament leads rather than in the plate and grid leads.

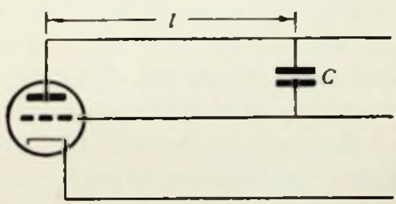


FIG. 84. Transmission-line oscillator.

An ordinary shorting bar can be used to adjust the line length instead if the d-c potential of the plate is isolated from the grid by coupling it to the line through a blocking condenser.

The loss due to the resistance of the leads is reduced considerably by the use of a double-ended tube. A resonant line is connected to the plate and grid terminals on the top of the tube and another line is connected to the extension of the terminals which comes through the bottom of the tube. Since the current in each of these lines is half what it would be with a single line which handled the same power, the loss in the leads is correspondingly reduced. The radiation loss is also decreased for the same reason.

Regardless of the specific physical connections, however, the equivalent circuit of this transmission-line oscillator takes the form indicated in Fig. 85. The positions of the plate, grid, and cathode are marked by the letters P, G, and K respectively. It will be noted that this circuit is very similar to the Colpitts circuit of Fig. 82*a*. The major difference is that the impedance $R + jX_L$ has been replaced by the parallel combination of the interelectrode capacity C_{gp} and the shorted transmission line. The tube introduces a negative resistance into the circuit by a process which is similar to that of the Colpitts oscillator.

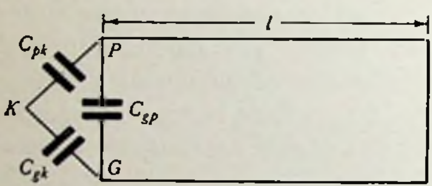


Fig. 85. Equivalent circuit of the transmission-line oscillator in Fig. 84.

The generated frequency is approximately equal to the frequency at which the equivalent interelectrode capacity of the tube resonates with the impedance of the line. The equivalent capacity of the tube is the grid-to-plate capacity C_{gp} in parallel with the series combination of grid-to-cathode capacitance C_{gk} and plate-to-cathode capacitance C_{pk}, and it is equal to

$$C = C_{gp} + \frac{C_{gk}C_{pk}}{C_{gk} + C_{pk}} \tag{215}$$

From equation 158, the impedance of a shorted length l of transmission line is $Z_0 \tan \frac{2\pi l}{\lambda}$; therefore, the generated wavelength λ is given by

$$\frac{1}{2\pi f C} = \frac{\lambda}{2\pi c C} = Z_0 \tan \frac{2\pi l}{\lambda} \tag{216}$$

where c is the velocity of light. A plot of λ as a function of the line length l is given in Fig. 86. For small values of l, $\tan \frac{2\pi l}{\lambda}$ is practically equal to $\frac{2\pi l}{\lambda}$. The substitution of this approximate relationship into equation 216 shows that the curve of Fig. 86 is parabolic in shape near the origin. On the other hand, when l is very

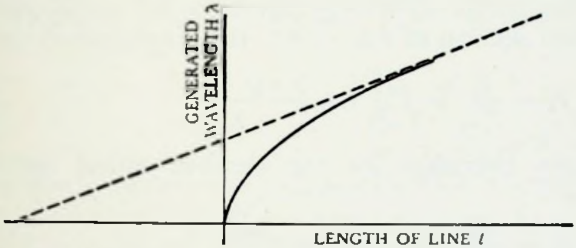
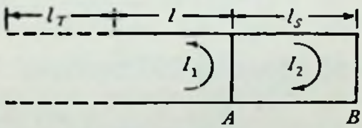


FIG. 86. Generated wavelength characteristic of the oscillator shown in Fig. 84.

large, the effect of the tube capacity is negligibly small in comparison with the impedance of the line. In this region, the wavelength produced becomes practically equal to the resonant wavelength of the line and is a linear function of the length l.

Further analysis of the wavelength characteristic of this transmission-line oscillator gives an insight to the wavelength characteristics of some of the other ultra-high-frequency oscillators. Consider the case where another shorting bar or condenser is placed on the line a distance l_s from the primary shorting bar or condenser and is followed by the usual quarter wavelength short-circuited section of line for isolation purposes. The equivalent

FIG. 87. Equivalent circuit of the oscillator in Fig. 84 when an additional condenser is shorted across the line.



circuit then takes the form shown in Fig. 87. In this diagram, the effect of the tube capacity has been replaced by an equivalent length of line l_T. The total length $(l_T + l)$ is one-quarter of the wavelength which is generated when the tube is operated without a secondary shorting bar. The primary shorting bar is at A and the secondary shorting bar is at B. The primary shorting bar sets up a mutual inductance M between the two circuits; therefore, a voltage E will be developed across it.

$$E = j\,2\,\pi f M(-\,I_1 - I_2) \tag{217}$$

This voltage can be considered to be the sending-end voltage for the two sections of line. One is the open-circuited section which extends to the left, and the other is the short-circuited section which extends to the right. From the equation for the impedance of an open-circuited section of line (159), the sending-end current I_1 is

$$I_1 = \frac{E}{Z_{oc}} = j\,\frac{E}{Z_0}\tan\frac{2\,\pi(l + l_T)}{\lambda} \tag{218}$$

and from the equation for the short-circuited line (158), the current I_2 is

$$I_2 = \frac{E}{Z_{sc}} = -\,j\,\frac{E}{Z_0}\cot\frac{2\,\pi l_s}{\lambda} \tag{219}$$

The substitution of equations 218 and 219 into equation 217 yields

$$1 = \frac{j\,2\,\pi f M}{Z_0}\left(-\,j\tan\frac{2\,\pi(l + l_T)}{\lambda} + j\cot\frac{2\,\pi l_s}{\lambda}\right)$$

$$\frac{Z_0\lambda}{2\,\pi M c} = \tan\frac{2\,\pi(l + l_T)}{\lambda} - \cot\frac{2\,\pi l_s}{\lambda} \tag{220}$$

where λ is the wavelength corresponding to the frequency f, and c is the velocity of light. If the mutual inductance M is assumed to be equal to the distributed inductance of the line L multiplied by a constant K,

$$\frac{Z_0\lambda}{2\,\pi M c} = \frac{\sqrt{\frac{L}{C}}\,\lambda}{2\,\pi K L c} = \frac{\lambda}{2\,\pi K} \tag{221}$$

and equation 220 becomes

$$\frac{\lambda}{2\,\pi K} = \tan\frac{2\,\pi(l + l_T)}{\lambda} - \cot\frac{2\,\pi l_s}{\lambda} \tag{222}$$

This equation is plotted in Fig. 88 for an assumed value of K.

The circuit resistance characteristic which is associated with curve 1 is greater than the maximum negative resistance which the tube is capable of developing except in the range between a and b. Similarly, the circuit resistance characteristic associated with curve 2 is greater than the maximum negative resistance of the tube except in the range between c and d. Consequently, oscillations are produced only along the solid portions of the curves. It will be noted that the length of the secondary line changes the generated wavelength slightly about a value which is established by the length of the primary line. There are two possible generated wavelengths in the neighborhood of the points where the secondary line is a multiple of a half wavelength long.

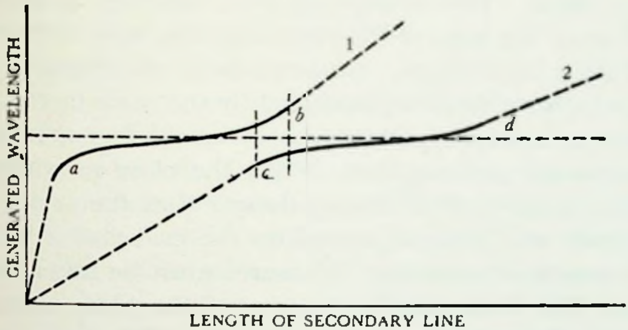


FIG. 88. Generated wavelength characteristic when a secondary line is added to the oscillator of Fig. 84.

As the length of the secondary line is gradually increased, the generated wavelength follows curve 1 to the end of the solid portion and then suddenly switches to curve 2. If the length of the secondary line is decreased from this point or any other point on curve 2, the generated wavelength follows that curve to the end of the solid portion and then suddenly switches to curve 1.

As a transmission-line oscillator is adjusted for higher and higher frequencies, a point is eventually reached where the oscillations cease. This effect is due to either a circuit limitation such as a greater shunting-current loss through the tube's interelectrode capacities, or the transit time of the electrons from cathode to plate becoming an appreciable portion of a cycle, or a combination of both of these effects. Higher frequencies are obtained by reducing the tube's interelectrode capacity and transit distance.

Compromise is necessary in the design of ultra-high-frequency triodes because the reduction of circuit limitations conflicts with the reduction of transit time. For example, low interelectrode capacity requires a large electrode spacing, but a large spacing means a longer transit time. If the interelectrode capacity is reduced by decreasing the size of the electrodes, then the power output of the tube becomes limited. When all of the linear dimensions of a given tube are reduced by the factor $\frac{1}{n}$ and the same operating potentials are used, the tube constants μ, r_p, and g_m are unchanged. Under these conditions, the electrode currents also remain the same. The current density, however, is increased by a factor n^2 since the area of the electrodes has been reduced by the square of their dimensions. Consequently, the emission per unit area of the cathode must be increased by the same factor. For this reason, the saturation current density establishes a fundamental limitation on the performance. With the close proximity of the cathode and grid, there is always danger that the active material of the cathode will become coated on the grid and it in turn will become a source of emission. Measures must be taken to prevent this effect. The reduction in the size of the plate means that it must handle n^2 more power per unit area. This factor also tends to limit the performance of the triode, but the power-dissipation problem can ordinarily be solved more easily than the emission problem.

Physically, the ultra-high-frequency triode is small in size. The WE-368A is a typical example. A triangular block with ample area for heat dissipation is used for the plate. It faces into one side of the filament and grid structure. The plate is supported by its lead which extends straight through both sides of the enclosing glass bulb. The grid, too, is attached directly to its lead which in turn acts as a support and terminates on both sides of the bulb. A straight single-wire filament passes through the center of the grid structure.

8–7. Positive-grid Oscillators. Barkhausen and Kurz showed that very short wavelengths could be produced by operating the grid of an ordinary triode with cylindrical symmetry at a positive potential with respect to cathode while keeping the plate at a zero or slightly negative potential. Some of the electrons which are accelerated toward the grid pass through it, and the field on the other side causes them to reverse the direction of their motion and swing back. After several such cycles the grid finally picks them up. This back-and-forth movement of the electrons through the grid structure induces charges on the cathode and plate which tend to coordinate their actions and make them move as groups rather than singly. The whole group moving together couples the energy of their motion to an external circuit through the capacity of the tube and constitutes a source of high-frequency power. The wavelength produced by the electronic oscillation is approximately

$$\lambda = \frac{1000\, d}{\sqrt{E_g}} \text{ cm}$$

where d is the diameter of the plate in centimeters and E_g is the voltage on the grid in volts.

The distance between the plate and grid terminals of the tube and the first shorting condenser across the transmission line supplying them exerts a certain amount of control over the generated wavelength. The effect of this line length is similar to the effect of the secondary line length on the performance of the negative-grid oscillator. The mass movement of the electrons in the tube acts like the primary or oscillating circuit of Fig. 87 and the wavelength is established by it. The attached line acts like the secondary line of Fig. 87. Figure 88 is, therefore, a typical wavelength characteristic curve for a positive-grid oscillator. These curves are also typical of all oscillators in which the generated wavelength is established by the natural period of an electronic motion. Some magnetron oscillators are of this type.

8-8. Magnetron. The theory of the magnetron is intimately associated with the motion of an electron in a combined magnetic and electrostatic field. A preliminary study of this motion is, therefore, essential.

It has already been shown by equation 28 that the deflecting force F on an electron passing through a magnetic field is

$$F = \frac{mv_x^2}{R} = \frac{Bev_x}{10} \text{ dynes}$$

where v_x is the velocity (cm/sec) with which the electron enters the field, B is the field strength (lines/cm²), and R the radius (cm) of the circular path which the electron follows. The charge and mass of the electron are e and m respectively. From the above equation, the angular velocity ω of this circular motion is found to be

$$\omega = \frac{v_x}{R} = \frac{Be}{10\,m} \text{ radians/sec} \tag{223}$$

In a crossed electrostatic and magnetic field, many different types of motion are possible. An electron moving toward the right along the x axis of Fig. 89 experiences both an upward vertical force $\frac{Bev_x}{10}$ dynes due to the magnetic field and a downward vertical force $\mathcal{E}e \times 10^7$ dynes due to the electrostatic field. When these two forces are equal, the mutual effect of the two fields reduces to zero and the path of the electron remains unchanged. The velocity at which this takes place is found by equating the two forces and solving for v_x or

$$v_x = \frac{\mathcal{E}}{B} \times 10^8 \text{ cm/sec} \tag{224}$$

The general solution of an electron's motion in crossed fields is greatly simplified by an application of this knowledge. Regardless of what the velocity of the electron might be, it is resolved into a component along the x axis equal to $\frac{\mathcal{E}}{B} \times 10^8$ and another component V. The first component of velocity reduces net effect of the two fields on the electron's motion to zero and it gives rise to a pure

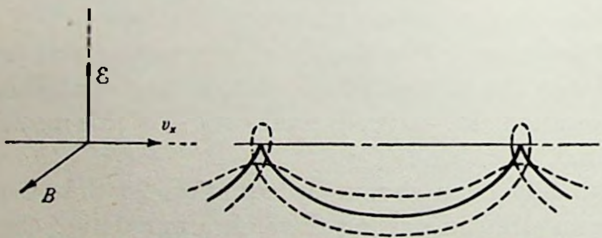


FIG. 89. Motion of an electron in crossed electrostatic and magnetic fields.

translational motion. The second component V gives the electron a circular motion which is due to the action of the magnetic field alone because it is the only field whose effect on the electron is a function of the electron's velocity. The general motion in a crossed electric and magnetic field is, therefore, a combination of circular and linear motion. In other words, it is a cycloid.

Example: Find the motion of an electron which enters a crossed field with zero velocity. The magnetic field intensity is B and the electric field intensity is $\mathcal{E}$.

The zero initial velocity is resolved into the components $\frac{\mathcal{E}}{B} \times 10^8$ cm/sec and $-\frac{\mathcal{E}}{B} \times 10^8$ cm/sec by the application of equation 224. The component $-\frac{\mathcal{E}}{B} \times 10^8$ reacts with the magnetic field and causes the electron to move in a circular path of radius

$$R = \frac{V}{\omega} = \frac{\mathcal{E}}{B} \times 10^8 \times \frac{10\,m}{Be}$$

The center of the circle moves to the right with a velocity $\frac{\mathcal{E}}{B} \times 10^8$, the translatorial velocity referred to above. Since the electron advances a distance $2\,\pi R$ in the same period of time as it takes to complete one revolution on the circular path, the net motion is a common cycloid. Other initial velocities give cycloidal paths of the other forms illustrated in Fig. 89.

In 1921, Hull discovered the magnetron, a tube which uses a magnetic field to control its anode current. The tube is essentially a cylindrical diode with a magnetic field directed along its axis

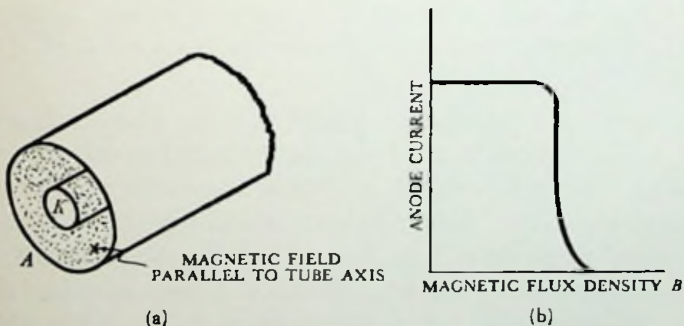


FIG. 90. (a) Hull magnetron (K = cathode and A = anode). (b) Characteristic curve of the Hull magnetron.

(Fig. 90a). Since the electrons leave the cathode with virtually zero velocity, their path is a common cycloid. (This assumes that the radial electrostatic field is uniform and that there is no space charge. Neither of these assumptions is strictly correct, but the error introduced is negligible.) The maximum distance the electrons can wander from the cathode is $2R$, a factor which varies inversely as the square of the magnetic field present. By adjusting the value of this field, the electrons can be made to either reach or miss the anode; consequently, the curve of anode current vs. magnetic flux density shows that the current is constant at the saturation value until the critical flux density is reached and then it drops suddenly to zero (Fig. 90*b*). The cutoff occurs when the electrons begin to just graze the anode or when $2R$ is equal to the distance between the cathode and the plate. The actual cutoff is not so sharp as the theory predicts because all of the electrons do not leave the cathode with zero velocity and no practical tubes have perfect symmetry.

In 1924, Habann introduced the negative-resistance magnetron oscillator. The Habann tube has the anode of the Hull magnetron

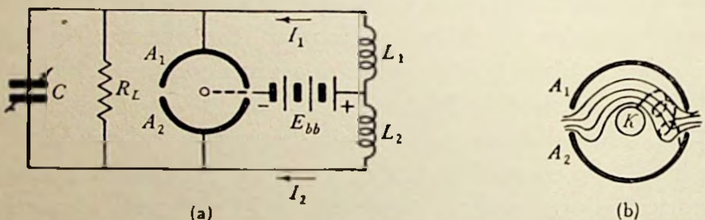


FIG. 91. Habann magnetron oscillator. (*a*) Circuit diagram and (*b*) path of an electron in the tube. The tank circuit consists of L_1, L_2, and C. R_L is the load on the oscillator.

split into two sections (A_1 and A_2 of Fig. 91). The combined electrostatic and magnetic fields force the electrons to do work against the applied electrostatic field, and thus cause this tube to act like a negative resistance. Assume that a radio-frequency voltage has been established in the circuit shown. During the course of a cycle, the upper plate will be more positive than the lower one and the equipotential lines will take the distribution shown in Fig. 91*b*. The potential gradient is relatively high on the upper side and relatively low on the under side of the tube for the instant of time considered. The radius of curvature of the electron's path will be greater on the upper side than on the lower. The whole path is bent toward the negative plate so that the electrons are attracted to it rather than the more positive plate. Since this makes them do work against the electrostatic field, the tube acts like a negative resistance in the circuit and sustains oscillations. The transit time in these tubes is large because the electrons must make several loops before they reach the plate; consequently, the Habann oscillator is only suitable for the generation of relatively low frequencies.

The negative-resistance property of these tubes can also be demonstrated from their static characteristic curve which is shown in Fig. 92. It is taken by measuring the current to each anode as the

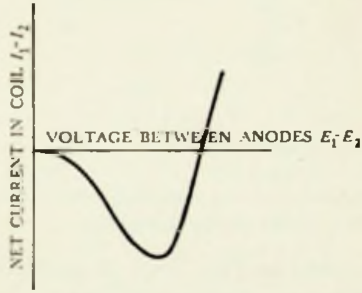


FIG. 92. Static characteristic curve of Habann oscillator. E_1 is the voltage between the cathode and A_1. E_2 is the voltage between the cathode and A_2 (Fig. 91*a*).

voltage on one is decreased in the same increments as the other is increased. The difference between these two voltages corresponds to the voltage which normally appears across the tank of the oscillator circuit. The anode-current difference corresponds to the tank current. The plot of one of these differences against the other, therefore, gives an index to the actual performance of the tube in operation. The current difference has a negative slope for quite a range of voltage differences. This shows that the tank current flows in the same direction as the tank voltage, or that the tube acts like a negative resistance and would sustain oscillations.

A Habann oscillator is put into operation by increasing the magnetic field strength slightly beyond critical so that the electrons will just miss the plate. The tube will then begin to oscillate and the magnetic field can be increased to a much higher and not particularly critical value.

Higher frequencies can be obtained from other forms of the magnetron which take advantage of the natural period of the electrons' motion. A single-anode magnetron connected to a circuit which resonates at a frequency

$$f = \frac{\omega}{2\pi} = \frac{Be}{20\,\pi m} = 2.79 \times 10^6\,B \text{ cps} \tag{225}$$

is an example of this sort. In pursuing their cycloidal paths, the electrons move up and down from the surface of the cathode at this frequency. Electrons which are in the right phase relationship with the radio-frequency field between the plate and cathode will give up more and more of their energy to the field with each successive cycle and thus maintain oscillations. Their motion is illus-

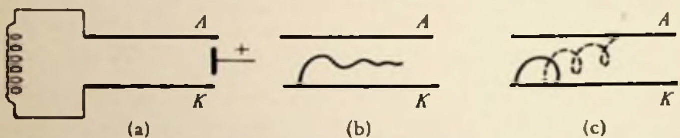


FIG. 93. Typical motion of electrons between the cathode K and the anode A of a magnetron when it is tuned to the natural frequency of the electrons' motion. The dotted line in (c) is the rest of the path which theoretically would have been followed by the electron were it not for the fact that it collides with the cathode at the end of the first cycle.

trated in Fig. 93b. A positive end-plate is used to remove the dead electrons from the field so that they will not accumulate as a space charge, or the same result can be achieved by tilting the magnetic field a few degrees from the axis. Since the electron motion is random, part of them are in phase with the field and are accelerated by it, absorbing energy rather than giving it up. As shown in Fig. 93c, these electrons fall back into the cathode at the end of their first cycle and are stopped; consequently, they do not constitute a serious loss. The tube is tuned by varying the active length of a short-circuited transmission line which is connected between the cathode and anode of the tube.

FIG. 94. Reversed cyclotron.

A cyclotron operated in reverse is another form of the natural-frequency magnetron oscillator. D-shaped electrodes form part of a resonant circuit and are mounted perpendicular to a magnetic field (Fig. 94). High-velocity electrons, introduced at the outer edges of the D's, start to move in a circular path due to the magnetic field. If these electrons are in the right phase relationship as they go through the space between the D's, they give up some of their energy to the radio-frequency field and tend to maintain oscillations. The complete path of such electrons is a decreasing spiral. An electrode in the center picks them up after they have lost all of their kinetic energy. As before, the circuit is tuned to resonate at $f = 2.79 \times 10^6\ B$ cps. Electrons which are in the wrong phase relationship travel in increasing spirals and are lost almost immediately by hitting the outer edge of the D's.

The main objection to the types of magnetrons thus far considered is that their physical size decreases very materially as their frequency increases. This difficulty is alleviated to some extent in a multisegment magnetron which utilizes the period of time required

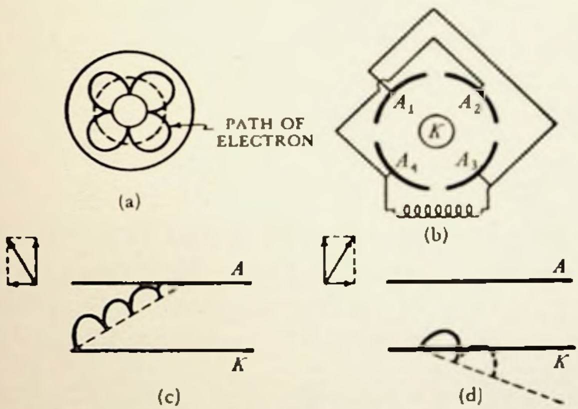


FIG. 95. Multisegment magnetron. (*a*) Path of the electron around the cathode. (*b*) Circuit connection for the magnetron. (*c*) Motion of the electrons which are in the proper phase relationship. (*d*) Motion of the electrons which are in the improper phase relationship.

for an electron to complete its journey around the cathode (Fig. 95*a*). The action of the tube can be likened to that of a polyphase motor. Consider a four-segment device for illustration (Fig. 95*b*). The anode segments are used as the capacity of the circuit and are so arranged that the adjacent plates are of opposite polarity. An electron which travels around the cathode at the proper speed and phase relationship will add its energy to the radio-frequency field as it passes through the field between the segments. The reason for this is that the field between successive segments changes its direction just before the electron comes along. Consequently, the electron loses some of its kinetic energy to the field every time it passes a region between segments. The resultant path is given in Fig. 95*c*. The vector diagram shows that the decelerating force between segments gradually moves the electron toward the anode. Again, the motion of some of the electrons will not be in the proper phase relationship with the radio-frequency field. They will be accelerated between segments. The vector diagram of their resultant motion is shown in Fig. 95*d*. This shows that the power-absorbing electrons will be lost at the end of their first cycle by falling back into the cathode and will not subtract materially from the radio-frequency power. The period of time required for a given electron to circle the cathode once is equal to the distance which the center of the circle producing the cycloidal path travels in one trip around the cathode divided by the translational velocity, or

$$T = \frac{2\pi(r + R)}{\frac{\mathcal{E}}{B} \times 10^8} \tag{226}$$

where r is the radius of the cathode and R is the radius of the electron motion in tracing the cycloid. The generated frequency is equal to the reciprocal of the time required for the electron to pass two anode segments; therefore, the generated frequency of a tube having N anode segments is

$$f = \frac{N}{4\pi(r + R)} \frac{\mathcal{E}}{B} \times 10^8 \tag{227}$$

The corresponding generated wavelength is found by dividing the velocity of light (3×10^{10} cm/sec) by the generated frequency.

The bombardment of the filament by the out-of-phase electrons may cause it to melt unless precautions are taken to avoid this difficulty. One solution to the problem is to provide the filament power supply with a regulator which reduces the power input to the filament as its temperature tends to increase. Back heating, as it is called, can also be avoided by the use of special tube constructions. For example, the usual filament can be replaced by a beam of electrons which are shot into the anode space by an electron gun mounted at the end of the cylindrical structure. The electrons enter the oscillation space through a hole in an end plate which separates the electron gun from the anode. The electrons which would otherwise cause back heating are picked up by this plate. This construction also offers an additional advantage. A grid can be mounted in the electron stream and the strength of the oscillations controlled or modulated by varying the potential on it. A magnetron equipped with a grid can be used as an amplifier as well as an oscillator. The generated frequency of these tubes is determined by equation 225 or a multiple thereof.

It must be remembered that the equations for frequency which have been developed in this discussion are based on assumptions which are not completely fulfilled in practice; hence, they are only approximate. The complication of more exact equations, however, is seldom justified in making ordinary calculations.

As previously mentioned, the wavelength characteristic curves of magnetrons which are tuned to the natural period of electronic motion take the form shown in Fig. 88. The discussion of the wavelength characteristic of the Barkhausen-Kurz oscillator also applies to them. The tuning of the external circuit has little effect on the generated wavelength, but it does have a considerable effect on the efficiency of oscillation. When the external circuit is not in tune with the natural frequency of the electronic motion, the tube must neutralize a higher resistance and is, therefore, less efficient.

8–9. Velocity-modulated Oscillators. The Klystron is a velocity-modulated tube. Essentially, it consists of an electron gun, a buncher, and a catcher (Fig. 96). As the name implies, the buncher gathers the electrons from the gun into small bunches spaced at intervals of time which are determined by the frequency generated. The bunches of electrons pass through the cavity-resonator catcher at those times when they are able to give up their kinetic energy to the radio-frequency field of the resonator. This negative resistance action sustains oscillations in the catcher. A portion of the energy from the catcher is fed back to the buncher to maintain its oscillation.

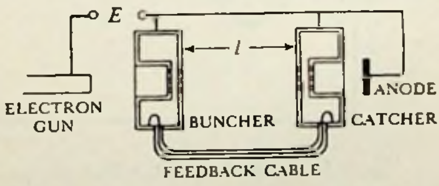


Fig. 96. Structural diagram of the Klystron.

Both the buncher and catcher are perforated cavity resonators so shaped that electrons can be shot through their radio-frequency field in much less time than that required for the direction of the field to change. During a portion of the cycle, the electrons passing

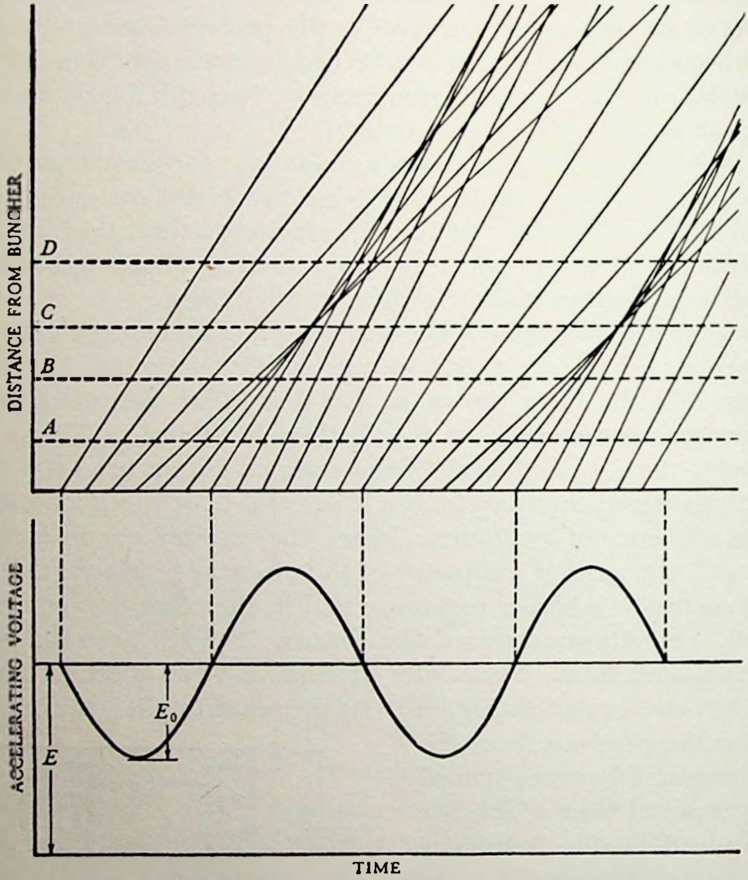


Fig. 97. An illustration which shows the nature of the bunching process.

through the field of the buncher are accelerated. A short time later, the next group of electrons passing through the field are decelerated. The effect of these differences of velocity is to bunch the electrons.

The best qualitative picture of the bunching process can be had from the diagram of Fig. 97. Here, the distance of an electron from the buncher is plotted as a function of time. The path of an electron is a line whose slope is proportional to the electron's velocity. The paths of subsequent electrons appear on the diagram as other lines to the right of the initial one. At the buncher, the electrons encounter an accelerating voltage which varies sinusoidally with time. Their velocity on emerging is proportional to the square root of the accelerating voltage E plus the voltage across the buncher $E_0 \sin \omega t$ or from equation 18 on page 50

$$v = 5.93 \times 10^7\sqrt{E + E_0 \sin \omega t} \text{ (neglecting relativistic mass)} \quad (228)$$

The depth of modulation is defined as $m = \dfrac{E_0}{E}$; therefore,

$$v = 5.93 \times 10^7\sqrt{E(1 + m \sin \omega t)} \text{ cm/sec} \quad (229)$$

The slope of each path is changed according to this formula. A horizontal line on the diagram represents a plane some distance

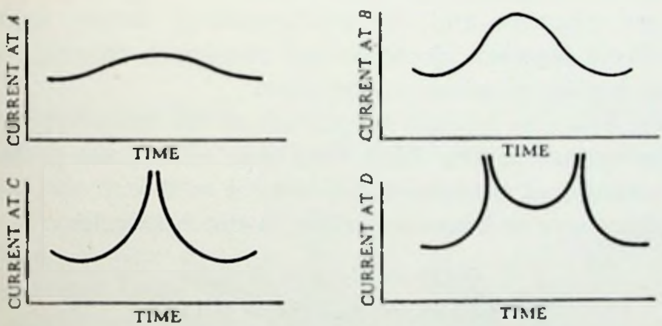


Fig. 98. Instantaneous variation in the current at the catcher for the several locations indicated in Fig. 97.

from the buncher. If the emission is uniform and the region between velocity lines is taken as representing equal numbers of electrons, the reciprocal of the distance between lines will be proportional to the current. A catcher located on one of the planes A, B, C, or D of the diagram would experience the variation of current shown in Fig. 98.

It should be noted that a change in the voltage on the buncher with respect to the cathode, which amounts to a change of the average velocity with which the electrons enter the field of the buncher, has the same effect on the current passing through the catcher as a change in the distance from buncher to catcher. It is further evident that some catcher distances yield a current which is very rich in harmonics. The catcher is easily tuned to any one of them; hence, a Klystron makes an excellent frequency multiplier.

A mathematical analysis by Webster of the current wave at the catcher shows that this current can be written as:

$$\begin{aligned} I(t) = I_0[1 + 2\{&J_1(k)\cos\omega t + J_2(2k)\cos 2\omega t \\ &+ J_3(3k)\cos 3\omega t + \cdots + J_n(nk)\cos n\omega t + \cdots\}] \end{aligned} \quad (230)$$

The J's are Bessel functions and the k's are equal to $\frac{\pi l E_0}{\lambda' E}$ where l is the distance between buncher and catcher, E_0 is the peak radio-frequency voltage across the grids of the buncher, E is the beam voltage and λ' is the distance an electron goes in one cycle. λ' is equal to the average velocity of the electrons $(5.93 \times 10^7\sqrt{E})$ divided by the frequency of oscillation.

Both the diagram and the mathematical theory neglect the mutual effects between electrons and the space charge, but these effects are usually of minor importance.

Example: Find the relative magnitude of the component currents up to and including the third harmonic which are present in a velocity modulated electron beam when $k = 0.5$.

From the curve of Bessel functions in the Appendix

$$\begin{aligned} J_1(k) &= J_1(0.5) = 0.24 \\ J_2(2k) &= J_2(1.0) = 0.115 \\ J_3(3k) &= J_3(1.5) = 0.06 \end{aligned}$$

The output current is therefore

$$I(t) = I_0[1 + 0.48\cos\omega t + 0.23\cos 2\omega t + 0.12\cos 3\omega t]\cdots$$

For a Klystron to oscillate, it is necessary that the voltage fed back to the buncher from the catcher be in phase with the buncher voltage. This condition is fulfilled when the total phase shift around the circuit (from buncher to catcher and back to buncher) is an integral multiple of 2π. This statement can be written as an equation:

$$T + \theta + \frac{\pi}{2} = 2\pi n \quad (231)$$

T is the angular equivalent of transit time. It is equal to the number of cycles in the transit time multiplied by 2π. θ is the phase angle of the transfer impedance which is the ratio of the voltage across the buncher to the beam current at the catcher. The $\frac{\pi}{2}$ comes from the fact that the electron at the center of the bunch at the catcher is the one which passed through the buncher when its voltage was zero; consequently, a quarter-cycle phase shift is introduced. The transit angle is given by the expression:

$$T = 2\pi \frac{l}{\lambda'} \tag{232}$$

The substitution of equation 232 into equation 231 and the solution of the resultant equation for $\frac{l}{\lambda'}$ yields

$$\frac{l}{\lambda'} = n - \frac{\theta}{2\pi} - \frac{1}{4} \tag{233}$$

For a given value of θ, there are a series of voltages or (λ')'s which satisfy the above equation. A curve of $\frac{l}{\lambda'}$ plotted against the integers n is a straight line. Usually, there are two possible values of θ since the feedback line acts like an overcoupled circuit. Its behavior is shown in Fig. 99. Oscillation will occur when the transfer impedance is large. There are two frequencies at which this occurs, and the θ of one differs from that of the other by about π. Consequently, there are two straight-line curves of $\frac{l}{\lambda'}$ vs. n — one each for the two values of θ.

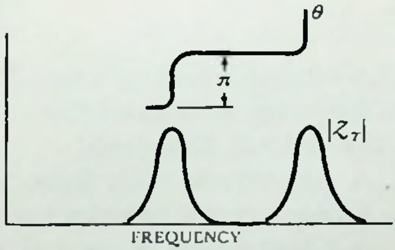


Fig. 99. The transfer impedance (magnitude and angle) of a typical Klystron feedback cable as a function of frequency.

A change in beam voltage changes the frequency of oscillation of a Klystron. This is shown by differentiation of equation 231. All the terms are constants except the transit angle T and the transfer impedance angle θ; hence,

$$dT = -\,d\theta \tag{234}$$

From equation 232,

$$T = 2\pi \frac{l}{\lambda'} = \frac{2\pi l f}{5.93 \times 10^7 \sqrt{E}}$$

If the change in frequency f is neglected for the time being, the derivative of T is

$$dT = -\frac{1}{2} \frac{2\pi l f\, dE}{5.93 \times 10^7 \sqrt{E} E} = -\frac{1}{2} T \frac{dE}{E} \tag{235}$$

The substitution of equation 235 into equation 234 gives

$$\frac{1}{2} T \frac{dE}{E} = d\theta \tag{236}$$

The change of impedance angle is associated with a change of frequency or $d\theta = \frac{d\theta}{df} df$. The factor $\frac{d\theta}{df}$ is difficult to calculate exactly, but since the circuit is decidedly overcoupled, it can be assumed that $\frac{d\theta}{df} = \frac{2Q}{f}$ or that the curve is almost exactly like that of a simple resonant circuit. When this is substituted in equation 236, we have

$$\frac{1}{2} T \frac{dE}{E} = \frac{2Q}{f} df$$

or

$$\frac{df}{dE} = \frac{fT}{4EQ} \tag{237}$$

The relationship of equation 237 can be used to calculate a change in frequency for a given change in accelerating voltage. An average value is 10 to 30 kc/volt.

A Klystron is easily frequency modulated. The audio voltage is simply inserted in series with the anode voltage.

A Klystron can be used as an amplifier as well as an oscillator. The tube is operated at a low level and the voltage to be amplified is fed to the buncher. High inherent noise level, however, limits the application of the tube as an amplifier to power stages. Most tubes now available are not suitable for use as detectors, but they can be especially constructed for this service. The different velocity electrons are separated by a magnetic field and an auxiliary electrode is used to identify those whose velocity was changed by passage through the tube. A multiplier Klystron has the buncher resonator tuned to a low multiple of the catcher frequency. The tube uses the electron's rich harmonic output current to obtain exact multiples of some control frequency. The three resonator tubes are like the ordinary Klystron plus an additional resonator beyond the usual catcher. They have the advantage of delivering varying amounts of power with no effect on oscillation. The single-resonator Klystron uses the same resonator as both buncher and catcher. It has the advantage of being simple to tune. The cavity resonators on a Klystron are tuned by deforming them slightly and thus changing their volume.

At first sight, it might seem like a nearly impossible task to adjust two resonators to the same frequency and at the same time find the proper length of feedback cable and accelerating voltage to produce oscillations. A logical procedure, however, eliminates the difficulties. The accelerating voltage is fluctuated through a range which is sure to include one of the oscillating points for the feedback cable used. One of the resonators is then tuned to the frequency of the other one by noting the point where weak oscillations are produced. The fluctuating anode voltage is then removed and replaced by the constant supply which is adjusted to the point where oscillations occur.

Problems

8-1. Calculate the input admittance of a 6J5 tube which is connected to a load consisting of 10 μh inductance and 80 ohms resistance. The plate voltage is 250 volts and the grid voltage is $-$ 8 volts. The frequency is 4 mc.

8-2. A resonant circuit consists of two coils and a condenser. One of the coils is connected between the grid and cathode of a triode, the other coil is connected between the grid and the plate, and the condenser is connected between the cathode and the plate. Show whether or not the tube is capable of introducing a negative resistance in the circuit.

8-3. A transmission-line oscillator of the type illustrated in Fig. 84 is set up using a WE-316A tube and a line which consists of two number 18 wires spaced $1\frac{1}{4}$ in. apart. Plot a curve of the generated wavelength as a function of the distance between the tube and the first shorting condenser across the line. The interelectrode capacities of the 316A tube are:

$$C_{gp} = 1.6\ \mu\mu\text{f}$$
$$C_{gk} = 1.2\ \mu\mu\text{f}$$
$$C_{pk} = 0.8\ \mu\mu\text{f}$$

8–4. An electron gun produces a stream of electrons at a 2000-volt potential. If this stream of electrons is used to generate radio frequency at a 20-cm wavelength in a reversed cyclotron, compute the strength of the magnetic field required and the radius of the *D*'s.

8–5. A certain Klystron has a 1.5-cm spacing between the buncher and catcher when it produces the greatest output in amplifying a given signal with no change in frequency.

(*a*) Find the optimum position for the catcher if the tube is used as a frequency doubler. Assume that the voltages are unchanged.

(*b*) Find the relative output current in each case.

8–6. A Klystron oscillator is to generate 10-cm waves.

(*a*) Calculate the approximate diameter and Q of copper cavity resonators which could be used for the buncher and the catcher.

(*b*) Find the optimum spacing between the buncher and catcher if the beam voltage is 400 volts and the depth of modulation is 0.2.

(*c*) From the results of (*a*) and (*b*), calculate the change in frequency which accompanies each volt change of the beam voltage.

Experiment 16

Object:

The object of this experiment is to demonstrate the negative resistance properties of some common oscillators.

Preliminary:

Solve problems 8–1 and 8–2.

Performance:

1. Assemble a crystal oscillator such as that of Fig. 80, using a 6J5 tube. Tune the circuit into oscillation as evidenced by a drop in the d-c plate current. Pick up the signal on a near-by communication receiver and make it audible by means of the beat frequency oscillator of the receiver.

2. Verify the fact that oscillations occur only when the plate circuit is tuned on the inductive side of resonance. Observe the change in frequency as evidenced by the change in tone of the beat note by varying the supply voltage and the tuning.

3. Substitute an ordinary tuned circuit for the crystal and repeat the observations of part 2. Explain the results.

4. Connect a 6J5 tube as an audio-frequency Hartley oscillator. Use Brooks inductometers or the equivalent for the inductances of the circuit. Connect batteries in series with these inductances to supply the proper plate and grid operating potentials and by-pass them with large con-

densers. Adjust the components of the circuit to give a frequency of approximately 1000 cps.

5. Measure the frequency and magnitude of oscillation by means of a cathode-ray oscillograph and a standard audio-frequency oscillator.

6. Measure the resistance of the tuned circuit. Measure the constants of the tube when it is operated over the same plate and grid swings as those measured in part 5. The tube constants can also be determined from the slope of a secant line drawn between the end points of operation on the characteristic curves of the tube used.

7. Substitute the results of the measurements of 5 and 6 in equations 213 and 214, and verify the fact that the net resistance and net reactance of an oscillating circuit are zero.

8. Insert a variable resistance in the tuned circuit and note the effect of the magnitude of this resistance on the magnitude and frequency of the oscillations generated.

Experiment 17

Object:

The object of this experiment is to determine the wavelength characteristic of a transmission-line oscillator.

Preliminary:

Solve problem 8–3.

Performance:

1. Connect a transmission line of the type described in problem 8–3 to the plate and grid terminals of a 316A tube.[1] Connect a small shorting condenser across the line a distance l from the tube. At a distance from this equal to one-fourth of the estimated wavelength produced by the tube, connect another condenser across the line.

2. Set up another transmission line about 1 ft from the one connected to the tube and parallel with it. Provide this line with a pair of shorting bars. Place a pick-up loop connected to crystal detector and microammeter near one of the shorting bars or place a vacuum-tube voltmeter of the type described in Experiment 15 midway between the two shorting bars. The connection which has just been described is known as a "Lecher wire system" and is useful in measuring the wavelength of radio waves which are generated in its immediate vicinity. The distance between the two shorting bars is changed until a point is reached where the meter

[1] The reader may prefer to use a line which is mechanically sturdier than this for these experiments. In that event, it is suggested that a line consisting of two $\frac{1}{4}$ in. diameter copper tubes or rods spaced 2 in. between centers be used instead of the wires.

suddenly reads a high value. The distance between the bars is then one-half wavelength (or multiple thereof) and the generated wavelength can be determined by measuring this distance with a meter stick.

3. Using the above apparatus, measure the generated wavelength of the tube for a series of different distances between the tube and first shorting condenser. Readjust the distance between the first and second shorting condenser to about one-quarter of the generated wavelength before making each reading. In the course of the readings, find the minimum wavelength at which the tube will oscillate.

4. Plot the results of part 3 in a contrasting color on the same coordinate axes as the curve of problem 8–3. Explain discrepancies between the curves.

5. Fix the position of the first shorting condenser somewhere on the line. At a distance l_s from the first condenser, connect a second condenser across the line. Connect a third condenser across the line one-quarter of the generated wavelength beyond the second condenser. Measure the generated wavelength by the Lecher wire system for a series of different distances l_s.

6. Plot the generated wavelength as a function of l_s.

Experiment 18

Object:

The object of this experiment is to determine the wavelength characteristic of a tube which depends on natural electronic oscillation for its operation.

Preliminary:

1. Solve problem 8–4.

2. A triode having a plate diameter of 2.0 cm is used as a Barkhausen-Kurz oscillator. The grid voltage is $+ 175$ volts and the plate voltage is -10 volts. Find the approximate generated wavelength.

Performance:

1. Substitute a magnetron or a triode operated as a Barkhausen-Kurz oscillator for the 316A tube of Experiment 17 and follow procedures 1, 2, and 3 of that experiment.

2. Plot the generated wavelength vs. the active length of attached transmission line.

3. Repeat the above test with a different magnetic field intensity on the magnetron or a different grid voltage on the Barkhausen-Kurz oscillator.

References

Barrow, W. L., "An Oscillator for Ultra-high Frequencies," *Review Scien. Instr.*, June 1938, pp. 170–174.

Dow, W. G., *Fundamentals of Engineering Electronics.* New York: John Wiley & Sons, Inc., 1937, pp. 141–145, 331–334.

Gavin, M. R., "Triode Oscillators for Ultra-short Wavelengths," *Wireless Engineer*, June 1939, pp. 287–295.

Hansell, C. W., and P. S. Carter, "Frequency Control by Low Power Factor Line Circuits," *Proc. I.R.E.*, April 1936, pp. 597–619.

Kilgore, G. R., "The Magnetron as a High-frequency Generator," *Jour. Applied Phys.*, October 1937, pp. 666–676.

King, R., "Wave-length Characteristics of Coupled Circuits Having Distributed Constants," *Proc. I.R.E.*, August 1932, pp. 1368–1374.

Kompfner, R., "Velocity Modulation," *Wireless Engineer*, November 1940, pp. 478–488.

Linder, E. G., "Description and Characteristics of the Endplate Magnetron," *Proc. I.R.E.*, April 1936, pp. 633–653.

Mason, W. P., and I. E. Fair, "A New Direct Crystal-controlled Oscillator for Ultra-short-wave Frequencies," *Proc. I.R.E.*, October 1942, pp. 464–472.

Okabe, K., "Electron Beam Magnetron," *Report of Radio Research in Japan*, June 1937, pp. 1–6.

Okabe, K., "Obtaining Dwarf-waves with Multi-split Anode Magnetrons," *Report of Radio Research in Japan*, June 1938, pp. 27–29.

Peters, L. J., *Theory of Thermionic Vacuum Tube Circuits.* New York: McGraw-Hill Book Company, Inc., 1927, pp. 62–103.

Ramo, S., "The Electronic Wave Theory of Velocity Modulation Tubes," *Proc. I.R.E.*, December 1939, pp. 757–763.

Samuel, A. L., "Extending the Frequency Range of the Negative-grid Tube," *Jour. Applied Phys.*, October 1937, pp. 677–688.

Terman, F. E., *Radio Engineering*, 2nd ed., New York: McGraw-Hill Book Company, Inc., 1937, pp. 349–391.

Tombs, D. M., "Velocity-modulated Beams," *Wireless Engineer*, February 1940, pp. 54–60.

Uda, S., "Sentron — A New Tube for Ultra Short Waves," *Report of Radio Research in Japan*, October 1938, pp. 47–56.

Varian, R. H., and S. F. Varian, "A High Frequency Oscillator and Amplifier," *Jour. Applied Phys.*, May 1939, pp. 321–327.

Webster, D. L., "Cathode-Ray Bunching," *Jour. Applied Phys.*, July 1939, pp. 501–508.

Webster, D. L., "The Theory of Klystron Oscillations," *Jour. Applied Phys.*, December 1939, pp. 864–872.

CHAPTER IX

MODULATION AND DETECTION

9–1. Types of Modulation. Modulation is the process of varying some characteristic of an electric wave in direct proportion to the instantaneous value of another which is called the modulating wave. The purpose of modulation is to facilitate the utilization of a radio wave as a carrier of information which is contained in the signal producing the modulation. Either the amplitude, frequency, or phase of the carrier is varied in the modulation process. For example, amplitude modulation is produced by increasing or decreasing the amplitude of the radio wave from a certain average value by an amount that is directly proportional to the instantaneous magnitude of the modulating wave. Consequently, a temporal plot of the deviation in the amplitude of the carrier from its average value is also a plot of the signal which produces the modulation. A receiver detects these changes in amplitude and reproduces the signal which modulated the carrier.

Frequency modulation is produced by increasing or decreasing the frequency of the carrier from an average value by an amount that is directly proportional to the instantaneous magnitude of the modulating wave. The maximum deviation in frequency from the carrier is determined by the peak value of the modulating wave and the number of frequency swings per second is determined by its frequency. In contrast, phase modulation advances or retards the phase of the carrier in direct proportion to the magnitude of the modulating wave. The maximum phase deviation from the mean value is determined by the peak value of the modulating wave and the number of phase swings per second is determined by its frequency. While a change in phase must necessarily accompany the change in frequency caused by frequency modulation, there is nevertheless a clear distinction between frequency and phase modulation. From definition, a plot of the deviation in frequency of a frequency-modulated carrier as a function of time is a plot of the modulating wave, but in general a plot of the accompanying deviation in the phase as a function of time is not a plot of the modulating wave; hence, the wave is not phase modulated. A similar distinction exists between true phase modulation and the frequency deviations which accompany it.

For the purpose of circuit analysis, it is convenient to resolve a modulated carrier into a series of simple constant-amplitude constant-frequency waves. Consider the general expression for the instantaneous current of an electric wave.

$$i = I \sin \phi \tag{238}$$

I is the amplitude of the rotating vector representing the wave, and ϕ is angular position of that vector at the particular instant of time under consideration. The angular velocity of the rotating vector is equal to the time rate of change of the angle or

$$\omega = 2\pi f = \frac{d\phi}{dt} \tag{239}$$

where f is the frequency of the wave. The angle ϕ at a particular instant of time t is given by the integral of equation 239

$$\phi = \int 2\pi f\, dt \tag{240}$$

When the frequency is constant, as it is in amplitude modulation, equation 240 becomes

$$\phi = (2\pi f t + \Phi) \tag{241}$$

Φ is the constant of integration and represents the phase angle of the rotating vector at the instant when t was zero. The substitution of equation 241 into equation 238 yields a familiar expression

$$i = I \sin (2\pi f t + \Phi) \tag{242}$$

While the above equation could have been assumed directly for the analysis of amplitude modulation, the equations which lead up to it must be considered in the analysis of frequency modulation and for that reason are included here. Amplitude modulation is produced by varying the factor I of equation 242 in direct proportion to the instantaneous value of the modulating wave.

For simplification of the analysis we will assume that the modulating wave is a pure sinusoid of frequency f_m. This assumption will not destroy the validity of the discussion because any complex wave can be resolved into a series of sinusoids by Fourier's series. If the amplitude of the radio wave is to deviate from the mean amplitude I_c by an amount which is directly proportional to the modulation, the instantaneous magnitude of the radio wave can be expressed as

$$I = I_c(1 + m_a \cos 2\,\pi f_m t) \tag{243}$$

The particular form of equation 243 was chosen for convenience in mathematical manipulation. The factor m_a is the amplitude modulation index. When this factor is expressed as a percentage, it is also called the percentage of modulation and indicates the relative completeness of the modulation process. The substitution of equation 243 into equation 242 and the selection of a reference axis which makes Φ zero gives the following equation:

$$i = I_c(\sin 2\,\pi f t + m_a \sin 2\,\pi f t \cos 2\,\pi f_m t) \tag{244}$$

But there is a trigonometric identity

$$\sin \alpha \cos \beta = \tfrac{1}{2} \sin (\alpha + \beta) + \tfrac{1}{2} \sin (\alpha - \beta) \tag{245}$$

which simplifies the last term of equation 244 and reduces that equation to

$$i = I_c[\sin 2\,\pi f t + \tfrac{1}{2}\, m_a \sin 2\,\pi(f + f_m)t + \tfrac{1}{2}\, m_a \sin 2\,\pi(f - f_m)t] \tag{246}$$

This equation shows that a radio wave which is amplitude modulated by a sinusoid consists of three constant-amplitude constant-frequency components. One of these components has the frequency and average amplitude of the carrier. The other two components lie on either side of the carrier frequency and are called the side bands of the wave. Figure 100*a* is a graphical representation of the frequency spectrum of an amplitude-modulated wave having a modulation index $m_a = 1.0$. The vertical lines are erected at points along the horizontal axis which correspond to the frequency of the components, and the height of these vertical lines corresponds to the magnitude of the components. The image formed by the cathode-ray tube of a highly selective panoramic receiver [1] is very

[1] A panoramic receiver is arranged so that it successively tunes over a band of frequencies. The horizontal sweep of a cathode-ray oscillograph is synchronized with

similar to Fig. 100*a* when the receiver picks up a signal from a station which is amplitude modulated with a pure tone. The voltage equation of an amplitude-modulated wave also has the form of equation 246.

The power in an electric wave varies as the square of the current or voltage; therefore, the power distribution between the carrier

FIG. 100. Amplitude and frequency position of the components making up (*a*) a wave which is amplitude modulated by a sinusoid to $m_a = 1.0$ and (*b*) a wave which is frequency modulated by a sinusoid to $m_f = 1.0$.

and the side bands is computed by squaring the coefficients of equation 246. The power required to produce amplitude modulation goes into the production of the side bands.

If a wave is frequency modulated with a sinusoid, the frequency at any instant can be expressed as a deviation from the mean or average frequency f_c thus:

$$f = f_c(1 + k_f \cos 2\,\pi f_m t) \tag{247}$$

The substitution of this expression in equation 240 yields

$$\phi = \int 2\,\pi f\,dt = \Phi + 2\,\pi f_c t + \frac{f_c k_f}{f_m} \sin 2\,\pi f_m t \tag{248}$$

The factor $\frac{f_c k_f}{f_m}$ is called the modulation index m_f and is equal to the variation in the radio frequency away from the mean divided by the modulating frequency. If equation 248 is substituted into equation 238 and the reference axis is selected so that Φ is zero, an equation for the current in a frequency-modulated wave is obtained.

$$i = I \sin (2\,\pi f_c t + m_f \sin 2\,\pi f_m t) \tag{249}$$

The trigonometric expression for the sine of the sum of two angles

$$\sin (\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta$$

the tuning sweep and the vertical deflection of the cathode-ray tube is controlled by the magnitude of the output from the receiver. As a result, the image formed on the cathode-ray screen is a plot of the signal strength received as a function of the frequency.

is used to reduce equation 249 to a more convenient form for analysis

$$i = I[\sin 2\pi f_c t \cos (m_f \sin 2\pi f_m t) + \cos 2\pi f_c t \sin (m_f \sin 2\pi f_m t)] \tag{250}$$

From the theory of Bessel functions

$$\sin (A \sin \alpha) = 2 J_1(A) \sin \alpha + 2 J_3(A) \sin 3\alpha + \cdots \tag{251}$$

and

$$\cos (A \sin \alpha) = J_0(A) + 2 J_2(A) \cos 2\alpha + 2 J_4(A) \cos 4\alpha + \cdots \tag{252}$$

where the J's are Bessel functions of the first kind. The substitution of equations 251 and 252 back into equation 250 gives the following result:

$$\begin{aligned} i = I[&J_0(m_f) \sin 2\pi f_c t \\ &+ 2 J_1(m_f) \sin 2\pi f_m t \cos 2\pi f_c t \\ &+ 2 J_2(m_f) \cos 2(2\pi f_m)t \sin 2\pi f_c t \\ &+ 2 J_3(m_f) \sin 3(2\pi f_m)t \cos 2\pi f_c t + \cdots] \end{aligned} \tag{253}$$

The above equation can be written in a more usable form by the application of equation 245

$$\begin{aligned} i = I[&J_0(m_f) \sin 2\pi f_c t \\ &+ J_1(m_f)\{\sin 2\pi(f_c + f_m)t - \sin 2\pi(f_c - f_m)t\} \\ &+ J_2(m_f)\{\sin 2\pi(f_c + 2f_m)t + \sin 2\pi(f_c - 2f_m)t\} \\ &+ J_3(m_f)\{\sin 2\pi(f_c + 3f_m)t - \sin 2\pi(f_c - 3f_m)t\} + \cdots] \end{aligned} \tag{254}$$

A comparison between equation 254 and equation 246 points out the major difference between the harmonic composition of an amplitude-modulated wave and a frequency-modulated wave. When a single frequency is used for modulation, an amplitude-modulated wave consists of the carrier plus a single upper side-band and a single lower side-band frequency, whereas a frequency-modulated wave has a whole series of side-band frequencies above and below the carrier frequency. Due to the multiplicity of its side-band frequencies, a broader band width is required to handle a frequency-modulated wave than is required to handle an amplitude-modulated wave.

Example: A 40.00-mc carrier is swung between 40.01 mc and 39.99 mc when it is frequency modulated by a 10,000 cps sine wave. Find the relative amplitudes and frequencies of the components of the wave.

$$f_c k_f = 40.01 - 40 = 0.01 \text{ mc} \quad \text{or} \quad 10 \text{ kc}$$

therefore the modulation index is

$$m_f = \frac{f_c k_f}{f_m} = \frac{10{,}000}{10{,}000} = 1.0$$

from the curve of Bessel functions in the Appendix

$$J_0(m_f) = 0.765$$
$$J_1(m_f) = 0.44$$
$$J_2(m_f) = 0.115$$
$$J_3(m_f) = 0.02 \quad \text{etc.}$$

Consequently, equation 254 indicates that the wave has the following component frequencies:

Frequency	*Relative amplitude*
40.00 mc	0.765
40.01 and 39.99 mc	0.44
40.02 and 39.93 mc	0.115
40.03 and 39.97 mc	0.02
etc.	

A plot of the magnitude of these components as a function of their frequency is given in Fig. 100*b*.

The contrast between the frequency composition of the two waves is evident from the figure.

When a wave is phase modulated by a sinusoid having a frequency f_m, the expression for the phase angle becomes

$$\phi = \phi_c(1 + k_p \sin 2\,\pi f_m t) \tag{255}$$

and the expression for the radio-frequency current is

$$i = I \sin\,(2\,\pi f_c t + k_p \phi_c \sin 2\,\pi f_m t) \tag{256}$$

The factor $k_p\phi_c$ is called the phase modulation index m_p and represents the maximum deviation in phase angle that is produced by the modulation process. Equation 256 differs from the frequency modulation equation 249 only in the definition of the modulation index; consequently, equation 254 gives the components of a phase-modulated wave as well as a frequency-modulated wave. It is apparent from this equation that wide-band circuits are required for the transmission of both phase and frequency modulation.

9–2. Methods of Producing Amplitude Modulation. It has already been shown (section 5–10) that a class C amplifier can be amplitude modulated. In the process known as "plate modulation," the plate voltage of a radio-frequency amplifier tube is varied about an average value by the modulating signal. Since the amplitude of the output wave is directly proportional to the voltage supplied to the plate, the change in amplitude is directly proportional to the change in plate voltage and amplitude modulation is produced. A variety of circuits can be employed for plate modulation. Essentially, however, the modulating signal is fed to a modulating amplifier, the output transformer of which is connected in series with the plate supply voltage to the radio-frequency amplifier. The modulating amplifier is isolated from the high-frequency currents of the carrier by a choke and by-pass condenser connection.

To develop 100 per cent modulation the plate supply voltage of a radio-frequency amplifier must be swung between zero and two times its normal value. In other words, the peak value of the wave produced by the modulator must be equal to the plate supply voltage of the radio-frequency amplifier. The modulating amplifier must also deliver a power output equal to the power in the side bands. Since the power in each side band is equal to one-fourth of the power in the carrier for 100 per cent modulation, the power output of the modulator must be equal to one-half of the power in the carrier. However, due to the plate loss in the amplifier tube, more power is actually required for modulation than this brief analysis would indicate. When a single tube is used to modulate a radio-frequency amplifier, the rating of the tube in the modulator must be greater than the rating of the tube in the amplifier. This is because the modulating tube must be operated class A with consequent low efficiency, whereas the amplifier tube is operated class C with consequent high efficiency.

The modulating signal can also be fed to the grid as indicated in the circuit diagram of Fig. 101. There are two possible methods of producing amplitude modulation with this connection. In the Van der Bijl modulator, the tube is operated class A near the curved portion of its characteristic curve. The addition of the high-frequency carrier to the low-frequency modulating wave produces the wave shape shown in Fig. 102. Due to the curvature in the characteristic, the output current from the tube takes the form shown. When this complex current wave is passed through a resonant circuit which is tuned to the frequency of the carrier, a voltage is produced across that circuit which takes the form shown at the extreme right in

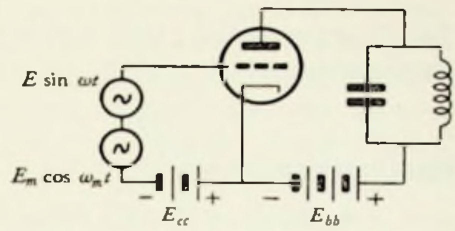


FIG. 101. Circuit connection for Van der Bijl modulator or grid modulator.

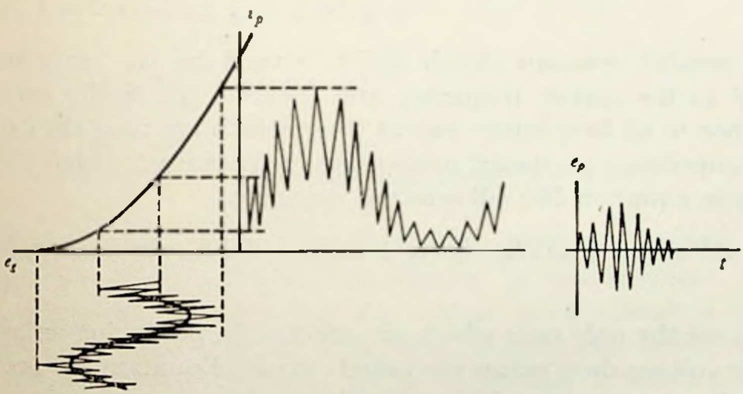


FIG. 102. Van der Bijl modulation. An illustration of the use of the parabolic portion of the dynamic characteristic curve for the production of amplitude modulation. The plate current i_p produces a voltage drop e_p across the tuned circuit in the output.

the figure. In other words, an amplitude-modulated wave is produced. Assume that the input to the tube is the carrier voltage

$$E \sin \omega t \tag{257}$$

and that the modulating voltage is

$$E_m \cos \omega_m t \tag{258}$$

The net alternating voltage which is applied to the grid is equal to the sum of equations 257 and 258 or

$$e_g = E \sin \omega t + E_m \cos \omega_m t \tag{259}$$

The plate current of a tube operated on the parabolic portion of its characteristic curve is given by equation 37 on page 67

$$i_p = A + Be_g + De_g^2$$

Substituting equation 259 in the above, we obtain

$$i_p = A + BE \sin \omega t + BE_m \cos \omega_m t + DE^2 \sin^2 \omega t \\ + 2\,DEE_m \sin \omega t \cos \omega_m t + DE_m^2 \cos^2 \omega_m t$$

Trigonometric identities reduce this equation to

$$i_p = A + BE \sin \omega t + BE_m \cos \omega_m t + \tfrac{1}{2}\,DE^2 - \tfrac{1}{2}\,DE^2 \cos 2\,\omega t \\ + DEE_m \sin (\omega + \omega_m)t + DEE_m \sin (\omega - \omega_m)t \\ + \tfrac{1}{2}\,DE_m^2 + \tfrac{1}{2}\,DE_m^2 \cos 2\,\omega_m t \tag{260}$$

The parallel resonant circuit which is used for the plate load is tuned to the carrier frequency and presents practically zero impedance to all frequencies except those which are near the carrier. The impedance presented to these frequencies is very high. Three terms in equation 260 fall into this category.

$$i_p' = BE \sin \omega t + DEE_m \sin (\omega + \omega_m)t + DEE_m \sin (\omega - \omega_m)t \tag{261}$$

They are the only ones which are effective in producing an appreciable voltage drop across the tuned circuit. Equation 261 is of the same form as equation 246 and indicates that an amplitude-modulated voltage wave will appear across the plate load. The presence of the other components of current introduces some distortion, but it is usually of negligible proportions.

In grid modulation, the tube is operated class C rather than class A. The input wave is clipped off so that the plate current consists of a series of pulses as shown in Fig. 103. Again the tuned circuit in the plate load presents practically zero impedance to the undesired components of the plate current, and the voltage which appears across the plate load takes the form of an amplitude-modulated wave as indicated at the extreme right of Fig. 103.

Cathode modulation is a combination of grid modulation and plate modulation. The combination of the two effects is produced by feeding the modulating voltage to a transformer which is wired in series with the cathode lead of the modulating tube. The voltage introduced in this manner is applied to both the grid and the plate.

The methods of amplitude modulation thus far considered vary the amplitude of the radio-frequency voltage appearing across a tuned circuit. Due to the inertia effect of the circuit (see the last part of section 6–6), these methods cannot be applied when the

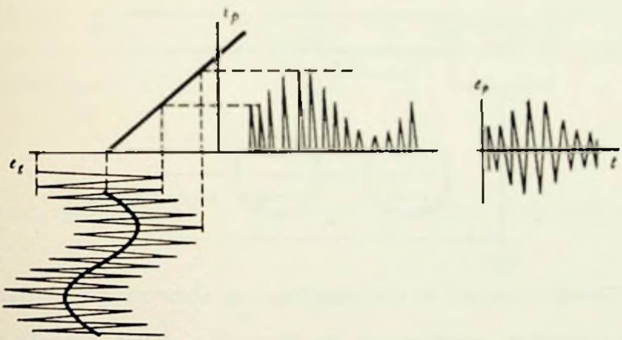


FIG. 103. Illustration of the grid-modulation process.

modulating voltage changes rapidly unless the Q is sacrificed. Absorption modulation provides a solution to this problem. The radio-frequency voltage across the output tuned circuit is maintained at a constant value. A network between this tuned circuit and the load varies the amount of absorption it introduces in accordance with the instantaneous magnitude of the modulating wave, and thus produces an amplitude modulation at the load. In addition to avoiding the inertia effect of high-Q circuits, absorption modulation is frequently preferred for ultra-high frequency because the method is also independent of the oscillator or amplifier characteristics.

A simple connection which has been used for modulating centimeter waves employs a glow-discharge tube for the absorption load. In one system, the radiating dipole antenna is sealed inside the helium-filled bulb which makes up the discharge tube. The anode of the tube is placed opposite the dipole and is operated at a positive potential with respect to it. An increase in the glow-discharge current causes a decrease in the radiated energy.

A transmission-line system for absorption modulation is shown in Fig. 104. The conductance of the tubes which are connected across the end of the stub line is varied by the modulating voltage which is applied to their grids. When there is no conduction through the tubes, the stub appears as a short circuit across the line at P. The distance between the generator and the point P is

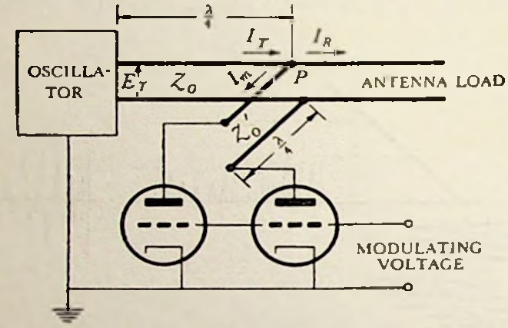


Fig. 104. Circuit diagram of transmission-line absorption modulation system.

made one-quarter wavelength so that the short circuit developed at P under the above conditions appears as an open circuit to the high-frequency generator; hence, no power is transmitted to the antenna load. When the tubes conduct a considerable current, the stub appears as an open circuit at P and has no effect on the transmission of energy to the antenna.

For an analysis of the transmission-line system, consider equation 171 of section 7–3.

$$Z_r Z_s = Z_0^2$$

where Z_s is the input impedance to a quarter-wave line having a characteristic impedance Z_0 and terminated in Z_r. The above equation can also be written as

$$\frac{E_r}{I_r}\frac{E_s}{I_s} = Z_0^2 \tag{262}$$

But the transfer impedance of a line is the same in both directions $\left(\frac{E_r}{I_s} = \frac{E_s}{I_r}\right)$ so it becomes

$$Z_t = Z_0 \tag{263}$$

That is, the transfer impedance of a section of line one-quarter wavelength long is equal to the characteristic impedance of the line. The antenna load in Fig. 104 can be replaced by a resistance R connected across the line at the junction point P. A portion (I_m) of the total current flowing in the junction P is diverted from the antenna load by the stub. The magnitude of this current depends on the magnitude of the conductance G which is developed by the tubes across the end of the stub. For convenience, the modulation factor m is defined as the fraction (I_R) of the total current at the junction which flows through the antenna load. It is also the fraction of voltage (E_R) which appears across the junction in terms of the total voltage (E_T) which would appear across the junction if no current were diverted.

$$m = \frac{I_R}{I_T} = \frac{E_R}{E_T} \tag{264}$$

The current and voltage at the sending end of the stub are

$$I_m = I_T - I_R = I_T(1 - m) \tag{265}$$

and

$$E_m = E_R = mE_T \tag{266}$$

From equation 263, the current at the junction in terms of the oscillator voltage (which is assumed to be constant and equal to E_T) is

$$I_T = \frac{E_T}{Z_0} \tag{267}$$

When there is no current diverted through the stub,

$$E_T = I_T R = \frac{E_T}{Z_0} R \tag{268}$$

The substitution of equation 267 into equation 265 and substitution of equation 268 into equation 266 shows that the sending-end current and voltage on the stub are

$$I_m = \frac{E_T}{Z_0}(1 - m) \tag{269}$$

and

$$E_m = \frac{E_T}{Z_0} Rm \tag{270}$$

The power diverted by the stub and lost in the modulation process is equal to the product of equations 269 and 270.

$$P_m = \frac{E_T^2}{Z_0^2} Rm(1 - m) \tag{271}$$

It is apparent from this equation that the power loss is zero both when the modulation factor is zero and when it is unity. The maximum power loss occurs when $m = 0.5$. The power delivered to the load when no diverting current is present is equal to the product of equations 267 and 268

$$P_T = \frac{E_T^2}{Z_0^2} R \tag{272}$$

A comparison between equation 272 and equation 271 with $m = 0.5$ shows that the maximum power lost by modulation is only one-fourth of the maximum power which can be delivered to the load. Since m takes on a range of values during the modulation process, the average power lost is even less than this.

No plate voltage is supplied to the modulating tubes. The cathodes are simply returned to a center tap on the section of the oscillator output coil which feeds the transmission line. The portion of the radio-frequency current rectified by the tubes is regulated by changing the grid voltage and thus changing their conductance G. From equation 262

$$\frac{1}{G}\frac{E_m}{I_m} = (Z_0')^2 \tag{273}$$

The substitution of equations 269 and 270 into equation 273 shows that the relationship between the tube conductance and the modulation factor m must be of the form

$$G = \frac{R}{(Z_0')^2}\frac{m}{1 - m} \tag{274}$$

Fortunately, the relationship between the grid voltage and the conductance is of this form over a wide range. In other words, there is a linear relationship between the modulation factor m and the grid voltage for all values of m except those close to zero or one; hence, the system can be successfully used for amplitude modulation.

The optimum operating conditions are obtained by a judicious choice of R, Z_0, and Z_0'. The antenna load resistance R must be made high enough so that the line resistance is negligible in proportion (an assumption made in the derivation), and yet it must be low enough so that the useful peak output will not be limited by the finite magnitude of the maximum conductance which can be produced by the modulator tubes.

9–3. Methods for Producing Frequency Modulation and Phase Modulation. A number of different devices have been invented for producing frequency and phase modulation. For example, special tubes built on the cathode-ray principle have been designed for this purpose. The most popular of all the methods used, however, is the reactance-tube modulator. A reactance tube is an ordinary vacuum tube which is connected so that it takes an alternating plate current which is 90° out of phase with the alternating plate voltage. The 90° phase relationship between the plate current and plate voltage is obtained by applying a voltage to the grid which is 90° out of phase with that on the plate. This relationship can be obtained by connecting a phase-shifting network, such as a resistance and capacity combination, between the plate and cathode of a tube and taking the grid voltage off the resistance element of that combination. When the reactance of the condenser is much larger (greater than 5 times) than the magnitude of the resistance, the alternating voltage on the grid is practically 90° ahead of the alternating voltage on the plate. If the reactance of the condenser is X_c, the resistance is R ohms, and the plate voltage is E_p, the grid voltage E_g is equal to the current which flows through the phase-shifting combination multiplied by R or it is approximately equal to the plate voltage divided by the capacitive reactance alone and multiplied by R.

$$E_g = \frac{E_p}{X_c} R \tag{275}$$

The plate current is equal to the grid voltage multiplied by the mutual conductance of the tube

$$I_p = \frac{E_p}{X_c} R g_m \tag{276}$$

Since this plate current leads the plate voltage by 90°, a capacitive reactance

$$X = \frac{E_p}{I_p} = \frac{X_c}{R g_m} \tag{277}$$

is produced by the tube. The mutual conductance of the tube can be changed by varying the grid bias; therefore, the amount of reactance which the tube produces can be changed in a like manner.

When a reactance tube is connected across the tuned circuit which controls the frequency of an oscillator, changes in the grid bias of the tube will change the frequency generated by the oscillator. If, on the other hand, the reactance tube is connected across the output tuned circuit of frequency-stabilized power amplifier or a crystal-controlled oscillator, a change in the bias on the reactance tube will change the phase of the output voltage rather than the frequency. The former connection is used in a frequency-modulated transmitter and the latter is used in a phase-modulated transmitter.

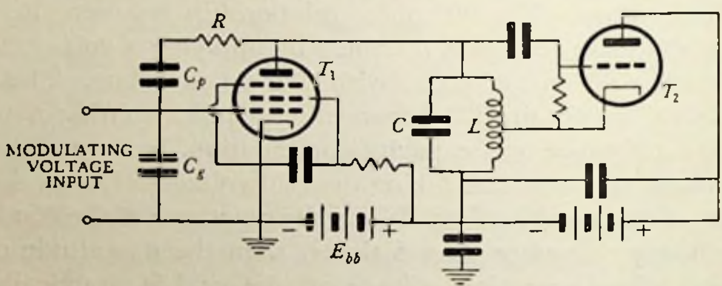


Fig. 105. Circuit diagram of a reactance tube used to frequency modulate a Hartley oscillator.

A typical circuit of a reactance tube employed as a frequency modulator is shown in Fig. 105. The phase-shifting network of the reactance tube T_1 consists of R and C_g. Capacity C_p simply isolates the grid from the plate voltage. It has a negligibly small reactance. The resistance R is made much larger than the reactance of C_g. A high resistance (not shown in the figure) provides the necessary conduction path between the grid and cathode. The grid bias and input signal is introduced across this resistance. Tube T_1 introduces a variable inductive reactance in parallel with the tuned circuit of the Hartley oscillator and thus controls its output frequency.

9–4. Transmitter Block Diagram. A complete transmitter consists of a series of basic circuits connected together in a manner similar to that indicated by the block diagram of Fig. 106. The primary element of the connection is the oscillator. As indicated in the preceding chapter, it may assume a variety of different forms. The modulator is the next element of the connection in importance. In simple amplitude or phase-modulated transmitters and in frequency-modulated transmitters, it is connected directly to the oscillator rather than occupying the position shown in Fig. 106. The oscillator may feed the antenna directly, but usually the energy is first passed through a buffer amplifier and then a power amplifier. The purpose of the buffer amplifier is to prevent any of the subsequent circuits from influencing the frequency generated by the

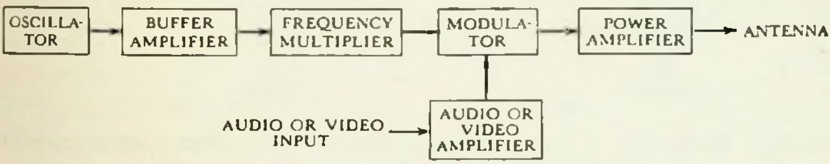


FIG. 106. Typical block diagram for a transmitter.

oscillator. The wave can be phase or amplitude modulated in one of the stages of radio-frequency amplification or it can be amplitude modulated between the final stage and the antenna. The former practice is indicated in Fig. 106.

9-5. Detectors for Amplitude Modulation. All amplitude-modulation detectors are essentially rectifiers. The rectified output current from the detector is fed into a simple resistance capacity filter with a time constant which is large enough so that the output voltage will not follow the high-frequency fluctuations of carrier and at the same time small enough so that the output voltage will follow the fluctuations of the modulating envelope. The rectified current can be considered to follow the high-frequency output current averaged over a few cycles at a time. Most practical detectors are not perfect rectifiers. Instead their characteristic curves are the same general shape as either Fig. 107*a* or 107*b*. These figures show the characteristic of the output current when an amplitude-modulated wave is passed through one of these devices. The rectified current is shown as a heavy dashed line on the right side of each figure. It is apparent in each case that this rectified current has the same general form as the wave which produced the modulation at the transmitter.

Detectors are classified as linear or squared law depending on the general shape of their characteristic curves. Figure 107*a* is the characteristic curve of a typical linear detector and Fig. 107*b* is the characteristic curve of a typical squared-law detector.

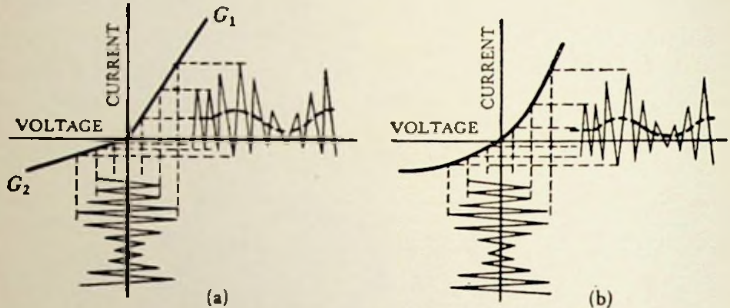


FIG. 107. Illustration of the rectification process on the characteristic curves of (*a*) a linear detector and (*b*) a squared-law detector.

As the name implies, the characteristic curve of a linear detector can be represented by two straight lines. One line has a slope G_1 and the other has a slope G_2. The equation of the envelope of the positive peaks of the modulated wave, that is the equation of the curve passing through the positive peaks, is

$$+ e_e = E_c(1 + m_a \sin 2\,\pi f_m t) \tag{278}$$

and the equation of the negative envelope is

$$- e_e = - E_c(1 + m_a \sin 2\,\pi f_m t) \tag{279}$$

On passing through the detector, the positive voltage envelope produces a positive current envelope equal to

$$+ i = G_1 E_c(1 + m_a \sin 2\,\pi f_m t) \tag{280}$$

and a negative current envelope

$$- i = - G_2 E_c(1 + m_a \sin 2\,\pi f_m t) \tag{281}$$

The average output current is the average of equations 280 and 281.

$$i = \left(\frac{G_1 - G_2}{2}\right) E_c(1 + m_a \sin 2\,\pi f_m t) \tag{282}$$

In other words, the output consists of a d-c component plus an alternating component which has exactly the same form as the modulating wave. As implied by equation 282, the effectiveness of a linear detector is measured by the magnitude of the difference between G_1 and G_2. The bend in the detector characteristic curve was set on the origin of the coordinate axes for convenience in analysis; however, the characteristic curve of an actual detector is not necessarily so placed. When the bend in the actual characteristic curve is not at the origin, the detector should be biased so that the axis of the input wave passes through the bend whereby the most efficient operation will be obtained.

A squared-law detector has a characteristic curve which is parabolic in form, and which can be represented by the equation

$$i = A + Be + De^2 \tag{283}$$

Assuming that the input voltage envelopes are of the same form as before, the positive output-current envelope is

$$+i = A + BE_c(1 + m_a \sin 2\,\pi f_m t) + DE_c^2(1 + m_a \sin 2\,\pi f_m t)^2 \tag{284}$$

and the negative current envelope is

$$-i = A - BE_c(1 + m_a \sin 2\,\pi f_m t) + DE_c^2(1 + m_a \sin 2\,\pi f_m t)^2 \tag{285}$$

The average current is the average of equations 284 and 285 or

$$\begin{aligned} i &= A + DE_c^2(1 + m_a \sin 2\,\pi f_m t)^2 \\ &= A + DE_c^2(1 + 2\,m_a \sin 2\,\pi f_m t + m_a^2 \sin^2 2\,\pi f_m t) \\ &= A + DE_c^2\left[1 + 2\,m_a \sin 2\,\pi f_m t + \frac{m_a^2}{2} - \frac{m_a^2}{2}\cos 2(2\,\pi f_m)t\right] \end{aligned} \tag{286}$$

Equation 286 consists of a d-c component plus an alternating component which has the same form as the modulating wave and in addition there is a second-harmonic distortion component. The magnitude of the distortion component is proportional to the modulation index squared; hence, it is negligible when low-level modulation is used. It will be noted that the useful portion of the output current has a magnitude which is proportional to the square of the input voltage. Consequently, a squared-law detector is more sensitive than an equivalent linear detector.

The electrical conductance of a point contact on the surface of certain minerals is greater when the current flows in one direction through the contact than when it flows in the other direction. A combination of this sort has rectifying properties and can be used for a detector. A typical characteristic curve is shown in Fig. 108.

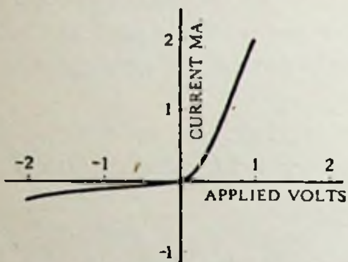


FIG. 108. Characteristic curve for a typical crystal detector.

The resemblance between this figure and Fig. 107*a* is quite apparent. Crystal detectors, as they are called, are particularly useful for the detection of very high-frequency radio waves because their shunting capacity and consequent loss of effectiveness can be made very low. The most important disadvantage of a crystal detector is its inability to handle overloads without permanent damage. When excessive currents flow through one of these detectors, the sharp point on the "cat's whisker" or contact electrode burns off, and the detector loses its sensitivity. This is because the sensitivity depends on making contact with a specific point on the surface of the crystal.

Rottgardt has tested the relative effectiveness of a large number of different combinations of materials for crystal detectors and found for wavelengths of 1.5 to 50 cm that a tungsten point on a silicon crystal gave the best results. A complete tabulation of the relative sensitivity for all the combinations tried is given in his paper.

Commercial crystal detectors are adjusted to their most sensitive position and then fixed in that position by enclosing both the crystal and the cat's whisker in a plastic. This unit is mounted in a cartridge which has convenient terminals for making external connections.

The diode is another very important linear detector. The connection for it is shown in Fig. 109*a*. The analysis of this detector is made by a consideration of the general form of the equation for the input voltage which is

$$e = E_c(1 + m_a \cos \omega_m t) \sin \omega t \tag{287}$$

where E_c is the peak value of the carrier voltage and m_a is the modulation index. When this voltage passes through the positive peak on its cycle, a pulse of current flows through the tube which charges the condenser C. As the input voltage starts on the downward portion of its cycle, the tube stops conducting, but the charge which was left on the condenser maintains the voltage across R until the next positive peak again charges the condenser. The voltage which appears across the resistance R as a consequence of

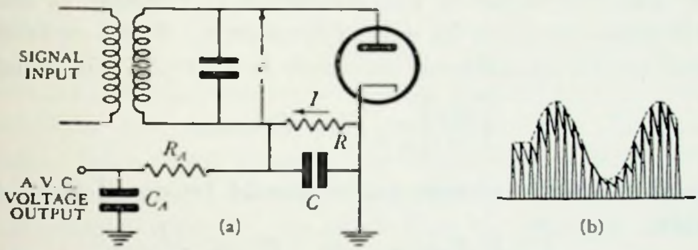


FIG. 109. (a) Diagram of connections for a diode detector. (b) Output voltage which is developed across the resistance R.

these events takes the form shown by the heavy line of Fig. 109b. It has practically the same shape as the modulation envelope. The time constant RC must be large enough so that the output wave is not excessively serrated, but it must also be small enough so that the output wave will follow the highest frequency in the modulating wave. Between the pulses of current, the condenser C discharges exponentially through R in accordance with equation 7 on page 12.

$$e = E_m \epsilon^{-\frac{t}{RC}}$$

The initial slope of the discharge curve is found by differentiating this equation and setting t equal to zero.

$$\left[\frac{de}{dt}\right]_{t=0} = -\frac{E_m}{RC} \tag{288}$$

For all practical purposes, it can be assumed that the voltage developed across R between pulses falls off linearly from the peak value of the preceding voltage pulse at the rate indicated by the above equation. With this idea in mind, an examination of Fig. 109b shows that the slope of each discharge curve (as given by equation 288) must be greater than the slope of the corresponding section of the modulation envelope (shown dotted in the figure) if the voltage across R is to follow the latter. The worst condition is encountered somewhere along the decreasing side of the modulation envelope. Its exact position depends on the magnitude of the modulation index. Experience has shown, however, that the voltage developed across R will follow the modulation envelope with a reasonably small amount of distortion if the time constant RC is made small enough so that the slope of the discharge curve is greater than the slope of the modulation envelope at the point where it passes through its axis of symmetry. From equation 287, the slope of the modulation envelope at this point is found to be

$$\left[\frac{de_e}{dt}\right]_{t=\frac{1}{4f}} = -E_c m_a \omega_m$$

The slope of the discharge curve should be equal to or greater than this.

$$\frac{E_c}{RC} \geqq E_c m_a \omega_m \quad \text{or} \quad RC \leqq \frac{1}{m_a \omega_m} \tag{289}$$

In order that an insignificant portion of the radio-frequency voltage applied to the input will appear across the output, the capacity C should be made about ten times greater in magnitude than the interelectrode capacity of the diode. Therefore, in the design of a diode detector circuit, the capacity C is selected by means of the above criterion and then an appropriate value for R is determined by equation 289. Since the efficiency of a detector increases with the magnitude of the load resistance, R should be made as large as possible.

Current does not begin to flow through the diode until the instant of time when the input voltage is equal to the output voltage e_o which is developed across the resistance R. Over a period of several cycles, the conduction can be assumed to start when the carrier passes through a certain angular position θ_0. At that time, the output voltage is equal to the input voltage.

$$e_o = E_c(1 + m_a \cos \omega_m t) \sin \theta_0$$

On the other hand, the equation of the envelope on the input modulated wave was

$$e_e = E_c(1 + m_a \cos \omega_m t)$$

The efficiency η of the detector is defined as the relative amplitude of the rectified output wave with respect to the amplitude of the envelope on the modulated wave. It is equal to e_o divided by e_e.

$$\eta = \frac{e_o}{e_e} = \sin \theta_0 \tag{290}$$

As the efficiency approaches 100 per cent, the rectified output voltage approaches the peak value of the alternating input voltage.

The efficiency of a diode with a particular load is practically constant over the usual operating range; consequently, there is a linear relationship between the input and output voltages. Since the efficiency is a function of the load, however, it is usually convenient to make calculations graphically. The characteristic curves of a diode are constant rms input voltage lines drawn on a coordinate axes of d-c output current vs d-c output voltage. The load line is a straight line drawn through the origin with a slope which is equal to the reciprocal of the load resistance R. As long as the output is utilized directly, the diode can be assumed to operate along this load line. Usually, however, the output is coupled through a condenser to the input grid resistor R_g of a succeeding audio or video amplifier. In that event, the dynamic load on the diode consists of the parallel combination of R and R_g. If the reactance of the coupling condenser is neglected, the tube can be assumed to operate along a straight line which has a slope equal to the reciprocal of the combined impedance of R and R_g and which passes through the quiescent point. The latter is at the intersection of the resistance R load line and the carrier voltage line.

When the resistance R_g is relatively small, the modulation index must remain below a certain value if distortion of the output wave is to be avoided. The effect can be demonstrated best by drawing a load line for the combined impedance of R and R_g on a diode characteristic curve in the manner indicated above and noting that there is a certain minimum value below which the envelope cannot fall without the output wave being clipped off.

One of the advantages of the diode detector is that it provides an output voltage of the proper polarity for the operation of automatic volume control (A.V.C.). Automatic volume control is a system whereby the output voltage of a receiver is maintained at a practically constant value regardless of the strength of the incoming signal. The d-c component of the output from the detector is used to control the bias on one or more variable-mu tubes in the radio-frequency amplifier which precedes it in the receiver. As shown in Fig. 110, the dynamic curve of a variable-mu tube is such that an increase of bias on the tube tends to reduce its amplification, and a decrease in the bias tends to increase the amplification. An increase in signal strength reduces the amplification of the preceding stages and maintains the output practically constant. The condenser C_A and the resistance R_A of Fig. 109 are used to filter out the modulating signal from the output of the detector. The time constant R_AC_A is made large enough so that the automatic volume control will not "wash out" the low-frequency components of the modulating signal and yet it is made small enough so that the control will adequately follow the changes in the strength of the signal being received. The values used for this constant vary from approximately 0.1 to 0.5 sec.

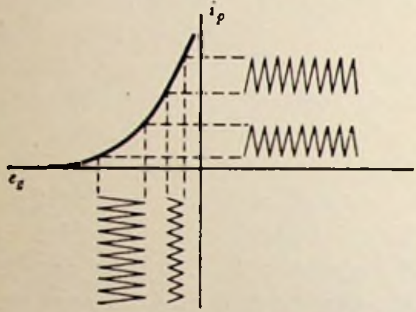


FIG. 110. Illustration of the automatic-volume-control process on the dynamic curve of a variable-mu tube.

Delayed automatic volume control is one which does not begin to operate until the incoming signal reaches a certain predetermined voltage level. The input signal is fed through a condenser to the plate of a diode section which is used exclusively for automatic volume control. A negative voltage (with respect to cathode) is applied to the plate of this diode section; consequently, it does not rectify any signal until the strength of that signal exceeds the magnitude of the negative plate voltage.

A triode can also be used as a detector. The tube can be operated on the parabolic portion of its dynamic curve and act as a squared-law detector. When a modulated signal is applied to the grid, the output plate current takes the form indicated in Fig. 107*b*. This current is passed through a parallel combination of resistance and capacity which filters out the radio frequency component. When used in this manner, the tube is known as an "anode detector." A resistance capacity combination is sometimes put in the grid lead of a triode and the tube operated at zero bias. In this case, the parabolic grid-current vs. grid-voltage curve is used to produce the detection. The average current flowing into the grid takes a form somewhat like that of the heavy dotted curve in Fig. 107*b* and the voltage built up across resistance and capacity combination follows the modulation envelope of the incoming signal. Since this voltage is applied to the grid, it is amplified by the tube before appearing in the plate circuit; consequently, a detector of this sort is very sensitive.

A grid detector can be given an even greater sensitivity by the application of a little positive feedback or regeneration. When a tube is used in this manner, it is known as a "regenerative detector." Since the tube acts in part like an amplifier, the discussion of section 5–3 applies. The addition of positive feedback causes an increase in the gain or sensitivity of the unit. If too much feedback is added, however, the tube will become unstable and begin to oscillate. Sometimes an excess of feedback is purposely applied to a regenerative detector and the oscillations produced by it are beat with the incoming signal to form a heterodyne (see section 9–6). Such an application is known as an "autodyne detector." The regenerative detector has its greatest sensitivity when it is operated on the verge of oscillation; but, it is difficult to adjust the feedback to exactly this level. High sensitivity is obtained and critical adjustments are avoided by the use of superregeneration. A low frequency (20 kc) "quench" voltage is added to the grid or plate of a regenerative detector. The tube operates with an optimum amount of feedback during the positive peaks of the quench-voltage cycle. At these times, the incoming signal triggers the tube into oscillation; but, the oscillations are immediately stopped by the reverse portion of the quench-voltage cycle so that the tube continues to operate in its normal manner. The quench voltage is supplied either from an external source or from a self oscillation of the tube. Superregenerative detectors tune very broadly; hence, they can be advantageously employed for the reception of very high-frequency signals.

A device which indicates the magnitude of a radio wave and is closely related to detectors is the bolometer. It consists of a fine resistance wire which is heated by the passage of the incoming radio-frequency current. Since the temperature of the wire is proportional to the I^2R loss in it, the resistance of the wire is also proportional to the power input. The change in resistance is measured on a bridge, and from this measurement the strength of the incoming electric wave can be calculated. The bolometer is particularly useful as a measuring instrument for very high-frequency fields because it can be accurately calibrated at low frequencies and maintain that calibration at high frequencies.

9–6. Heterodyne Detection. Everyone has observed the beat note which is formed when two sound waves having slightly different frequencies are combined. The magnitude of the resultant wave varies at a fundamental frequency which is equal to the difference between the frequencies of the other two. The same kind of beat can be formed between two electric waves of slightly different frequency. Suppose that the equation of one of the waves is

$$e_1 = E_1 \sin \omega t \tag{291}$$

and the equation for the other wave is

$$e_2 = E_2 \sin (\omega \pm \omega_d)t = E_2[\sin \omega t \cos \omega_d t \pm \cos \omega t \sin \omega_d t] \tag{292}$$

where ω_d is the difference in frequency of the two waves. The sum of equations 291 and 292 is

$$e = e_1 + e_2 = [E_1 + E_2 \cos \omega_d t] \sin \omega t \pm [E_2 \sin \omega_d t] \cos \omega t \tag{293}$$

This equation consists of two components which are 90° out of phase. Using the familiar rules of vector addition, we find that the resultant wave is

$$e = \sqrt{(E_1 + E_2 \cos \omega_d t)^2 + (E_2 \sin \omega_d t)^2} \sin (\omega t + \phi) \tag{294}$$

The angle ϕ is a phase angle which depends on the magnitude of the two components. It is practically constant when $E_2 \ll E_1$. The coefficient of equation 294 or amplitude of the resultant wave reduces to

$$E = \sqrt{(E_1^2 + E_2^2) + 2\, E_1 E_2 \cos \omega_d t} \tag{295}$$

The expansion of equation 295 by the binomial theorem shows that this coefficient can also be expressed as

$$E = (E_1^2 + E_2^2)^{\frac{1}{2}} + \frac{1}{2} \frac{2\, E_1 E_2 \cos \omega_d t}{(E_1^2 + E_2^2)^{\frac{1}{2}}} + \cdots \tag{296}$$

If the magnitude of one of the two beating waves is relatively small, the terms which were omitted from equation 296 become negligible. Consequently, it is seen that the resultant wave is amplitude modulated at a frequency which is equal to the difference frequency between the two waves. When such a wave is applied to a detector, the output is a wave having a frequency equal to the difference frequency. This process of heterodyne detection, as it is called, is often used to reduce the frequency of an incoming wave so that it can be amplified more advantageously.

A typical circuit for producing a difference frequency or heterodyne is shown in Fig. 111. The incoming signal is applied to the

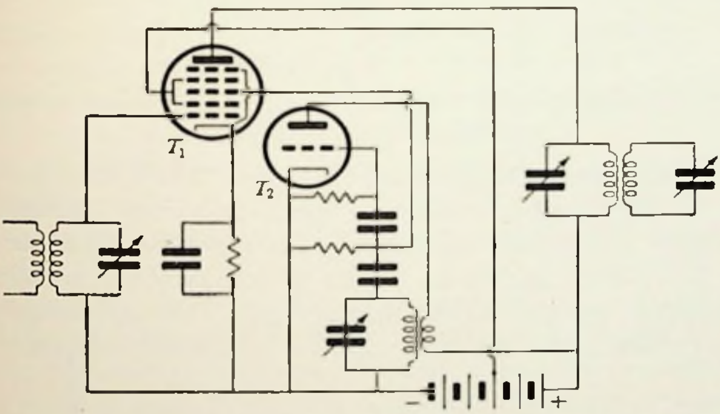


FIG. 111. Circuit diagram of a heterodyne detector or frequency converter. The local oscillator T_2 is of the tuned-grid type.

first grid of the mixer tube T_1 and accordingly controls the electron stream of the tube. The output from the local oscillator T_2 which is operating at a slightly different frequency is applied to the third grid of T_1. Its effect is added to the effect of the first grid so that the alternating current flowing in the plate circuit is amplitude modulated at the difference frequency. Since this current flows through a circuit which is tuned to the difference frequency, the only appreciable voltage drop appearing across the tuned circuit is at the difference frequency. The screen grid of T_1 prevents interaction between the input and oscillator grid.

Sometimes the local oscillator and the mixer are combined into one tube as shown in Fig. 112. Here the first grid is the grid of the oscillator, and the second grid corresponds to the plate of the oscillator. The input signal is applied to the fourth grid which is isolated from the others by a screen. In other respects this circuit is exactly the same as the one in Fig. 111.

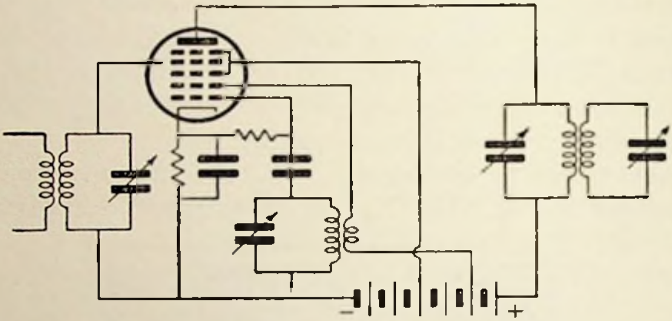


Fig. 112. A variation of Fig. 111 using a single tube (pentagrid converter) for both oscillator and mixer.

9–7. Frequency Discriminator. Figure 113*a* is the circuit diagram of one possible connection for a frequency discriminator. Essentially this circuit consists of two ordinary diode detectors (Fig. 109) which are fed by a pair of doubly tuned transformers connected in parallel on the primary side. The secondary of the upper transformer is tuned to resonate at a higher frequency than

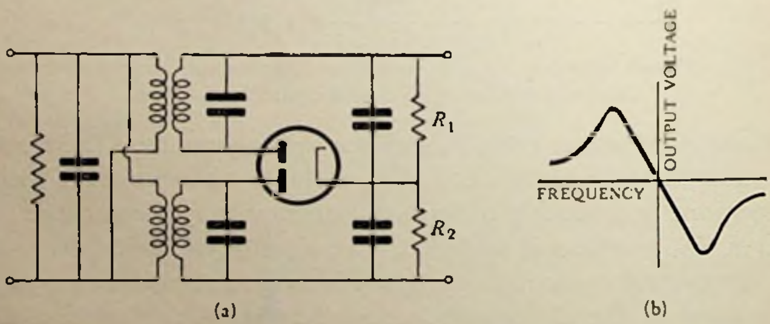


Fig. 113. (*a*) Frequency-discriminator circuit and (*b*) its corresponding performance curve.

that at which the secondary of the lower transformer is resonated. When the frequency of the input is midway between these two frequencies, the current passed by each half of the diode is the same, the voltages developed across R_1 and R_2 are the same, and the net output voltage is zero. If a lower frequency is applied to the input, however, the voltage drop developed across R_2 is greater than the voltage drop developed across R_1, and the net voltage at the output is positive. On the other hand, the application of a higher frequency to the input will develop a higher voltage drop across R_1 than that which is developed across R_2 and the net output voltage will be negative. A curve of the output voltage as a function of frequency is shown in Fig. 113*b*. The resonant peaks of the two secondary circuits can be clearly distinguished in this figure. Since the output voltage is directly proportional to the deviation in frequency from some mean value, this device can be used to demodulate a frequency-modulated wave. Other circuits are available which accomplish the same end, but a detailed discussion of them will be omitted here.

A phase discriminator similar to the one described in section 3-8 is used to demodulate a phase-modulated wave.

9-8. Receiver Block Diagram. A receiver performs two essential functions. First it selects the desired signal from all the others which may be present, and second it demodulates that signal. The desired frequency is selected by the resonant properties of tuned circuits and the demodulation is accomplished by a detector or a discriminator. Since the incoming signals are frequently too weak to make the detector function properly, it may be preceded by a tuned radio-frequency (T.R.F.) amplifier. Such an arrangement is known as a "T.R.F. receiver" and has greater sensitivity and selectivity than the detector having just a tuned input.

A greater gain and more flexible response characteristic can be realized when a radio-frequency amplifier is constructed for operation at a single fixed frequency. To incorporate this advantage in a receiver and at the same time make it respond to a variety of signals having different frequencies, a heterodyne detector is used which converts these incoming signals to a fixed intermediate frequency (I.F.). The receiver is then tuned by changing the frequency of its local oscillator. As pointed out in section 9-6, however, a given intermediate frequency can be produced by beating the local oscillator with either a higher or lower frequency input. The only requirement is that the difference between the frequency of the local oscillator and the incoming signal must be equal to the intermediate frequency. The interference which results from the simultaneous reception of these two signals is limited by feeding the mixer tube by a circuit which resonates at one of them (usually the lower). The other which is at a frequency equal to twice the intermediate frequency away from the former is accordingly attenuated. When the heterodyne detector is also preceded by a tuned-radio-frequency amplifier, the attenuation of this spurious response, called the image, is even more effective. A measure of the effectiveness with which it is suppressed is the image response ratio. This factor is defined as the ratio of the signal strength of the image required to produce a given output to the signal strength of the desired frequency which produces the same output.

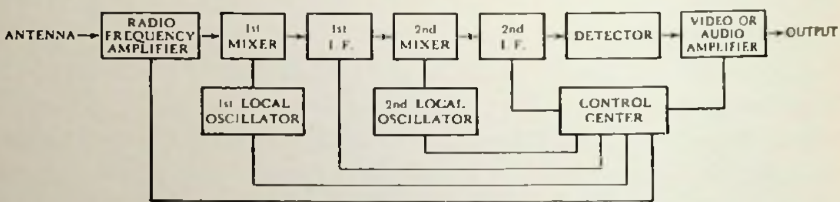


FIG. 114. Typical block diagram of a receiver.

A receiver which utilizes heterodyne detection in the manner described above is called a superheterodyne. Figure 114 is a generalized block diagram for a typical receiver of this sort. As indicated in the figure, the frequency of the signal may be changed several times before it eventually reaches the detector. An audio or video amplifier follows the detector so that the necessary power output can be produced. The control center includes such things as manual volume control, automatic volume control, and automatic frequency control. The latter detects changes in the intermediate frequency produced by the heterodyne detector and compensates for them by controlling the frequency generated by the local oscillator.

Occasionally an additional local oscillator which has an adjustable frequency is incorporated in a receiver. It is arranged so that it can produce an audible beat note with the intermediate frequency signal. This device gives C.W. (continuous wave) telegraph signals a tone so that they can be read.

The major difference between a receiver intended for frequency or phase modulation and one intended for amplitude modulation is in the use of a discriminator for demodulation instead of a detector. In addition, however, frequency or phase-modulation receivers are also equipped with a limiter which cuts off any amplitude modulation which may be present. In general terms, it can be said that static and interference are usually manifest as a change in amplitude and are automatically eliminated by this device. The limiter has no effect on the modulation because it passes changes in frequency or phase without appreciable distortion.

The sensitivity of a receiver is defined in terms of the input signal strength required to develop a certain standard output. At first sight it may seem that the sensitivity could be increased indefinitely by simply increasing the number of stages of amplification. This cannot be done because signals below a certain strength are masked by the inherent thermal noise of the receiver's resistance components. The rms value of this noise voltage E is given by the equation

$$E^2 = 4\,kT\int_{f_1}^{f_2} R\,df \tag{297}$$

where k is Boltzmann's constant (1.38×10^{-23}), T is the absolute temperature in degrees Kelvin, R is the resistance in ohms, and f is the frequency in cycles per second. By integration over a uniform resistance pass band of Δf width, we find that the noise voltage is

$$E = \sqrt{4\,kTR\,\Delta f} \tag{298}$$

This represents the theoretical minimum noise level which can be obtained in a receiver. In practice, however, it is found that other circuit elements contribute much more noise than the true circuit resistances; therefore, the actual noise level is somewhat higher than indicated by this figure.

Problems

9-1. (*a*) A carrier is frequency modulated with a symmetrical square wave. Sketch the variation in frequency as a function of time. Directly below this figure sketch the phase variation which accompanies the frequency modulation.

(*b*) A carrier is phase modulated with a symmetrical square wave. Sketch the frequency variation which accompanies the phase modulation.

9–2. A 1000-cps 100-volt carrier is 60 per cent amplitude modulated with a 60-cps wave. Add the sine curves of the carrier and the two side bands point by point and show that the resultant curve has the proper shape.

9–3. Find the power required to amplitude modulate a 1-kw carrier 80 per cent.

9–4. Plot a series of diagrams similar to Fig. 100 for a carrier which is frequency modulated by swinging it through a 1-kc change in frequency by a modulating frequency of

(*a*) 250 cps
(*b*) 500 cps
(*c*) 1000 cps
(*d*) 2000 cps

Construct all of the diagrams to the same scale and arrange them so that they can be readily compared.

9–5. (*a*) Derive an equation for the inductive reactance produced by tube T_1 of Fig. 105.

(*b*) Assume that T_1 is a 6J7 operating from a supply voltage $E_{bb} = 150$ volts. The voltage developed across the tuned circuit of the oscillator is 25 volts peak at a frequency of 2 mc. If R is 200,000 ohms, C_g is 0.0001 μf, and C_p is 0.01 μf, find the inductance introduced into the oscillator circuit when the grid bias is $-$ 3.0 volts.

(*c*) Repeat part (*b*) for a grid bias of $-$ 1.0 volts.

Experiment 19

Object:

The object of this experiment is to produce amplitude modulation and to study some of its characteristics.

Preliminary:

Solve problems 9–2 and 9–3.

Performance:

1. Set up a modulator like Fig. 101 using a 6J5 tube operated on the curved portion of its characteristic curve. Supply a 1000-cps carrier to the grid from a beat-frequency oscillator and supply a 60-cps modulating voltage to the grid from a low-voltage power transformer. Connect a cathode-ray oscillograph, vacuum-tube voltmeter, and wave analyzer across the tuned circuit.

2. Tune the plate-load circuit into resonance at the frequency of the carrier with the modulation source turned off.

3. Turn on the modulation and adjust its voltage to give approximately 60 per cent modulation as indicated by the peak reading of the vacuum-tube voltmeter. Observe the wave form produced on the cathode-ray oscillograph and compare it with the results of problem 9–2.

4. Measure the amplitude of all the component frequencies in the output voltage by means of the wave analyzer. Compare the frequencies found with those predicted by equation 260. Plot the measured amplitude of the components as a function of frequency in the manner of Fig. 100.

Experiment 20

Object:

The object of this experiment is to set up a reactance-tube modulator and study the characteristics of frequency modulation.

Preliminary:

Solve problem 9–5.

Performance:

1. Assemble a circuit like Fig. 105 making T_1 a 6J7 and T_2 a 6J5. Set the bias on the 6J7 at -2.5 volts and adjust the oscillator frequency to 2 mc. Adjust the frequency by beating the output from the oscillator with the output from a standard signal generator.

2. Leave the oscillator tuning alone and measure the change in output frequency as a function of the bias on the 6J7. Plot a curve of frequency vs. grid bias. Compare this curve with the results which you would expect on the basis of the answers to problem 9–5.

3. With the bias on T_1 again set at -2.5 volts, pick up the oscillator output on a communication receiver which is equipped with an S meter (calibrated in decibels or microvolts) and a crystal filter. Apply a 500-cps modulating signal to the grid of T_1 and with the receiver pick up and measure the strength of the resultant carrier and all the side bands. Plot the results of this test in the form of Fig. 100.

4. Compute the modulation index for the signal used and compare the theoretical magnitudes of the side bands and carrier with the results of part 3.

Experiment 21

Object:

The object of this experiment is to study crystal, diode, and heterodyne detection.

Preliminary:

1. A 6H6 diode detector connected as shown in Fig. 109 has a load resistance R equal to 200,000 ohms. From the characteristic curves in a tube manual, obtain data and plot a curve of rectified current as a function of the applied input voltage.

2. If the by-pass condenser C is 0.0005 μf, calculate the maximum frequency which can be used to modulate a 10-volt (rms) carrier input to the detector without distorting the peaks of the modulation in the rectified output. Assume 75 per cent modulation.

Performance:

1. Connect an adjustable crystal detector across a low-voltage 60-cps supply through a dropping resistance (250 ohms). Apply the supply voltage to the horizontal deflecting plates of a cathode-ray oscillograph and apply the voltage across the dropping resistor to the vertical plates. The image formed on the cathode-ray screen will be a plot of the current passed through the detector as a function of the applied voltage. Adjust the crystal to the point where the optimum characteristic curve appears on the screen. Leave this adjustment fixed and measure the resistance of the crystal in each direction.

2. Substitute a section of a 6H6 diode for the crystal detector in the above circuit, and observe its characteristic curve. Compare the curve of the diode with that which was obtained with the crystal.

3. Connect a 6H6 diode in the circuit of the preliminary problems. Use a 465-kc doubly tuned transformer for the input transformer of Fig. 109 and make $C_A = 0.1$ μf and $R_A = 1$ megohm. Place a d-c microammeter in series with R and connect a vacuum-tube voltmeter across the secondary side of the input transformer. Apply a 465-kc carrier to the input. Vary the input voltage and measure the rectified output current. Plot the results in a contrasting color on the same coordinate axis as part 1 of the preliminary.

4. Apply amplitude modulation to the carrier. Observe the wave form developed across the secondary of the input transformer, across the resistance R, and across the capacity C_A. Sketch these wave forms and explain their differences.

5. Set the voltage of the carrier input at 10 volts and the percentage modulation at 75 per cent. Increase the modulating frequency until the wave form developed across R is distorted. Compare this frequency with the answer to part 2 of the preliminary.

6. Assemble a circuit like Fig. 111 or 112 and connect its output to the input transformer of the diode detector used in the previous experi-

ments. Substitute a 100,000-ohm resistance for the tuned circuit connected to the input grid of the converter or mixer tube. Make certain the local oscillator is working and set its frequency at 2 mc. Apply an amplitude-modulated signal across the 100,000-ohm input resistor and find the two frequencies which produce an output in the detector as observed on a cathode-ray oscillograph connected across the resistance R. Compare these frequencies with the expected results.

7. Replace the 100,000-ohm input resistance by the original resonant circuit and tune it to resonate at the lower frequency observed in part 6. Measure the image response ratio.

References

Armstrong, E. H., *et al.*, *Frequency Modulation* (Notes on Lecture Course Presented During October, November, and December 1940). New York: American Institute of Electrical Engineers, 1940.

Born, H., "Indirect Modulation of Centimeter Waves," *Hochf-tech. u. Elec-akus.*, October 1940, pp. 112–118. (In German.)

Hund, A., "Reactance Tubes in F–M Applications," *Electronics*, October 1942, pp. 68–71, 143.

Klumb, H., "Bolometers for Short Electric Waves," *Zeits. f. Techn. Phys.*, March 1940, pp. 71–75. (In German.)

Klumb, H., "Observations and Investigations of Crystal Detectors," *Physik. Zeits.*, October 1939, pp. 640–643. (In German.)

M.I.T. Staff, *Applied Electronics.* New York: John Wiley & Sons, Inc., 1943, pp. 624–715.

Parker, W. N., "A Unique Method of Modulation for High-fidelity Television Transmitters," *Proc. I.R.E.*, August 1938, pp. 946–962.

Roder, H., "Amplitude, Phase, and Frequency Modulation," *Proc. I.R.E.*, December 1931, pp. 2145–2175.

Rottgardt, J., "Investigations on Detectors in the Region of Very Short Electric Waves," *Zeits. f. Techn. Phys.*, August 1938, pp. 262–264. (In German.)

Tibbs, C., "Phase and Frequency Modulation," *Wireless World*, September 1942, pp. 210–213.

CHAPTER X

RADIATION

10–1. Basic Concepts. The fundamentals of radiation are based on principles which are well known to every student of electrical engineering. The first of these principles is concerned with the magnitude of induced voltage generated by a changing magnetic field. One volt is induced in a circuit when the magnetic flux threading it changes at the rate of one weber per second. Faraday's law is a formulation of this concept. The other familiar principle of importance in the study of radiation is concerned with the magnitude of the magnetic-field intensity which is produced by the flow of an electric current. The magnetic-field intensity multiplied by the length of the magnetic path is equal to the ampere turns. Ampere's law is a formulation of this relationship. The vector equations which express Ampere's law and Faraday's law are collectively known as "Maxwell's equations."

To formulate a vector expression for Faraday's law, consider a small rectangular area dx by dy lying in the xy plane. Assume that it is threaded by a magnetic flux perpendicular to the area and having a density of B_z webers per square meter. If a conductor marks off the boundary of the area and t is time, a change in the magnetic flux density would induce a voltage in the conductor.

$$E = -\frac{\partial B_z}{\partial t}\,dx\,dy \tag{299}$$

Even if the conductor were not present, a voltage gradient would nevertheless be set up along the sides of the area under consideration. If the electric-field intensity along the side closest to the y axis is $\mathcal{E}_y$ volts per meter, the electric-field intensity along the opposite side of the rectangle would differ from it by the rate of change in electric-field intensity multiplied by the distance between the two sides

$$\mathcal{E}_y + \frac{\partial \mathcal{E}_y}{\partial x}\,dx$$

Similarly, the electric-field intensity along the other two sides of the rectangle could be expressed as $\mathcal{E}_x$ and $\mathcal{E}_x + \dfrac{\partial \mathcal{E}_x}{\partial y} dy$. The voltage along a side of the rectangle is equal to the field intensity multiplied by the length of the side. The net voltage developed in the loop is equal to the algebraic sum of the voltages developed along each side of the rectangle.

$$E = \mathcal{E}_x\, dx + \left(\mathcal{E}_y + \frac{\partial \mathcal{E}_y}{\partial x} dx\right) dy - \left(\mathcal{E}_x + \frac{\partial \mathcal{E}_x}{\partial y} dy\right) dx - \mathcal{E}_y\, dy \quad (300)$$

The voltage E in equations 299 and 300 are the same. Equating them and simplifying gives the result

$$\frac{\partial \mathcal{E}_y}{\partial x} - \frac{\partial \mathcal{E}_x}{\partial y} = -\frac{\partial B_z}{\partial t} \quad (301)$$

A change in the x or y component flux density would not induce any voltage in the above circuit; hence, B_x and B_y do not appear as factors in the above equation. Changing these factors, however, does induce voltages in circuits lying parallel to the other two planes; consequently, two other equations besides this one are necessary for the complete expression of Faraday's law. It is convenient to write all three expressions in the form of a single equation. For this purpose the unit vectors i lying along the x axis, j lying along the y axis, and k lying along the z axis are introduced. These vectors identify the direction of the component associated with them in the same manner as the factor j identifies the direction of the component associated with it in a complex number. With this notation the three equations which express Faraday's law can be written in the form of a determinant.

$$\begin{vmatrix} i & j & k \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ \mathcal{E}_x & \mathcal{E}_y & \mathcal{E}_z \end{vmatrix} = -i\frac{\partial B_x}{\partial t} - j\frac{\partial B_y}{\partial t} - k\frac{\partial B_z}{\partial t} \quad (302)$$

Equation 301 is simply the k component of equation 302. The other two equations are given by i and j components. The symmetrical form of equation 302 makes it easy to remember.

Usually equation 302 is written in terms of the magnetic-field intensity H rather than the magnetic-flux density. The units of magnetic-field intensity in the rationalized MKS system are amperes per meter. A long solenoid has a magnetic field intensity of 1 amp/meter when it is operated with 1 amp turn/meter. The field intensity at the center of such a solenoid is

$$H \text{ (oersteds)} = \frac{0.4\,\pi NI \text{ (amp turns)}}{l \text{ (cm)}} \tag{303}$$

therefore, a field intensity of 1 amp/meter is equal to $4\,\pi \times 10^{-3}$ oersteds. In a nonmagnetic medium such as air, a 1-amp/meter field intensity produces a magnetic-flux density of $4\,\pi \times 10^{-3}$ gauss or $4\,\pi \times 10^{-7}$ webers/meter^2. In the rationalized MKS system of units, the permeability μ_0 of free space is taken as being equal to the factor $4\,\pi \times 10^{-7}$. As a result, bothersome conversion factors which are usually found in equations of this kind are avoided. The magnetic-flux density becomes

$$B(\text{webers/meter}^2) = \mu(\text{henries/meter})\ H(\text{amp/meter})$$

where the permeability μ is expressed in terms of the rationalized system of units. It is equal to the permeability of the medium when referred to free space multiplied by μ_0. Equation 302 becomes

$$\begin{vmatrix} i & j & k \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ \mathcal{E}_x & \mathcal{E}_y & \mathcal{E}_z \end{vmatrix} = -\,\mu \frac{\partial \mathbf{H}}{\partial t} \tag{304}$$

where **H** is the vector corresponding to $(iH_x + jH_y + kH_z)$. For convenience in writing, the determinant on the left side of equation 304 is abbreviated by mathematical shorthand to either "$\nabla \times \mathcal{E}$" or "curl $\mathcal{E}$" where $\mathcal{E}$ is the vector $(i\mathcal{E}_x + j\mathcal{E}_y + k\mathcal{E}_z)$.

Following a procedure similar to the above, a vector expression for Ampere's law can also be formulated. First, however, an understanding of displacement current is necessary. By recognizing the necessity for the existence of this current, Maxwell was able to predict the existence of radio waves. One of Kirchhoff's laws states that the current flowing into a closed area must be equal to the current flowing out. Suppose that a closed area is arranged to surround one plate of a simple parallel plate condenser, but not the other. If the condenser is charged, a *conduction* current flows into the area under consideration, but no *conduction* current flows out.

Current flowing out of the area must be of a different form. The nature of this current can be found in the following way: The capacity C of a parallel plate condenser with air dielectric is

$$C(\text{farads}) = \frac{KA(\text{meters}^2)}{d(\text{meters})} \tag{305}$$

where A is the area of the plates, d is the distance between them, and K is the dielectric coefficient. For free space, the latter is

$$K_0 = \frac{1}{36\,\pi \times 10^9}\text{ farads/meter}$$

If the voltage across the plates is E, the electric-field intensity between them is

$$\mathcal{E} = \frac{E}{d} = \frac{Q}{Cd} = \frac{q}{K} \tag{306}$$

where Q is the total charge and q is the charge per unit area on the condenser. Differentiating equation 306 with respect to time shows that a current density $\dfrac{\partial q}{\partial t}$ exists in space when the electric-field intensity changes with time.

$$\frac{\partial q}{\partial t} = K\frac{\partial \mathcal{E}}{\partial t} = \frac{\partial D}{\partial t} \tag{307}$$

The factor $K\mathcal{E}$ is called the electric displacement or electrostatic-flux density D, and the current represented by equation 307 is called the *displacement* current to distinguish it from the more familiar *conduction* current.

For a moment consider a solenoid in the form of a torus. The magnetic-field intensity along the center line of the magnetic path when multiplied by the length of the path is equal to the ampere turns on the solenoid. It is also equal to the net current flowing through the flat area bounded by the center line. The above is an illustration of the general principle that the magnetic-field intensity along the boundary of an area multiplied by the length of the boundary is equal to the net current flowing through the area. A vector expression for Ampere's law is obtained by applying this principle to differential areas in the three coordinate planes.

The net current flowing through a small rectangular area dx by dy in the xy plane can be written as

$$I = \left(J_z + \frac{\partial D_z}{\partial t}\right) dx\, dy \tag{308}$$

where J_z is the conduction current density (amp/meter²) and $\frac{\partial D_z}{\partial t}$ is the displacement current density. Assuming that the magnetic-field intensity along the side of the area closest to the y axis is H_y amperes per meter, the magnetic-field intensity along the opposite side is $H_y + \frac{\partial H_y}{\partial x} dx$. Similarly the magnetic-field intensities along the other two sides are H_x and $H_x + \frac{\partial H_x}{\partial y} dy$. The total of the products of the field intensities and the length of the bounding sides is equal to the current flowing through the area or

$$I = H_x dx + \left(H_y + \frac{\partial H_y}{\partial x} dx\right) dy - \left(H_x + \frac{\partial H_x}{\partial y} dy\right) dx - H_y dy \tag{309}$$

From equations 308 and 309

$$\left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y}\right) = J_z + \frac{\partial D_z}{\partial t} \tag{310}$$

Again two more equations like 310 are required to account for the other two planes. The complete vector equation for Ampere's law is

$$\begin{vmatrix} i & j & k \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ H_x & H_y & H_z \end{vmatrix} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} = \mathbf{J} + K\frac{\partial \mathcal{E}}{\partial t} \tag{311}$$

where $\mathbf{J}$ is the vector $(iJ_x + jJ_y + kJ_z)$ and $\mathbf{D}$ is $(iD_x + jD_y + kD_z)$. As with Faraday's law, the determinant portion of this equation can also be written as "$\nabla \times \mathbf{H}$" or "curl $\mathbf{H}$."

10–2. Characteristics of Radiation. From Maxwell's equations, it can be shown that a flat conducting sheet of infinite extent generates plane electromagnetic waves when it carries an alternating current. Suppose that the sheet lies on the yz plane passing through the origin and that the alternating current flows parallel to the y axis. The magnetic-field intensity due to this current is all in the

z direction; therefore, H_x and H_y together with all of their derivatives are zero. Furthermore, the field intensity is uniform in any given zy plane so $\dfrac{\partial H_z}{\partial y} = 0$. The electric-field intensity does not vary in either the y or the z direction; hence,

$$\frac{\partial \mathcal{E}_z}{\partial y} = \frac{\partial \mathcal{E}_y}{\partial z} = \frac{\partial \mathcal{E}_x}{\partial z} = \frac{\partial \mathcal{E}_x}{\partial y} = 0 \tag{312}$$

Also there is no conduction current in the open space in front of the conducting sheet. The substitution of the above conditions into the vector equation of Faraday's law (equation 304) gives the following results for the three component equations

$$(i)\ 0 = 0 \quad (j)\ -\frac{\partial \mathcal{E}_z}{\partial x} = 0 \quad (k)\ \frac{\partial \mathcal{E}_y}{\partial x} = -\mu \frac{\partial H_z}{\partial t} \tag{313}$$

Similarly the substitution of the above conditions into the vector equation of Ampere's law (equation 311) gives the following component equations:

$$(i)\ 0 = K\frac{\partial \mathcal{E}_x}{\partial t} \quad (j)\ -\frac{\partial H_z}{\partial x} = K\frac{\partial \mathcal{E}_y}{\partial t} \quad (k)\ 0 = K\frac{\partial \mathcal{E}_z}{\partial t} \tag{314}$$

Differentiating the (k) equation from Faraday's law (313) with respect to time t, and differentiating the (j) equation from Ampere's law (equation 314) with respect to x, we have

$$\frac{\partial^2 \mathcal{E}_y}{\partial t\, \partial x} = -\mu \frac{\partial^2 H_z}{\partial t^2} \tag{315}$$

and

$$-\frac{\partial^2 H_z}{\partial x^2} = K\frac{\partial^2 \mathcal{E}_y}{\partial x\, \partial t} \tag{316}$$

The combination of equations 315 and 316 gives the wave equation

$$\frac{\partial^2 H_z}{\partial t^2} = \frac{1}{\mu K}\frac{\partial^2 H_z}{\partial x^2} \tag{317}$$

This expression is called the wave equation because its solution has the form of a wave moving with a velocity equal to $v = \dfrac{1}{\sqrt{\mu K}}$.

A solution to equation 317 is obtained by following the usual method for handling a partial differential equation which has two variables. The equation is written as the product of a function of x alone and a function of t alone

$$H_z = XT \tag{318}$$

If derivatives are indicated by primes, the substitution of this equation into equation 317 yields

$$XT'' = v^2TX'' \quad \text{or} \quad \frac{T''}{T} = \frac{X''}{X}v^2 \tag{319}$$

Since each side is a function of t or x alone, variations on one side of the equation cannot introduce variations on the other side, and each side is independently equal to a constant A. Consequently, each side of equation 319 yields an ordinary differential equation:

$$\frac{d^2T}{dt^2} - AT = 0 \tag{320}$$

and

$$\frac{d^2X}{dx^2} - \frac{AX}{v^2} = 0 \tag{321}$$

The solutions of these equations are

$$T = A_1 \cos(\sqrt{A}t + \phi) \tag{322}$$

and

$$X = A_2 \cos\left(\sqrt{A}\frac{x}{v} + \theta\right) \tag{323}$$

Since H_z was assumed to be the product of X and T in equation 318, the product of the above equations gives the desired solution.

$$H_z = A_3 \cos(\sqrt{A}t + \phi)\cos\left(\sqrt{A}\frac{x}{v} + \theta\right) \tag{324}$$

If the wave starts at a positive maximum when $t = 0$ and $x = 0$, the arbitrary phase angles ϕ and θ are zero and

$$H_z = A_3 \cos\sqrt{A}t \cos\sqrt{A}\frac{x}{v}$$

From trigonometry this equation expands into

$$H_z = H_0\left[\cos\sqrt{A}\left(t - \frac{x}{v}\right) + \cos\sqrt{A}\left(t + \frac{x}{v}\right)\right] \tag{325}$$

where H_0 is equal to the constant $\frac{A_3}{2}$.

The arguments of the cosine terms in 325 are $\left(t - \frac{x}{v}\right)$ and $\left(t + \frac{x}{v}\right)$. Due to the presence of these factors, the waves associated with them are functions of both time and space. The wave corresponding to the first term moves to the right with a velocity v and represents the generated wave. The wave corresponding to the second term moves to the left with the same velocity and represents the reflected wave. When the generated wave is not reflected, the latter term drops out of the equation. From the conditions of the problem, the constant $\sqrt{A}$ has to be equal to the frequency ω of the alternating current in the sheet. The generated wave portion of equation 325 becomes

$$H_z = H_0 \cos \omega \left(t - \frac{x}{v}\right) \text{ amp/meter} \tag{326}$$

The equation for the electrostatic-field intensity is obtained by differentiating the above equation with respect to x, substituting the result in the (j) portion of equation 314, and integrating. The result is

$$\mathcal{E}_y = \sqrt{\frac{\mu}{K}}\, H_0 \cos \omega \left(t - \frac{x}{v}\right) \text{ volts/meter} \tag{327}$$

From these results, it is apparent that the electric-field intensity and magnetic-field intensity are at right angles with each other in space and are in time phase. The combined wave motion of the electric and magnetic fields constitute the radiated electromagnetic wave. A diagrammatic picture of the fields in the wave is shown in Fig. 115*a*. The electric-field intensity lines are drawn in red and the magnetic-field intensity lines are drawn in black. It will be noted that the density of the lines varies as a cosine function along the x axis. As time passes the whole picture moves to the right with the velocity of v. Using the permeability and dielectric constant of free space, we see that this velocity is

$$v = \frac{1}{\sqrt{\mu_0 K_0}} = \frac{1}{\sqrt{\dfrac{4\pi \times 10^{-7}}{36\pi \times 10^{9}}}} = 3 \times 10^8 \text{ meters/sec}$$

which is the velocity of light.

The direction of the electrostatic-field intensity or flux lines determines the polarization of the wave. The wave of Fig. 115*a*, for example, is vertically polarized. A horizontally polarized wave would have been produced if the current in the conducting sheet had flowed parallel to the z axis instead of flowing parallel to the y axis.

A conventionalized diagram of the electromagnetic wave is given in Fig. 115*b*, and a vector diagram of the wave is given in Fig. 115*c*.

A receiving antenna picks up energy because voltages are induced in it by the passing wave. Consider the effect of the wave in Fig. 115 on a short length of wire placed in front of the conducting sheet and parallel to the z axis. The motion of the wave can be simulated by moving the conductor to the left. Under these circumstances, the conductor does not cut magnetic lines of force and has no voltage induced in it as a result, nor does it lie along an electric-field intensity line which would set up voltage difference between the ends. Similarly, an antenna which is short compared with a wavelength and which is mounted parallel to the x axis will pick up no signal from either the magnetic field or the electric field. On the other hand, an antenna mounted parallel to the y axis would meet with better success. It will cut a maximum possible number of magnetic lines of force and will have a maximum voltage generated in it. At the same time, it will also pick up a voltage from the electric-field intensity lines. Since they represent lines of voltage gradient, a voltage difference is established between the ends of the conductor lying along one of them. If the antenna were mounted at some angle with respect to the y axis, its pickup voltage would be correspondingly less. As a matter of fact, if we assume that the antenna is short and lies in a xy plane, its pickup is proportional to the cosine of the angle it makes with the y axis. A polar coordinate plot of the relative amount of pickup as a function of the angle the antenna makes with the oncoming wave is known as its "radiation pattern" or "directional diagram." For this particular antenna, the Hertz doublet, the radiation pattern is two circles in the form of a figure eight.

The theorem of reciprocity states that the radiation pattern of a receiving antenna is of the same form as its radiation pattern when

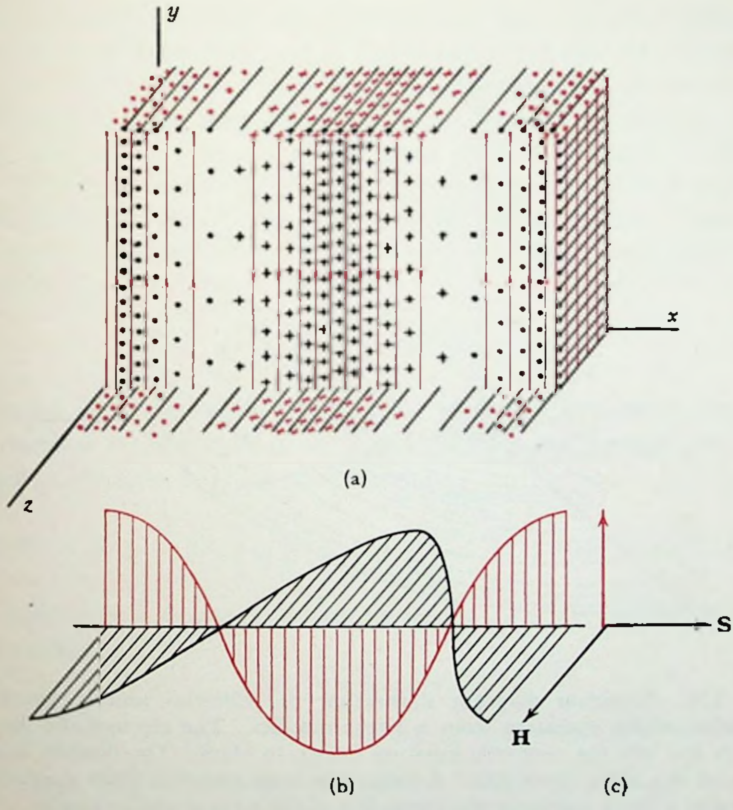


FIG. 115. (*a*) Electric- and magnetic-field distribution of a plane electromagnetic wave. The electrostatic lines of force are drawn in red and the magnetic lines of force are drawn in black. A conventionalized diagram of the wave is given in (*b*) and a vector diagram is given in (*c*).

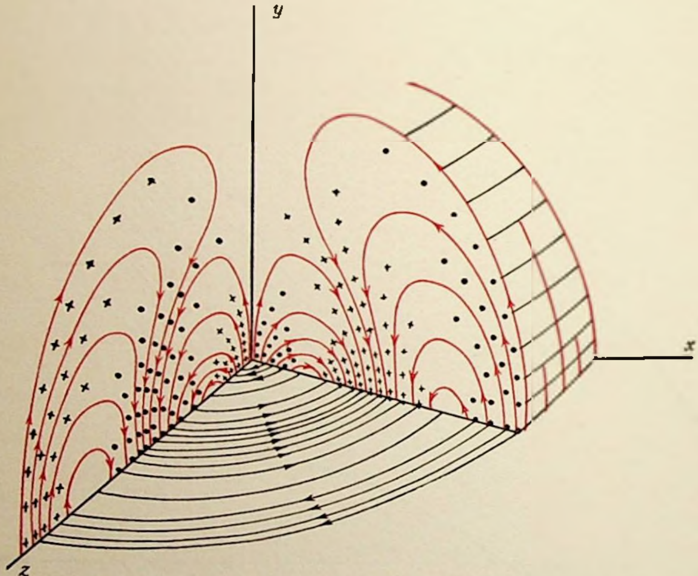


FIG. 116. Sectional diagram illustrating the electric- and magnetic-field distribution of the radiation from a Hertz doublet. The electrostatic lines are drawn in red and the magnetic lines are drawn in black. The doublet is at the origin and lies along the y axis. A sector has been removed from the radiation field so that a more complete representation of the wave could be made.

used as a transmitting antenna. A little later, the equation of the field intensity radiated from a Hertz doublet will be given. It should be noted that it leads to a radiation pattern of this same form.

The Poynting vector **S** of an electromagnetic wave lies along the direction of propagation and is the measure of the rate of energy flow due to the wave (Fig. 115*c*). A consideration of the energy in the electric and magnetic fields lead to its formulation. The energy stored in a uniform electrostatic field of intensity $\mathcal{E}$ is $W_e = \frac{1}{2} K\mathcal{E}^2$ joules per cubic meter, and the energy stored in a uniform magnetic field is $W_m = \frac{1}{2} \mu H^2$ joules per cubic meter. From the equations of the magnetic and electrostatic field intensities (326 and 327), we have

$$\mathcal{E}_y = \sqrt{\frac{\mu}{K}} H_z \quad \text{or} \quad \frac{1}{2} K\mathcal{E}_y^2 = \frac{1}{2} \mu H_z^2$$

In other words, the energy carried by the electric wave is equal to that carried by the magnetic wave. The total energy per unit volume is the sum of these two energies

$$W = K\mathcal{E}_y^2 = K\mathcal{E}_y \left(\sqrt{\frac{\mu}{K}} H_z \right) = \sqrt{K\mu}\mathcal{E}_y H_z \text{ joules/meter}^3 \qquad (328)$$

The energy in a square volume of space 1 meter on edge and dx meters thick is

$$dW = \sqrt{K\mu}\mathcal{E}_y H_z \, dx \qquad (329)$$

The energy of the wave moves along with a velocity v; therefore, it will have passed through the above volume in a time $dt = \frac{dx}{v}$. The rate of energy flow or the power in the wave is

$$S_x = \frac{dW}{dt} = v\sqrt{K\mu}\mathcal{E}_y H_z = \mathcal{E}_y H_z \text{ watts/meter}^2 \qquad (330)$$

This is a formulation of Poynting's theorem. The power in an electromagnetic wave is equal to the product of the electric- and magnetic-field intensities. The Poynting vector for a plane wave is given by the product of equations 326 and 327.

$$S_x = \sqrt{\frac{\mu}{K}} H_0^2 \cos^2 \omega \left(t - \frac{x}{v} \right) \qquad (331)$$

10–3. Antennas. The Hertz doublet is the fundamental element of all ordinary antennas. As indicated in the previous section, it consists simply of a short straight length of wire. If such a doublet of length dl lies along the y axis at the origin of a rectangular coordinate system and carries an alternating current

$$i = I_m \sin \omega t$$

the intensities of the fields in the surrounding space are given by the following expressions which are derived from Maxwell's equations.

$$d\mathcal{E}_\theta = \frac{30\, dlI_m}{2\,\pi r^3}\lambda \sin\theta \left[\cos\omega\left(t - \frac{r}{v}\right) - \frac{2\,\pi r}{\lambda}\sin\omega\left(t - \frac{r}{v}\right) - \left(\frac{2\,\pi r}{\lambda}\right)^2 \cos\omega\left(t - \frac{r}{v}\right)\right] \text{volts/meter} \tag{332}$$

and

$$dH_\phi = -\frac{dlI_m}{4\,\pi r^2}\sin\theta\left[\sin\omega\left(t - \frac{r}{v}\right) + \frac{2\,\pi r}{\lambda}\cos\omega\left(t - \frac{r}{v}\right)\right] \text{amp/meter} \tag{333}$$

Since a doublet has a differential length, the fields arising from it are expressed as differentials in these equations. The factor r is the distance from the origin to the point in question, θ is the angle measured from the y axis to the line r, ϕ is the azimuth angle measured in the xz plane from the x axis. The subscripts on $\mathcal{E}$ and H indicate the direction of these fields. The electrostatic lines lie along the meridians and the magnetic fields lie along the latitudes. At every point in space, the two fields are mutually perpendicular. The ordinary induction fields, which are most pronounced where $r \ll \frac{\lambda}{2\,\pi}$, as well as the radiation fields which predominate when $r \gg \frac{\lambda}{2\,\pi}$, are included in these equations. At relatively great distances, the induction fields become negligibly small so that the only field which need be considered is the radiation field.

$$d\mathcal{E}_\theta = -\frac{60\,\pi dlI_m}{r\lambda}\sin\theta\left[\cos\omega\left(t - \frac{r}{v}\right)\right] \text{volts/meter} \tag{334}$$

$$dH_\phi = -\frac{dlI_m}{2\,r\lambda}\sin\theta\left[\cos\omega\left(t - \frac{r}{v}\right)\right] \text{amp/meter} \tag{335}$$

The factor $\cos \omega \left(t - \frac{r}{v}\right)$ in each expression indicates the presence of a traveling wave. The coefficients of these factors give the magnitude and direction of the electric and magnetic fields as a function of the direction of propagation.

Field intensities drop off inversely as the distance from the doublet. They have a maximum value along a line drawn perpendicular to the doublet and are zero along its axis. In between, the variation is a sine function. This result was predicted from a consideration of the receiving doublet in the discussion of the plane wave. At a given angle θ and distance r, the field intensity is the same for all azimuth angles ϕ. Consequently, a three-dimensional plot of the radiation pattern for a doublet is in the form of a torus with an infinitely small center hole. In the xy plane alone, the radiation pattern is in the form of two circles which are tangent to each other at the origin. The magnetic and electric fields of this wave are shown diagrammatically in Fig. 116. Again the electrostatic flux lines are drawn in red. They lie along meridians. The magnetic lines are drawn in black and are in the form of circles around the y axis.

A half-wavelength antenna (i.e., a wire having a physical length of 96–98 per cent or even 90 per cent of $\frac{\lambda}{2}$ due to the end effects) can be considered as being made up of a number of Hertz doublets of length dl (Fig. 117). The current distribution along the antenna is approximately sinusoidal. If $I_{\max}$ is the peak current at the center of the antenna, the current distribution is

$$I_m = I_{\max} \cos \frac{2\,\pi l}{\lambda}$$

The elemental electric-field intensity of each increment length of the antenna is found from equation 334.

$$d\mathcal{E}_\theta = -\frac{60\,\pi dl\, I_{\max}}{\lambda r} \cos \frac{2\,\pi l}{\lambda} \sin \theta \left[\cos \omega \left(t - \frac{r - l\cos\theta}{v}\right)\right] \qquad (336)$$

The distance to a remote point is essentially the same for all parts of the antenna, and lines between parts of the antenna and the point are parallel. There is, however, a phase difference in the arrival time of the radiation from the different parts. Radiation from an element dl at l meters above the center of the antenna arrives at the distant point (Fig. 117) earlier than that from the

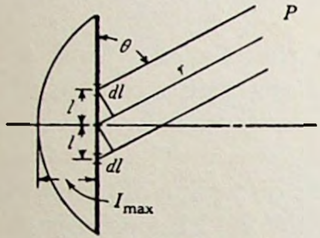


FIG. 117. Diagram of the current distribution along a half-wavelength antenna together with the notation used for deriving the equation of the radiation pattern.

center. Similarly, the radiation from an element dl at l meters below the center arrives at the point an equal amount of time later. Considering both of these effects together

$$d\mathcal{E}_\theta = \frac{-60\,\pi I_{max}}{\lambda r} \cos\frac{2\,\pi l}{\lambda} \sin\theta \left[\cos\omega\left(\left\{t - \frac{r}{v}\right\} + \frac{l\cos\theta}{v}\right) + \cos\omega\left(\left\{t - \frac{r}{v}\right\} - \frac{l\cos\theta}{v}\right)\right] dl \quad (337)$$

The application of the trigonometric formula for the cosine of the sum and difference of two angles reduces the above equation to a form which can be integrated.

$$\mathcal{E}_\theta = \int_0^{\frac{\lambda}{4}} d\mathcal{E}_\theta$$

$$= \frac{-120\,\pi I_{max}}{\lambda r} \sin\theta \cos\omega\left(t - \frac{r}{v}\right) \int_0^{\frac{\lambda}{4}} \cos\frac{2\,\pi l}{\lambda} \cos\left(\frac{2\,\pi l\cos\theta}{\lambda}\right) dl \quad (338)$$

Since $$\int_0^{\frac{\lambda}{4}} \cos\frac{2\,\pi l}{\lambda} \cos\left(\frac{2\,\pi l\cos\theta}{\lambda}\right) dl = \frac{\lambda\cos\left(\frac{\pi}{2}\cos\theta\right)}{2\,\pi\sin^2\theta}$$

the above equation reduces to the following form:

$$\mathcal{E}_\theta = \frac{-60\,I_{max}\cos\left(\frac{\pi}{2}\cos\theta\right)}{r\sin\theta} \cos\omega\left(t - \frac{r}{v}\right) \text{ volts/meter} \quad (339)$$

From equations 334 and 335, the relationship between the electric-field intensity and magnetic-field intensity in free space is

$$\mathcal{E}(\text{volts/meter}) = 120\,\pi H(\text{amp/meter})$$

therefore

$$H_\phi = -\frac{I_{\text{max}} \cos\left(\frac{\pi}{2}\cos\theta\right)}{2\,\pi r \sin\theta} \cos\omega\left(t - \frac{r}{v}\right) \text{ amp/meter} \qquad (340)$$

The radiation pattern of a half-wave antenna is given by a polar coordinate plot of the coefficient of $\cos\omega\left(t - \frac{r}{v}\right)$ in either equation 339 or 340. In this case, the diagram is again in the form of a figure eight revolved about the axis of the antenna, but it is somewhat more elongated than the circular figure-eight pattern of the Hertz doublet.

The relative effectiveness of an antenna in concentrating its radiation in a given direction is indicated by the value of its "absolute directivity." This factor is defined as the ratio of the power radiated per unit solid angle in a desired direction to the average power radiated in all directions. A hypothetical antenna which is capable of radiating its energy equally in all directions would have an absolute directivity of one. The directivity of the Hertz doublet is $\frac{3}{2}$ and of the half-wave dipole, it is 1.64. The absolute directivity of an antenna system is sometimes referred to as the gain of the system. Usually, however, the gain of an antenna system is defined as the ratio of the energy radiated per unit solid angle in the desired direction from the system to the energy which would be radiated in the maximum direction from a single half-wave dipole antenna which is supplied with the same power input. This value is more easily checked directly with field measurements.

The radiation patterns of other length antennas can be determined by the same procedure as was used for the half-wave antenna. When an antenna is grounded or placed just above the ground, the earth acts like a big reflecting surface to it; hence, its radiation appears to come from an underground image as well as the antenna itself. Consequently, a quarter-wave vertical antenna mounted just above the ground has the same radiation pattern as a free-space half-wave antenna. The current distributions along antennas of various lengths as well as their radiation patterns are given in Fig. 118.

The power radiated by a half-wave antenna can be found by integrating the power per unit area passing through a spherical surface whose center is coincident with the center of the antenna.

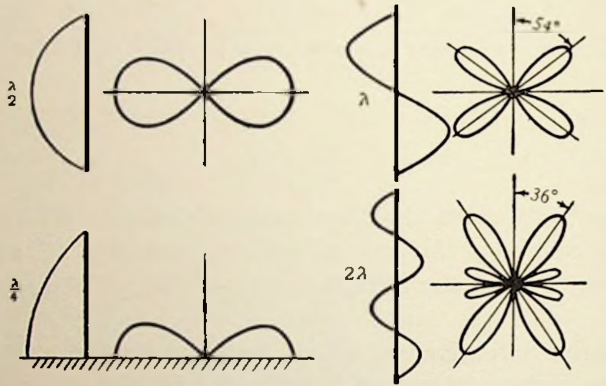


Fig. 118. Current distribution along antennas of various lengths (left) and the corresponding radiation patterns (right).

The energy passing through a differential area on this surface is essentially in the form of a plane wave. Consequently, from equation 330, the power per unit area is

$$S = \mathcal{E}H \text{ watts/meter}^2$$

The field intensity is the same at all points on a zonal band having $r\,d\theta$ width and passing around the spherical surface at a constant angle θ, and this band which has an area $r\,d\theta\ 2\,\pi r \sin\theta$ can be used as the element of area. From equations 339 and 340 the power passing through this elemental area is

$$dS = \frac{60\, I_{\max}^2 \cos^2\left(\frac{\pi}{2}\cos\theta\right)}{2\,\pi r^2 \sin^2\theta} \cos^2 \omega\left(t - \frac{r}{v}\right) 2\,\pi r^2 \sin\theta\, d\theta \qquad (341)$$

The power passing through the entire sphere is two times the integral of equation 341 between the limits $\frac{\pi}{2}$ and 0.

$$P = 120\, I_{\max}^2 \cos^2 \omega\left(t - \frac{r}{v}\right) \int_0^{\frac{\pi}{2}} \frac{\cos^2\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta}\, d\theta \qquad (342)$$

The definite integral has a value of 0.6095, and the average value of the cosine squared term over a period of several cycles is one-half; therefore, equation 342 becomes

$$P = 36.57\ I_{\max}^2 \text{ watts} \tag{343}$$

The radiation resistance of an antenna is defined as the total power being radiated divided by the rms current being supplied. For a half-wave dipole, the radiation resistance is

$$R = \frac{P}{I^2} = \frac{2\,P}{I_{\max}^2} = 73.14 \text{ ohms}$$

With a given current, the amount of energy radiated by a quarter-wave vertical antenna is just one-half of that radiated by its corresponding free-space half-wave antenna. Its radiation resistance is just half as much or 36.57 ohms.

As the length of an antenna which is being operated at a certain frequency is varied, its radiation resistance varies too. The same thing also happens when the length of the antenna is kept the same and the applied frequency is varied over a wide range. This change in radiation resistance, together with the simultaneous change of antenna reactance, results in mismatch of impedance with the transmission line and causes a considerable loss in radiation effectiveness at the deviation frequencies. In television, frequency modulation, and in other high-frequency applications, a very wide band of frequencies must be transmitted; therefore, antennas for these services must be designed to have a constant impedance over a very wide range. Some circuits whose input impedance is independent of the applied frequency are shown in Figs. 119*a* and 119*b*.

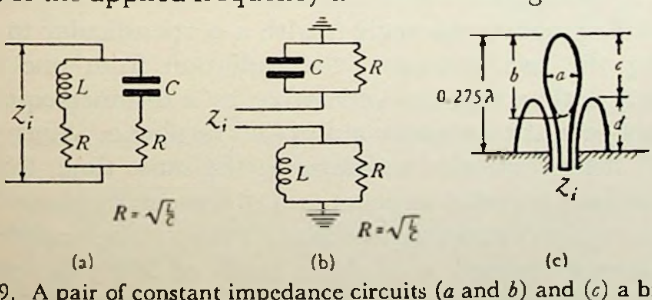


FIG. 119. A pair of constant impedance circuits (*a* and *b*) and (*c*) a broad-band constant-impedance antenna derived from them. For proper operation, the antenna should be dimensioned as follows: $\frac{b}{a} = \frac{15}{6}$, $\frac{c}{d} = \frac{7}{5}$, and $\frac{c}{a} = \frac{3}{2}$.

If an antenna is built of coaxial transmission-line sections, it can be made to simulate one of these constant-impedance circuits (Fig. 119*c*). The upper portion *c*, in the figure, has a capacitive reactance component, and the lower portion *d* has an inductive reactance component. The length of each can be adjusted so that they carry equal portions of the radiation load. The square root of their L over C ratio can be made equal to their radiation-resistance component by shaping the electrodes properly. The best shape is an ellipsoidal top element surrounded by an ellipsoidal collar as in Fig. 119*c*. The dimension ratios shown give the best operation. This antenna has an input impedance of 110 ohms which remains constant within 5 per cent over a 20 per cent range of frequency deviation.

10–4. Antenna Arrays. As previously mentioned, the radiation all around a single-wire antenna at a given angle with the axis is absolutely uniform. In order to obtain directional characteristics in a plane perpendicular to the length of the antenna, more than one must be used. An array of antennas can be so arranged that the radiation in some particular direction from each antenna is in phase with the radiation from all the others. A pronounced reinforcement of the energy traveling in that direction occurs. At the same time, the energies from the antennas in another direction may be so out of phase that they completely cancel.

A broadside array is a group of antennas so excited that a maximum field intensity is radiated perpendicular to the line of array. The simplest array of this sort consists of two vertical half-wave antennas spaced one-half wavelength apart and excited in phase (Fig. 120). When the waves travel to a distant point which lies along a radius making an angle α with a perpendicular to the line connecting the two antennas, the radiation from one antenna travels further than that from the other by a distance equal to the spacing between the antennas multiplied by sine α. Since the two waves left their respective antennas at the same time, this difference in distance traveled appears as a difference in phase between the waves or the vectors representing them. One wavelength of time difference appears as a phase angle of 360°; therefore, the phase angle between the field intensities at a distant point is $\beta = 360° \frac{d}{\lambda} \sin \alpha$ where d is the distance between the antennas and λ is the wavelength produced. For the half wavelength spacing mentioned above, $\beta = 180° \sin \alpha$. Figure 120 shows the vector addition of the fields for this case together with the resultant radiation

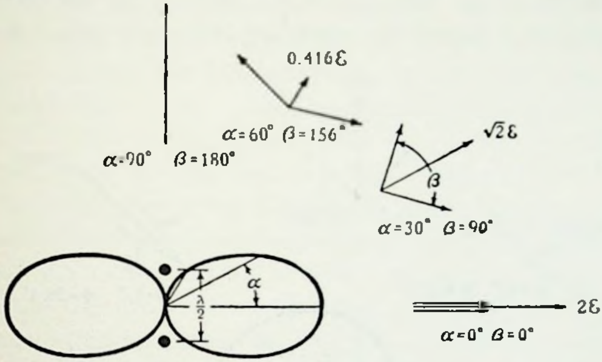


FIG. 120. An illustration of the method for finding the radiation pattern of a simple broadside array.

pattern for the array. The radiation pattern for other spacings is given in Fig. 121. A maximum gain is obtained when the spacing is 0.7 λ. The radiation pattern then consists of two major lobes perpendicular to the line connecting the antennas and two minor lobes along this line.

An end-fire array consists of a line of antennas so arranged and excited that a maximum field intensity is radiated along the line

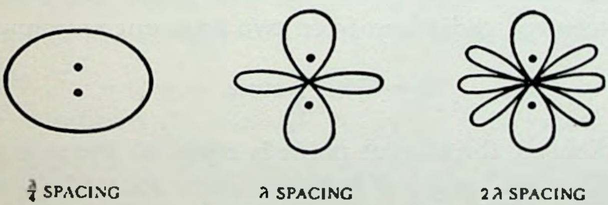


FIG. 121. Radiation patterns for two antennas which are excited in phase and are spaced various distances apart.

of the array. If we have two antennas spaced a quarter wavelength apart and excited 90° out of phase, the phase angle due to the spacing is $\beta = 90° \sin \alpha$; however, the radiation leaves the two antennas with a phase difference due to the mode of excitation so the total angle between the vectors representing the fields is $\Phi =$

$90° + 90° \sin \alpha$. The vector addition for this case, together with the resultant radiation pattern, is shown in Fig. 122.

A higher gain for an antenna system is obtained by using more antennas in the array. The radiation pattern for one of these more complex systems is found by applying the same principles as were

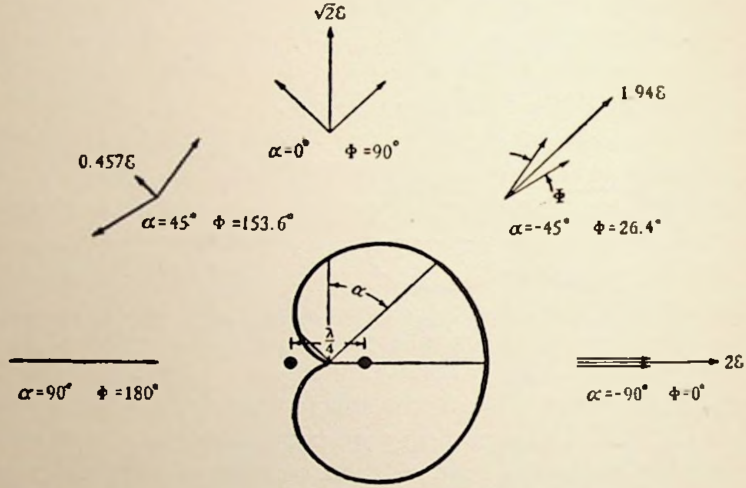


Fig. 122. Construction of the radiation pattern for a simple end-fire array.

used above. Suppose that an array consists of n elements, a distance d apart, and excited so that there is a phase angle γ between the currents in adjacent elements. The phase angle at a distant point between the radiations from two adjacent antennas is

$$\Phi = 360° \frac{d}{\lambda} \sin \alpha + \gamma \tag{344}$$

The total field at the distant point is equal to the sum of n equal vectors having an angle Φ between each succeeding vector. A similar addition of vectors is encountered in the calculation of the breadth factor for a-c machines. The formula for this addition applied to the addition of a field intensity $\mathcal{E}$ radiated from each of n antennas is

$$\mathcal{E}_T = \mathcal{E} \frac{\sin n \dfrac{\Phi}{2}}{\sin \dfrac{\Phi}{2}} \tag{345}$$

where $\mathcal{E}_T$ is the total field intensity at the distant point. The radiation pattern for any array of antennas can be determined by solving equations 344 and 345 for a series of different directional angles α. A general picture of the radiation diagram, however, can be obtained in a much simpler manner. At directional angles which make Φ equal to zero, the field intensity at a distant point will be a maximum and equal to $n\mathcal{E}$. Nulls on the radiation pattern occur at points where Φ has a value such that the n vectors add to zero. For example, four vectors add to zero when $\Phi = 90°$ or $180°$. This can be readily seen by drawing a vector diagram with four vectors passing through a common origin and each spaced from its predecessor by an angle of 90° or 180°. Lobes on the diagram will occur between the null points, and so a rough sketch of the radiation pattern — good enough for the solution of some problems — can be made. The maxima on the lobes, however, occur at the points where equation 345 passes through a maximum. An equation for finding these points is obtained by differentiating equation 345 with respect to $\frac{\Phi}{2}$ and equating this derivative to zero. The result is

$$\tan\frac{\Phi}{2} = \frac{\tan\frac{n\Phi}{2}}{n} \qquad (346)$$

An explicit expression for Φ cannot be obtained from equation 346, but the equation can nevertheless be solved graphically. The solutions to the equation occur at the points where the curve $y_1 = \dfrac{\tan\frac{n\Phi}{2}}{n}$ intersects the curve $y_2 = \tan\frac{\Phi}{2}$. A convenient way to find these intersections is to plot the latter curve between $\frac{\Phi}{2} = 0$ and 90° on a sheet of paper. Plot this same curve on a sheet of tracing paper to a scale which is smaller than the latter by the factor $\frac{1}{n}$. This curve is a plot of the first section of the equation for y_1. Place the curve on the tracing paper over the curve on the other paper and by shifting the tracing paper to the points along the $\frac{\Phi}{2}$ axis which it would occupy if it were plotted out completely, pick off the value of $\frac{\Phi}{2}$ at the intersections. The magnitude of the field intensity at these maxima points is obtained by substituting the values of $\frac{\Phi}{2}$ from the above back into equation 345. From this data a complete sketch of the radiation pattern can be made.

Example: Find the radiation pattern for an array of four half-wave antennas which are excited in phase and which are spaced one-half wavelength apart.

From the data of the problem

$$\beta = 180^\circ \sin \alpha \quad \text{and} \quad \gamma = 0$$

therefore,

$$\Phi = 180^\circ \sin \alpha$$

A maximum field intensity equal to $4\,\mathcal{E}$ occurs at the point where $\Phi = 0$ and $\alpha = 0^\circ$ or 180°. Nulls occur at the points where $\Phi = 90^\circ$ or 180°; hence,

$$180^\circ \sin \alpha = 90^\circ$$
$$\alpha = \pm\, 30^\circ \quad \text{or} \quad \pm\, 150^\circ$$

and

$$180^\circ \sin \alpha = 180^\circ$$
$$\alpha = \pm\, 90^\circ$$

Using the tracing-paper procedure outlined above, the first intersection of the curves is at $\Phi = 0$ which gives the two lobes of intensity $4\,\mathcal{E}$ which have already been computed. Another intersection is found at $\frac{\Phi}{2} = 65.9^\circ$. The angle at which this maximum occurs is given by

$$2 \times 65.9^\circ = 180^\circ \sin \alpha \quad \text{or} \quad \alpha = \pm\, 47.1^\circ \quad \text{or} \quad \pm\, 132.9^\circ$$

The magnitude of the lobe is found by substituting the value for $\frac{\Phi}{2}$ into equation 345. The result is

$$\mathcal{E}_T = \mathcal{E}\,\frac{\sin 4 \times 65.9^\circ}{\sin 65.9^\circ} = 1.09\,\mathcal{E}$$

A sketch of the radiation pattern can be plotted from this data. It consists of two major lobes having a magnitude $4\,\mathcal{E}$ and four minor lobes each having a magnitude $1.09\,\mathcal{E}$.

The phase angle between the currents in the adjacent elements of an array is controlled by the length of transmission line between them which is used for feeding. For example, to feed an array of antennas in phase when they are spaced a half wavelength apart, the transmission line between the elements could be one-half wavelength long, giving 180° phase difference. In between the two elements, the line is transposed to give the other 180° phase difference necessary to bring the two currents back in phase with each other. Other phase relationships can be obtained by using different lengths of transmission line.

Half-wave antenna elements can be either current or voltage fed. A half-wave dipole is an example of a current-fed element. In this antenna the transmission line is connected at the center of the half-wave element which has been divided into two poles (dipole). One wire from the transmission line is connected to each of the poles. A voltage-fed half-wave element is a single wire one-half wavelength long. One end of the element is connected to one of the wires in the transmission line feeding it. The other wire of the transmission line is left unconnected. The terms *voltage* and *current feed* come from the nature of the current and voltage distribution along the half-wave element. Voltage maxima occur at the ends of the antenna and a voltage minimum is at the middle. The current distribution, on the other hand, is as illustrated in Fig. 117.

In the case of a two-element end-fire array, it is not necessary to run a transmission line to both of the elements. If one element is fed, the other element has a current induced in it which is in phase with the passing wave or 90° behind the current in the excited antenna. Due to the current flowing in the unexcited antenna, energy is reradiated from it so that the net field pattern from the two antennas is like that shown in Fig. 122. Due to the effect of the unexcited element on the radiation pattern, it is called a reflector.

The phase relationship between the current in an antenna and the incident wave depends on the nature of the impedance of the element. An antenna which is slightly longer than one-half wavelength is inductive, and one shorter than a half wavelength is capacitive in nature, but an antenna which is a half wavelength long is a pure resistance. A reflector should be a half wavelength long.

A multielement broadside array is often used in conjunction with a reflector array a quarter wavelength behind it so that radiated energy is concentrated in a single direction. The resultant pattern is equal to the pattern of the broadside array multiplied by the cardioid pattern of a single antenna and its reflector (Fig. 122). Since a fed reflector at a quarter wavelength spacing is difficult to adjust due to a high mutual effect between the antennas, spacings of $\frac{3\lambda}{4}$ and $\frac{5\lambda}{4}$ are often used instead.

Sometimes a director is added to a two-element end-fire array. It is an unexcited antenna a little less than a half wavelength long which is placed in front of the excited antenna opposite the reflector. Since this short antenna has a capacitive reactance, the current induced in it leads the incident wave. If the spacing between the excited antenna and the director is adjusted to the proper value, the leading current flowing in the director will reradiate energy in the proper phase relationship to produce a pronounced reinforcement of the energy traveling in the direction of the array. As a result of this process, a more directional pattern is obtained than that shown in Fig. 122.

A parabolic reflector is especially useful in concentrating ultra-high-frequency radiation into a narrow beam. Physically, it can consist of a multitude of parasitic reflecting antennas arranged in the form of a parabola or it can be a solid sheet of metal. In the former case, the antenna used is usually a half wavelength long and the parasitic antennas are set parallel to it along the line of the parabola. The excited antenna is placed at the focus of the parabola. This arrangement only attempts to concentrate the radiation coming off the antenna perpendicular to its axis. A solid sheet of metal in the form of a parabola can be used rather than the individual reflecting elements. If the solid sheet is made in the form of a paraboloid of revolution instead of a parabolic sheet, the radiation coming off the antenna at angles near to its axis will be more effectively concentrated in a narrow beam. In this case, the action of the parabola is a great deal like that of a parabolic searchlight reflector. To limit diffraction effects, the diameter of the parabola should be several times greater than the wavelength radiated.

The antenna used in conjunction with the parabolic reflector is generally a half-wave dipole fed with a coaxial line. The dipole is balanced with respect to ground by isolating it, insofar as the high-frequency currents are concerned, from the grounded outer conductor of the coaxial line. This is done by placing a quarter-wavelength isolating section of line between the end of the outer conductor where it feeds the antenna and the point where it is grounded. The outer conductor of the feed line is used for the inner conductor of the short-circuited section. An oversize tube one-quarter wavelength long makes up the outer conductor of this section. The line formed by these two tubes is left open on the side next to the antenna and shorted on the other end. The impedance of this section as seen from the antenna end is infinite so that the latter is effectively isolated from ground, and currents are prevented from flowing back along the outside of the line.

Another means of concentrating ultra-high-frequency energy into a beam is by the use of a wave-guide horn. This device matches the impedance of a wave guide to the surrounding space in a somewhat analogous manner to the acoustical matching which is made by an exponential horn speaker. More will be said about horns in the discussion of wave guides.

10–5. Propagation. Since electromagnetic waves in free space travel along straight lines, their apparent ability to follow the curvature of the earth in long-distance radio communication needs an explanation. Kennelly and Heaviside independently suggested that an ionized layer of atmosphere high above the earth's surface acted like a reflector to radio waves and permitted them to travel around the curvature of the world by a series of reflections between this ionosphere and the ground.

Ionospheric reflections are very similar to the underwater reflections of light which occur between the surface of a pond and the air above it. Since the velocity of light in water is less than the velocity of light in air, a beam of light traveling from the water into the air is refracted away from the normal (Fig. 123). In other words, the angle between the normal and the refracted beam θ_2 is greater than that between the normal and the incident beam θ_1.

From Snell's law $\frac{\sin \theta_2}{\sin \theta_1} = \frac{v_1}{v_2}$ (where v_1 and v_2 are the velocities of light in the two media), we see that θ_2 becomes 90° when

$$\sin \theta_1 = \frac{v_1}{v_2} \tag{347}$$

This value of θ_1 is known as the critical angle. All beams at greater angles of incidence are totally reflected from the water's surface. It is this phenomenon which explains why a fish's view of the world is a peculiar combination of some things above the water and some things below.

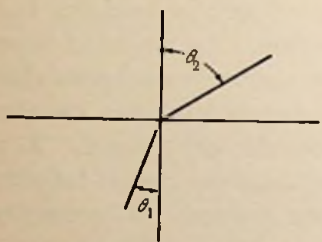


FIG. 123. Refraction at the boundary between two media having different properties.

A radio wave follows the same laws in its reflection at the ionosphere as does the light wave in the above analogy. The phase velocity of a radio wave is higher in an ionized region than it is in free space and it is the velocity referred to in Snell's law. The phase velocity is the velocity at which the crests of the waves move and should be distinguished from group velocity which is the rate at which the energy of the wave travels along. Group velocity is the rate at which the modulations or individual groups of waves move through space. The phase velocity of an electromagnetic wave is

$$v = \frac{1}{\sqrt{\mu K}}$$

Since the effective dielectric constant in an ionized region is less than it is in free space, the phase velocity of a radio wave in such a region is greater than it is in free space.

The ions in space are accelerated by a force equal to their charge q (coulombs) multiplied by the intensity of the electric field $\mathcal{E}$ (volts/meter) which is present. From Newton's law of motion, this force is equal to the mass m (kilograms) of the ions multiplied by their acceleration (meters/sec/sec) or

$$q\mathcal{E} = m\frac{dv}{dt} \tag{348}$$

If the field intensity varies sinusoidally, the field intensity at any instant is given by

$$\mathcal{E} = \mathcal{E}_0 \sin \omega t \tag{349}$$

The velocity of the ions can be found by integration and is equal to

$$v = -\frac{q}{m\omega}\mathcal{E}_0 \cos \omega t \tag{350}$$

The motion of N ions per cubic meter adds a current

$$i_1 = Nqv = -\frac{Nq^2}{m\omega}\mathcal{E}_0 \cos \omega t \tag{351}$$

to the ordinary displacement current

$$i_2 = K_0 \frac{\partial \mathcal{E}}{\partial t} = K_0 \omega \mathcal{E}_0 \cos \omega t \tag{352}$$

The total current in space is equal to the sum of i_1 and i_2.

$$i_1 + i_2 = \left(K_0 - \frac{Nq^2}{m\omega^2}\right)\omega\mathcal{E}_0 \cos \omega t = \left(1 - \frac{Nq^2}{K_0 m\omega^2}\right)K_0 \frac{\partial \mathcal{E}}{\partial t} \tag{353}$$

Consequently, the effective dielectric constant is equal to the dielectric constant of free space multiplied by the factor

$$1 - \frac{Nq^2}{K_0 m\omega^2}$$

In other words, the effective dielectric constant decreases as the intensity of ionization increases. The ratio of the phase velocity in the ionosphere to that in vacuum becomes

$$\frac{v_1}{v_2} = \sqrt{1 - \frac{Nq^2}{K_0 m\omega^2}} = \sqrt{1 - \frac{f_0^2}{f^2}} \tag{354}$$

where $f_0^2 = \dfrac{Nq^2}{K_0 m(2\pi)^2}$ a constant for a particular ion density. The critical angle for a wave of any frequency striking the ionized layer is given by the combination of equations 347 and 354.

$$\sin \theta_1 = \sqrt{1 - \frac{f_0^2}{f^2}} \tag{355}$$

If the frequency of the electromagnetic wave f is below the critical frequency f_0, it is reflected back from the ionosphere regardless of its angle of incidence.

The height of the ionosphere is determined by measuring the time required for a wave to make the trip up to it and back to earth. A short pulse of carrier is transmitted vertically upward and its leaving is marked on an oscillograph. When the echo returns, it is also indicated on the oscillograph. The distance between the two marks is a measure of time required for the trip which in turn is a measure of the distance.

A typical record of ionosphere height vs. frequency of the carrier is shown in Fig. 124*a*. The abrupt rise on the record indicates that

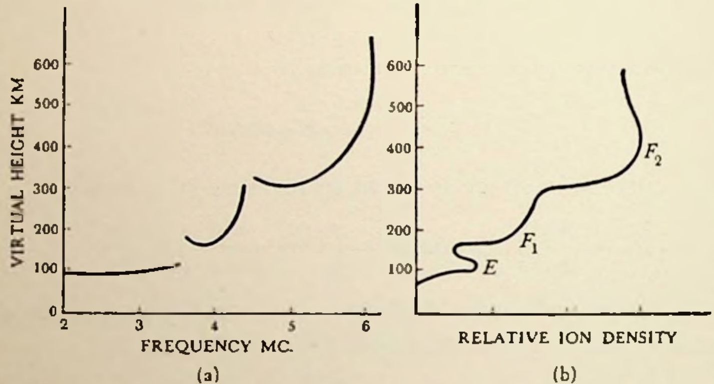


FIG. 124. (*a*) Oscillographic record of the virtual height of the ionosphere as a function of the frequency of the radio waves being directed toward it and (*b*) the corresponding ion density as a function of height.

the critical frequency has been reached and a penetration of the layer has taken place. Frequencies near the critical get lost in the region between the ionized layers so that there is a discontinuity in the curve. From an experimental curve like 124*a*, the relative ion density above the earth can be calculated. It has a distribution somewhat like that shown in Fig. 124*b*. There are two distinct regions. The lowest is the E region and the highest is the F region. The F region is usually composed of two layers, F_1 and F_2, but the E region is usually just one layer. A weakly ionized region also occurs at about a 50-km height during the daytime. It is called the D region. The existence of several layers in the ionosphere is probably due to the change in the composition of the atmosphere with height.

As shown in Fig. 125*a*, very low-frequency waves are essentially guided between the ionosphere and the surface of the earth as they move out from the transmitter. These wavelengths are very reliable but require large amounts of power to span great distances. At higher frequencies, the radiation which leaves the antenna at an angle smaller than critical escapes into space. As a result, there is a region on the earth's surface which is skipped by the signal and a dead spot results (Fig. 125*b*). The action of these

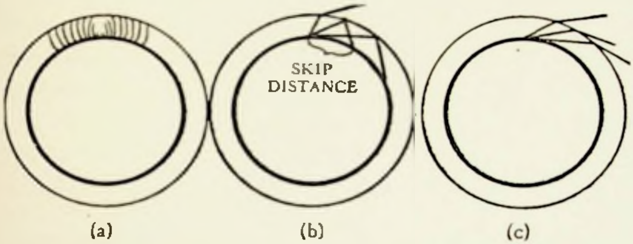


Fig. 125. Diagram illustrating the propagation of (*a*) low, (*b*) high, and (*c*) ultra-high-frequency waves along the surface of the earth.

short waves is very similar to that of light in the pond analogy given above. Good communication over long distances can be obtained with these waves, but the quality of the communication is very dependent on the condition of the ionosphere. On the other hand, the critical angle for the ultra-high frequencies is so large that no radiation from the surface of the earth can exceed it, and no reflection from the ionosphere takes place. This phenomenon is illustrated in Fig. 125*c*. These waves act most nearly like light, and communication is essentially limited to line of sight distances. The longer microwaves travel somewhat greater distances, however, due to diffraction effects. The transmission of microwaves is not hindered by the presence of fog or mist or rain or snow.

Although they are not reflected by the ionosphere, ultra-high-frequency waves can be reflected from an ionized region if the ionization is great enough. They have been modulated by the use of a gas discharge tube in the path of their propagation.

Problems

10-1. (*a*) Show that the absolute directivity of a half-wave antenna is 1.64.

(*b*) Show that the absolute directivity of a Hertz doublet is $\frac{3}{2}$.

10–2. Prove that the input impedance to the circuit shown in Fig. 119*b* is independent of frequency.

10–3. Find the radiation pattern for a broadside array consisting of 8 dipole antennas excited in phase and spaced one-half wavelength apart.

10–4. An antenna system consists of three elements. There is a reflector one-quarter wavelength behind an excited half-wave antenna. One-eighth wavelength in front of the excited antenna is a director. Assume that the reradiated field intensity from the director is one-half of the field intensity from each of the other two elements. Also assume that the current in the reflector is in phase with the incident wave whereas the current in the director is 90° ahead of the incident wave. Plot the radiation pattern for the system.

Experiment 22

Object:

The object of this experiment is to study antenna patterns.

Preliminary:

1. Plot the theoretical radiation patterns for parts 1, 2, and 4 of Performance.
2. Solve problem 10–4.

Performance:

1. Feed a horizontal half-wave antenna with a 1-meter oscillator. Pick up the signal on a horizontal half-wave doublet which is connected to a thermal galvanometer. Rotate the receiving antenna in the horizontal plane around a line perpendicular to its axis and note the readings on the meter. The person making the reading should be as far from the antenna as possible and should occupy the same position for each reading.
2. Place a half-wave reflecting element a quarter wavelength from the vertical antenna of the oscillator. With the receiving antenna in a fixed position, read the galvanometer for various positions of the reflector around the transmitting antenna.
3. Add a director to the array of part 2 one-eighth wavelength in front of the excited antenna. Rotate this combination about the axis of the excited antenna and obtain data for an experimental radiation pattern. Compare the experimental pattern with the results of part 2 in the preliminary and explain the discrepancies.
4. Set up a broadside array of two vertical half-wave antennas fed in phase and spaced a half wavelength apart. Rotate this structure about a vertical axis midway between the two antennas and record the pickup on a fixed receiving antenna.

5. Plot the radiation patterns for parts 1, 2, and 4 in a contrasting color on the same coordinates as the preliminary patterns. Multiply the experimental data by a factor which will make the maxima of the two diagrams lie on the same point. Discuss the results.

References

Bradbury, N. E., "Fundamental Mechanisms in the Ionosphere," *Jour. Applied Phys.*, November 1937, pp. 709–717.

Brainerd, J. G., *et al.*, *Ultra-high-frequency Techniques*. New York: D. Van Nostrand Company, Inc., 1942, pp. 369–454.

Breit, G., and M. A. Tuve, "A Test of the Existence of the Conducting Layer," *Phys. Review*, September 1926, pp. 554–575.

Glasgow, R. S., *Principles of Radio Engineering*. New York: McGraw-Hill Book Company, Inc., 1936, pp. 413–502.

Grammer, G., and B. Goodman, *The ARRL Antenna Book*. West Hartford, Conn.: American Radio Relay League, Inc., 1939.

Green, E., "Extended Aerial Systems (Calculating the Polar Diagrams)," *Wireless Engineer*, May 1942, pp. 195–199.

Hara, G., "The Circuit Theory of Antennas," *Report of Radio Research in Japan*, June 1937, pp. 57–73.

Howe, G. W. O., "The Polar Diagram of a Simple Broadside Array," *Wireless Engineer*, May 1942, pp. 193–194.

Lindenblad, N. E., "Television Transmitting Antenna for Empire State Building," *R.C.A. Review*, April 1939, pp. 387–408.

Page, L., and N. I. Adams, *Principles of Electricity*. New York: D. Van Nostrand Company, Inc., 1931, pp. 551–591.

Pierce, G. W., *Electric Oscillations and Electric Waves*. New York: McGraw-Hill Book Company, Inc., 1920.

Skilling, H. H., *Fundamentals of Electric Waves*. New York: John Wiley & Sons, Inc., 1942.

Slater, J. C., *Microwave Transmission*. New York: McGraw-Hill Book Company, Inc., 1942.

CHAPTER XI

WAVE GUIDES

11–1. General. In its most general sense, the term "wave guide" includes all structures which guide electric waves from one point to another (transmission lines, hollow tubes, etc.); however, the tendency has been to restrict its meaning to metal pipes filled with low-loss dielectric. Lord Rayleigh wrote the pioneer paper on hollow-tube transmission in 1897. The subject then rested rather quietly until recently when it was discovered that the attenuation of a traveling wave in a conducting tube could be made extremely low. Since then, many papers have been written and the development of this new science has proceeded at a rapid pace.

Suppose that a round metal pipe having an internal radius of about a quarter wavelength is placed over the antenna which produced the field pattern of Fig. 116 and is arranged so that the axis of the pipe coincides with the y coordinate axis in the figure. The generated fields would be confined inside the pipe and would have a distribution like that shown in Fig. 126*a*. The electric-field intensity lines would intersect the edge of the pipe at right angles. The magnetic lines would be circles which are concentric with the pipe. Instead of being propagated in space, the electromagnetic wave would be confined to the inside of the pipe and propagated down it. Due to the action of the pipe, it is called a wave guide. When the field distribution inside the guide is like that shown in Fig. 126*a* it is known as a "transverse magnetic" (abbreviated TM) or an "E" wave. It is called a TM wave because the magnetic-field intensity lines are all transverse to the direction of propagation. The designation as an E wave comes from the fact that there is a component of electric-field intensity in the direction of propagation. The subscripts refer to the number of cyclic variations in electric-field intensity which occur along some particular dimension of the pipe. In a round guide, the first subscript refers to the number of full sinusoidal variations in electric-field intensity found

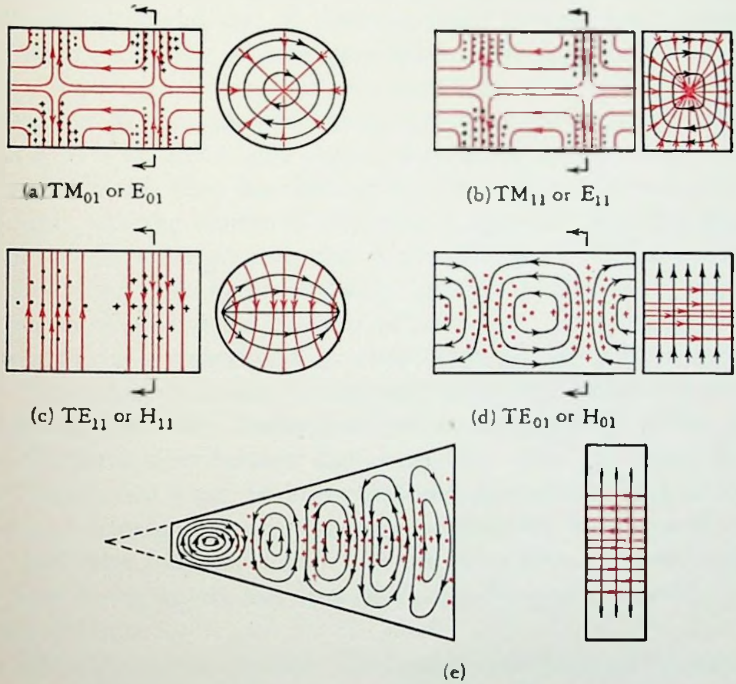


FIG. 126. An illustration of the field distribution of various modes of wave guide transmission together with (*e*) the field distribution in a sectoral horn carrying TE_{01} waves. The electrostatic lines are drawn in red and the magnetic lines are drawn in black.

around the circumference or θ component, and the second subscript refers to the number of half sinusoidal variations in electric-field intensity found across a diameter of the pipe. The wave in the round pipe of Fig. 126*a* is accordingly an E_{01} wave or a TM_{01} wave.

In a rectangular guide, the first subscript gives the number of half sinusoidal variations in electric-field intensity encountered along a path across the narrow dimension of the guide. The second subscript refers to the number of variations encountered in going from one side of the guide to the other along the wide dimension. Referring to Fig. 126*b* and going across the guide from top to bottom, it is noted that the electric-field intensity is zero at the top of the pipe. As the center of the pipe is approached, the density of the lines becomes greater and a maximum of field intensity is reached. Then the field intensity gradually decreases until it again becomes zero at the bottom of the pipe. The distribution of field intensity as a function of position is in the form of a half sine wave. A similar situation is encountered going from one side of the guide to the other; consequently, the designation of the wave giving the field distribution shown in Fig. 126*b* is either E_{11} or TM_{11}. When two axial antennas instead of one are used to excite a wave in a rectangular guide, a field distribution is obtained which in the end view looks like two end views of Fig. 126*b* set side by side. The wave which has this field distribution is called an E_{12} or TM_{12} wave.

The H or transverse electric (TE) mode is defined as the mode of wave propagation which has an axial component of magnetic-field intensity, or one in which the electric-field intensity is all transverse to the axis of the guide. The field configuration of dominant TE mode in both the round and the rectangular pipe is shown in Figs. 126*c* and 126*d*. The designations are made in the same manner as that indicated above and are TE_{11} or H_{11} and TE_{01} or H_{01} waves, respectively. These waves are excited by means of an antenna which lies across the pipe or by a loop which is placed so that it is perpendicular to the magnetic lines. The transverse electric mode of transmission is of particular importance because it has a lower attenuation in the guide than any other useable configuration. Henceforth, we shall confine our attention to it alone.

The plane of polarization of a guided wave is defined as the direction of the electric-field intensity lines. For example, the polarization of the wave in Fig. 126*c* is vertical and in 126*d* it is horizontal. Usually, the plane of polarization will remain the same once it has been established; however, ellipticity in a round pipe may cause it to shift. The polarization in a rectangular pipe can be strictly maintained if the dimension of the pipe along the line of polarization is made less than the cutoff limit (see below). The plane of polarization can be purposely twisted in a round pipe by means of a corkscrew conducting membrane across a diameter of

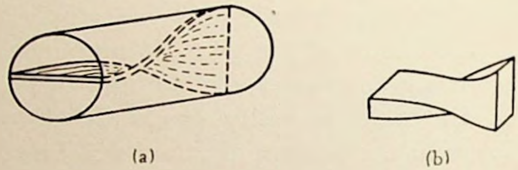


FIG. 127. Sections of wave guide which are used to change the plane of polarization.

the pipe such as that illustrated in Fig. 127*a*. The electric-field intensity will always remain perpendicular to this conducting surface. A twisted section of rectangular guide such as that shown in Fig. 127*b* accomplishes the same result in rectangular pipes.

11–2. Phase Velocity and Attenuation. Expressions for the propagation of electric waves through pipes are obtained by solving Maxwell's equations under the restrictions which are imposed by the boundary conditions set up by the conducting surfaces of the pipe. The derivation of these expressions is given in a number of references and will not be repeated here because it is rather lengthy. Instead the important end results of the solutions will be discussed.

The phase velocity of wave propagation in a guide is greater than it is in free space. It is a function of the size of the pipe as well as a function of the frequency of the wave being transmitted. For a round pipe which has a diameter of d and which is filled with material having a dielectric constant $\frac{K}{K_0}$ with respect to free space, the phase velocity of the TE_{11} wave is

$$v_o = \frac{c}{\sqrt{\frac{K}{K_0} - \left(\frac{\lambda_a}{1.708\,d}\right)^2}} \tag{356}$$

where c is the velocity of light and λ_a is the wavelength of the energy being carried by the guide when it is propagated in free space. The phase velocity of a TE_{01} wave in a rectangular pipe which is a centimeters along the edge parallel with the electric flux lines and b centimeters along the edge perpendicular to them

$$v_r = \frac{c}{\sqrt{\frac{K}{K_0} - \left(\frac{\lambda_a}{2\,b}\right)^2}} \tag{357}$$

The dimensions of the guide and the wavelength must be in the same units.

If we assume that the dielectric filling the guide is air, the attenuation to the transmission of a TE_{11} wave in a round guide is

$$\alpha_o = \frac{159\sqrt{\frac{\mu}{\mu_0}\rho}}{\sqrt{\lambda_a}} \frac{\frac{1}{d}\left[0.418 + \left(\frac{\lambda_a}{1.708\,d}\right)^2\right]}{\sqrt{1 - \left(\frac{\lambda_a}{1.078\,d}\right)^2}} \text{ db/meter} \tag{358}$$

where $\frac{\mu}{\mu_0}$ is the permeability of the material making up the pipe relative to that of free space and ρ is its resistivity in ohms per centimeter cube. The wavelength in free space λ_a is expressed in centimeters. Similarly, a rectangular guide filled with air dielectric and transmitting a TE_{01} wave introduces an attenuation of

$$\alpha_r = \frac{79.3\sqrt{\frac{\mu}{\mu_0}\rho}}{\sqrt{\lambda_a}} \frac{\left[\frac{1}{a} + \frac{2}{b}\left(\frac{\lambda_a}{2\,b}\right)^2\right]}{\sqrt{1 - \left(\frac{\lambda_a}{2\,b}\right)^2}} \text{ db/meter} \tag{359}$$

As before, a and b are the dimensions of the guide in centimeters and λ_a is the wavelength (centimeters) in free space of the energy being transmitted.

If equation 359 is plotted as a function of the impressed wavelength, it is observed that the attenuation approaches infinity as the wavelength λ_a approaches $2\,b$. Larger values of λ_a make the attenuation imaginary. This means that a hollow pipe acts like a high-pass filter for electromagnetic waves. Only those frequencies which are higher than the critical value will be transmitted through the guide and all others are attenuated with an infinite attenuation. The cutoff wavelength λ_c (referred to free space) of a rectangular wave guide transmitting by the TE dominant mode is

$$\lambda_c = 2\,b \tag{360}$$

When the guide is filled with some other dielectric besides air the cutoff wavelength becomes

$$\lambda_c = 2\,b\sqrt{\frac{K}{K_0}} \tag{361}$$

In a similar manner, the cutoff wavelength for a round pipe is

$$\lambda_c = 1.708\,d\sqrt{\frac{K}{K_0}} \tag{362}$$

Some of the other modes of propagation besides the dominant TE mode may be unintentionally set up in a wave guide if it is made too large. To avoid the occurrence of these spurious modes which have a high attenuation, the diameter of a round wave guide should satisfy the following equation:

$$d \leqq 0.763\,\lambda_a\sqrt{\frac{K_0}{K}} \tag{363}$$

The limits on an appropriate size for a round wave guide which is intended to transmit TE_{11} waves are thus established by equations 362 and 363. The diameter of the guide used in a particular application should lie somewhere between these two limits. A similar situation exists in the use of rectangular guides for the transmission of TE_{01} waves. The width of the guide should lie between the following limits

$$\frac{\lambda_a}{2}\sqrt{\frac{K_0}{K}} < b < \lambda_a\sqrt{\frac{K_0}{K}} \tag{364}$$

The dimension a of the guide is not critical, but it should be somewhat less than the dimension b so that the proper polarization will be established and maintained.

Equations 356 and 357 for the phase velocity in a guide reduce to the same form by substituting the value of the cutoff wavelength (361 and 362) in them.

$$v = \frac{c}{\sqrt{\frac{K}{K_0}\left[1 - \left(\frac{\lambda_a}{\lambda_c}\right)^2\right]}} \tag{365}$$

The wavelength in the guide λ_g divided by the phase velocity is equal to the frequency of the wave. The frequency is also equal to the wavelength in free space divided by the velocity of light; consequently,

$$\lambda_g = \lambda_a \frac{v}{c} \tag{366}$$

Assuming that the guide is filled with air dielectric, a convenient relationship between the wavelength in the guide and the corresponding wavelength in free space is given by the substitution of equation 365 into equation 366.

$$\lambda_g = \frac{\lambda_a \lambda_c}{\sqrt{\lambda_c^2 - \lambda_a^2}} \tag{367}$$

Equation 367 can also be solved explicitly for the free-space wavelength λ_a. This result is as follows:

$$\lambda_a = \frac{\lambda_g \lambda_c}{\sqrt{\lambda_g^2 + \lambda_c^2}} \tag{368}$$

11–3. Characteristic Impedance and Matching. The definition of the characteristic impedance of a wave guide is the same as the definition of the characteristic impedance of a transmission line. It is equal to the voltage across an infinitely long guide divided by the input current. Space voltage and space current must be used in making this calculation. The space voltage is the integral of the electric-field intensity taken across the guide and the space current is the integral of the displacement current density taken over the cross-sectional area of the guide. When these mathematical operations are carried out, the characteristic impedance of a circular guide carrying TE_{11} waves is found to be equal to

$$Z_0 = \frac{353}{\sqrt{\frac{K}{K_0} - \left(\frac{\lambda_a}{1.708\,d}\right)^2}} \text{ ohms} \tag{369}$$

Similarly, the characteristic impedance of a rectangular guide carrying TE_{01} waves is

$$Z_r = \frac{465\,a}{b\sqrt{\dfrac{K}{K_0} - \left(\dfrac{\lambda_a}{2\,b}\right)^2}} \text{ ohms} \tag{370}$$

The characteristic impedance for a round pipe is never less than 353 ohms, but the impedance of a rectangular pipe can be made equal to most any value by simply adjusting the dimension a.

Standing waves are set up in wave guides whenever they are terminated in an impedance different from their characteristic impedance. A piece of apparatus known as a "traveling detector" is used to measure these standing waves and to show when the guide has been properly terminated. An absence of standing waves,

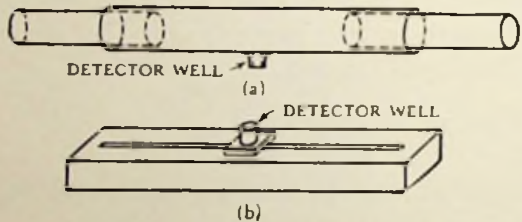


Fig. 128. Typical traveling detectors (*a*) for round wave guides and (*b*) for rectangular wave guides.

of course, indicates a proper impedance match. A traveling detector is a pickup and indicating device which can be moved along a length of guide which measures the relative field intensity as a function of distance. Two forms of the apparatus are shown in Fig. 128. One of these is intended for use on round wave guides, and the other is intended for use on rectangular wave guides. The slot which is cut in the rectangular traveling detector and which allows movement of the detector unit itself should be cut in the face of the guide on which the electrostatic-field lines terminate. If the slot were cut in one of the sides which are parallel to the polarization of the wave, it would interrupt the conduction currents which flow in these sides as a result of the wave propagation. A slot arrangement is also sometimes used in round wave guides to adapt them for traveling detector uses.

Figure 129 is an illustration of two possible detector units which can be used in conjunction with the apparatus in Fig. 128. The unit in Fig. 129*a* consists of a pickup loop, crystal detector, by-pass condenser, and microammeter. It is primarily intended for picking up the magnetic component of the field in the traveling wave; consequently, the loop should be oriented in the guide so that the maximum possible number of magnetic lines will thread through it. The unit shown in Fig. 129*b* is primarily intended for picking up the electrostatic field which is in the traveling wave. The pickup antenna must be oriented so that it lies along the lines of

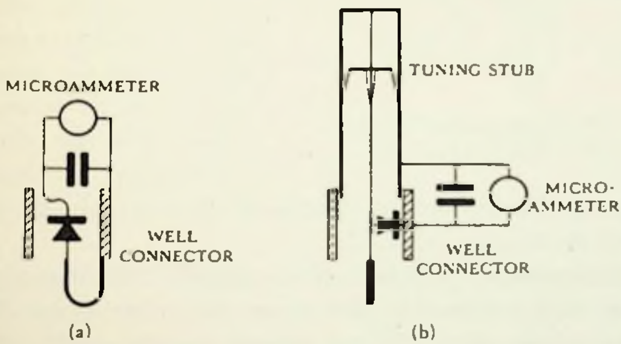


FIG. 129. Detector units for use with traveling detectors and resonant chambers — (*a*) magnetic pickup and (*b*) electrostatic pickup.

electrostatic-field intensity. For proper operation, the antenna is tuned into resonance by means of the shorted transmission line which is attached to it. The other elements in the assembly are the same as those in Fig. 129*a*. Other detectors besides crystals can also be used in the assemblies shown in the figure.

As in transmission lines, a quarter wavelength section of guide can be used to match two guides having different impedances. The characteristic impedance of the matching section should be the geometric mean of the two impedances to be matched. Two wave guides of different impedances can also be matched by a tapered section of guide between them. For example, a round guide can be matched to a rectangular guide by a length of guide which gradually changes its cross section from a circle to a rectangle.

A coaxial line can be matched to a wave guide by the arrangement shown in Fig. 130. The length of shorted coaxial line opposite the feed line is adjusted until a maximum power is obtained. A further adjustment is made by moving the piston to a point about a quarter wavelength from the center conductor. If the matching section is made with a fixed back plate, the latter adjustment can be made by trimming screws which are inserted in the sides of the wave guide and which project an adjustable amount into that space between the back plate and the coaxial cable. The shorted coaxial-line adjustment adds the reactance necessary to make the match, and the movable piston brings the waves reflected from it into the proper phase relationship to reinforce the waves leaving the center conductor of the coaxial line and being propagated down the guide.

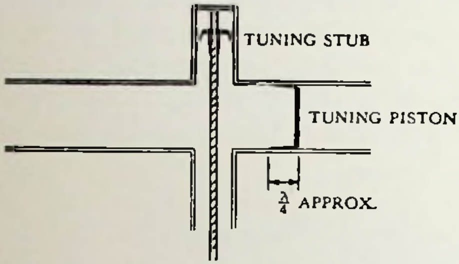


FIG. 130. Matching section between a coaxial line and a wave guide.

11–4. Resonance. A length of wave guide terminated by an iris on one end and a movable piston on the other is one form of a resonant chamber (Fig. 131) or resonant wave guide. If the distance between the iris and the piston is made some multiple of a half wavelength, traveling waves which enter the hole in the iris will produce large standing waves inside the chamber. The action is very similar to the resonance observed on a transmission line which is being used as a Lecher wire system. Rather than having the form indicated in Fig. 131, the resonator can be completely enclosed and have the energy fed to it by a small coupling loop rather than through the hole in an iris. In either event, the movable piston should be provided with a contact sleeve which extends into the chamber about a quarter wavelength from the piston's face. The sleeve is essential to assure a good electrical connection between the piston and the guide. The edge of the sleeve will be on a voltage maximum; hence, contact resistance is more easily broken down. The cavity resonators illustrated in Fig. 75 of Chap. VII are variations of this resonant chamber. According to Fig. 75, the resonant wavelength of a rectangular

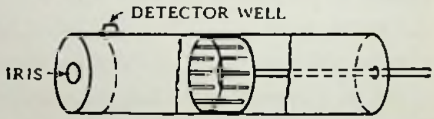


FIG. 131. Resonant Chamber.

cavity 2 a cm wide and 2 a cm long was $\lambda = 2\sqrt{2}\,a$ cm. This cavity is simply a half wave section of rectangular wave guide closed on both ends. From the cutoff equation 361 and the width of this guide,

$$\lambda_c = 2 \times 2\,a = 4\,a \tag{371}$$

The length of the wave in the guide is equal to the length of the cavity multiplied by 2 or

$$\lambda_g = 2 \times 2\,a = 4\,a \tag{372}$$

From equation 368, the wavelength in air for a wave which would resonate in such a cavity is

$$\lambda_a = \frac{\lambda_g \lambda_c}{\sqrt{\lambda_g^2 + \lambda_c^2}} = \frac{16\,a^2}{\sqrt{2 \times 16\,a^2}} = 2\sqrt{2}\,a \tag{373}$$

which is the same result as before.

11–5. Horns. Functionally, electromagnetic horns radiate ultra-high-frequency energy in very restricted directions. They are built in a variety of different forms. The simplest of all is the sectoral horn. It is of rectangular cross section and consists of two sides which flare outward and two sides which are parallel to each other. The pyramidal horn differs from the sectoral horn in that the other two sides are also flared. A round horn is a pyramidal horn which is circular rather than rectangular in section and is intended for use in conjunction with round wave guides. When energy is fed into the throat of both sectoral and pyramidal forms, the radiation from the mouth of the horn is concentrated into a single unidirectional beam. A somewhat different radiation pattern is obtained with the biconical horn. In its most elementary construction, this unit consists of two concentric cones which have a common apex, but which have different apex angles. Energy is fed into the horn at the apex and is radiated from the annular space between the cones. The radiation in a plane perpendicular to the axis of the cones is uniform in all directions, but the radiation in any plane which contains the axis is very directional.

The field distribution in a simple sectoral horn carrying TE_{01} waves is illustrated in Fig. 126*e*. The wavelength of the electromagnetic energy is longer near the throat of the horn where it joins the wave-guide feed than it is near the mouth. One would expect this because the wavelength in free space is shorter than it is in a wave guide. The horn assists in making the gradual transition between the two wavelengths.

For design purposes, the flare angle ϕ_0, of a sectoral horn is defined as the angle between the two flared sides measured at the vertex of the horn, which is at the point where the two sides would intersect if they were extended. The length of the horn l is measured along one of the flared sides from the apex to the mouth. The throat of the horn must be made large enough so that it will transmit TE_{01} waves with low attenuation, but it must also be made small enough so that the higher modes of transmission cannot be set up in the horn. Undesirable minor lobes appear on the radiation pattern when these higher modes are present. Equation 364 can be used to determine the proper size to make the throat.

The beam angle of a horn is defined as the angle through which the radiation takes place; i.e., it is the angle between the zero points on the directional diagram of the radiation. If the horn is infinitely long, the beam angle is equal to the flare angle because all of the radiation is confined to this angle by the sides of the horn. It has already been pointed out that the distribution of the electric-field intensity across the face of a rectangular wave guide carrying a TE_{01} wave is in the form of a half sinusoid. A similar sinusoidal distribution of electric-field intensity exists along an arc across the horn which is centered at the vertex. The electric-field intensity radiated in a particular direction from the end of an extremely long sectoral horn is practically proportional to the field intensity on that directional radius at the mouth of the horn. If the directional angle ϕ is measured from the center line of the horn, the electric-field intensity across an arc in the horn and therefore the field intensity at a distant point from the horn radiating TE_{01} waves has the following form:

$$\mathcal{E} = \mathcal{E}_T \cos \frac{\pi\phi}{\phi_0} \tag{374}$$

where $\mathcal{E}_T$ is the field intensity on the center line. The above equation is valid only inside the flare angle because there is no radiation outside of that sector.

The length of practical horns is limited; consequently, the beam angle produced by them is somewhat broader than the flare angle.

The curve of the beam angle from a horn of given length as a function of the flare angle is asymptotic to both the vertical axis and the diagonal line which represents the characteristic of an infinitely long horn (beam angle equal to flare angle). The minimum point on the curve gives the optimum flare-angle design for that particular length. For a horn 10 wavelengths long the minimum beam angle which can be produced is about 50°. The flare angle corresponding to this minimum is 30°. Longer horns are capable of producing narrower beam angles. For example, the minimum beam angle which can be produced by a horn 50 wavelengths long is 23°. The corresponding flare angle is about 15°.

Formulas have been derived for computing the radiation pattern from a horn of finite length, but they are rather complex. In lieu of them, however, equation 374 can be used to get a fair idea of the radiation pattern which can be expected, provided that the horn is rather long and has a flare angle which is near or greater than the optimum value.

A rough design for a pyramidal horn is obtained by following the procedure which has been outlined for the sectoral horn. The power gain from a pyramidal horn is proportional to the height of the throat. The gain is also proportional to a factor which is practically equal to unity for a small flare angle ϕ_0' on the other sides of the horn, but this factor is greatly reduced when the angle ϕ_0' becomes moderately large. If a sharp pattern is to be obtained in the ϕ_0' plane with TE_{01} excitation, the angle ϕ_0' should be kept small.

The radiation from an array of horns follows exactly the same laws as the radiation from an array of antennas. Equation 345 is also applicable to them. The field intensity from each horn at every angular position is added according to equation 345. In other words, the net radiation pattern for the array is obtained by multiplying the radiation pattern of a single horn by the characteristic of the array as given by equation 345.

Problems

11-1. Suppose that standard rectangular pipe is 1 in. deep and is available in $\frac{1}{4}$ in. increments up to any desired width.

(*a*) Select the size pipe which would give a minimum attenuation to the transmission of 10-cm (wavelength in free space) TE_{01} waves and at the same time would not allow any undesirable modes of transmission to appear.

(*b*) Calculate the attenuation of the pipe selected, assuming that the material from which it is made is sheet steel with $\frac{\mu}{\mu_0}\rho = 630$.

11–2. Calculate the characteristic impedance of a round wave guide 3 in. in diameter when it carries 10-cm (wavelength in free space) TE_{11} waves. Dielectric is air.

11–3. A $1\frac{3}{4}$-in. × 1-in. rectangular wave guide is used to feed 5-cm (wavelength in free space) TE_{01} waves into a $1\frac{3}{4}$-in. × $\frac{1}{2}$-in. wave guide (air dielectric). Compute the dimensions of the wave guide of constant cross section which would match the impedance of these two guides. Sketch the connection.

11–4. A certain sectoral horn for TE_{01} transmission has a flare angle of 30°.

(*a*) Assuming that the horn is very long, plot the directional diagram of its radiation in the plane which is parallel with the two parallel sides and passes through the center of the horn.

(*b*) Plot the radiation pattern from two of the above horns if they are mounted side by side one-half wavelength apart and are excited 5° out of phase with each other.

11–5. Making the same assumptions as those which were used in the derivation of equation 374, plot the directional diagram of the radiation from a very long sectoral horn having a 30° flare angle if it is fed with TE_{02} waves.

Experiment 23

Object:

The object of this experiment is to study the properties of wave guides.

Preliminary:

1. Solve problem 11–1.

2. Calculate the cutoff frequency to TE_{11} waves for a 3-in. inside diameter pipe.

Performance:

1. Feed the output from a 10-cm oscillator into a section of round wave guide. Measure the approximate cutoff diameter for a round guide by noting the point at which a variable diameter pipe ceases to pass the radiation.

2. Connect a detector unit such as one of those illustrated in Fig. 129 to a resonant chamber like Fig. 131. Feed 10-cm energy into the resonant

chamber through the iris. Obtain a resonance curve for the chamber by plotting position of the piston vs. detector current.

3. Repeat part 2 with a different size iris. Give the reason for the difference between these results and those of part 2.

4. Measure the wavelength in the guide by measuring the distance between the positions of the piston which give two successive maxima readings on the milliammeter.

5. Using a wave meter, measure the wavelength (in air) of the energy which was used for part 4. From this data and the diameter of the chamber, compute the theoretical wavelength in the guide and compare with the results of part 4.

6. Couple a traveling detector to a wave guide carrying 10-cm waves. Terminate the detector with a solid closure and measure the resultant standing wave. If a circular traveling detector like Fig. 128*a* is used, it may be necessary to rotate the detector for a maximum indication on each reading and thus compensate for changes in the plane of polarization which may occur along the length of the pipe. Plot the results.

7. Repeat part 6 with the following terminations: an iris, an open end, and a horn. Explain the reason for the difference in standing wave ratios which are measured with each of the terminations.

Experiment 24

Object:

The object of this experiment is to study the field distribution in a circular wave guide and to measure the radiation patterns of horns and parabolas.

Preliminary:

Solve problem 11–4*a*.

Performance:

1. Connect a wave guide carrying TE_{11} waves to a circular horn. Cover the face of the horn with a sheet of cardboard and explore the field at the cardboard surface by use of a dipole antenna connected to a crystal detector and microammeter. Be sure that there is a quarter wavelength short-circuited line "insulator" on the line feeding the detector so that a return path is provided for the rectified direct current. When the reading on the microammeter is a maximum, the antenna is along a line of electric-field intensity. Sketch these lines on the cardboard. Substitute a pickup loop for the dipole and demonstrate that the magnetic lines are perpendicular to the electric lines. Compare the field pattern obtained by exploration with Fig. 126*c*.

2. Pick up energy from a 30° sectoral horn 10 wavelengths long on the dipole antenna and crystal detector when the latter is placed some distance from the horn. Rotate the horn about a vertical axis and read the received current as a function of the horn's position. Plot the data in a contrasting color on the same coordinate axes as part 1 of the preliminary.

3. Excite a half-wave dipole at the face of a reflecting paraboloid and obtain the strength of the current picked up from this combination as a function of the parabola's position.

4. Obtain a radiation pattern on the above with the antenna placed behind the parabola's focus and also obtain another with the radiating antenna in front of the focal point.

5. Plot the radiation pattern of parts 3 and 4.

References

Barrow, W. L., and C. Shulman, "Transmission of Electromagnetic Waves in Hollow Tubes of Metal," *Proc. I.R.E.*, October 1936, pp. 1298–1328.

Barrow, W. L., and L. J. Chu, "Theory of the Electromagnetic Horn," *Proc. I.R.E.*, January 1939, pp. 51–64.

Barrow, W. L., L. J., Chu, and J. J. Jansen, "Biconical Electromagnetic Horns," *Proc. I.R.E.*, December 1939, pp. 769–779.

Barrow, W. L., and C. Shulman, "Multiunit Electromagnetic Horns," *Proc. I.R.E.*, March 1940, pp. 130–136.

Carson, J. R., S. P. Meade, and S. A. Schelkunoff, "Hyper-frequency Wave Guides — Mathematical Theory," *Bell Sys. Tech. Jour.*, April 1936, pp. 310–333.

Chu, L. J., and W. L. Barrow, "Electromagnetic Waves in Hollow Metal Tubes of Rectangular Cross-Section," *Proc. I.R.E.*, December 1938, pp. 1520–1555.

Chu, L. J., and W. L. Barrow, "Electromagnetic Horn Design," *Elec. Eng.*, July 1939, pp. 333–338.

Sarbacher, R. I., and W. A. Edson, *Hyper and Ultrahigh Frequency Engineering*. New York: John Wiley & Sons, Inc., 1943, pp. 1–318, 364–424.

Southworth, G. C., "Hyper-frequency Wave Guides — General Considerations and Experimental Results," *Bell Sys. Tech. Jour.*, April 1936, pp. 284–309.

Southworth, G. C., "Some Fundamental Experiments with Wave Guides," *Proc. I.R.E.*, July 1937, pp. 807–822.

Ware, L. A., and H. R. Reed, *Communication Circuits*. New York: John Wiley & Sons, Inc., 1942, pp. 167–213.

APPENDIX AND INDEX

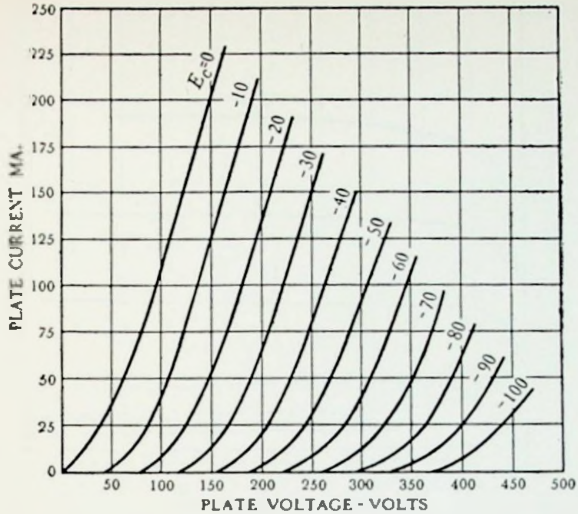


FIG. 132. Plate characteristics of a typical 2A3 triode. Interelectrode capacities: $C_{gp} = 16.5$ $\mu\mu$f, $C_{gk} = 7.5$ $\mu\mu$f, $C_{pk} = 5.5$ $\mu\mu$f. Operating characteristics: $E_p = 250$ volts, $E_c = -45$ volts, $\mu = 4.2$, $r_p = 800$ ohms.

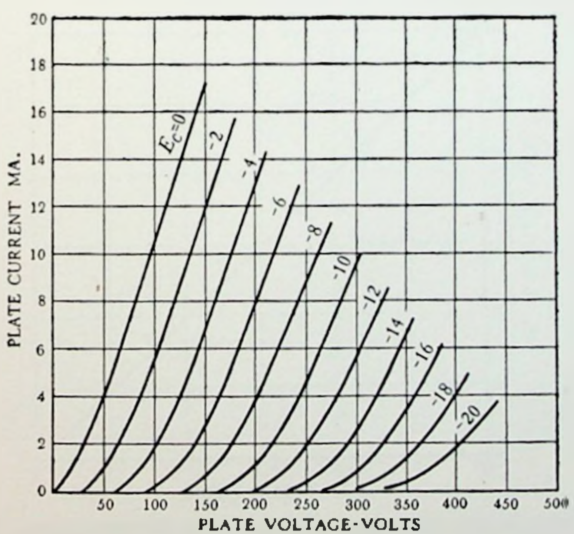


FIG. 133. Plate characteristics of a typical 6J5 triode. Interelectrode capacities: $C_{gp} = 3.4$ $\mu\mu$f, $C_{gk} = 3.4$ $\mu\mu$f, $C_{pk} = 3.6$ $\mu\mu$f. Operating characteristics: $E_p = 250$ volts, $E_c = -8$ volts, $\mu = 20$, $r_p = 7700$ ohms.

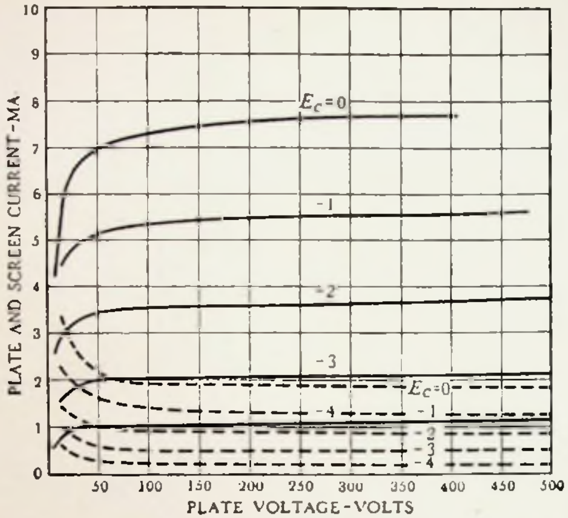


FIG. 134. Plate and screen-grid characteristics of a typical 6J7 pentode (screen-grid voltage $E_{c2} = 100$ volts, suppressor-grid voltage $E_{c3} = 0$ volts). The solid lines are plate current and the dotted lines are screen current. Interelectrode capacities: Input $= 7\ \mu\mu$f. Output $= 12\ \mu\mu$f. Operating characteristics: $E_p = 250$ volts, $E_c = -3$ volts, $E_{c2} = 100$ volts, $E_{c3} = 0$ volts, $g_m = 1225$ micromhos.

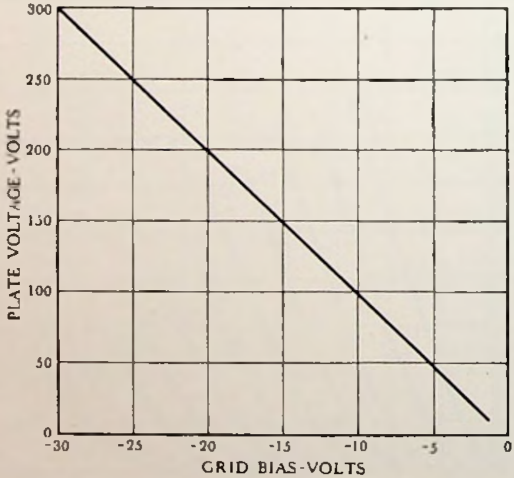


FIG. 135. Critical grid characteristic of a typical 884 gas triode. Interelectrode capacities: $C_{gp} = 3.5\ \mu\mu$f, $C_{gk} = 3.5\ \mu\mu$f, $C_{pk} = 2.5\ \mu\mu$f. Operating characteristics: $E_p = 300$ volts max., $I_p = 300$ ma max., tube drop $= 15$ volts approx. The resistance in series with the grid should be 1000 ohms for each volt the grid is driven positive.

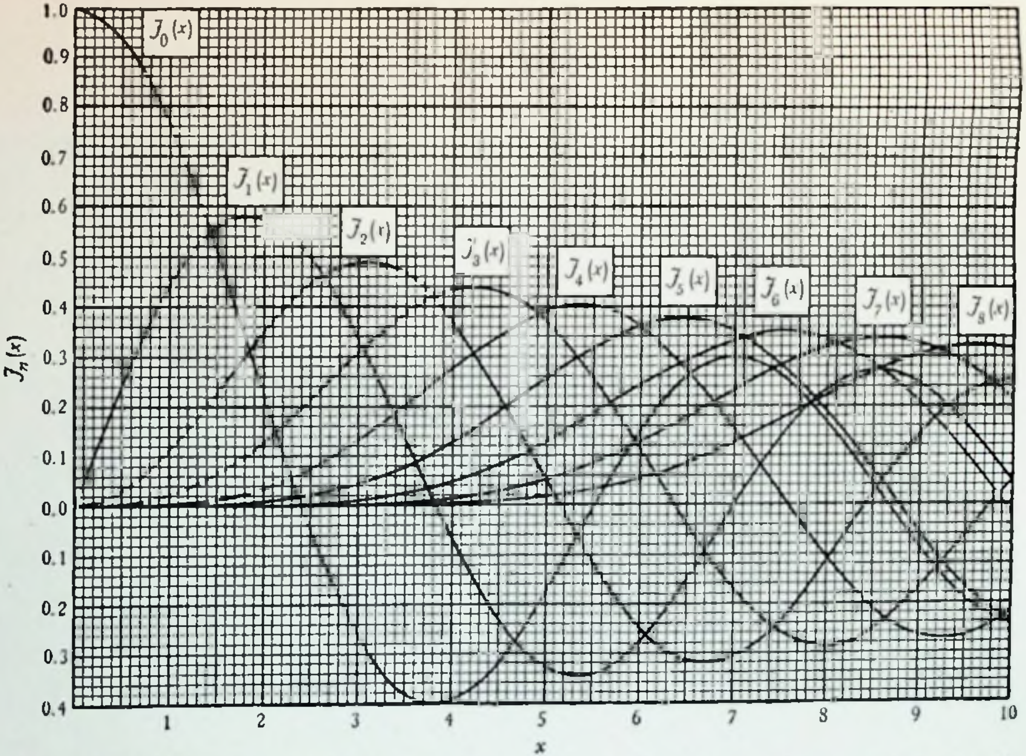


FIG. 136. Bessel functions of the first kind. Functions beyond the range of the curves can be computed by the following series:

$$J_n(x) = \frac{x^n}{2^n n!} - \frac{x^{n+2}}{2^{n+2}(n+1)!} + \frac{x^{n+4}}{2^{n+4}\,2!\,(n+2)!} - \cdots + (-1)^k \frac{x^{n+2k}}{2^{n+2k}\,k!\,(n+k)!} - \cdots$$

INDEX

Q

R

S