ELECTRON OPTICS IN TELEVISION

ELECTRON OPTICS IN TELEVISION

With Theory and Application of Television Cathode-ray Tubes

BY

I. G. MALOFF

Research Division, RCA Manufacturing Co., Inc., Camden, N. J.

AND

D. W. EPSTEIN

Research Division, RCA Manufacturing Co., Inc., Camden, N. J.

FIRST EDITION
FOURTH IMPRESSION

McGRAW-HILL BOOK COMPANY, INC.
NEW YORK AND LONDON
1938

PREFACE

Electron optics and Cathode-ray television are two hardly separable subjects. The extensive development of the former was prompted by the needs of the latter. The extensive development of the latter was due to the great progress of the former. Although there are applications of electron optics other than television, they are of lesser practical importance. This book has been written for the purpose of developing the theory of electron optics and its most useful application—the television cathode-ray tube. Since the tube has to meet a number of exacting requirements imposed by the associated apparatus, a treatment of the latter is included in the text.

This book should not be considered an exhaustive treatment of the subject, but rather an account of that part of it with which the authors have had first-hand experience at the Research Laboratories of the RCA Manufacturing Co., Inc., Camden, N. J.

The Introduction, which presents a brief description of a complete cathode-ray television system and a few applications of electron optics, is for the purpose of informing readers not already familiar with such systems and applications, and also for the purpose of clarifying certain concepts and terms used in television which are referred to later.

The body of the book is divided into two parts. In Part I the theory of electron emission and electron optics is developed. Although the theory of pure magnetostatic and combined electrostatic-magnetostatic focusing is developed, by far the greater part of the treatment of electron optics is limited to pure electrostatic lenses, particularly those involving two coaxial cylindrical electrodes. Part II deals for the most part with the problems encountered in designing tubes, practical and economical to construct and capable of producing satisfactory television pictures when used with practical associated apparatus. Included in Part II are several approximate methods for solving in a practical way some of the nonlinear-circuit problems connected

with the design of apparatus associated with the television cathode-ray tube.

The authors are very grateful to the management of the RCA Manufacturing Co., Inc., for permission to publish the results of work performed at the laboratories of that company. They are especially indebted to L. M. Clement and E. W. Engstrom for their interest and encouragement. They also acknowledge the full cooperation and valuable help of the entire research staff of RCA Manufacturing Co., Inc. W. A. Tolson has contributed greatly to the development of the deflecting systems described herein, and C. M. Burrill, Dr. H. L. Donley and Dr. E. R. Piore have helped by reading and correcting the manuscript.

I. G. MALOFF.
D. W. EPSTEIN.

CAMDEN, N. J.,
February, 1938.

CONTENTS

PART I

ELECTRON OPTICS

CHAPTER 1

ELECTRON OPTICS IN TELEVISION

INTRODUCTION

I.1. Electron Optics.—Electron optics deals with the trajectories of electrons in electric and magnetic fields and considers them from the point of view of geometric optics. More than 100 years ago, Sir William Hamilton showed that there exists a direct analogy between the path of a ray of light through refracting media and the path of a particle through conservative fields of force. When the electron was discovered and its properties understood, it became apparent that the electron moving in vacuum and subject to action of electric and magnetic fields obeyed the laws and the principles formulated by Hamilton.

In general, an electron moving in an electrostatic field will follow a curved path. In studying this path we may adopt a point of view of electron dynamics: that the path is curved because an electrostatic force, which is a function of position, acts on the electron. According to electron optics the path is curved because the electron is passing through a medium in which the index of refraction is a function of position. The reason for adopting the point of view of electron optics is that in many cases this point of view clarifies the problem and simplifies its solution. This is particularly true of electric and magnetic fields having axial symmetry. In electron optics an electrostatic field having axial symmetry is considered as an electron lens, similar to a glass lens for light, and, knowing the constants of the lens, the focusing action of the lens is uniquely and simply determined for paraxial electrons. Thus, given a source of electrons (real or virtual) and an electron lens, the position and size of the electron image of the source may be simply determined (as with light) from the knowledge of the size of the source and the distance between the lens and source.

Electrostatic electron lenses having axial symmetry are formed by applying various potentials to electrodes having axial symmetry, such as coaxial cylinders, aperture disks, cones, etc. An axially symmetric magnetostatic lens is produced by sending direct current through a coil whose axis coincides with the axis of symmetry. It is often misleading to consider electron lenses as similar to the simple glass lenses of optical instruments, since the electron lenses are much more involved. Electrostatic electron lenses are constructed of media in which the index of refraction is a function of position. Further, the ratio of the indexes of refraction on the two sides of the lens may be as high as several hundred. Magnetostatic electron lenses are constructed of media in which the index of refraction is a function of position and direction, so that the electron does not remain in the same meridian plane throughout its path in the lens.

The principal problems of electron optics are (1) lens analysis and (2) lens synthesis. Lens analysis includes the determination of the positions of the lens constants, *i.e.*, the focal and principal points, for a given electrostatic and/or magnetostatic field (or a given electrode and/or coil arrangement) and the determination of the lens aberrations. Since both the lens constants and aberrations may be determined from electron trajectories through the electrostatic and/or magnetostatic field it is seen that the major task of lens analysis is electron-trajectory determination. Lens synthesis involves the determination of electric and/or magnetic fields, or, better, the electrode and/or coil arrangement necessary to produce a lens of given constants and aberrations. As is to be expected, lens synthesis is by far the more difficult of the two. Until recently, electron optics has been occupied almost exclusively with lens analysis. Recently, some work has been done on the determination of lenses having a minimum of aberration.

The treatment of electron optics presented in this book is limited to axially symmetric lens analysis, and particularly to axially symmetric lens systems useful in television cathode-ray apparatus. The electron-optics point of view has been developed to such an extent as to be useful for technical and engineering purposes.

I.2. Scanning Principle.—Since cathode-ray television is one of the most useful results of the advances of electron optics, a

brief description of its principles and essential components seems worth while. The main television principle was born when, in 1884, Nipkow proposed to send pictures point by point in a rapid and orderly succession instead of sending all parts at once. This orderly splitting of the picture into points or elements is now called *scanning*, and the original scanning means employed by Nipkow is termed the *scanning* or *Nipkow disk*.

Nipkow's system, with its mechanically driven rotating disk, remained idle for nearly 40 years. Not until thermionic tubes were developed did Nipkow's system become a practical reality. It had many defects, the greatest being its limitations in transmitting scenes of high detail with sufficient brilliancy. However, it played a great part in the general development of television, chiefly by encouraging the advances in the associated apparatus.

I.3. Nipkow's Scanning Scheme.—While cathode-ray television systems of high definition are very far removed in appearance and in underlying principles from the original Nipkow system, any television system employing one wire or one communication channel can be reduced in its basic performance to Nipkow's scanning scheme. According to this scheme, a square aperture is passed along straight lines over the entire area of a rectangular picture to be transmitted. The picture is divided into a definite number of rectangular strips, usually horizontal, and these strips are scanned in orderly succession by the aperture. The brightness of the part of the picture under the aperture is impressed on a photoelectric cell by means of an optical system. The photoelectric cell generates an electric current proportional to the brightness of the area under the diaphragm. The electric current thus generated is amplified and is sent either over wires or by means of a high-frequency carrier to a receiver.

At the receiver there is a similar aperture moving in synchronism with the one at the transmitter. The electric current received by wire or radio is applied to a source of light, the luminous intensity of which follows precisely the intensity of the received current. This light is sent through the diaphragm. In this way, the total amount of light which falls on a given part of the picture at the receiver is approximately proportional to the illumination of the corresponding part of the original picture scanned at the transmitter. The light at each point of the receiver picture falls there only when the diaphragm is passing

over it. Owing to the retentiveness of the human eye, the received picture will appear similar to the picture at the transmitter if the rate of scanning is high and a sufficient number of pictures is sent per second.

The order in which the scanning is done, *i.e.*, whether it is done from left to right or from top to bottom, is not essential. The only thing that matters is that the picture is scanned completely each time it is transmitted.

I.4. General Scheme of Television System.—Figure I.1(*a*) shows a block diagram of a complete television system. Both high- and low-definition television systems can be divided into

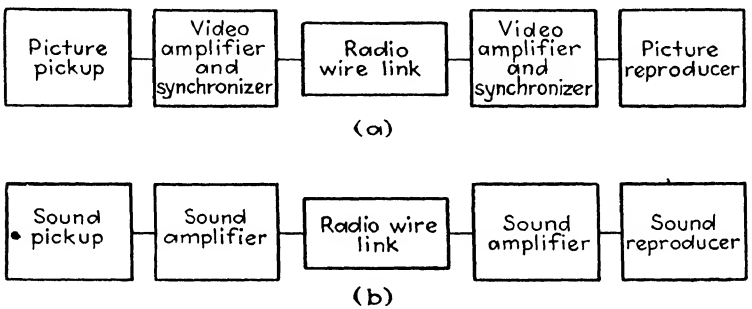


FIG. I.1.—(*a*) Block diagram of television-communication system. (*b*) Block diagram of sound-communication system.

the parts shown. The only difference between the television systems of high and low definition is the degree of detail transmitted and received with satisfactory illumination levels. The old mechanical systems are satisfactory for transmitting pictures that have a detail equivalent to 100 scanning lines, and reach their useful limit of definition somewhere between 100 and 200 lines. The high-definition systems employing no moving parts are almost unlimited in their capability of definition. The limitations on the latter class of television systems are imposed chiefly by the associated communication apparatus and by commercial considerations.

In Fig. I.1(*b*) a block diagram of a complete sound-communication system is shown. The only essential difference between the latter and the system shown in Fig. I.1(*a*) is that the television system requires synchronizing arrangements while the sound system does not. Apart from synchronization, the similarity is great. Of course, the amplifiers have to be different; they

have to pass frequency bands of greatly differing widths. But in all other essentials both systems are alike. Both systems pick up forces affecting our senses, and both translate them into intensity-*vs.*-time variations of electric currents. In both systems the electric current at the pickup stage is of very low intensity, and both have to use amplifiers to bring it up to levels suitable for transmission over distances. Again, both systems may or may not use high-frequency carriers to transmit the intelligence over distance by means of either radio or cable.

At the receiving end both systems have to demodulate and amplify the received signals to bring them to the required power levels, and then, by means of suitable transducers, to reproduce forces similar to those which were originally picked up and which affect our senses.

I.5. Electron Beam.—An electron beam in cathode-ray tubes may be visualized as a narrow pencil of negatively charged particles moving with velocities of the order of 30,000 miles/sec. While electron beams were discovered as early as the end of the last century, it is only in the last decade that their properties have been understood and means for their generation and control developed. The discovery of the electron, of electron emission, and the development of means for their control opened the way which led to present-day radio broadcasting and communication systems in general. In exactly the same way the gradual development of means for concentrating electrons into narrow beams and the development of means for controlling the intensity and direction of these beams were directly responsible for the development of the present-day high-definition television systems.

In a high-definition television system, both the picture-pickup device and the picture-reproducing device employ the cathode-ray beam. They employ the beam for the scanning of both the transmitted and the received pictures. V. K. Zworykin, who is responsible for the development of a successful and complete television system, named the picture-pickup device the *iconoscope*, a combination of the Greek words "eikon," meaning image, and "skopein," meaning to observe. The cathode-ray receiving tube he named *kinescope*. In this book, an abbreviation, "TCR tube," will be used to denote "television cathode-ray tube." It is now in order to describe briefly the salient features of both of these devices before describing the performance of a modern

cathode-ray television system as a whole. The term "television system," as used below will refer to a modern high-definition television system employing the electron beam for scanning both at the transmitter and at the receiver.

I.6. TCR Tube.—The cathode-ray tube in a television receiver corresponds to the combination of a power-output tube and a loudspeaker in a sound radio receiver. A cross section of a typical TCR tube is shown in Fig. I.2. It consists essentially of five component parts: first, a glass envelope, sealed for the maintenance of high vacuum; second, a hot cathode from which the electron beam originates; third, a device for controlling and focusing the electron beam; fourth, an arrangement (either internal or external) for deflecting the beam for scanning purposes; and fifth, a fluorescent screen on which the image is reproduced.

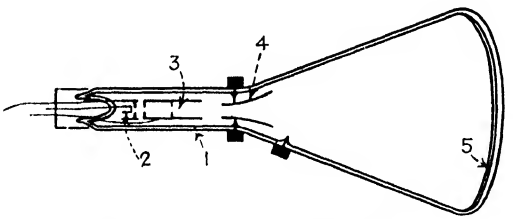


FIG. I.2.—A typical TCR tube.

The envelope, usually of glass, is made strong enough electrically and mechanically, being designed to withstand the atmospheric pressure with a large margin of safety. The tube, in its processing, is carefully baked and evacuated to obtain a high vacuum of the order of 10^{-7} cm. of mercury.

The early tubes were partly gas-filled and utilized for beam concentration, the space charge resulting from the collisions of electrons with molecules of residual gas. This so-called gas focusing is rather uncertain, since the gas pressure in any tube varies with life. If in a tube the focusing of the electron beam is critically dependent on pressure, the useful life of such a tube is comparatively short.

Most of the modern cathode-ray receiving tubes are of the high-vacuum type. In this type, the focusing of the beam is accomplished by the action of electric and magnetic fields. The vacuum treatment of modern TCR tubes is such that even at the

end of a tube's life, *i.e.*, when the cathode emission has fallen off, the vacuum still is so high that the collisions of electrons with molecules of residual gas produce negligible effects.

The cathode in a TCR tube is of a tubular form, with a flat emitting surface covered with a preparation of barium and strontium oxides. Only the flat end facing the luminescent screen is covered by the electron-emitting material. A tungsten heater, noninductively wound and insulated from the cathode by means of a heat-resisting material, is located inside of the tubular cathode.

The electron gun, or the device for concentrating, controlling and focusing of the electron beam, consists of a grid sleeve and a first anode. Sometimes it includes another electrode, usually called screen grid, which is not essential for operation and is not shown on the figure. The grid sleeve is of a tubular form with a disk parallel to the flat emitting surface of the cathode. A circular hole in the center of the disk is coaxial with the cathode sleeve. The first anode is also of tubular form coaxial with the rest of the gun. It usually carries diaphragms or aperture disks on the inside to stop or limit the beam angle and the penetrations of electrostatic fields. The glass envelope of the TCR tube carries a black conductive coating on its inner side and has a sealed-in conductor leading to this coating. The conductive coating forms the second anode. The final electron-accelerating potential is applied between the cathode and the second anode. The purpose of the first anode is: first, to stop the beam, which is similar to the action of an optical stop in a lens; and, second, together with the second anode, to create an axially symmetric electrostatic field or an electron lens which starts the initially divergent electrons of the beam toward the axis. By adjusting the voltage on the first anode the luminous spot on the screen can be brought to a minimum diameter or focus. The voltage on the first anode for the best focus is usually about one-fourth or one-fifth of that of the second anode. Figure I.3 shows a photograph of a TCR tube having a screen 5 in. in diameter and capable of reproducing a picture of about 3 by 4 in.

Instead of an electrostatic system for focusing the beam, a magnetostatic system may be used. A magnetostatic field having axial symmetry acts on an electron beam similarly to an electrostatic field. The general matter of focusing is treated at

much greater length in Part I. The details of this action will be left until then, since for an understanding of the general operation of a high-definition television system a description of an electrostatic TCR tube will be sufficient.

FIG. I.3.—Five-inch TCR tube.

Just as soon as the beam leaves the first anode it is subjected to the action of either magnetic or electric fields for the purpose of deflecting the beam in a predetermined manner. This deflecting is scanning in television. There are three ways of scanning: first, by means of two electric fields at right angles to each other;

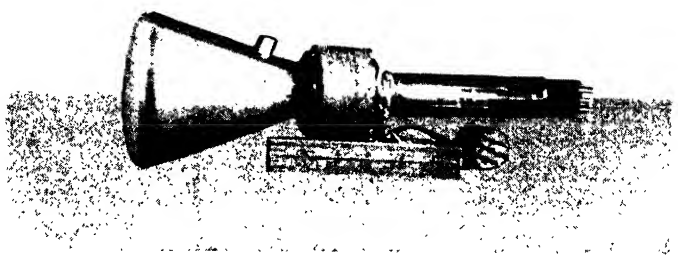


FIG. I.4.—TCR tube with deflecting yoke.

second, with two magnetic fields at right angles to each other; and third, with electric and magnetic fields parallel to each other. It may be mentioned at this point that an electrostatic field deflects the electrons along the lines of force and an electromagnetic field deflects them perpendicularly to the lines of force. The deflecting plates are usually placed inside the glass envelope, while the magnetic deflecting coils are usually outside of the

envelope. In Fig. I.4 is shown a TCR tube with a typical all-magnetic deflecting arrangement. This arrangement is called the *deflecting yoke.*

A luminescent screen is placed on the inner side of the front face of the tube. This face is as flat as the mechanical strength permits. The material of the screen is usually either a sulfide or silicate of zinc and is deposited in a very thin, translucent layer. Under electron bombardment, this material emits light which quickly vanishes when the bombardment ceases.

I.7. Operation of TCR Tube.—The operation of a TCR tube is as follows:

The electrostatic field created by the potential applied to the first anode penetrates the grid opening and draws the emitted electrons into a well defined beam. The grid is usually at somewhat lower potential than the cathode and in this way limits and controls the beam intensity. By applying a more negative potential to the grid, the beam can be completely cut off. The beam scans the luminescent screen, and its intensity is varied just as the intensity of the neon light in the Nipkow's scheme is varied. The received picture is reproduced on the luminescent screen. The voltage controlling the intensity of the beam is impressed between the grid and cathode of the TCR tube. After entering the first anode, the beam passes through a masking diaphragm, which cuts off some of the irregular peripheral portions of the beam. Then the beam enters the region of the field produced by the difference of potentials between the first and second anodes. In this field, a strong focusing action takes place, which gives the electrons a radial velocity component directed towards the axis of symmetry of the beam. The radial momentum acquired by the electrons is sufficient to bring them, after a flight through the equipotential space of the main body of the tube, to a focus on the screen.

Some of the electrons originally drawn from the cathode are cut off by the masking aperture disk. They return to the source of the e.m.f. through the first anode lead. The rest of the electrons strike the fluorescent screen. They excite the screen and dissipate most of their kinetic energy there. This kinetic energy has been acquired by the electrons through acceleration from the very small velocities of emission to that corresponding to the second-anode potential.

Some of the energy of the beam is transformed into light, some goes into heat, raising the temperature of the glass, while the rest is spent in knocking out the secondary electrons from the screen material. These low-velocity secondaries flow in a steady stream to the conductive coating of the second anode. An equilibrium condition is quickly established, the conductive coating maintaining the voltage equal to that of the direct-current supply, while the fluorescent screen slides a few score of volts below the potential of the coating. The difference of potentials between the fluorescent screen and the conductive coating adjusts itself so that it draws the secondaries to the coating at exactly the same rate as the rate of arrival of the primaries to the screen.

I.8. Image Pickup.- The key to a high-definition television system is an image-pickup device capable of translating the light image into variations of electric current. Many devices have been developed and proposed for this purpose. Without exception, all the devices used utilize the well known property of certain materials of emitting electrons while exposed to light. Image pickups may be divided into two groups; *viz.*, storage pickups and non-storage pickups. The difference between the two is as follows:

In the former the photoelectric current from an element[1] of the picture charges an individual condenser for a period of time equal to the scanning time of one complete picture. This condenser is discharged once during a scanning time of a complete picture, the time of discharge being only the time of scanning of one picture element. In the non-storage pickup the current from the photoelectric cell flows only during time of scanning, does not charge any condenser, and therefore in the second type of image pickup no storage of the charge caused by the photoelectric effect takes place.

At present one image pickup of the storage type seems to have come into universal use. It is the iconoscope developed by V. K. Zworykin.[2] The iconoscope consists of a photosensitive

[1] The concept of the picture element and of the scanning line is treated separately in Sec. I.9.

[2] A large part of the material in this section is taken from the articles, The Iconoscope, by V. K. Zworykin, *Proc. I.R.E.*, **22** (1934), and Theory and Performance of Iconoscope, by V. K. Zworykin, G. A. Morton and L. E.

mosaic and an electron gun, assembled in a glass bulb which is highly evacuated. The electron gun is of the same general type as the gun described in Sec. I.6. The gun assembly is shown diagrammatically in Fig. I.5. It functions as an electron optical system for producing a narrow bundle of electrons which is made to scan the mosaic by means of magnetic deflecting coils.

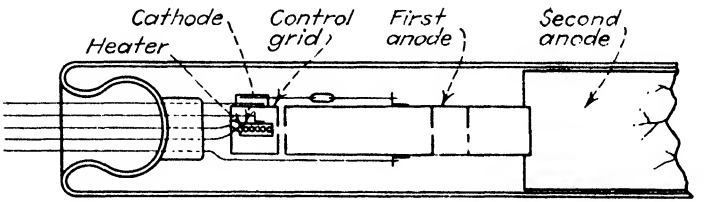


FIG. I.5.—Iconoscope electron gun.

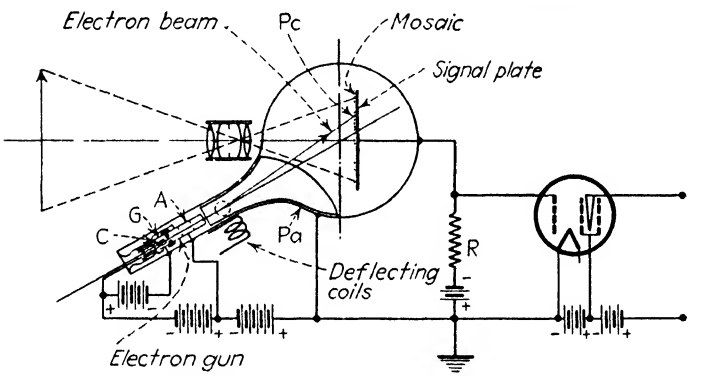


FIG. I.6.—Iconoscope and its circuit.

The mosaic consists of a thin sheet of mica coated with a conducting metal film on one side, and covered on the other with a vast number of tiny, photosensitized silver globules. This mosaic is mounted in the blank in such a position that the electron beam strikes the photosensitized side at an angle of 30 deg. from the normal, and the optical image to be transmitted is projected normal to the surface on the same side. The arrange-

Flory, *Proc. I.R.E.*, **25** (1937). Two more papers on the theory of storage image pickups are of importance: On Light Storing Devices, by V. I. Krasovsky, *I.E.S.T.*, Feb., 1936, (in Russian) and The Operation of Cathode Ray Scanning Devices with Storage, by R. Urtel, *H.u.E.*, **48**, 150 (1936) (in German).

ment of these elements in the tube is shown in Fig. I.6, while Fig. I.7 shows a photograph of an iconoscope and its deflecting yoke.

Briefly, the iconoscope mosaic may be thought of as a two-dimensional array of tiny photoelectric cells, each shunted by a condenser which couples it to a common signal lead. When the mosaic is illuminated, these condensers are charged positively with respect to their equilibrium potential due to the emission of photoelectrons from the photosensitive elements. For any

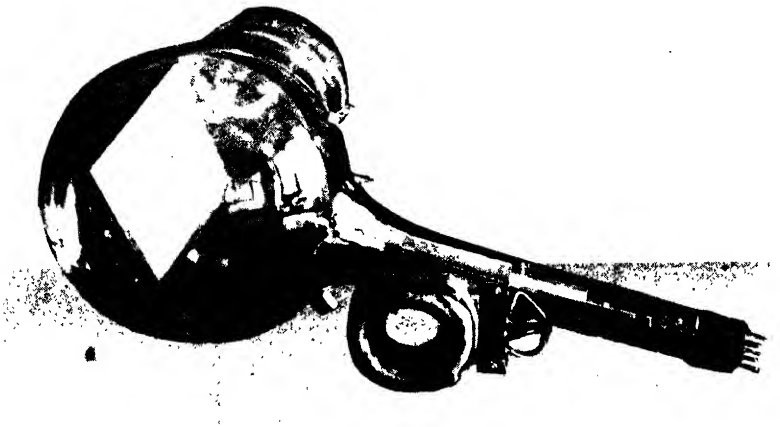


FIG. I.7.—Iconoscope with deflecting yoke.

particular element, this charging process continues for a time equal to the picture-repetition interval, *i.e.*, until the beam in the process of scanning returns to the element. When the beam strikes it, it is driven to equilibrium, releasing its charge and inducing a current impulse in the signal lead. The train of impulses thus generated constitutes the picture-signal output of the iconoscope. This description serves to illustrate the general principles of the iconoscope but is not sufficiently accurate to form the basis of an analysis of its operation.

In considering the operation of the iconoscope, because of the fact that the silver globules are so small that a great number of them are under the beam at any instant, the mosaic may be treated as a continuous surface which has infinite transverse

resistance, and which has both a high secondary-emission ratio (in the neighborhood of 5 or 7) and photosensitivity. This surface has a capacity of about 100 $\mu\mu$f/sq. cm. in the case of a standard tube.

The average potential of the mosaic under bombardment, while no light is falling upon it, is between 0 and -1 volt with respect to the elements which collect the secondary emission from the mosaic, *i.e.*, the second anode. However, the potential is not uniform over the entire surface. The area directly under the scanning beam will be at a potential of $+3$ volts with respect to the second anode, this being the potential at which the secondary-emission ratio from caesiated silver becomes unity. Away from the point under immediate bombardment in the region which has just been traversed by the scanning beam, the potential will be found to decrease until at a distance equal to 25 or 30 per cent of the vertical scanning distance, the potential reaches about $-1\frac{1}{2}$ volts with respect to the second anode. The rest of the mosaic is at this potential. This decrease in potential is caused by electrons which leave the point under bombardment and return to the mosaic as a more or less uniform rain of low-velocity electrons.

The scanning beam sweeping over the mosaic acts like a resistive commutator. The resistance in this case is determined by the current-voltage relation of the secondary electrons from the bombarded point. While this resistance is actually non-ohmic, appreciable error is not introduced in assuming it ohmic over the small voltage range dealt with in the case of the iconoscope. Experimental measurements show this beam impedance Z to be given by the relation

$$Z = \frac{Z_0}{i_b} \tag{I.1}$$

where i_b is the beam current and Z_0, the coefficient of beam impedance, having a value between 1 and 2 ohm-amp.

Any small element of area of the mosaic when it is swept over by the scanning beam must change its potential from $-1\frac{1}{2}$ volts to its equilibrium of approximately $+3$ volts. However, even with no light on the mosaic, the equilibrium, both under the beam and away from the beam, is not constant over the surface of the mosaic.

The electron beam itself is not a commutator. It sprays the same amount of electrons on all portions of the mosaic, irrespective of the small variations in the potentials of the individual elements resulting from the photoelectric effect and redistribution of charges. The beam establishes, however, a conductance between the bombarded spot and the second anode, and this conductance acts as a resistant commutator.

The entire secondary current, however, will not reach the second anode because, on the average, as much current must reach the mosaic as leaves it. Approximative calculations show that out of the total charge released by the element when struck by the beam, only about 25 per cent will reach the second anode, the remainder being returned to the mosaic. In other words, only about 25 per cent of the stored charge is available for producing the picture signal.

As was mentioned above, the equilibrium potentials are not constant over the mosaic, even when it is in darkness. As a consequence, there is a variation in the current reaching the second anode as the mosaic is scanned. This gives rise to a spurious signal which produces irregular shading over the picture. In practice, this spurious signal is compensated by means of electrical correcting networks.

Besides the inefficiency consequent on the redistribution losses, the photoelectric emission is also inefficient due to the small fields drawing the electrons away from the mosaic. The effective photoelectric emission of the mosaic in the standard iconoscope is only about 20 to 30 per cent of its saturated value. This means that the overall efficiency of the iconoscope is only 5 or 10 per cent. In spite of this inefficiency, the very great advantage resulting from the use of the storage principle makes the iconoscope a very effective pickup tube.

The photosensitivity of the mosaic is of the same order as that of high-vacuum caesiated-silver photoelectric cells. The same is true of the color response.

The electron gun producing the beam is quite an important factor in the performance of the iconoscope. Since the resolution of the iconoscope is determined by the size of the scanning spot, the gun should be designed accordingly. In its essential details, it is very similar to the gun used in kinescopes. The gun is mounted in the long, narrow neck attached to the spherical

bulb housing the mosaic screen. The inner surface of the neck and part of the sphere are metallized and serve as the second anode for the gun and also as a collector for the electrons from the mosaic. The first anode operates at a fraction of the potential applied to the second anode, which operates at about 1,000 volts. The deflection of the electron beam for scanning the mosaic is accomplished by a magnetic field. The deflection coils are arranged in a yoke which slips over the neck of the iconoscope. An assembled deflecting unit is shown beside the tube in Fig. I.7. The scanning is linear in both vertical and horizontal directions and is produced by saw-tooth impulses of current passing through the deflecting coils.

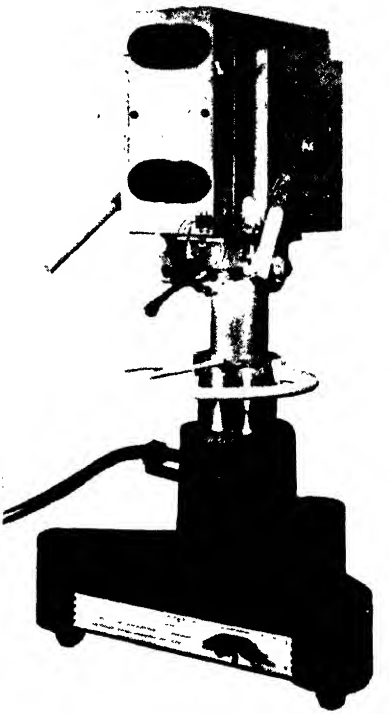


FIG. I.8.—Iconoscope camera.

Since the iconoscope is practically a self-contained pickup unit, it is possible to design a very compact camera containing the iconoscope and a pair of amplifying stages connected to the main amplifier by means of a long cable. Since the camera is portable, it can be taken to any point of interest for the transmission of a television picture. The photograph of such a unit is shown in Fig. I.8.

I.9. Scanning Line and Picture Element.—Under the heading, "Nipkow's Scanning Scheme," it was stated that according to this scheme a square aperture is passed along straight lines from left to right (or from right to left) over the entire area of a rectangular picture to be transmitted. Also, that the picture is thus divided in a certain number of horizontal rectangular strips, and these strips are scanned in orderly succession by the aperture.

The horizontal rectangular strip along which the scanning is done is called the scanning line, and the number of these strips

in a complete television picture is called the number of scanning lines in the television picture. This term has been already used in the preceding paragraphs of this book but has not been explained. A complete theory of scanning is beyond the scope of this book,[1] but a few elementary explanations are deemed necessary.

Suppose our rectangular television picture is scanned horizontally by means of a square aperture. Let n be the number of scanning lines in the picture and k the aspect ratio (ratio of the width of the picture to its height). The product n^2k is called the *theoretical number* of picture elements in the scanned picture. A theoretical picture element thus can be defined as the area of the scanning aperture. The roundabout way in which the definition of the picture element is given is justified by the fact that the picture element in the Nipkow's scanning scheme is fictitious. This fiction is very helpful, however, for a rough estimation of the degree of detail of which a given television system is capable.

Take for a picture to be scanned a black and white checkerboard made of squares of the same size as our theoretical picture element. For an ideal scanning system a periodic current of rectangular wave shape in the photoelectric cell should result. A fair approximation of this shape is a sine wave having the duration of a complete cycle equal to the time the spot passes over two complete picture elements. To transmit such a checkerboard would then require $n^2k/2$ complete cycles of a sine wave. Continuing the same line of reasoning, if we desire to send m complete pictures (or frames) per second we will need a sine wave of frequency equal to $n^2mk/2$ cycles/sec. Thus for a transmission of a checker-board picture of 343 lines, with 4:3 aspect ratio, at 30 frames per second the required theoretical frequency, f is

$$\frac{n^2mk}{2} = 2{,}350{,}000 \text{ cycles/sec.} \qquad \text{(I.2)}$$

This quantity is called the *maximum theoretical frequency* required to transmit a television picture under given scanning conditions. As will be shown in what follows, the synchronizing requirements

[1] For a rather extensive treatment of the subject of scanning in television the reader is referred to an article by Erich Hudec, *E.N.T.*, June, 1931, p. 229.

and also the requirements for balanced directional resolution somewhat modify the results just obtained.

I.10. Directional Resolution.[1]—Suppose that at the transmitter we scan a luminous spot of negligible dimensions (we shall call it a point) by a square spot of the same dimensions as at the receiver, and moving with the same velocity, of course. Also

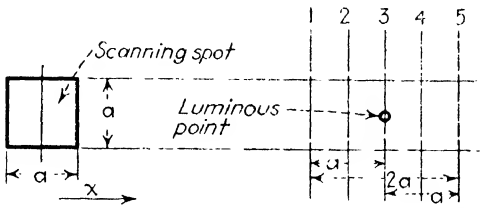


FIG. I.9.—Square spot scanning a luminous point.

suppose that the rest of the picture is dark and that the luminous spot is of such an intensity as to make the current in the pickup device jump to a certain value. This value will remain constant for all the time that the transmitter's rectangular spot is covering the point.

Figure I.9 shows such a point. When the center of the scanning spot reaches a position indicated as 2 in the figure, the current at the pickup device rises to a certain value and stays at that value until the center of the scanning spot reaches position 4. The current at the pickup is shown in Fig. I.10. Assuming an ideal transmitting and receiving system, the current at the receiver will also look like the one in Fig. I.10. This means that at the receiver the beam is out until the center of it reaches position 2. At this point, the light is turned on, reaching the final and constant value instantly. It will stay at that value until its center reaches position 4, where it is turned off again. So, in the received picture there will be light thrown on all points along the scanning line from line 1 and up to line 5. The distance between these two lines is equal to twice the width of the scanning spot. In other words, some light will be thrown on all points along the scanning

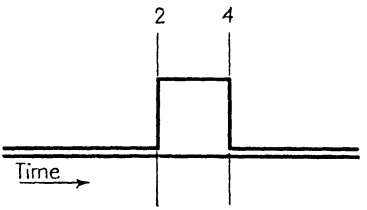


FIG. I.10.—Photo-current corresponding to conditions in Fig. I.9.

[1] MALOFF, I. G., Problems of Cathode Ray Television, *Electronics*, January, 1934.

line from one spot width of scanning line before the true position of the original spot until one spot width after this position. The apparent brightness of the various points of this illuminated strip will not be uniform, however.

Let I be the brightness of the spot in the reproducer while the center of the spot travels from position 2 to position 4. For all other locations of the center of the spot the brightness will be zero. Also let T be the period of one frame, v the velocity of the scanning spot and a the width of it. Since we are interested only in the apparent illumination of points along the x axis between positions 1 and 5, let us count x from position 1.

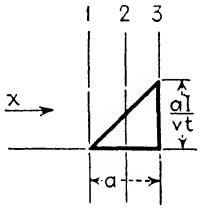


FIG. I.11.—Apparent illumination at receiver corresponding to Fig. I.9.

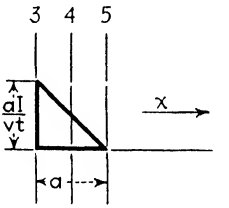


FIG. I.12.—Apparent illumination at receiver corresponding to Fig. I.9.

An interval of time for which a point, or rather a vertical line through any point x, between $x = 0$ and $x = a$ is illuminated, is x/v, which is equal to the time during which the left edge of the scanning spot travels from position 1 to position x. The apparent average intensity of illumination of this line is

$$I_{\text{apparent}} = \frac{xI}{vT}. \tag{I.3}$$

At point 1, x is equal to 0, so that the apparent illumination at point 1 is 0; similarly, at point 3, x is equal to a, so that the apparent illumination at point 3 is aI/vT, and between these points it varies along a straight line, as shown in Fig. I.11.

Now let us see what happens after the left edge of the scanning spot leaves position 3. Let us start measuring anew from position 3. Here the time of illumination of a vertical line through any point x, is $\frac{a - x}{v}$; therefore, the apparent illumination may be expressed as follows:

$$I_{\text{apparent}} = \frac{(a - x)I}{vT}. \tag{I.4}$$

At point 3, x is equal to 0, so that the apparent illumination at this point is aI/vT; similarly, at point 5, x is equal to a, and apparent illumination at point 5 is 0. Between points 3 and 5, it varies along a straight line, as shown in Fig. I.12. In this way we arrive at a distribution of apparent brightness of illumination along the scanning line, as shown in Fig. I.13. The appearance

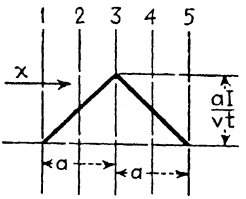


FIG. I.13.—Apparent illumination at receiver having ideal response corresponding to Fig. I.9.

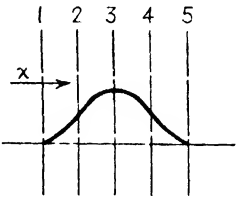


FIG. I.14.—Same as Fig. I.13 except for receiver with limited frequency response.

of the reproduced image of an illuminated point is shown in Fig. I.15. If the system has a limited instead of an ideal response, and limited to the frequency calculated by the approximative formula already described, the received intensity of illumination will be distributed as shown in Fig. I.14, and the appearance of the reproduced image still will be close to that shown in Fig. I.15.

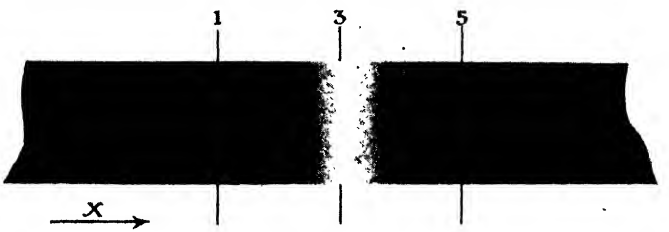


FIG. I.15.—Reproduced image of illuminated point.

Let us consider what has happened to the image of the original point. Along the horizontal we have a maximum at position 3, the exact location of the original point. Along the vertical our point could be anywhere along the vertical line 3 inside the scanning line. It is true that the image along the horizontal is somewhat blurred, but the general character and the location

are correctly reproduced, while along the vertical an illuminated strip would give us the same image as the illuminated point.

I.11. Frequency Band Required by Television Picture.—After trying several objects other than the point, one will find that for nearly all of them the vertical resolution is poorer than the horizontal. It may also be mentioned that the diagonal resolution will appear poorest of all.

When viewing a television picture, any picture for that matter, the observer tends to adjust his viewing distance in such a way that the structure of the picture disappears. With it disappear the imperfections in reproduction of detail. When there is unequal directional detail the observer will tend to adjust himself for the poorest and will lose some of the best. Therefore, the excess of detail along one direction over the other cannot be utilized and is lost.

We may say, therefore, that for a given frequency band from 0 to f, the number of scanning lines should be considerably higher than that derived from the theoretical relation previously given

$$n = \sqrt{\frac{2f}{mk}}. \tag{I.5}$$

It can be shown either theoretically by computing various patterns, or experimentally by trying various pictures, that the received picture improves in overall resolution for a given frequency band with the number of lines increasing until this number reaches a value between

$$n = 1.25\sqrt{\frac{2f}{mk}} \quad \text{and} \quad n = 1.4\sqrt{\frac{2f}{mk}}. \tag{I.6}$$

This means that a band required to reproduce a picture of m frames and n lines is only 65 per cent of the theoretical. For 343 lines at 30 frames/sec. the required frequency band (or the highest frequency required) is only 1,530,000 cycles/sec. instead of 2,350,000 cycles. The formula for f becomes

$$f = 0.325n^2mk \text{ cycles/sec.}$$

I.12. Frame- and Line-frequency Requirements.—Television images consist of rapidly superimposed, individual frames much

the same as motion pictures.[1] In the case of motion pictures, a group of time-related "stills" is projected at a uniform rate, rapid enough to form a continuous picture through persistence of vision. By present methods each frame of a television image is built up element by element in a definite order, and these time-related frames are reproduced at a rapid rate.

In motion pictures the taking or camera frame frequency determines how well the system will reproduce objects in motion This has been standardized at 24 frames per second. In television it is assumed that we shall use a frame frequency of 24 per second or greater. Since this is satisfactory for motion pictures, it is also satisfactory for television, and this characteristic of frame frequency will, therefore, not be considered further.

If we qualify and limit the ability to tell a desired story to specific conditions, the experience we have had with television allows us to make some interesting approximate generalizations. If we take as a standard the information and entertainment capabilities of 16-mm. home-movie film and equipment, we may estimate the television images in comparison.

TABLE I.1.—NUMBER OF SCANNING LINES AS FUNCTION OF PERFORMANCE

No. scanning lines	Performance
60	Entirely inadequate
120	Hardly passable
180	Minimum acceptable
240	Minimum satisfactory
360	Excellent
480	Equivalent for practical conditions

The comparison in Table I.1 assumes advanced stages of development for each of the line structures. We may say, therefore, that a number of scanning lines in excess of 360, say 440, will give a very good performance comparable with 16-mm. home-movie film.

I.13. Flicker Characteristics.—In his two papers[2] on the study of television-image characteristics, E. W. Engstrom treats in great detail the problem of determination of frame and line frequencies in terms of detail, illumination and flicker character-

[1] For a detailed treatment of this subject see E. W. Engstrom, A Study of Television Image Characteristics, Pt. I, *I.R.E. Proc.*, **21** (1933), and Pt. II, *I.R.E. Proc.*, **23** (1935).

[2] *Loc. cit.*

istics. We shall attempt to give here in a condensed form some of his findings, which will be useful in understanding television reception.

In the reproduced image there is another effect of the frame frequency, which has been the subject of lengthy theoretical and experimental investigations. This is the effect of frame frequency on flicker. Motion-picture projectors commonly used are of intermittent type. The usual cycle of such projectors is that, at the end of each projection period, the projection light is cut off by a light cutter. The film is then moved and stopped so that the succeeding frame registers with the picture aperture. At this time the light cutter opens, starting the next projection period. The above procedure is repeated for each frame (24 per second). Since a projection at 24 light stoppages per second with the illumination levels used in motion pictures causes too great a flicker effect, the light is also cut off at the middle of the projection period for each frame for a time equal to the period that it is cut off while the film is moved from one frame to the next. This results in projection at 24 frames per second with 48 equal and equally spaced light impulses. Such an arrangement provides a satisfactory condition as regards flicker.

In television we also may have a reproduced image at 24 frames per second, but because of the manner in which the image is reconstructed, a continuous scanning process, it is not practicable further to break the light impulses by means of a light chopper in a manner similar to that used in the projection of motion pictures. We, therefore, have for the usual systems of television a flicker frequency which corresponds to the actual frame frequency (24 per second, for example). This would be satisfactory at very low levels of illumination, but becomes increasingly objectionable as the illumination is increased.

If an eye is exposed to a source of rapidly varying intensity, the effects of the finite rates of growth and decay of sensation may prevent flicker from being noticeable. This is true provided the total cycle of variations is regular and of high enough frequency. If the frequency of the varying source is sufficiently high so that flicker is imperceptible, the eye is able to integrate the brightness over the cycle of variation. The effect is as if the light for each cycle were uniformly distributed over the period of the cycle.

The highest frequency at which the flicker can just be detected is called the critical frequency. It has been shown that the critical frequency is practically a linear function of the logarithm of the brightness of the field of vision (over the range of brightness of interest in television). The sensitivity of the eye to flicker is noticeably increased when increasing the field of view from a few degrees to an image of the size and viewing distances encountered in motion-picture practice. The sensitivity to flicker is also greater for averted vision when viewing large fields of varying brightness.

In television there are a great number of factors that contribute to flicker effects. These in general are:

Number of frames (light impulses) per second.

Brightness of the image.

Percentage of time the image is illuminated for one frame cycle.

Wave form of rise and decay of light impulse.

Size of the image in terms of angle subtended at the eye of the observer.

I.14. Flicker in Television.—In a television system using a TCR tube in the receiver, each element of the image on the luminescent screen, when excited by the electron beam, fluoresces and assumes a value of brightness corresponding with the excitation. Upon removing the excitation this brightness decays (phosphoresces) in an exponential manner dependent upon the screen material. The persistence of the image screen aids the persistence of the eye in viewing the reproduced image. The persistence characteristic of a TCR tube screen of the type generally used (willemite) is shown in Fig. I.16.

In order to obtain data on flicker from a kinescope under several conditions of operation, a series of observations were made by Engstrom[1] and his associates. The deflecting circuits were so arranged that the vertical speed (frames per second) could be varied. With the picture height 6 in. and the viewing distance 3 ft., the ratio of picture height to viewing distance was 1:6. The stray room illumination was of the order of 0.1 footcandle. The screen illumination was measured with a direct-reading illuminometer at the kinescope glass, looking toward the screen. A sufficient number of lines of even distribution was used to fill completely the picture height.

[1] *Loc. cit.*

Data were taken by a number of observers for three conditions of flicker; *viz.*, just noticeable, noticeable but satisfactory, disagreeably objectionable. The average results of these observations are shown in Fig. I.17.

These observations indicate that for even one foot-candle of screen illumination 38 frames per second is required for just

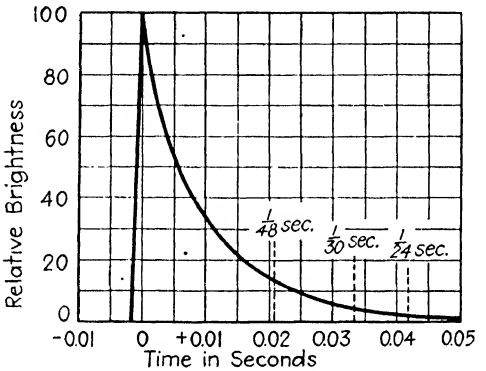


FIG. I.16.—Persistence characteristic of willemite screen.

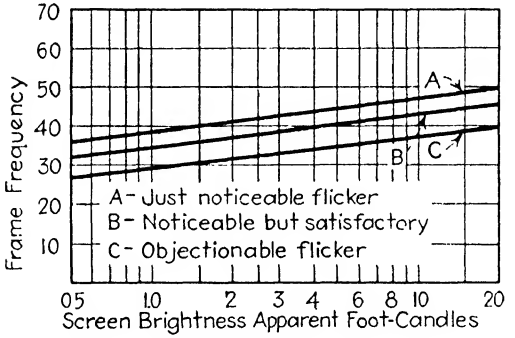


FIG. I.17.—Flicker of willemite screen.

noticeable flicker, 35 frames per second for noticeable but satisfactory flicker and 28 frames per second results in disagreeably objectionable flicker. Thus a standard of 24 frames per second cannot be justified. These data also indicate that 48 frames per second will be satisfactory from the standpoint of flicker for values encountered in television.

I.15. Interlaced Scanning.—In motion-picture practice the projected light is broken up at the rate of two times the frame frequency to reduce the flicker effect. In ordinary television

scanning, the object is divided into equal horizontal strips or lines, and these lines are scanned in their regular order: 1, 2, 3, 4, etc. (progressive order). This results in one overall light impulse for each frame repetition in the reproduced image. If this procedure is modified so that the scanning is, for the first half of the frame period, in the order of lines 1, 3, 5, 7, 9, . . . from top to bottom of the frame, and for the second half of the frame period in order of lines 2, 4, 6, 8, . . . from top to bottom of the frame (interlaced scanning), then the flicker effect of the reproduced image is changed. Each frame period now consists of two portions with respect to time, the first of alternate lines and the second of remaining alternate lines, properly staggered to form a complete interlaced pattern. This results in an overall effect of two light impulses for each frame repetition, twice that of the ordinary method of scanning.

Observations made on systems using an interlaced scanning pattern with a television receiver using a kinescope indicate that this method provides a satisfactory solution of the flicker problem with frame frequencies equal to or greater than 24 per second. This, of course, results in a pattern having 48 or more equivalent light impulses. This frequency of light impulses is called *field frequency*.

I.16. Effect of Alternating-current Supply Ripple on Interlaced Scanning.—Since in television systems the power supply is provided by the rectifiers of alternating current, the residual ripples find their way into transmitted and reproduced images from numerous sources. In progressive scanning the adjacent lines are closely related in time, and, therefore, the relative displacement of the adjacent lines is small. Unless the frame frequency is a multiple or a submultiple of the supply frequency, a ripple will move across the image. If the frame frequency is a multiple or a submultiple of the power-supply frequency, such as 30 frames for a 60-cycle source, then the effect of this ripple is stationary on the image and much less pronounced. The moving-ripple pattern is almost as disturbing as the true flicker, and the visual effects of it are almost the same.

With the interlaced-scanning pattern, the adjacent lines are separated in time by one-half of one frame, and in case the power-supply frequency is not a multiple or submultiple of the frame frequency, the scanning lines first draw together and then separate

in correspondence with the time difference between the frame and supply frequencies. The visual effect is a loss of one-half of the total number of lines. The solution is found in choosing a 30-frame interlaced-scanning pattern for a 60-cycle source and a 25-frame pattern for a 50-cycle source. The ripple effect is then stationary, and with reasonably well designed power-supply systems is quite unnoticeable.

I.17. Synchronization.—Synchronization in television is a very important matter, a detailed treatment of which might easily make a volume of the size of this book. It is recommended that the reader study the series of several papers, An Experimental Television System.[1] A description of synchronizing arrange-

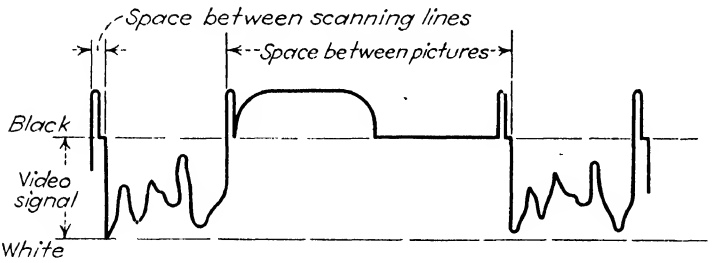


FIG. I.18.—Composite television signal.

ments in a television system utilizing the progressive (not interlaced) scanning is therein given. After the reader becomes familiar with the synchronizing practice in a progressive system, a later paper published by Kell, Bedford and Trainer, Scanning Sequence and Repetition Rate of Television Images,[2] will be of help in understanding the synchronizing practice in an interlaced system. In the present treatise only a brief mention of some general factors pertaining to the synchronization problem will be made.

In order to reproduce a picture received by means of an electrical signal over radio or wires, it is necessary that the scanning luminous spot on the screen of the kinescope be maintained in synchronism with scanning spot on the mosaic of the iconoscope. A variation of synchronism at the receiver to a linear equivalent of one picture element will result in a loss of 50 per cent of the normal resolution. In a modern television system such distortion

[1] *I.R.E. Proc.*, **22**, 1241 (1934).
[2] *I.R.E. Proc.*, **24**, 559 (1936).

is less than one-half of the picture element. Figure I.18 shows a section of the received-signal voltage wave taken over a period of time equivalent to that of four scanning lines. The section is taken at the bottom of the picture in order to include the vertical synchronizing impulse.

It should be noted that the horizontal synchronizing impulses are superimposed upon the "black" signal between the scanning lines, and the vertical impulse is superimposed upon the black signal between the pictures. The fact that the video signal can-

FIG. I.19. Signal after removal of video component.

not go beyond the black amplitude assures that the video signal will not interfere with the synchronizing action. The video signal is then removed from the synchronizing signal by means of a suitable filter. The resultant synchronizing signal is shown in Fig. I.19. Owing to the wide difference between the time duration of the vertical and horizontal synchronizing impulses, either may be removed from the composite synchronizing signal by means of very simple frequency-selective circuits. After the separation, the respective signals have the appearance shown in Fig. I.20. The horizontal and vertical synchronizing impulses

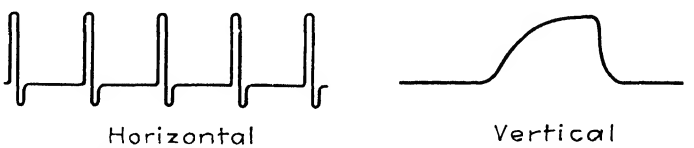


FIG. I.20.—Synchronizing signals after separation.

are then applied, respectively, to the horizontal and vertical deflection oscillators. The subject of deflection oscillators is treated in Part II. The signals shown above are those of the ordinary progressive scanning system. The interlaced system imposes many additional requirements; synchronization of such a system will not be treated here because of its complexity and the fact that it is nonessential for the subject matter of this book.

I.18. Television Receivers.—One paper of the series mentioned above is devoted to the receivers.[1] A large part of this section is

[1] HOLMES, R. S., W. L. CARLSON AND W. A. TOLSON, *I.R.E., Proc.*, **22**, 1266 (1934).

taken from the above paper. A present-day receiver does not greatly differ from those described in the paper mentioned. The video-frequency band is wider at present. It is around

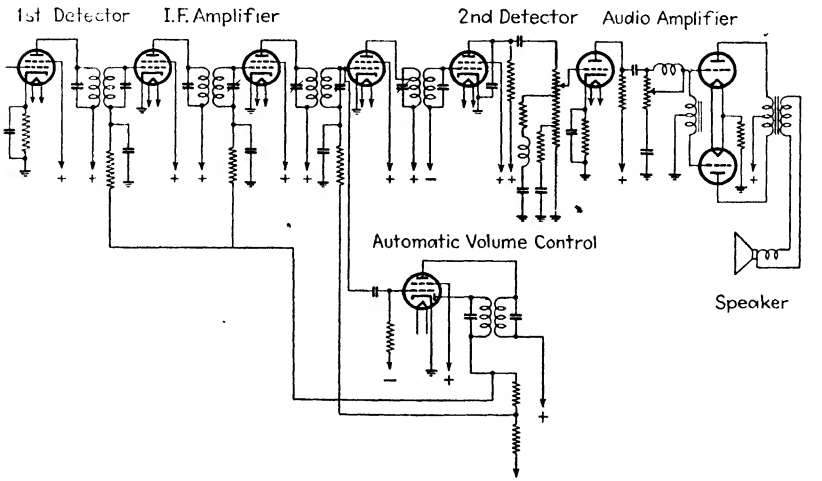


FIG. I.21.—Receiver sound channel.

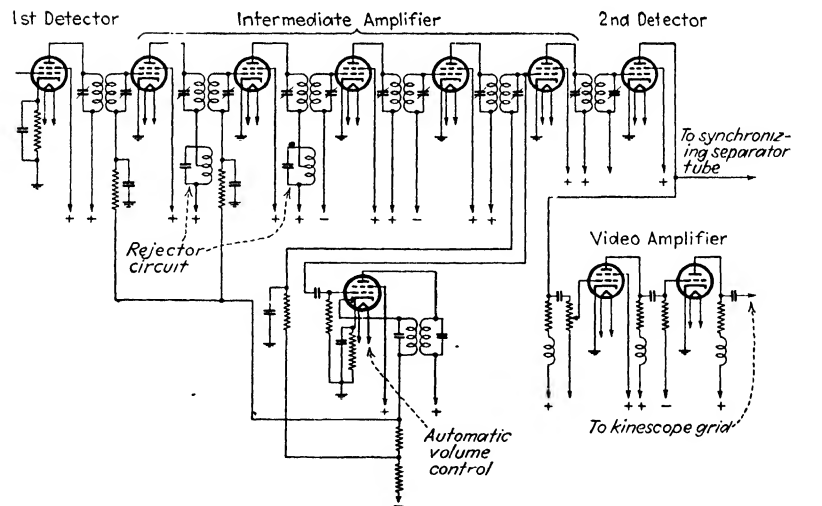


FIG. I.22.—Receiver picture channel.

1,500,000 cycles and even goes to 2,500,000 cycles, as compared with the 1,000,000-cycle-band previously used. The sound and the picture are usually transmitted on two separate carriers

in the frequency band between 40 and 80 megacycles. One form of receiver provides a single radio-frequency tuning system, consisting of two coupled radio-frequency circuits having a sufficient band width to accept both carriers and their side bands simultaneously. A heterodyne oscillator beats the two carriers

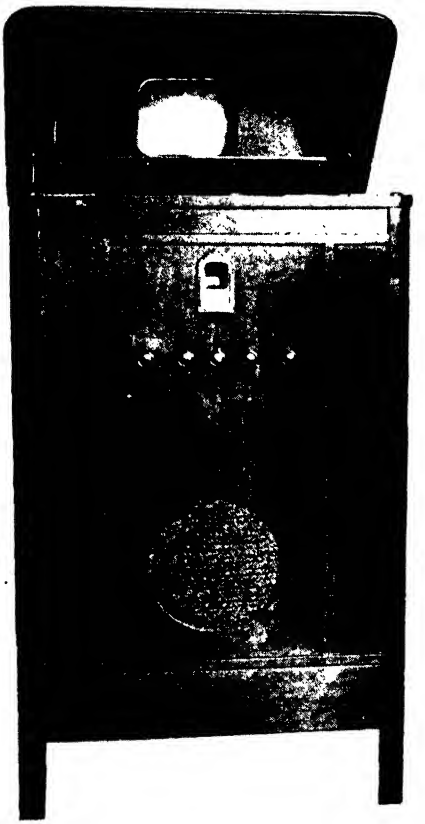


FIG. I.23.—(*a*) Television receiver, front view.

and produces two intermediate frequencies. Two first-detector tubes supply the resulting intermediate frequencies to two separate intermediate-frequency amplifiers. Since the tuning of the sound intermediate amplifier is relatively sharp, it furnishes a sharp reference point for tuning the receiver, and assures that when the sound is tuned in, the picture receiving circuit is also

properly tuned. A schematic diagram of the sound channel of such a receiver is shown in Fig. I.21. A corresponding diagram of the picture-receiving channel is shown in Fig. I.22.

The particular television and sound receiver here described had its intermediate-frequency amplifiers tuned to 6 and 7 megacycles

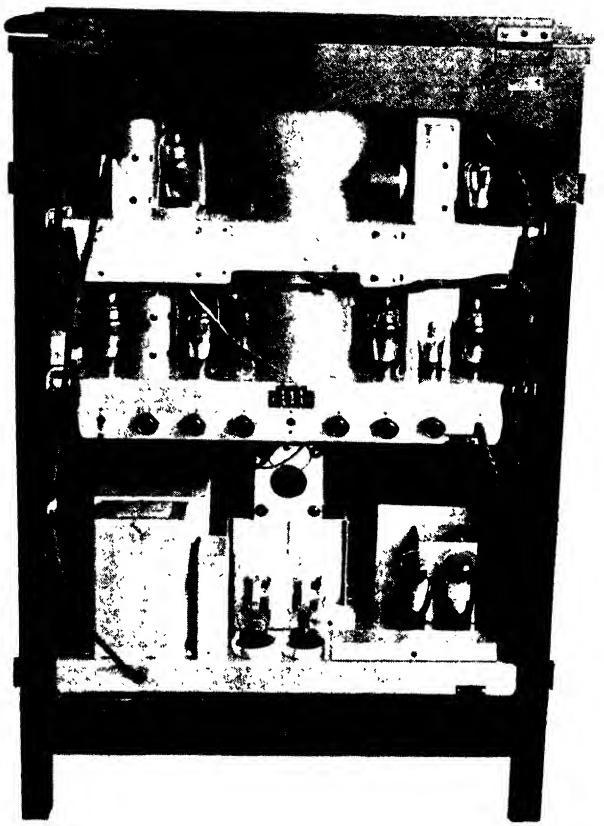


FIG. I.23.—(*b*) Television receiver, rear view.

for sound and picture, respectively. As the schematic diagrams show, both the sound and the video channel are provided with automatic-volume control. These controls maintain the signals on the grids of the two second-detector tubes constant when the signals on the antenna posts of the receiver exceed 100 microvolts.

The video amplifier which follows the second detector of the picture-channel part of the receiver had a flat frequency response

from 30 to 1,000,000 cycles. A present-day receiver has this range extended to at least 1,500,000 cycles.

In Fig. I.23 a complete television receiver is shown.

Deflecting circuits for scanning purposes form a more or less separate subject related to the use of TCR tubes and are treated in detail in Part II of this book.

I.19. Television Transmitters.[1]—In Fig. I.24 a block diagram of a modern television transmitter is shown. The film and studio cameras are nearly identical. A motion-picture projector forms an image on the iconoscope plate of the film camera, while a

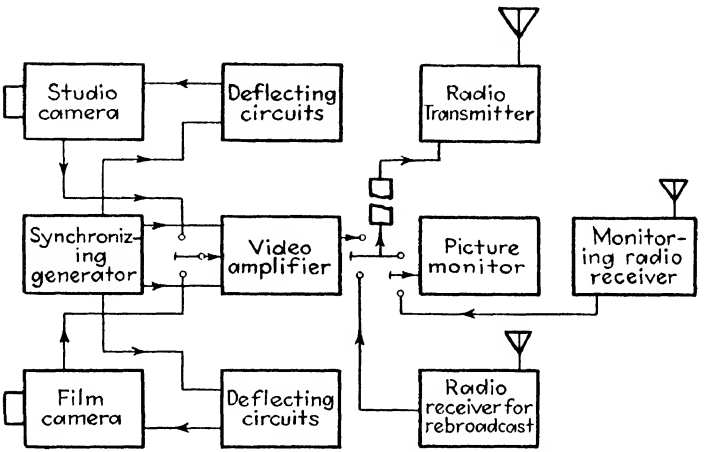


FIG. I.24.—Block diagram of television transmitter.

photographic lens forms an image in the studio camera. Figure I.7 is a photograph showing the general appearance of the studio camera.

Figure 1.25 shows the general arrangement of the parts in the camera. An electrostatic shield (*a*) separates the video-frequency amplifier (*b*) from the rest of the wiring of the iconoscope. The deflecting yoke (*c*) contains windings for both the vertical and horizontal deflections. The amplifier (*b*) has an output impedance sufficiently low for transmitting the signal to the control room over a reasonable length of special cable. At the end of the cable additional video amplifiers are provided. The signal obtained directly from the iconoscope is of the order of 1

[1] KELL, R. D., A. V. BEDFORD AND M. A. TRAINER, An Experimental Television System, Part-II, *I.R.E. Proc.*, **22**, 1246 (1934).

millivolt across 10,000 ohms. With this input, an amplifier having a voltage gain of 2,000 is sufficient to supply a signal at a 2-volt level.

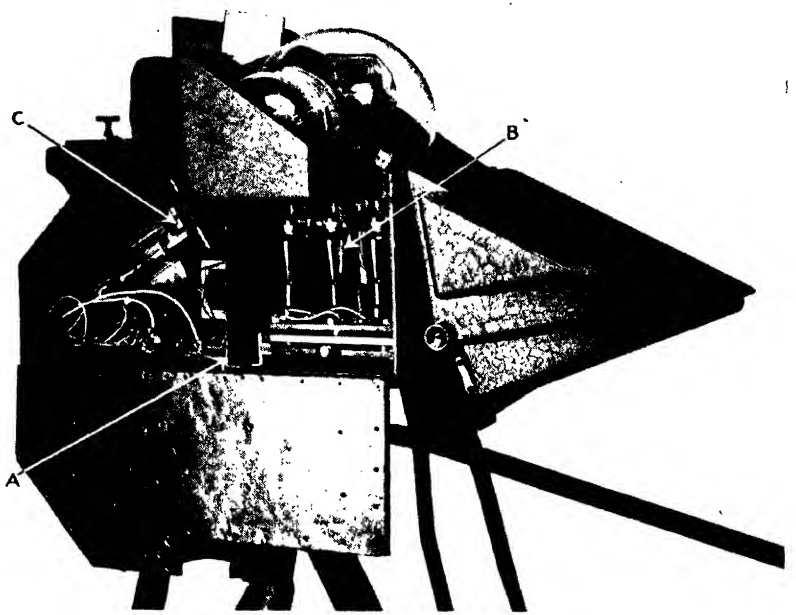


FIG. I.25.—General arrangement of parts in a television camera.

A typical amplifier stage is shown in Fig. I.26. The response of the amplifier at high frequencies is equalized by placing a small inductance L in series with each plate resistor R. The value of the plate inductance for a given circuit is determined by a

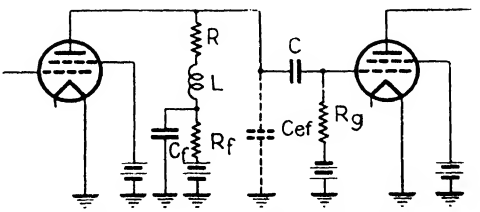


FIG. I.26.—A typical video-amplifier stage.

simple rule which gives a flat amplifier characteristic and a negligible phase distortion. The plate resistor in an amplifier stage should be equal in ohms to the effective reactance of the tube and distributed capacity at the highest frequency which it

is desired to pass. The reactance of the inductance in the plate circuit at this frequency should be equal to one-half the value of the plate resistor. Plate filters are used in order to get flat frequency response at the low end.

Figure I.27 shows a main television signal-control room. These racks contain the circuits for amplifying and mixing picture

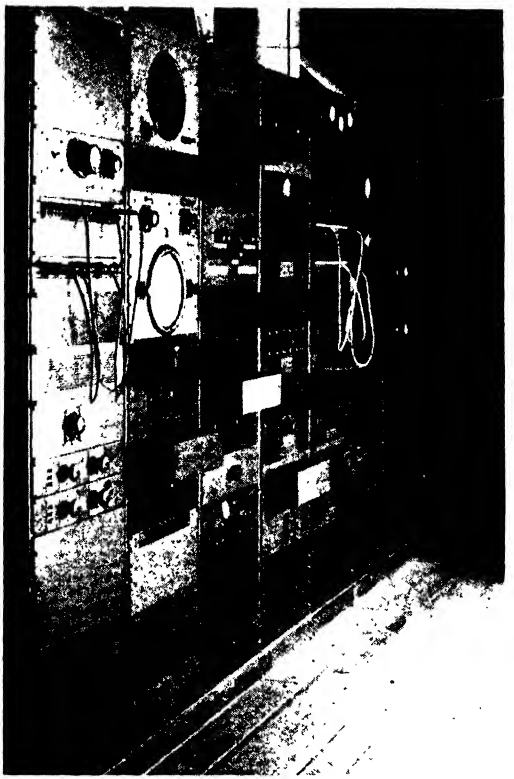


FIG. I.27.—Television signal-control room.

signals and synchronizing impulses. They contain the deflecting and control circuits for both studio and motion-picture cameras, a monitor for checking the operation of the system, all switching facilities to change the input to the radio transmitters from studio to motion picture and, of course, the complete installation for controlling and monitoring of sound signals.

The video-frequency signal generated in this part of the system is at a level of the order of 2 volts across 30 ohms. The

low impedance output is such as to match the impedance of a special cable which carries it to the radio transmitter, which may be a considerable distance away. At the transmitter the picture signal is stepped up in power in order to modulate the power-output stage. Except that the rather broad frequency band of the picture signal requires wide, flat tops in the tuned stages of the radio-frequency amplifiers, the general technique of the picture transmitter is similar to that of any ultra-short-wave transmitter operating with high power at frequencies between 30 and 100 megacycles. The reason for going to such short waves is that the frequency band between the extreme side bands of the modulated video carrier is approximately 4 to 6 megacycles, minimum, and the ultra-short-wave band mentioned appears the most convenient from the standpoint of the number of channels available and the reasonably well developed technique of realization of high-power outputs.

I.20. Image Tube.—Two important applications of electron optics are the image tube and electron microscope, which are discussed in this and the following sections.

The image tube is a device for producing an electron image of any extended source of electrons. The image tube has been used for converting an optical picture projected onto a photosensitive surface into an electron image by focusing the emitted electrons with magnetostatic or electrostatic lenses. An optical image focused upon a photosensitive surface produces an electron emission from each point proportional to the intensity of the light at the point, thus forming an electron object similar to the optical image. The electrons which leave each point of the photosensitive surface as a divergent cone are then focused by the magnetostatic or electrostatic lenses into an electron image.

The electron image may be made visible by placing a fluorescent screen in the plane of the electron image. Loss in intensity due to energy exchange at the photosensitive surface may be compensated by suitable choice of accelerating potential. The picture projected on the photosensitive surface need not be of visible light; it may, for example, be of infrared light.

The following discussion is limited to the electrostatic image tube as developed by Drs. Zworykin,[1] Morton and Ramberg.[2]

[1] ZWORYKIN, V. K., AND G. A. MORTON, *Jour. Optical Soc. Amer.*, **26**, 181–189 (1936).

[2] MORTON, G. A., AND E. G. RAMBERG, *Physics*, **7**, 451–459 (1936).

Figure I.28 is a cross section through the axis of an image tube, and Fig. I.29 is a photograph of an image tube with a 9-in. viewing screen. The electrostatic lens system of this tube consists of a cathode, a series of rings and an anode cylinder, each at a

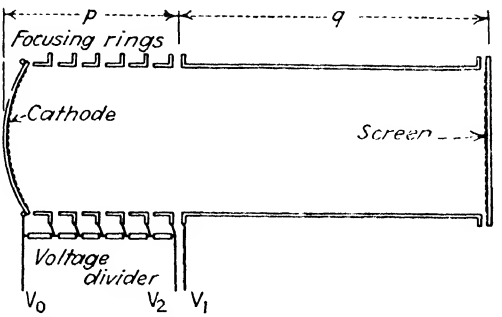


FIG. I.28.—Image tube, fixed magnification.

progressively higher potential. It is worth noting that it is unnecessary to bring out separate leads for each of the rings but that the potential divider supplying the voltage to these rings may be mounted within the tube. The construction of the

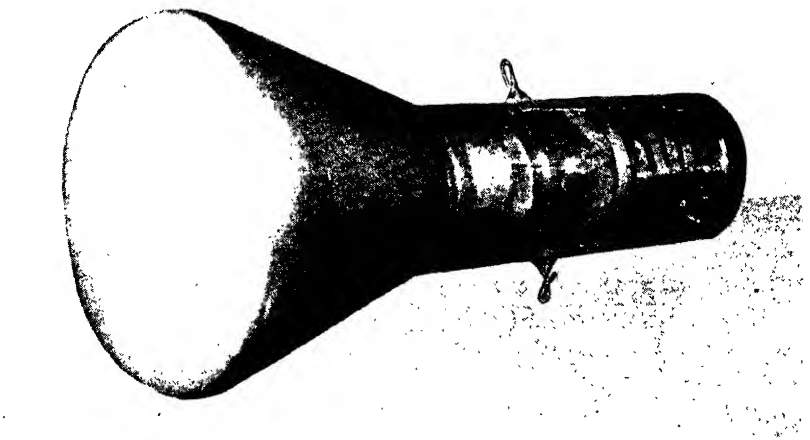


FIG. I.29.—Photograph of a 9-in. image tube.

photosensitive cathode depends upon the spectral region in which maximum sensitivity is required. Cathodes having a photosensitivity of 5 μa. per lumen or better and a spectral response extending well into the infrared may be formed as follows:

A metal layer is sprayed onto the curved glass disk to be used as the cathode. This layer serves to form an inert conducting backing for the photosensitive layer. A layer of silver is then evaporated on the backing metal. The cathode is then mounted and the tube baked and exhausted in the usual way. The activation of the cathode is similar to that used for the ordinary caesium photoelectric cell. The silver is oxidized by an electric discharge in oxygen at a pressure of 1 or 2 mm. Caesium is then admitted from a side tube, and the cathode is baked to promote a reaction between the silver oxide and the caesium.

The cathode is curved in order to correct for the aberrations which are present when a plane cathode is used. It is found

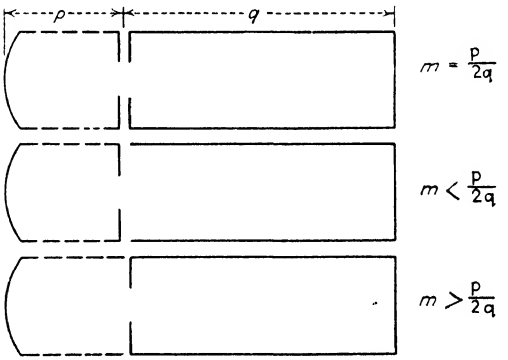


FIG. I.30.—Arrangement of apertures in image tubes.

experimentally that if the cathode is spherical and of radius equal to p (see Fig. I.28), very satisfactory results are obtained. The magnification of a tube such as that shown in Fig. I.28 is $p/2q$. Higher or lower magnifications may be obtained by using apertures as shown in Fig. I.30. The magnification of the uppermost arrangement is also $p/2q$, but the focal length of the lens is shorter for a given voltage ratio V_1/V_2. In practice, this reduces the voltage difference between the anode cylinder and the final focusing ring, which simplifies the insulation problem and decreases the danger of cold-cathode discharge. In the case of the next two arrangements of Fig. I.30, the magnification is less than $p/2q$ if the aperture is on the final focusing ring and greater than $p/2q$ if the aperture is on the anode. With the arrangement shown in Fig. I.31 it is possible to construct a tube which gives a very satisfactory electron image and whose magnification can

be controlled electrically. If the potential V_3 equals the anode voltage, the magnification is greater than $p/2q$, and if V_3 is made equal to the final-ring potential, the magnification is less than $p/2q$.

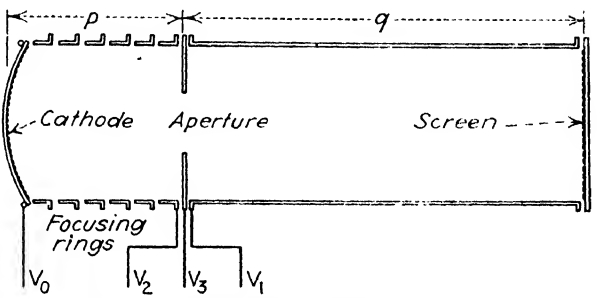


FIG. I.31.—Image tube, variable magnification.

An application of the image tube for television is the incorporation of it into an iconoscope. Such a device is known as an image-multiplier iconoscope[1] and is shown in diagrammatic form in Fig. I.32. The image tube focuses onto a secondary emissive mosaic an electron reproduction of an optical image formed on

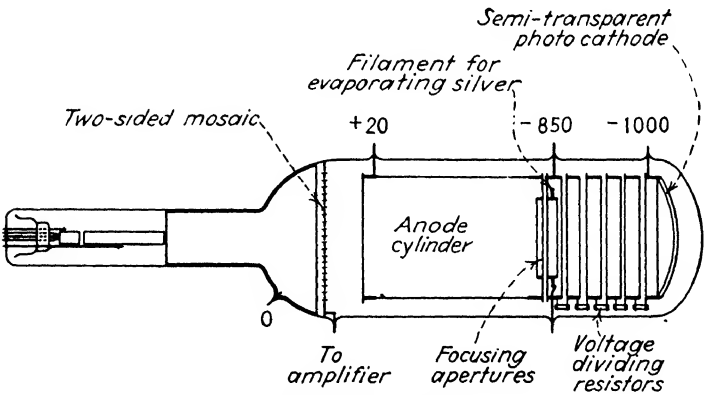


FIG. I.32.—Image-multiplier iconoscope.

the photocathode. This mosaic differs from that in the standard iconoscope in that the mosaic elements extend through the insulating matrix in such a way that the electron image can be formed on one side while the mosaic is scanned on the other.

[1] ZWORYKIN, MORTON AND FLORY, *I.R.E. Proc.*, **25**, 1090–1092 (1937).

The image-multiplier iconoscope, which at present must be considered as a purely experimental device, has a sensitivity about ten times greater than that of a standard iconoscope. Figure I.33 represents a photograph of a laboratory-built image-multiplier iconoscope.

I.21. Electron Microscope.—The electron microscope is a device for producing a highly magnified electron image of any small source of electrons. A fluorescent screen placed in the plane of the electron image permits a visual study of the highly magnified electron source. Primary requirements of a microscope are good resolution and high magnification. Theoretically,

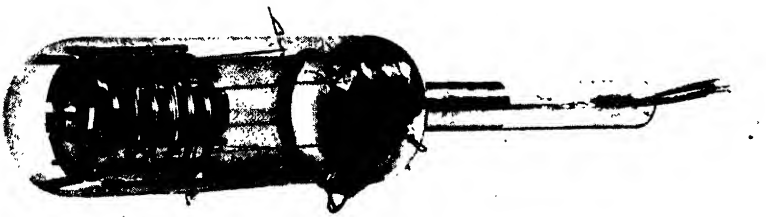


FIG. I.33.—Photograph of image-multiplier iconoscope.

the resolution of the electron microscope is much superior[1] to that of the light microscope. The electron optical problem of the electron microscope is to obtain electrostatic or magnetostatic lenses of short focal lengths capable of producing high magnifications with low image distortions.

Figure I.34 shows a schematic cross section through the axis of the electrode system of an electrostatic electron microscope. Figure I.35 shows a similar cross section of a magnetostatic microscope. The choice between the electrostatic and magnetostatic microscope depends upon many factors, a very important one being the type of source of electrons which it is desired to image. In general, the electrostatic microscope is more useful for studying objects from which the electrons issue with low

[1] The wave length of an electron is given by $\lambda = \sqrt{\frac{150}{V}} \times 10^{-8}$ cm., where V is the potential in volts. Thus, a 150-voit electron has a wave length of one angstrom unit, which is more than 10^{-3} times smaller than the wave length of visible or ultraviolet light.

velocities, such as thermionic cathodes, while the magnetostatic microscope is more useful for studying objects from which the electrons issue with high velocities, such as thin foil through which high-velocity electrons are passing. A factor in favor of the magnetostatic microscope is the fact that the position and power of the lens may be adjusted and experimented with from

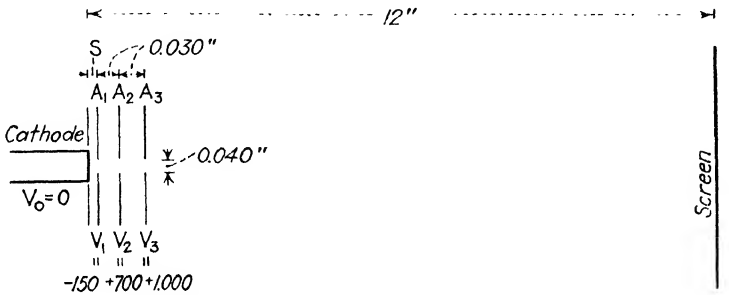


FIG. I.34.—Electrostatic electron microscope.

the outside of the tube, by moving and changing the focusing coil and by varying the current through the coil.

The strength of the magnetostatic lens is increased; *i.e.*, the focal length is decreased by encompassing the coil with an iron mantel but leaving a narrow gap (see coil of Fig. I.35). The increased strength of the lens is accompanied by stronger aberrations, to counteract which the lens aperture (the beam diam-

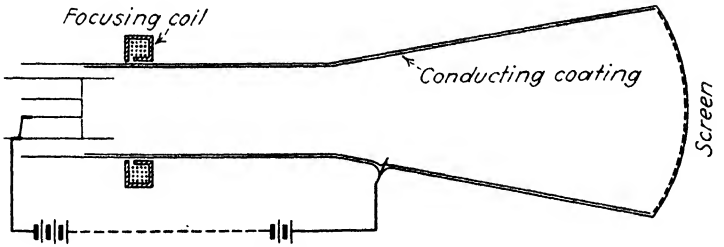


FIG. I.35.—Magnetostatic electron microscope.

eter) is reduced as much as light-intensity considerations on the screen will permit. With the magnifications obtained, the decrease in lens aperture does not decrease the resolution (as it does in light) on account of the very small wave length of the electron. Using a magnetostatic microscope composed of two lenses, Ruska[1] claims to have obtained magnifications as high

[1] RUSKA, E., *Zeits. f. Physik*, **87**, 580–602 (1934).

as 12,000 and a resolution of 5×10^{-6} cm., which is about ten times better than the maximum resolution possible with a light microscope. Ruska's microscope, however, is a rather complicated and expensive device.

The electrostatic microscope shown in Fig. I.34, which is quite simple and satisfactory, has been described by Johannson.[1] Magnifications of the order of 300 may be obtained with such a microscope in a tube with an overall length of about 15 in. The magnification increases as the distance between the cathode and first aperture decreases. For high magnification it is necessary that the spacings be about 0.008 in. or less. With the approximate electrode dimensions and voltages as given in Fig. I.34, this microscope produces a satisfactory magnified image of a cathode area of about 0.010 in. in diameter. The size of the cathode area that is accurately imaged is limited by the aberrations of the lens.

[1] JOHANNSON, H., *Ann. d. Physik*, **21**, 274–284 (1934).

PART I

ELECTRON OPTICS

CHAPTER 1

FUNDAMENTAL CONCEPTS

1.1. Electron.—An understanding of the electron is necessary for the comprehension of the theory of cathode-ray tubes. However, the conception of the electron of modern quantum or wave mechanics is not essential. The classical conception of the electron is sufficient for an engineering understanding of the theory of cathode-ray tubes.

The electron is the smallest bit of negative electric charge yet discovered. The charge of the electron has been determined as

$$\begin{aligned} -e &= -4.77 \times 10^{-10} \text{ e.s.u.}^1 \\ &= -1.59 \times 10^{-20} \text{ e.m.u.} \\ &= -1.59 \times 10^{-19} \text{ coulomb.} \end{aligned}$$

The ratio of charge to mass has been determined as

$$\frac{-e}{m_0} = -1.77 \times 10^7 \text{ e.m.u./g.}$$

so that the mass of the electron is

$$m_0 = 9.0 \times 10^{-28} \text{ g.}$$

If the electron is assumed to be a sphere of radius a, then as shown by classical electrodynamics, the mass of the electron is

$$m_0 = \frac{2e^2}{3ac^2}$$

where $c = 3 \times 10^{10}$ cm./sec. is the velocity of light, and e, the charge on the electron, measured in e.s.u.

This expression for the mass enables one to calculate the value of the radius of the electron. Substituting the values of e, m and c given above, there results for the radius of the electron

$$a = 1.9 \times 10^{-13} \text{ cm.}$$

[1] The latest experiments lead to a higher value for e ($= 4.80 \times 10^{-10}$ e.s.u.).

An interesting point to notice is that if in the above expression for mass, one sets $e = 0$, then the mass of the electron also becomes zero. This indicates that the mass of the electron is entirely electrical in nature. So from the above considerations the electron is to be considered not as a material particle charged with electricity, but simply as a bit of electricity.

According to the theory of relativity, the mass of the electron should vary with velocity according to the equation

$$m = m_0\left[1 - \left(\frac{v}{c}\right)^2\right]^{-\frac{1}{2}}$$

where m is the mass of the electron when moving with velocity v relative to the observer, m_0 is the rest mass of the electron. This equation has been checked experimentally.

For the nonrelativity case the velocity v of an electron, corresponding to an accelerating voltage V is given by the law of conservation of energy as:

$$\frac{1}{2}m_0v^2 = \frac{eV}{300} \tag{1.1}$$

and for the relativity case the equation becomes

$$m_0c^2\left(\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} - 1\right) = \frac{eV}{300}. \tag{1.2}$$

TABLE 1.1.—ELECTRON VELOCITY AS FUNCTION OF ACCELERATING VOLTAGE

Accelerating volts	1	10	50	100	200	400	600	800	1,000	2,000
Nonrelativity velocity in 10^7 cm./sec	5.95	18.8	42.1	59.5	84.1	119	146	168	188	266
Actual velocity in 10^7 cm./sec	5.95	18.8	42.1	59.5	84.1	119	146	168	188	265

Accelerating volts	3,000	4,000	5,000	6,000	7,000	8,000	9,000	10,000	20,000
Nonrelativity velocity in 10^7 cm./sec	326	376	421	461	498	532	564	595	841
Actual velocity in 10^7 cm./sec	324	374	418	457	493	526	557	586	817

Accelerating volts	100,000	1,000,000
Nonrelativity velocity in 10^7 cm./sec	1,880	5,950
Actual velocity in 10^7 cm./sec	1,640	2,820

Table 1.1 gives the electron velocities as functions of accelerating voltages calculated from Eqs. (1.1) and (1.2). It is seen that the change of mass of an electron does not become appreciable unless the electron has fallen through a potential drop of about 10,000 volts or more. As TCR tubes operate at voltages around 10,000 volts or less, it is unnecessary to take into consideration the variation of the mass of the electron with its velocity.

1.2. Proton and Neutron.—The proton is a small positively charged particle. The charge of the proton is

$$e = +4.774 \times 10^{-10} \text{ e.s.u.}$$

The mass of the proton is

$$m = 1.661 \times 10^{-24} \text{ g.}$$

Hence the mass of the proton is 1,840 times that of the electron. The radius of the proton is less than 10^{-13} cm.

The neutron is a small uncharged (neutral) particle having about the same mass and dimensions as the proton.

1.3. Atom and Molecule.—The atom may be considered as a miniature solar system, in which electrons rotate in definite orbits about a positively charged structure called the *nucleus*. The nuclei of atoms consist of closely packed neutrons and protons, the attraction between the neutrons and protons counterbalancing the repulsion between the protons.

The number of protons in the nucleus is called the *atomic number*, usually denoted by Z. The atomic weight of an atom is the weight of the protons and neutrons in its nucleus (the weight of the planetary electrons is relatively negligible), measured by an arbitrary scale based on an atomic weight of 16 for oxygen. The nucleus of oxygen contains eight neutrons and eight protons, and by definition has an atomic weight of 16; the atomic number of oxygen is 8. Hydrogen has one proton in its nucleus and so has the atomic number 1 and an atomic weight of almost 1.

A normal atom of atomic number Z has Z planetary electrons describing various orbits about the nucleus. Thus, helium of atomic weight 4 and atomic number 2 has two neutrons and two protons in its nucleus and two planetary electrons. The atom as a whole is therefore neutral.

It is the behavior of these planetary electrons that determines the chemical and most of the physical properties of the atom. It

is, therefore, the atomic number, and not the atomic weight, which is the important number for the atom. Accordingly, in the more recent periodic tables the atoms are arranged in increasing atomic numbers from unity to 92.

Atoms are of a very open nature as the central nucleus has an effective radius of the order of 10^{-13} cm. The effective diameter of the atom as a whole is of the order of 10^{-8} cm., so that the nucleus is only about 1/100,000 of the dimensions of the atom.

Atoms are able to exert forces on each other which are in some cases attractive and in some cases repulsive. These interatomic forces hold atoms together to form molecules, and they also hold molecules together to form the different states of matter such as gases, liquids and solids.

The molecule consists of one or more atoms. A molecule consisting of but a single atom is called *monatomic;* a molecule composed of two atoms is called *diatomic*, etc. The molecules of most metals and of the inert gases are monatomic while the molecules of the active gases such as oxygen, hydrogen, etc., are diatomic.

1.4. Gas.—A gas is an assemblage of a very large number of molecules moving in all directions and continually colliding with one another and with the walls of the containing vessel. The collisions are perfectly elastic. In gases under low pressures the molecules are so far apart and occupy such a small fraction of the total volume of the containing vessel that they may be considered as independent of one another.

By a series of favorable impacts it is possible that a molecule will acquire a very high velocity. However, it is extremely improbable that a long series of favorable impacts will occur without the occurrence of an impact resulting in a diminution of speed. Similarly, it is extremely improbable that a molecule will come to a standstill as a result of a series of unfavorable impacts. Actually, the velocity of a molecule varies about a mean value, rising only rarely to a very high velocity and falling only rarely to a value near zero.

The velocity of a single molecule varies continually, but when a large number of molecules is considered the fraction of the total number which have velocities between given limits is constant. By applying the laws of probability, Maxwell calculated that the number of molecules having speeds between c and $c + dc$ is

$$dN = N4\pi\left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}} c^2\epsilon^{-\frac{mc^2}{2kT}}dc \tag{1.3}$$

where N is the total number of molecules, T is the absolute temperature and k is a universal constant. This is known as Maxwell's law for the distribution of speeds of the molecules in a gas. There is one speed which is most probable; designating this speed by C_m, the number of molecules in the range of speeds dc at C_m is greater than the number of molecules in the same range

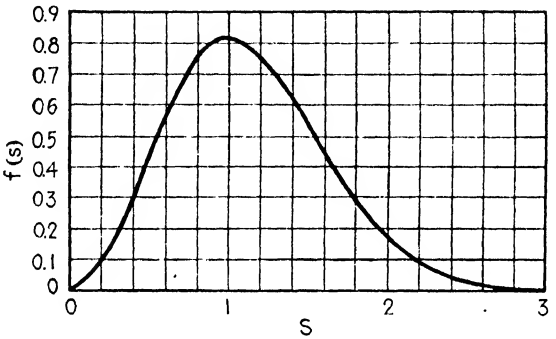


FIG. 1.1.—Maxwell's law for the distribution of speeds.

dc but at any other speed c. It may be deduced from Eq. (1.3) that the most probable speed is

$$C_m = \sqrt{\frac{2kT}{m}} = \sqrt{\frac{2RT}{M}} = 1.29 \times 10^4\sqrt{\frac{T}{M}} \text{ cm./sec.} \tag{1.4}$$

where M is the molecular weight of the gas and R is the gas constant. With the aid of (1.4) Eq. (1.3) may be rewritten as

$$f(s) = \frac{1}{N}\frac{dN}{ds} = \frac{4}{\sqrt{\pi}}s^2\epsilon^{-s^2} \tag{1.5}$$

where s is c/C_m, *i.e.*, s is the speed of a molecule measured to the scale on which the most probable speed is unity. Figure 1.1 shows $f(s)$ plotted against s; $f(s)$ is the fraction of the total number of molecules per unit range at the speed s. The number of molecules in the range ds at s is then $f(s)ds$.

Let G^2 be the mean value of the square of the velocity of a molecule and m its mass, then the mean kinetic energy of a molecule is the same for all gases at a given temperature and is

proportional to the temperature. Thus

$$\tfrac{1}{2}mG^2 = \tfrac{3}{2}kT \tag{1.6}$$

where k is a universal constant known as the *Boltzmann gas constant* ($k = 1.372 \times 10^{-16}$ erg/deg.) and T is the absolute temperature.

The pressure P exerted by a gas on the enclosing walls is

$$P = \tfrac{1}{3}mnG^2 \tag{1.7}$$

where n is the number of molecules per $cm.^3$

Equation (1.6) permits one to calculate the root mean square velocity of the molecules as

$$G = \sqrt{\frac{3kT}{m}} = \sqrt{\frac{3RT}{M}} = 1.58 \times 10^4 \sqrt{\frac{T}{M}} \text{ cm./sec.} \tag{1.8}$$

Thus at room temperature the root mean square velocity of a nitrogen molecule is about 5×10^4 cm./sec.

The average distance traversed by the molecules between successive collisions is known as the *mean free path* of the molecules. The mean free path λ of the molecules is given by the equation

$$\lambda = \frac{1}{\sqrt{2}\pi n d_m^2} \tag{1.9}$$

where d_m is molecular diameter.

TABLE 1.2.—CONSTANTS OF VARIOUS GASES AT ATMOSPHERIC PRESSURE AND 20°C.

Gas	M	m 10^{-24} g.	G 10^2 cm./sec.	λ 10^{-6} cm.	ν 10^5 $cm.^{-1}$	d_m 10^{-8} cm.	Ionizing potential, volts
Argon	39.9	65.24	428	9.9	.10	2.9	15 7
Helium	4.0	6.54	1,358	27.5	.36	1.9	24.5
Hydrogen	2.0	1.65	1,904	17.4	.58	2.4	15.4
Mercury	201.0	656.4	191	14.7	.67	3.0	10.4
Nitrogen	28.0	45.84	511	9.1	1.1	3.2	15.8
Oxygen	32.0	52.34	478	9.9	1.0	3.0	12.5

The average number of collisions per centimeter of path is

$$\nu = \frac{1}{\lambda} = \sqrt{2}\pi n d_m^2. \tag{1.10}$$

Table 1.2 gives the values of M, m, G, λ, ν and d_m of various gases at atmospheric pressure and 20°C.

1.5. Vacuum.—By vacuum is usually meant a gas under low pressure; thus one speaks of a high vacuum when one means a gas at very low pressure.

The centimeter-gram-second (c.g.s.) unit of pressure is the bar or dyne per square centimeter. Other units of pressure used in vacuum work are the pressure exerted by a column of mercury 1 mm. high, and the unit of atmospheric pressure, called *one atmosphere*. The relation between these units and the degree of vacuum of various tubes is shown in the vacuum scale of Table 1.3. This table also gives the boiling points of mercury and water at various pressures.

It is of interest to see the number of molecules per cubic centimeter present in a high vacuum. The number of molecules per cubic centimeter at the absolute temperature T and pressure P (in bars) is

$$n = 7.3 \times 10^{15}\frac{P}{T}. \tag{1.11}$$

The highest vacua in cathode-ray tubes range around 10^{-3} bar. Even at this low pressure the number of molecules per cubic centimeter at 20°C. is about 25,000,000,000.

Inserting the value of n from Eq. (1.11) into Eqs. (1.9) and (1.10), it results that the mean free path is

$$\lambda = 3.1 \times 10^{-17}\frac{T}{d_m^2 P} \tag{1.12}$$

and that the average number of collisions per centimeter of path is

$$\nu = 32.3 \times 10^{15}\frac{d_m^2 P}{T}. \tag{1.13}$$

Table 1.4 gives the mean free path and number of collisions per centimeter for nitrogen at various pressures. It is seen that the mean free path at a pressure of 10^{-3} bar is large, hence the majority of the molecules are traveling in straight lines as far

as the dimensions of the tube will allow, and the number of intermolecular collisions per second is negligible in comparison with the rate at which the molecules bombard the walls of the tube.

The number of molecules of a gas that strike a unit area per second is given by

$$n = 2.65 \times 10^{19} \frac{P}{\sqrt{MT}} \tag{1.14}$$

where P is the pressure in bars, M the molecular weight and T the absolute temperature. Thus for air at room temperature and under atmospheric pressure (10^6 bars) about 3×10^{23} molecules strike one square centimeter per second, while for a vacuum of 10^{-3} bar the number of molecules striking a square centimeter is about 3×10^{14}.

TABLE 1.3.—VACUUM SCALE

	Atm.	Bars or dynes/cm.^{2}	Mercury column, mm.	Boiling points, °C.	
				Hg	H_2O
Atmospheric pressure	1	10^6	7.6×10^2	357	100
Gas-filled lamps	10^{-1}	10^5	7.6×10	243	45
Neon lamps	10^{-2}	10^4	7.6×1	172	7
	10^{-3}	10^3	7.6×10^{-1}	120	−26
	10^{-4}	10^2	7.6×10^{-2}	80	
Gas-focused cathode-ray tubes	10^{-5}	10	7.6×10^{-3}	43	
	10^{-6}	1	7.6×10^{-4}	17	
	10^{-7}	10^{-1}	7.6×10^{-5}	− 7	
Hard cathode-ray tubes	10^{-8}	10^{-2}	7.6×10^{-6}	−28	
	10^{-9}	10^{-3}	7.6×10^{-7}	−45	
	10^{-10}	10^{-4}	7.6×10^{-8}	−58	
Highest attained	10^{-11}	10^{-5}	7.6×10^{-9}	−75	

TABLE 1.4.—MEAN FREE PATH AND NUMBER OF COLLISIONS OF NITROGEN AT 20°C. AND AT VARIOUS PRESSURES

P, in bars	10^6	10^4	10^2	1	10^{-1}	10^{-2}	10^{-3}
λ, in cm.	9.1×10^{-6}	9.1×10^{-4}	9.1×10^{-2}	9.1	9.1×10	9.1×10^2	9.1×10^3
ν, in cm.$^{-1}$	1.1×10^5	1.1×10^3	1.1×10	1.1×10^{-1}	1.1×10^{-2}	1.1×10^{-3}	1.1×10^{-4}

1.6. Collision between Electrons and Molecules.—Consider an electron acting as a projectile colliding with a molecule acting as a target. If the kinetic energy of the electron is in excess of a certain minimum amount it may knock out an electron from the molecule. The process of knocking out one or more electrons from the molecule is known as *ionization.* The remaining positively charged part of the molecule is known as a *positive ion.*

The kinetic energy of electrons is usually measured in electron-volts. An electron-volt, which is equivalent to 1.59×10^{-12} erg or 1.59×10^{-19} watt-sec., is the energy acquired by an electron in falling through a potential difference of one volt. The minimum kinetic energy measured in electron-volts required of an electron for ionization is known as the *minimum ionization potential.* There are also higher ionizing potentials corresponding to the removal of two or more electrons. Table 1.2 shows the minimum ionizing potential for various gases.

The probability that an electron, in passing through a gas, will ionize a molecule increases from zero for electrons having energies less than the minimum ionizing energy up to a maximum for electrons having about twice the minimum energy. For electrons having still greater energy, the probability of ionization decreases.

The mean free path of an electron traveling through a gas is

$$\lambda \text{ electron} = 5.66\lambda \text{ molecule}$$

where λ molecule may be determined by means of Eq. (1.12). In hard cathode-ray tubes, the mean free path of the electron is large in comparison with the dimensions of the tube, so there are very few collisions between electrons and molecules.

CHAPTER 2

ELECTRON EMISSION

2.1. General.—Electron emission occurs under various circumstances, and it has been classified as (1) thermionic, (2) secondary, (3) cold, (4) photoelectric, etc. Of these, thermionic emission is by far the most important in TCR tubes. Next in importance is secondary emission. Photoelectric emission is of great importance in the iconoscope. Cold emission is usually of negligible importance.

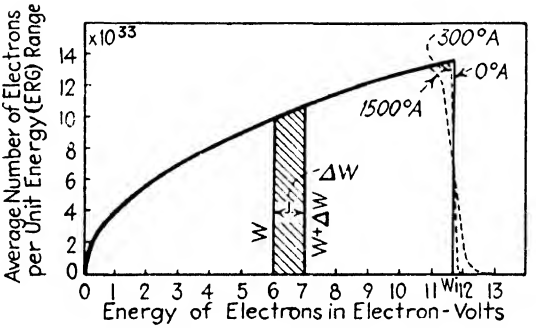


FIG. 2.1.—Distribution of energy of free electrons in nickel at 0°A, 300°A, and 1500°A.

2.2. Electron Theory of Metals.—The mechanism of electron emission follows from the electron theory of metals. According to this theory a metal consists of a three-dimensional lattice of positive ions with electrons wandering about among these relatively fixed ions. Some of these electrons are relatively free, *i.e.*, the attractions of the ions are nearly cancelled by the repulsions of the other electrons. The same electron need not stay free for a long time, and may after some time come under the control of one of the ions. But on the average there is always a fixed number of free electrons. It appears that the number of free electrons ranges from one to two per atom in a metal. It is this very large number of free electrons that is responsible for the large electrical and thermal conductivities of metals.

The kinetic energies of these free electrons in metals are not all the same but are distributed according to a law illustrated in Fig. 2.1. This distribution, known as the *Fermi-Dirac distribution,* is considerably different from the Maxwellian distribution of molecules in a gas. The solid curve of Fig. 2.1 gives the distribution of energy of the free electrons in nickel at 0°A. and the dotted curves at 300°A. and 1500°A. To find the number of electrons having energies between W and $W + \Delta W$ it is necessary to obtain the value of the integral

$$\int_W^{W+\Delta W} F(W)dW$$

which is the area of a strip such as that shown shaded in Fig. 2.1.

It is to be noted that whereas the molecules of a gas have zero kinetic energies at 0°A. (see Eq. 1.1), the free electrons inside a metal have kinetic energies at 0°A. varying from zero up to a maximum Wi. Curves similar to Fig. 2.1 hold for all metals. The values of Wi for various metals range from about 3 to 12 electron-volts.

2.3. Potential Barrier.—In nickel there are about two free electrons per atom so that a cubic centimeter of nickel contains about 18×10^{22} free electrons. The average kinetic energy of a free electron is $\frac{3}{5}Wi$. Hence the total kinetic energy of the free electrons in a cubic centimeter of nickel is

$$E = \tfrac{3}{5} \times 18 \times 10^{22} \times 11.7 = 12.7 \times 10^{23} \text{ electron-volts}$$
$$= 12.7 \times 10^{23} \times 1.59 \times 10^{-12} = 20 \times 10^{11} \text{ ergs.}$$

The pressure that these free electrons exert on the walls of the cubic centimeter of metal is

$$P = \tfrac{2}{3}E = 13.3 \times 10^{11} \text{ bars} = 13 \times 10^{5} \text{ atm.}$$

Under such tremendous pressures the electrons tend to explode out of the metal. However, there are large surface forces that prevent the electrons from leaving the surface. These large surface forces are ascribed to a surface-potential barrier such as that shown in Fig. 2.2. The potential barrier shows that the potential energy of an electron inside the metal is less than that outside the metal, the difference being the work (or kinetic energy converted into potential) that an electron has to do against the electrical surface forces in order to escape.

The potential barrier at the surface of a metal is accounted for by the facts that (1) the atoms near the surface are polarized and so produce a steep potential gradient very close to the surface, and (2) when an electron leaves the metal surface it induces an equal and opposite charge on the surface, attracting it to the surface.

From the above it follows that electron emission will occur if by any means (1) some of the electrons inside the metal are supplied with sufficient energy to enable them to overcome the potential barrier and escape or (2) the potential barrier is lowered, permitting the faster electrons to escape. Thermionic emission, secondary emission and photoelectric emission occur primarily because of (1) and cold emission, primarily because of (2). All types of emission are, of course, increased by lowering the potential barrier.

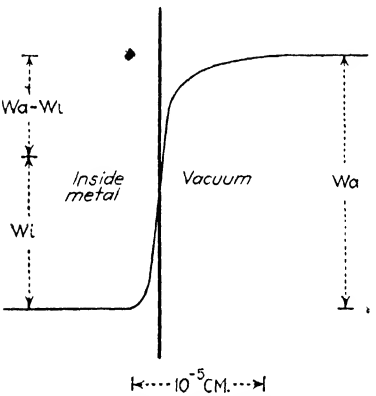


FIG. 2.2.—Potential barrier of a clean metal surface.

2.4. Thermionic Emission.—The magnitude of the potential barrier for nickel is about 16 volts, so that if an electron is to escape from nickel, it must have a component of velocity normal to the surface sufficiently large to overcome a barrier of 16 volts. Figure 2.1 shows that no electrons will escape at 0°A., but as the temperature is increased a few electrons will have sufficiently high energy to escape. The higher the temperature the more electrons will have the energy necessary to escape. The number of electrons that escape from the metal per unit time per unit area at any temperature, T, is

$$i = AT^2\epsilon^{-\frac{Wa-Wi}{kT}} \tag{2.1}$$

where A is a constant and Wa is the height of the potential barrier. The quantity

$$\phi = \frac{300(Wa - Wi)}{e}$$

is known as the *work function* and is the additional energy which must be supplied to an electron inside the metal in order to enable it to escape from the metal. The measured work function for nickel is about 5 volts.

The simple theory of emission just outlined holds only for perfectly clean surfaces. For contaminated, composite or coated surfaces, it is much more complicated but follows the same general lines.

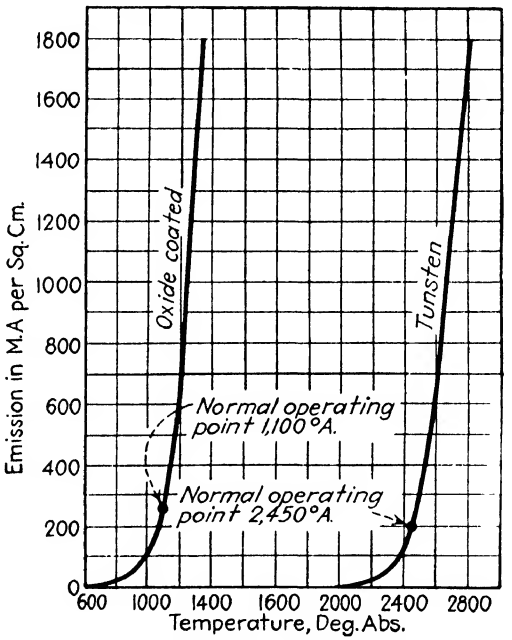


FIG. 2.3.—Variation of emission with absolute temperature.

2.5. Oxide Cathodes.—Most practical thermionic emitters are of the coated- or composite-surface types. The electron emitter or cathode of cathode-ray tubes is of the so-called oxide-coated type. Oxide-coated cathodes are extremely good emitters as compared with pure-metal emitters operated at the same temperature. This is true because of the fact that the work function of an oxide-coated cathode is only about 1 volt while the work function of pure metals is about 5 volts. Figure 2.3 gives the variation of emission with temperature of an oxide-coated and pure-tungsten cathode. These curves were obtained from Eq. (2.1) assuming that for the oxide-coated cathode $A = 10^{-2}$

and $\frac{Wa - Wi}{k} = 12{,}000$, and that for the tungsten cathode $A = 60$ and $\frac{Wa - Wi}{k} = 52{,}400$. As shown in Fig. 2.3, the normal operating temperatures for oxide-coated and pure-tungsten cathodes are about 1100°A. and 2450°A., respectively. The normal operating temperatures are such as to allow a reasonable life. The life of a cathode is greatly reduced as the operating temperature is increased.

Oxide-coated cathodes consist of a core metal upon which is deposited a coating material in the form of alkaline earth oxides, carbonates or nitrates. In TCR tubes the core material is usually nickel, and the coating is a mixture of the carbonates of barium and strontium. The coating is normally deposited as a spray.

A nickel core upon which barium and strontium carbonates have been sprayed will not act as a good emitter. The cathode is still "inactive," and in order to make it a good emitter it is necessary to "activate" the cathode. A cathode is activated by heating it in vacuum to a temperature considerably higher than the normal operating temperature and drawing a large electron current at this temperature. The ease with which a cathode is activated depends upon the cathode temperature, the electron current drawn, the thickness of the coating and the type and amount of residual gas in the tube. Different activation schedules are necessary for cathodes differently prepared.

The activity of oxide-coated cathodes is ascribed to a thin layer of free barium on the surface of the cathode, and the entire object of activation is to produce this thin surface layer of barium. By suitable heat treatment, in activation, the barium and strontium carbonates are broken down to the oxides, and then sufficient barium oxide is reduced to barium to produce the thin surface layer. If the activated cathode be exposed to air, the free barium on the surface is oxidized and the cathode loses its activity; similarly, if an activated cathode is overheated, the free barium evaporates and the cathode again loses its activity. In both of these cases the cathode may be partially or wholly revived by reactivation. After a certain time of life the cathode begins to lose its emission; this is supposedly due to the fact that the free barium has evaporated and most of the BaO has been

used. Figure 2.4 shows the interesting fact that the emission of a BaO-SrO-coated cathode has its maximum activity when the composition of the coating is about 40 per cent BaO and 60 per cent SrO. Hence the activity of a cathode with the composition of 50 per cent BaO and 50 per cent SrO will first improve with life, and then as the quantity of BaO decreases considerably, the activity of the cathode will decrease.

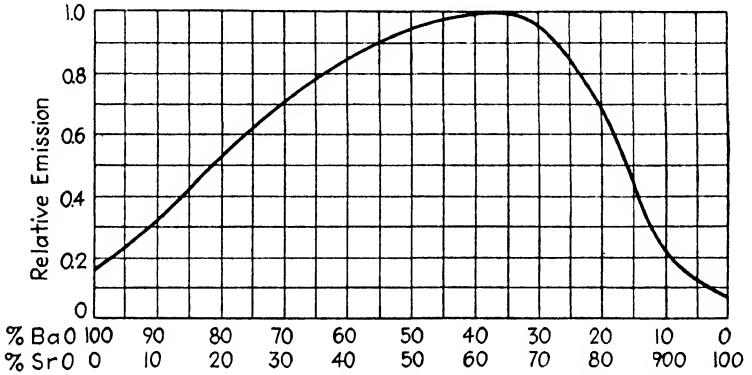


FIG. 2.4.—Variation of emission with composition of BaO-SrO cathode.

2.6. Velocity Distribution of Thermionically Emitted Electrons.—As the velocities of the electrons inside the metal are not all the same and as all electrons are retarded by the same surface-potential barrier, the velocities of the emitted electrons are likewise not all the same. It has been shown theoretically[1] and verified experimentally that the initial velocities of the electrons emitted from a cathode are distributed according to the Maxwellian distribution of velocities. Owing to the form of the cathode and gun, it is most convenient to express the Maxwellian distribution in cylindrical coordinates. Thus, let the axis of z be the axis of the tube, r the radius perpendicular to the axis of z, and ϕ the angle that $\dot{r}(= dr/dt)$—the radial component of the initial velocity of an electron—makes with a fixed plane passing through the axis of the tube. Then the number of electrons emitted per unit area per unit time which have velocity components along the z axis (normal to the cathode) lying between $\dot{z}(= dz/dt)$ and $\dot{z} + d\dot{z}$ is

[1] The Fermi-Dirac distribution reduces to the Maxwellian distribution for the densities encountered in thermionic emission.

$$dN_{\dot{z}} = N\frac{m\dot{z}}{kT}\epsilon^{-\frac{m\dot{z}^2}{2kT}}d\dot{z} \tag{2.2}$$

while the number for which $\dot{r}$ is between $\dot{r}$ and $\dot{r} + d\dot{r}$ and ϕ *simultaneously* between ϕ and $\phi + d\phi$ is

$$dN_{r,\phi} = N\frac{m\dot{r}}{2\pi kT}\epsilon^{-\frac{m\dot{r}^2}{2kT}}d\dot{r}d\phi. \tag{2.3}$$

In this expression, N is the total number of electrons emitted per unit area per second (the saturation current density),

$$k = 1.372 \times 10^{-16} \text{ erg/deg.}$$

is Boltzmann's constant, $m = 9.01 \times 10^{-28}$ g. is the mass of the electron, and T is the temperature on the absolute scale.

The number of electrons for which $\dot{r}$ is between $\dot{r}$ and $\dot{r} + d\dot{r}$ and ϕ simultaneously between $\phi = 0$ and $\phi = 2\pi$ is

$$dN_{\dot{r}} = \int_0^{2\pi} dN_{r\phi} = \frac{Nm\dot{r}}{2\pi kT}\epsilon^{-\frac{m\dot{r}^2}{2kT}}d\dot{r}\int_0^{2\pi} d\phi = \frac{Nm\dot{r}}{kT}\epsilon^{-\frac{m\dot{r}^2}{2kT}}d\dot{r}. \tag{2.4}$$

It is interesting to note that Eq. (2.4) is of exactly the same form as Eq. (2.2). In other words, the number of electrons having velocity components along the z axis lying between, say, 0.05 and 0.1 equivalent volt is equal to the number of electrons having radial-velocity components *for all* ϕ lying between 0.05 and 0.1 equivalent volt.

Since Eqs. (2.2) and (2.4) are similar, it is only necessary to analyze one equation, instead of two, in order to obtain the velocity distribution of the initial electrons.

The analysis of Eq. (2.2) or (2.4) is quite simple. Thus, the number of electrons having velocity components along the z axis and having velocities between $\dot{z}_1$ and $\dot{z}_2$ is

$$\Delta N_{\dot{z}_2-\dot{z}_1} = -N\int_{\dot{z}_1}^{\dot{z}_2}\epsilon^{-\frac{m}{2kT}\dot{z}^2}\left(-\frac{m}{kT}\dot{z}d\dot{z}\right)$$

which becomes after integration

$$\Delta N_{\dot{z}_2-\dot{z}_1} = N\left[\epsilon^{-\frac{m}{2kT}\dot{z}_1^2} - \epsilon^{-\frac{m}{2kT}\dot{z}_2^2}\right]. \tag{2.5}$$

The fraction of the total number of emitted electrons having velocities between $\dot{z}_2$ and $\dot{z}_1$ is then given by

$$\frac{\Delta N_{\dot{z}_2-\dot{z}_1}}{N} = \epsilon^{-\frac{m}{2kT}\dot{z}_1^2} - \epsilon^{-\frac{m}{2kT}\dot{z}_2^2}$$

$$= \epsilon^{-\frac{3.3\times 10^{12}}{T}\dot{z}_1^2} - \epsilon^{-\frac{3.3\times 10^{12}}{T}\dot{z}_2^2}. \qquad (2.6)$$

If it is convenient to express the velocities in equivalent volts, then as

$$\dot{z}^2 = 2\frac{e}{m}\frac{V}{300}$$

$$\frac{m}{2kT}\dot{z}^2 = \frac{eV}{300kT} = \frac{1.16 \times 10^4 V}{T}$$

and

$$\frac{\Delta N_{(V_2-V_1)}}{N} = \epsilon^{-\frac{1.16\times 10^4 V_1}{T}} - \epsilon^{-\frac{1.16}{T}\times 10^4 V_2}. \qquad (2.7)$$

As either Eq. (2.6) or Eq. (2.7) can also be used to calculate the radial distribution of the emitted electrons, these equations give the complete velocity distribution of the initial velocities emitted by a plane cathode at temperature T.

Equation (2.6) was used to calculate the velocity distribution for temperatures of 800°A., 1000°A., 1200°A. and 2400°A. The results are shown in Table 2.1. The last row of Table 2.1 shows that:

99.9 per cent of all the electrons emitted from a plane cathode at	have velocities lower than 0.45 equivalent volt;	800°A.
99.5 per cent	"	1000°A.
98.8 per cent	"	1200°A.
88.8 per cent	"	2400°A.

If the two velocities $\dot{z}_1$ and $\dot{z}_2$ are taken close in magnitude so that $\dot{z}_2 - \dot{z}_1$ is small in comparison with either $\dot{z}_1$ or $\dot{z}_2$, then Eq. (2.6) gives the fraction of the total emitted electrons having the velocity $\frac{\dot{z}_1 + \dot{z}_2}{2}$.

In this way the total number of electrons emitted by a cathode is broken up into many groups of electrons, all the electrons of one group possessing but one velocity. The larger the number of groups the closer is the distribution to the continuous Maxwellian distribution. For many purposes a sufficiently close approximation can be obtained by considering only five or six groups.

TABLE 2.1.—APPROXIMATE VELOCITY DISTRIBUTION OF THERMIONICALLY EMITTED ELECTRONS

$\dot{z}_1$ or $\dot{r}_1$, cm./sec.	$\dot{z}_2$, or $\dot{r}_2$, cm./sec.	$\frac{dN}{N}$				Voltage approximately corresponding to $\frac{\dot{z}_2 + \dot{z}_1}{2}$
		800°A	1000°A	1200°A	2400°A	
0.0×10^7	0.2×10^7	0.016	0.013	0.011	0.005	0.001
0.2	0.4	0.048	0.039	0.027	0.017	0.005
0.4	0.6	0.074	0.060	0.051	0.026	0.010
0.6	0.8	0.094	0.078	0.067	0.036	0.018
0.8	1.0	0.106	0.091	0.079	0.044	0.028
1.0	1.2	0.110	0.097	0.087	0.051	0.041
1.2	1.4	0.107	0.099	0.090	0.056	0.055
1.4	1.6	0.101	0.095	0.088	0.060	0.072
1.6	1.8	0.082	0.087	0.084	0.064	0.091
1.8	2.0	0.070	0.076	0.078	0.063	0.113
2.0	2.2	0.056	0.065	0.069	0.063	0.137
2.2	2.4	0.043	0.052	0.059	0.061	0.163
2.4	2.6	0.032	0.042	0.049	0.058	0.191
2.6	2.8	0.021	0.032	0.040	0.053	0.222
2.8	3.0	0.015	0.025	0.032	0.051	0.237
3.0	3.2	0.010	0.017	0.024	0.045	0.271
3.2	3.4	0.007	0.012	0.019	0.042	0.307
3.4	3.6	0.003	0.008	0.013	0.035	0.345
3.6	3.8	0.002	0.005	0.009	0.031	0.386
3.8	4.0	0.001	0.004	0.007	0.027	0.430
4.0	∞	0.001	0.005	0.012	0.112	>0.450

Thus the initial velocity distribution of electrons emitted from a plane cathode at 1200°A. can be roughly assumed to be that enumerated below:

TABLE 2.2.—AXIAL OR RADIAL VELOCITY DISTRIBUTION AT 1200°A.

Per cent	Cm./sec.
About 4 have velocities of about	0.3×10^7
About 20 have velocities of about	0.8
About 27 have velocities of about	1.4
About 23 have velocities of about	2.0
About 15 have velocities of about	2.7
About 11 have velocities of about	3.0

The distribution is shown more clearly by means of Table 2.3, derived from the approximate distribution just given.

TABLE 2.3.—VELOCITY DISTRIBUTION AT 1200°A.

Radial component cm./sec.	Axial component, cm./sec.					
	0.3×10^7	0.8×10^7	1.4×10^7	2.0×10^7	2.7×10^7	3.0×10^7
0.3×10^7	0.16%	0.80%	1.08%	0.92%	0.60%	0.44%
0.8	0.80	4.00	5.40	4.60	3.00	2.20
1.4	1.08	5.40	7.29	6.21	4.05	2.97
2.0	0.92	4.60	6.21	5.29	3.45	2.53
2.7	0.60	3.00	4.05	3.45	2.25	1.65
3.0	0.44	2.20	2.97	2.53	1.65	1.21

The meaning of Table 2.3 may be made clearer by an example. Thus Table 2.3 shows that 3 per cent of the emitted electrons are emitted with a velocity having an axial component of 2.7×10^7 cm./sec. and a radial component of 0.8×10^7 cm./sec. Otherwise stated, 3 per cent of the electrons are emitted at 16.5° $\left(= \tan^{-1} \frac{0.8 \times 10^7}{2.7 \times 10^7}\right)$ with respect to the axis and with a speed equal to $[(0.8)^2 + (2.7)^2]^{\frac{1}{2}} 10^7 = 2.8 \times 10^7$ cm./sec.

2.7. Secondary Emission.—If an electron strikes a surface with a sufficiently high velocity it may impart to the electrons near the surface sufficient energy to enable them to overcome the surface-potential barrier and escape. The electron striking the surface is known as the *primary electron* and the electrons emitted, as the *secondary electrons*. The number of secondary electrons emitted per primary varies with the velocity of the primary electrons and the size and type of surface-potential barrier. The latter depends upon the chemical nature and physical condition of the surface.

The curves of Fig. 2.5 give the ratio of the number of secondary electrons to primary electrons as a function of the velocity of the primary electrons. The top curve is for a surface composed of caesium on silver oxide, and the bottom curve is for a pure-nickel surface. The secondary emission from a contaminated metal surface is usually greater than for a clean surface. Figure 2.6 gives the typical velocity distribution of the secondary electrons for a given primary velocity. The first hump in the velocity-distribution curve changes but slightly with the velocity of the primaries, and its area represents the total low-velocity secondaries. The

area of the second hump represents the number of reflected primary electrons. It is seen that by far the greater number of secondaries have velocities between 0 and 25 volts.

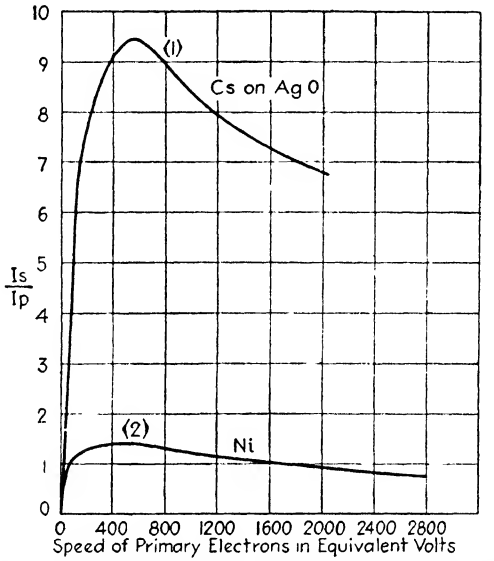


FIG. 2.5.—Secondary emission of Cs on AgO and Ni surfaces.

Secondary emission in cathode-ray tubes is of particular significance in regard to the potential assumed by the luminescent screen and glass walls when bombarded with electrons. As significant experiments on secondary emission of insulators are

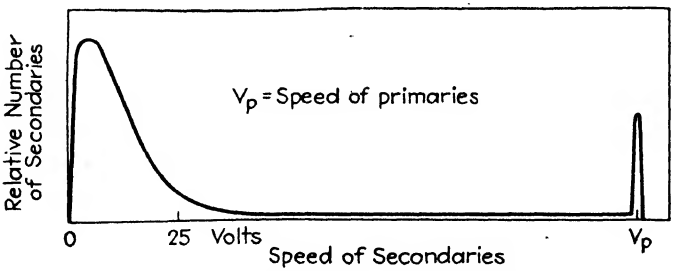


FIG. 2.6.—Speed distribution of secondary electrons.

very difficult to perform, there exists very little quantitative information concerning the secondary emission of glass and luminescent screens. Various portions of a luminescent screen may differ considerably in their secondary-emission character-

istics, and so there exists the tendency for various portions to be charged to different potentials. However, as soon as one portion becomes slightly positively charged with respect to a neighboring portion it will attract the slow secondary electrons, thus preventing any two neighboring portions from ever becoming charged to any great difference in potential. From the fact that most of the secondary electrons are of rather low velocity, it follows that the greatest difference in potential between the neighboring portions of the screen is of the order of a few volts.

An idea as to the manner in which a luminescent screen is charged to a certain potential may be gained by analogy with the charging of an insulated electrode. Thus suppose the luminescent screen is replaced by a pure-nickel plate insulated from the electrodes in the tube and bombarded with 2,000 volt-electrons (2,000 volts on second anode). Just prior to turning on the beam, the potential of the insulated plate will be that of the space (which is practically equipotential) in which the plate finds itself, *i.e.*, 2,000 volts. Then as the beam is turned on, curve 2 of Fig. 2.5 shows that for a primary velocity of 2,000 volts there are only 90 secondaries emitted per 100 primaries, and hence the plate will accumulate electrons. The plate will continue collecting electrons until the potential of the plate has been reduced to 1,700 volts, and when this has occurred the velocity of the electron striking the plate has also been reduced to 1,700 volts. Curve 2 of Fig. 2.5 then shows that for 1,700 volt-electrons the number of secondaries emitted by the nickel plate just equals the number of primaries, and so the metal plate is at equilibrium.

It is to be noted that the equilibrium potential of the nickel plate is 1,700 volts for any primary velocity of 1,700 volts and above. If the primary velocity is between 50 and 1,700 volts, there is the tendency for the plate to charge positively with respect to the second anode, since the number of secondaries is greater than the number of primaries. However, as soon as the plate becomes positive by a few volts, it attracts the slow secondaries, and so in this case the plate potential will practically be the same as that of the second anode. If the primary velocity is between 0 and 50 volts, the number of primaries exceeds the number of secondaries, and so the nickel plate will collect electrons until its potential has been reduced to such a value that no primary electrons will reach the plate. If instead of pure, clean

nickel, a contaminated or impure nickel is used for the target in this experiment, the equilibrium potential of the target may be several times larger than the value (1,700 volts) given.

As the secondary emission characteristics of a luminescent screen are not the same as those of a pure metal, the manner in which a fluorescent screen is charged to a certain potential is much more involved than in the case of the insulated metal plate.

2.8. Cold Emission.—If the voltage between two electrodes in a high-vacuum tube is sufficiently high, electrons are emitted from the negative electrode. This is because of the fact that the field at the electrode surface becomes sufficiently large to appreciably lower the potential barrier and thus allow some of the faster electrons to escape from the metal. The order of magnitude of the field necessary may be surmised from Fig. 2.2, where it is seen that in order to decrease the potential barrier appreciably it is necessary to have at the surface a field of the order of 10^6 volts/cm.

Fowler and Nordheim have shown that the cold-emission current density in amperes per square centimeter is given by

$$I = 6.2 \times 10^{-6} \frac{Wi^{\frac{1}{2}}}{(Wi + \phi)\phi^{\frac{1}{2}}} F^2 \epsilon^{-\frac{2.1 \times 10^7 \phi^{\frac{3}{2}}}{F}}$$

where ϕ is the work function in volts, Wi is in volts and F is the field at the surface in volts per centimeter. The calculation of the field at a surface is difficult, since the strongest fields occur at the sharp irregularities of the surface, and the exact form of these irregularities is seldom known. For perfectly clean surfaces, cold-emission currents are quite small for normal field strengths; thus for a field strength of 10^6 volts/cm., the current is less than 10^{-7} amp. With contaminated surfaces the currents are much larger; thus a field strength of 10^6 volts/cm. may cause currents of the order of 10^{-3} amp. These larger currents are presumably a result of the ionization of the adsorbed gas on the electrodes.

In cathode-ray tubes the edges of all electrodes, and especially those of the first anode, are carefully polished to remove the sharpest surface irregularities and thus reduce cold emission to a negligible quantity.

2.9. Photoelectric Emission.—From the modern point of view light may be considered as consisting of corpuscles, called

photons, each of energy $h\nu$ and momentum $h\nu/c$, where

$$h = 6.55 \times 10^{-27} \text{ erg sec.},$$

ν is the frequency of the light and c ($= 3 \times 10^{10}$ cm./sec.) is the velocity of light. When a surface is illuminated by light, the photons impinging upon the surface may impart to the electrons near the surface sufficient energy to enable them to overcome the surface-potential barrier and escape.

The mechanism of photoelectric emission is that a "free" electron in the metal of kinetic energy W receives from a photon the energy $h\nu$, and if $W + h\nu \geq Wa$ (the surface-potential barrier) then the electron will escape from the surface with the kinetic energy

$$\tfrac{1}{2}mv^2 = h\nu + W - Wa. \tag{2.8}$$

Figure 2.1 shows that at room temperatures the maximum energy of the free electrons is very nearly Wi, and so the maximum kinetic energy of an emitted electron is

$$\frac{1}{2}mv_m^2 = h\nu - (Wa - Wi) = h\nu - \frac{e\phi}{300}. \tag{2.9}$$

Equation (2.9) is known as Einstein's *photoelectric equation.* The frequency ν_0 for which $\frac{1}{2}mv_m^2 = 0$ is known as the *threshold frequency;* it is the lowest frequency that will cause electrons to escape from the metal. By letting $\frac{1}{2}mv_m^2 = 0$ in Eq. (2.9) it follows that

$$h\nu_0 = \frac{e\phi}{300} \tag{2.10}$$

and so Eq. (2.9) may be rewritten as

$$\tfrac{1}{2}mv_m^2 = h\nu - h\nu_0. \tag{2.11}$$

The threshold frequency ν_0 is a very important parameter used in comparing photoelectric properties of surfaces. Instead of using ν_0 it is more convenient to use the threshold wave length λ_0, which is the longest wave length that will cause the emission of electrons from a surface. Table 2.4 gives the threshold wave lengths for a number of substances. As 4000Å. is the shortest wave length that is visible, many substances such as nickel, copper, and iron will not emit electrons when illuminated by visible radiation.

TABLE 2.4.—THRESHOLD WAVE LENGTH OF PARTIALLY OUTGASSED SURFACES

Material	Nickel	Copper	Iron	Calcium	Barium	Lithium	Sodium
λ_0, in Å.	2700	2800	3000	4500	6000	5000	5500

Material	Potassium	Caesium	Cs on CsO on Ag
λ_0, in Å.	6000	7200	10,000

Of great practical significance is the photoelectric sensitivity of a surface, which is defined as the photoelectric current in microamperes drawn from a surface illuminated by one lumen of white light. Thus the photoelectric sensitivity of a nickel

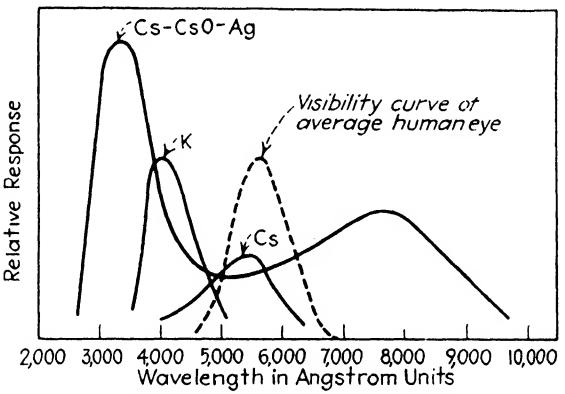


FIG. 2.7.—Spectral-distribution curves for Cs-CsO-Ag, K and Cs surfaces.

surface is zero, and the sensitivity of a potassium surface (the most sensitive of all metal surfaces) is about one microampere per lumen. The composite surfaces are among the most sensitive known; thus a sensitivity as high as 40 microamperes per lumen has been noted for a Cs on CsO on Ag surface. The photosensitive mosaic of the iconoscope has a photoelectric sensitivity of about 5 or 10 μa per lumen. The photoelectric sensitivity of a surface is deducible from its spectral-distribution curve; this gives the frequency range over which the surface is photoelectrically active. Figure 2.7 gives the spectral-distribution curves of several substances and also the visibility curve of the average human eye.

CHAPTER 3

ANALOGY BETWEEN ELECTRON OPTICS AND LIGHT

3.1. Geometrical Optics.—Electron optics has nothing to do with optics in the ordinary sense of the word[1] but is the name given to the subject dealing with the paths of electrons in electrostatic and magnetostatic fields. This name has been chosen because of the analogy existing between the track of an electron moving through electrostatic and magnetostatic fields and the track of a ray of light in passing through refracting media.

That there exists an analogy between the path of a ray of light passing through refracting media and the path of a particle, say an electron, passing through conservative fields of force has been known for a long time. Sir William Hamilton formulated this analogy in great detail in the early part of the nineteenth century.

The subject of optics (light) is usually divided into two parts; *viz.*, geometrical and physical optics. Geometrical optics treats only of the geometrical relations of the propagation of light, while physical optics deals with other than purely geometrical properties. A similar division may be made of the subject of electron optics. In cathode-ray tubes it is only necessary to deal with geometrical electron optics.

The following laws form the customary starting point for geometrical optics:

1. Rectilinear propagation of light.
2. The law of refraction.
3. The law of reflection.
4. The independence of the different rays of a beam of light.

From these laws the entire subject of geometric optics may be deduced.

3.2. Analogy in Particular Case.—In order to visualize the analogy between electron optics and light, it will now be shown

[1] Ordinarily, optics pertains to the eye, vision and light, and electron optics is suggestive that somehow electrons are directly connected with vision; this suggestiveness is especially apt to cause misunderstanding when electron optics is applied to cathode-ray tubes because of the light emitted by the fluorescent screen when struck with electrons.

that in a particular case the first three laws of geometrical optics hold in electron optics.

Figure 3.1 shows a thin plane surface separating two regions of different electrostatic potentials. Let the potential above the surface be V_1 and that below the surface, V_2. An electron traveling either above or below the surface will travel in a straight line since it will find itself in an equipotential region where no forces will be acting on it. At the surface, however, there is a very rapid change in potential, and an electron at the surface will have a

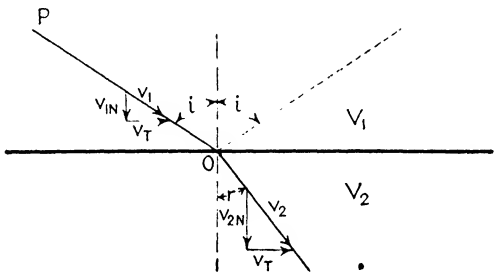


FIG. 3.1.—Electron refraction and reflection at plane surface.

very strong force acting on it in the direction of the normal to the surface.

Now consider an electron moving in the upper region in the direction PO with the speed v_1. When it arrives at the surface a force normal to the surface will act on it. After it passes through the surface its speed will have changed to, say, v_2. The force being normal to the surface, only the component of initial velocity[1] v_N normal to the surface will change; the tangential component v_T of velocity will be the same on both sides of the surface. From Fig. 3.1 there results

$$v_T = v_2 \sin r = v_1 \sin i$$

where i and r are the angles of incidence and refraction, respectively. Hence

$$\frac{\sin i}{\sin r} = \frac{v_2}{v_1} = \mu \text{ (a constant)}. \tag{3.1}$$

This formula expresses the well-known law of refraction, that the ratio of the sine of the angle of incidence to that of refraction is a constant.

[1] It should be carefully noted that velocity is a vector quantity while speed is a scalar representing the absolute value of the velocity.

Equation (3.1) may be expressed in a different form. The work done by the field on the electron when it goes from the first into the second medium is $e(V_2 - V_1)$. Then from the law of conservation of energy it follows that

$$\tfrac{1}{2}mv_2^2 = \tfrac{1}{2}mv_1^2 + e(V_2 - V_1)$$

and so

$$\mu = \frac{v_2}{v_1} = \sqrt{1 + \frac{(V_2 - V_1)e}{\tfrac{1}{2}mv_1^2}}. \tag{3.2}$$

If, further, the incident electron velocity v_1 is that corresponding to the voltage V_1 then $\frac{1}{2}mv_1^2 = eV_1$ and

$$\mu = \frac{v_2}{v_1} = \sqrt{1 + \frac{(V_2 - V_1)}{V_1}} = \sqrt{\frac{V_2}{V_1}} \tag{3.3}$$

If $V_2 < V_1$ so that $V_2 - V_1$ is negative, and if in absolute magnitude it is larger than $\frac{1}{2}m(v_1 \cos i)^2$—the part of the kinetic energy of the electron corresponding to the normal component of its velocity—then the electron will be shot back from the surface with its normal velocity component reversed. The direction of the reflected electron OP' makes with the normal to the surface the same angle i as that of the incident ray.

It has thus been shown that in the special case considered the first three laws of geometric optics hold in electron optics. The fourth law of geometric optics does not hold *strictly* in electron optics, since the path of one electron *is* influenced by the presence of the other electrons in the beam. This is, of course, the well known effect of space charge. However, for low beam intensities the effect of space charge is negligible. If the effect of space charge is not small it can be taken into consideration, as will be shown later.

3.3. Analogy in General Electrostatic Case.—The analogy in the general electrostatic case may be seen from the principles of least time for geometrical optics and of least action for electron optics. The principle of least time states that the path of a ray of light from point A to point B is always such as to make the integral

$$\int_A^B \mu(\nu,x,y,z)ds$$

μ = index of refraction
ν = frequency of light
ds = element of path

an extreme (usually a minimum) with respect to all neighboring paths for rays of the same frequency. The principle is usually stated as

$$\delta\int_A^B \mu(\nu,x,y,z)ds = 0. \qquad (\nu = \text{constant}) \tag{3.4}$$

The principle of least action for electron velocities less than one-tenth the velocity of light states that an electron of total energy E, kinetic energy T and mass m moves through an electrostatic field with potential energy $V(x,y,z)$ in such a way as to make the action integral

$$S = \int_A^B 2T dt = \int_A^B [2m(E - V)]^{\frac{1}{2}} ds$$

over the actual path between the two points A and B an extreme as compared with its value for all adjacent paths for the same value of E. As the integrand $[2m(E - V)]^{\frac{1}{2}}$ is identical with the *absolute* value of the momentum p which the electron would assume at (x,y,z), the principle may be stated as

$$\delta\int_A^B p(E,x,y,z)ds = 0. \qquad (E = \text{constant}) \tag{3.5}$$

A comparison of Eqs. (3.4) and (3.5) shows that the path of an electron in an electrostatic field may be identified with the rays of light in geometrical optics if the index of refraction is chosen to be

$$\mu = k'(E - V)^{\frac{1}{2}} = k''p = kv \tag{3.6}$$

where k is a constant of proportionality and v is the speed of the electron. So the index of refraction at any point of an electrostatic field is proportional to the speed of the electron at the point. If, as is customarily done, the index of refraction is taken as a pure numeric, then k must have the dimension $1/v$. The value assigned to k is of no importance since only the ratio of μ at two different places is used, so that if μ_1, and μ_2 are the indexes of refraction at two different places, the relative index of refraction is

$$\mu = \frac{\mu_2}{\mu_1} = \frac{kv_2}{kv_1} = \frac{v_2}{v_1}. \tag{3.7}$$

Since the potential function $V(x,y,z)$ is a continuous scalar function of position, it follows from Eq. (3.6) that the index of refraction of an electrostatic field is also a continuous function

of position. Optically speaking this means that an electrostatic field constitutes an isotropic, nonhomogeneous medium for electrons (corresponding to a medium of continuously variable density for light rays).

3.4. Analogy in General Case.—In the general case, electrostatic and magnetostatic fields may be present. The analogy between electron optics and light in this case may be shown with the aid of the Lagrangian function. The Lagrangian[1] function for an electron moving through superimposed electrostatic and magnetostatic fields with a speed less than about one-tenth that of light is

$$L = \frac{1}{2}mv^2 - eV - \frac{e}{c}(\overline{vA}) \tag{3.8}$$

where $(\overline{vA})$ stands for the scalar product of the vectors $\bar{v}$ and $\bar{A}$, $\bar{v}$ being the velocity of the electron and $\bar{A}$ the vector potential of the magnetic field defined by the relation $\bar{H} = \text{curl } \bar{A}$, $\bar{H}$ being the intensity of the magnetostatic field.

The momentum of the electron is determined from Eq. (3.8) as

$$p = \frac{\partial L}{\partial v} = \left[mv - \frac{e}{vc}(\overline{vA}) \right]$$

and the principle of least action becomes (see Eq. (3.5))

$$\delta \int_A^B \left[mv - \frac{e}{vc}(\overline{vA}) \right] ds = 0. \tag{3.9}$$

A comparison of Eq. (3.9) with (3.4) brings out the fact that the path of an electron in electrostatic and magnetostatic fields is analogous to the tracks of light rays in geometrical optics, if the index of refraction of the electrostatic and magnetostatic fields be identified with

$$\mu = k\left[v - \frac{e}{mcv}(\overline{vA}) \right]. \tag{3.10}$$

It is to be noted that the first term of Eq. (3.10) is identical with Eq. (3.6) and hence represents the contribution of the electrostatic field to the index of refraction. The second term gives the contribution of the magnetostatic field to the index of refraction. It is seen that the index of refraction of a magnetostatic field is a function not only of position but also of direction.

[1] BORN, MAX, The Mechanics of the Atom, p. 208.

Hence a magnetostatic field constitutes a nonhomogeneous anisotropic medium for electrons (corresponding to a variable-density crystalline medium for light).

3.5. Axially Symmetric Focusing Systems.—So it is to be expected that certain forms of electrostatic and magnetostatic fields will act as focusing systems or "lenses" for electron beams, just as certain forms of refracting media act as focusing systems for light beams. The forms of fields required will depend upon the type of focusing.

In cathode-ray tubes one is interested in electron-focusing systems having axial symmetry. Most optical systems for light consist of a series of spherical refracting surfaces having a common

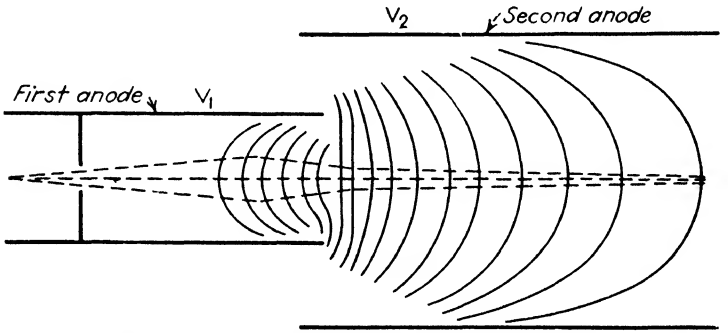


FIG. 3.2.—Equipotential line plot of two cylinders.

axis of symmetry called the *optic axis*. In the case of light, however, the optical systems are usually such that the index of refraction changes abruptly as light passes from one to the other medium. In the case of electron optics, as is shown above, the index of refraction is a continuous function of position.

Figure 3.2 represents a cross section through the axis of an electron-focusing system; the heavy lines represent two cylindrical, metallic electrodes at the potentials V_1 and V_2, and the light lines represent the equipotential surfaces in the space (vacuum) between the electrodes. From Eq. (3.6) it follows that each equipotential surface represents a surface of constant index of refraction. In Fig. 3.2 there are shown only a few of the equipotential surfaces; actually there are, of course, an infinite series of equipotential surfaces having a common axis. The electron-focusing system of Fig. 3.2 may therefore be considered as a very large number of coaxial refracting surfaces.

CHAPTER 4

MOTION OF ELECTRONS IN AXIALLY SYMMETRIC ELECTROSTATIC FIELDS

4.1. Vector and Scalar Fields.—An electrostatic field is a region wherein a charge, placed in the field, would experience an electric force. An electrostatic field may be described either by a vector field or by a scalar field. A vector field is a region to every point of which there corresponds a vector. Thus to every point of an electrostatic field one may assign the force (a vector) that would act on a unit positive charge placed at the point. A scalar field is a region to every point of which there corresponds a scalar (a number). Thus, to every point of an electrostatic field there corresponds the electrostatic potential (a scalar) at the point.

The scalar field does not involve directions and is very much simpler of representation on a diagram than the vector field. Thus, although the representation of an electrostatic field by a vector or a scalar field is equivalent in the sense that from either representation the other may be derived, the scalar (potential) field method of describing an electrostatic field is by far the simpler and is the one usually used. The potential function is diagrammatically represented by giving the surfaces on every point of which the potential has the same value, *i.e.*, by the equipotential surfaces.

4.2. Electrostatic Potential.—The electrostatic potential at a point in space is customarily defined as the work done on a unit *positive* charge in bringing it by any path from infinity to that point. The potential is a relative quantity, *i.e.*, the potential of a given point has meaning only when referred to the potential of some other point (line, surface or volume). Thus in the above definition infinity is taken as the reference point.

In the case of a system of conductors it is customary to choose one of the conductors as the reference point for potential. In the case of the cathode-ray tube the cathode is usually taken as the reference point and is said to be at zero potential so that

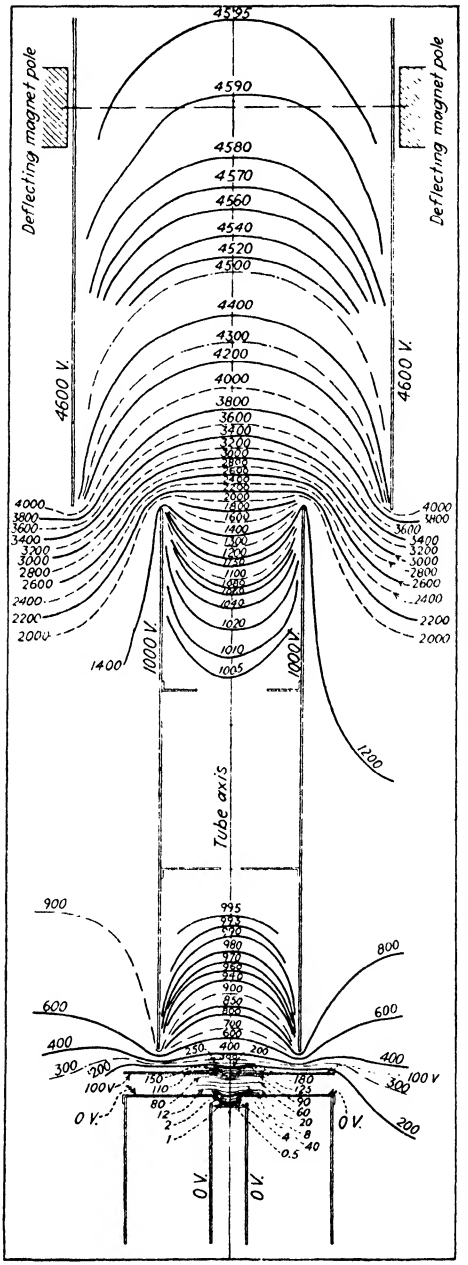


FIG. 4.1.—Electrode system and equipotential lines of a TCR tube.

the potential of all other electrodes and of the space between the electrodes is the difference in potential between them and the cathode. The double lines of Fig. 4.1 represent a cross section through the axis of the system of conductors composing the electron gun of a cathode-ray tube.

On Fig. 4.1 is also shown the cross section through the axis of the equipotential surfaces in the space between the electrodes. The meaning of an equipotential surface is that the work performed against the electrostatic field in bringing a unit positive charge by any path from any point on the cathode to any point on the equipotential surface is the same, no work being required to bring a charge *by any path* from one point on the equipotential surface to any other point on the same surface. Thus, $900/300 = 3$ ergs is required to bring a unit positive charge from the cathode to the 900-volt equipotential surface, and no work is required to take a charge from one point to any other point on the same surface.

The potential at a given point may also be considered as the potential energy which a unit positive charge would possess at the point; so an equipotential surface is the surface over which the potential energy of a charge is constant. Thus, a unit positive charge possesses the maximum potential energy at the highest equipotential surface.

In the cathode-ray tube, one deals with electrons (negative charges) rather than with positive charges. In terms of electrons, the potential at a given point is the work performed *by the field* in bringing a unit negative charge (2.095×10^9 electrons) from the cathode to the point. In dealing with electrons, it has become customary to use as a unit of energy the electron-volt ($= 4.77/300 \times 10^{-10} = 1.59 \times 10^{-12}$ erg). It is the work performed by the field in bringing an electron from one equipotential surface to another equipotential surface one volt higher. It is seen that the electron has its maximum potential energy at the cathode.

4.3. Relation between Potential and Force.—A complete knowledge of the value of the potential everywhere, or diagrammatically of the equipotential surfaces, gives one a complete knowledge of the electrostatic forces everywhere in the field. Thus, let $V(r,z)$ represent the potential function of an axially symmetric electrostatic field such as that shown in Fig. 4.1.

Then each equipotential surface may be represented by the equation

$$V(r,z) = C.$$

The magnitude and direction of the force that would act on a unit positive charge placed at any point in the field are

$$F = -\frac{dV}{dn}$$

where dV/dn is the derivative, at the point in question, of the potential function taken along the direction of the normal to the equipotential surface.

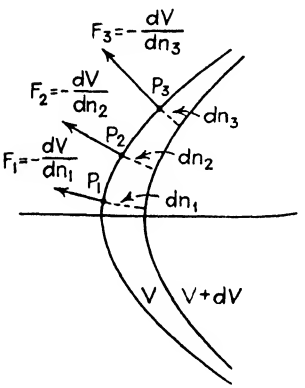


FIG. 4.2.—Relation between potential and force.

Figure 4.2 shows two adjacent equipotential surfaces having a difference in potential of dV. The force on a positive charge $-dV/dn$ at the points P_1, P_2 and P_3 is indicated by the vectors F_1, F_2, and F_3, respectively. The force is $-dV/dn$ instead of dV/dn as the positive direction of the normal to an equipotential is taken in the direction of increasing potential, while the force on a positive charge is along the normal in the direction of decreasing potential.

The r and z components of the vector dV/dn are $\partial V/\partial r$ and $\partial V/\partial z$, respectively, so that the components of force on a positive charge of magnitude q are given by the equations

$$F_r = -q\frac{\partial V}{\partial r}$$

$$F_z = -q\frac{\partial V}{\partial z}$$

and the components of the force on an electron ($q = -e$)

$$F_r = e\frac{\partial V}{\partial r}$$

$$F_z = e\frac{\partial V}{\partial z} \qquad (4.1)$$

4.4. Axial Distribution of Potential.—One of the major electrostatic problems of the cathode-ray tube is to determine the

potential distribution in space (as given by the equipotential surfaces of Fig. 4.1) for given voltages on the electrodes. This problem is solved if one can solve the reduced Laplace equation

$$\frac{\partial^2 V}{\partial r^2} + \frac{1}{r}\frac{\partial V}{\partial r} + \frac{\partial^2 V}{\partial z^2} = 0 \tag{4.2}$$

subject to the boundary conditions that V assume the given values of potential on the electrodes. In general, it has not been found possible to obtain a simple analytical solution of Eq. (4.2) subject to the actually existing boundary conditions. However, the required solution of Eq. (4.2) subject to the existing boundary conditions is easily obtained experimentally.[1] The equipotential line plot shown in Fig. 4.1 was thus obtained.

Of great significance is the fact that the potential distribution in space is uniquely determined if the distribution of potential along the axis, together with its even derivatives, is known. This may be shown as follows: Let $V(r,z)$ be the required solution of Eq. (4.2). Then owing to the axial symmetry one can expand $V(r,z)$ into an infinite series containing only even powers of r, *i.e.*,

$$V(r,z) = V_0(z) + r^2 V_2(z) + r^4 V_4(z) + \cdots + r^{2n} V_{2n}(z) + \cdots \tag{4.3}$$

Differentiating (4.3), substituting into (4.2) and equating the coefficients of equal powers of r to zero, there results

$$V(r,z) = V_0(z) - \frac{r^2}{2^2} V_0''(z) + \frac{r^4}{2^2 \cdot 4^2} V_0^{(4)}(z) + \cdots + \frac{(-1)^n r^{2n}}{2^2 \cdot 4^2 \cdots (2n)^2} V_0^{(2n)}(z) + \cdots \tag{4.4}$$

where the primes and the superscripts in parentheses denote differentiation with respect to z.

By setting $r = 0$ in Eq. (4.4) note that $V(0,z) = V_0(z)$, *i.e.*, $V_0(z)$ represents the distribution of potential along the axis. If the function $V_0(z)$, together with all its even derivatives, is known, then the potential distribution off the axis can be found by means of Eq. (4.4).

[1] For a method of experimentally determining the potential distribution, see E. D. McArthur, *Electronics*, June, 1932, p. 192.

The r component of the force on unit negative charge becomes, from Eq. (4.4),

$$\frac{\partial V(r,z)}{\partial r} = -\frac{r}{2}V_0''(z) + \frac{r^3}{2^2 \cdot 4}V_0^{(4)}(z) - \cdots + \frac{(-1)^n r^{2n-1}}{2^2 \cdot 4^2 \cdots (2n)}V_0^{(2n)}(z) + \cdots \quad (4.5)$$

Similarly, the z component of the force is

$$\frac{\partial V(r,z)}{\partial z} = V_0'(z) - \frac{r^2}{2^2}V_0^{(3)}(z) + \cdots + \frac{(-1)^n r^{2n}}{2^2 \cdot 4^2 \cdots (2n)^2}V_0^{(2n+1)}(z) + \cdots \quad (4.6)$$

4.5. Equations of Motion of Electron.—The equations of motion of an electron moving in a meridian plane of an axially symmetric electrostatic field are

$$m\frac{d^2z}{dt^2} = e\frac{\partial V}{\partial z}$$
$$m\frac{d^2r}{dt^2} = e\frac{\partial V}{\partial r}. \quad (4.7)$$

In terms of the axial distribution these equations become, from Eqs. (4.5) and (4.6),

$$m\frac{d^2z}{dt^2} = e\left[V_0'(z) - \frac{r^2}{2^2}V_0^{(3)}(z) + \cdots + \frac{(-1)^n r^{2n}}{2^2 \cdot 4^2 \cdots (2n)^2}V_0^{(2n+1)}(z) + \cdots\right]$$
$$m\frac{d^2r}{dt^2} = e\left[-\frac{r}{2}V_0''(z) + \frac{r^3}{2^2 \cdot 4}V_0^{(4)}(z) + \cdots + \frac{(-1)^n r^{2n-1}}{2^2 \cdot 4^2 \cdots (2n-2)^2 \cdot 2n}V_0^{(2n)} + \cdots\right]. \quad (4.8)$$

4.6. Energy Equation.—Equations (4.7) may be rewritten as

$$m\frac{d^2z}{dt^2} = m\frac{d}{dt}\left(\frac{dz}{dt}\right) = m\left(\frac{dz}{dt}\right)\frac{d}{dz}\left(\frac{dz}{dt}\right) = e\frac{\partial V}{\partial z}$$
$$m\frac{d^2r}{dt^2} = m\frac{d}{dt}\left(\frac{dr}{dt}\right) = m\left(\frac{dr}{dt}\right)\frac{d}{dr}\left(\frac{dr}{dt}\right) = e\frac{\partial V}{\partial r}. \quad (4.9)$$

Multiplying the first of the Eqs. (4.9) by dz, the second by dr and adding, there results

$$m\left[\frac{dz}{dt}d\left(\frac{dz}{dt}\right) + \frac{dr}{dt}d\left(\frac{dr}{dt}\right)\right] = e\left(\frac{\partial V}{\partial z}dz + \frac{\partial V}{\partial r}dr\right) = edV. \quad (4.10)$$

Equation (4.10) is exact, and its solution is

$$\frac{1}{2}mv^2 = \frac{1}{2}m\left[\left(\frac{dz}{dt}\right)^2 + \left(\frac{dr}{dt}\right)^2\right] = eV + C \quad (4.11)$$

where C is the constant of integration. If the electron speed v is zero when $V = 0$ (in the cathode-ray tube this occurs for an electron having zero speed at the cathode), then $C = 0$ and

$$\frac{1}{2}mv^2 = \frac{1}{2}m\left[\left(\frac{dz}{dt}\right)^2 + \left(\frac{dr}{dt}\right)^2\right] = eV, \quad (4.12)$$

so the speed of the electron becomes

$$v = \sqrt{2\frac{e}{m}V} = 5.95 \times 10^7\sqrt{V_{\text{volts}}}\ \text{cm./sec.} \quad (4.13)$$

If the electron speed is not zero but equal to v_0 when $V = 0$ (this case occurring if the electron at the cathode has an initial speed v_0), then $C = \frac{1}{2}mv_0^2$ and

$$\tfrac{1}{2}m(v^2 - v_0^2) = eV. \quad (4.14)$$

It has become customary to express the speed of electrons in equivalent volts. An electron is said to have a speed of V equivalent volts if its speed is $v = 5.95 \times 10^7\sqrt{V}$ cm./sec. Thus an electron is said to have a speed of 100 equivalent volts if its speed is 5.95×10^8 cm./sec.

If the initial speed v_0 be given in equivalent volts V_0 then $\frac{1}{2}mv_0^2 = eV_0$, and Eq. (4.14) becomes

$$\tfrac{1}{2}mv^2 = e(V + V_0), \quad (4.15)$$

and the expression for the speed of the electron becomes

$$v = \sqrt{2\frac{e}{m}(V + V_0)} = 5.95 \times 10^7\sqrt{(V + V_0)}\ \text{cm./sec.} \quad (4.16)$$

The energy equation (4.12) shows that an electrostatic field of force is conservative, *i.e.*, the total energy of an electron moving

through the field is constant. Thus an electron moving from cathode to second anode loses the potential energy eV and gains the exactly equivalent amount of kinetic energy $\frac{1}{2}mv^2$, and so the total amount of energy (kinetic + potential) remains constant.

4.7. Differential Equation of Trajectory of Electron.—It will be now shown that the trajectory of an electron traversing an axially symmetric electrostatic field described by the potential function $V(r,z)$ satisfies the differential equation

$$\frac{d^2r}{dz^2} + \frac{\left[1 + \left(\frac{dr}{dz}\right)^2\right]}{2V}\frac{\partial V}{\partial z}\frac{dr}{dz} - \frac{\left[1 + \left(\frac{dr}{dz}\right)^2\right]}{2V}\frac{\partial V}{\partial r} = 0. \quad (4.17)$$

To show this note that

$$\frac{d^2r}{dt^2} = \frac{d}{dt}\left(\frac{dr}{dt}\right) = \frac{dz}{dt}\frac{d}{dz}\left(\frac{dr}{dz}\frac{dz}{dt}\right) = \left(\frac{dz}{dt}\right)^2\frac{d^2r}{dz^2} + \frac{dr}{dz}\frac{dz}{dt}\frac{d}{dz}\left(\frac{dz}{dt}\right) \quad (4.18)$$

$$\frac{d^2z}{dt^2} = \frac{d}{dt}\left(\frac{dz}{dt}\right) = \frac{dz}{dt}\frac{d}{dz}\left(\frac{dz}{dt}\right) \quad (4.19)$$

and that (4.12) may be written as

$$\frac{1}{2}m\left(\frac{dz}{dt}\right)^2\left[1 + \left(\frac{dr}{dz}\right)^2\right] = eV. \quad (4.20)$$

From Eqs. (4.18), (4.19) and (4.20) it follows that

$$\frac{d^2r}{dt^2} = \frac{2\frac{e}{m}V}{\left[1 + \left(\frac{dr}{dz}\right)^2\right]}\frac{d^2r}{dz^2} + \frac{dr}{dz}\frac{d^2z}{dt^2}. \quad (4.21)$$

Inserting into (4.21) the values of d^2r/dt^2 and d^2z/dt^2 as given by Eq. (4.7) there results Eq. (4.17).

It is of great interest to note that e/m does not appear in Eq. (4.17), signifying that the trajectory is the same for *any* charged particle. Further, it is to be carefully noted that Eq. (4.17) is homogeneous in V so that if the voltages on the electrodes are all increased by a constant factor the trajectory of the electron will remain unaltered. Equation (4.17) is also homogeneous in r,z so that if all dimensions are increased by a constant factor, then the trajectory is also increased by the same factor.

4.8. Paraxial Electrons.[1]—An optical system is usually described in terms of paraxial or first-order imagery. Actual imagery departs from paraxial imagery. These departures are described as *aberrations*.

The focusing action of an electrostatic field is similarly described to a first approximation by considering only paraxial electrons. Paraxial electrons are characterized by the fact that in calculating their paths it is assumed that their distances r from the axis and their inclination dr/dz toward the axis are so small that the second and higher powers of r and dr/dz are negligible.

Limiting the discussion to paraxial electrons, Eqs. (4.4), (4.9) and (4.12) become

$$V(r,z) = V_0(z) \tag{4.4 p}$$

$$\left.\begin{aligned} m\frac{d^2r}{dt^2} &= -e\frac{r}{2}V_0''(z) \\ m\frac{d^2z}{dt^2} &= eV_0'(z) \end{aligned}\right\} \tag{4.9 p}$$

$$\frac{1}{2}mv^2 = \frac{1}{2}m\left(\frac{dz}{dt}\right)^2 = eV_0(z) \tag{4.12 p}$$

and that the differential equation for the trajectory traversed by a paraxial electron becomes, from Eq. (4.17),

$$\frac{d^2r}{dz^2} + \frac{V_0'}{2V_0}\frac{dr}{dz} + \frac{V_0''}{4V_0}r = 0. \tag{4.17 p}$$

Equation (4.17 p) (or Eqs. (4.9 p)) may be taken as the fundamental equation of the electron optics of axially symmetric electrostatic fields.

It is interesting to note that Eq. (4.17 p) is double edged, *i.e.*, given the axial distribution of potential V_0 and its first and second derivatives V_0' and V_0'', it allows one to determine the trajectories of paraxial electrons; or given the trajectory $r(z)$ and its first and second derivatives dr/dz and d^2r/dz, it permits one to determine the axial distribution of potential that will produce the given trajectory.

4.9. Determination of Trajectories of Paraxial Electrons.—In order to determine the trajectory of a paraxial electron it is

[1] The letter "p" after the equation number is to indicate that the equations so lettered are valid for paraxial electrons only.

necessary to solve Eq. (4.17 p) or its equivalent, Eqs. (4.9 p). An exact solution of Eq. (4.17 p) is obtainable only for relatively simple expressions of $V(z)$.[1] One such expression is

$$V(z) = Ce^{2kz} \tag{4.22 p}$$

where C and k are constants. For this expression of $V(z)$ Eq. (4.17 p) becomes

$$\frac{d^2r}{dz^2} + k\frac{dr}{dz} + k^2r = 0. \tag{4.23 p}$$

The solution of (4.23 p) is simply

$$r = Ae^{-\frac{k}{2}(1+\sqrt{3}j)z} + Be^{-\frac{k}{2}(1-\sqrt{3}j)z} \tag{4.24 p}$$

where $j = \sqrt{-1}$ and A and B are arbitrary constants to be evaluated by the initial conditions. Another such expression is

$$V(z) = C(a + bz)^n \tag{4.25 p}$$

where C, a, b and n are constants. With this expression of $V(z)$ Eq. (4.17 p) becomes

$$\frac{d^2r}{dz^2} + \frac{nb}{2(a + bz)}\frac{dr}{dz} + \frac{n(n-1)b^2}{4(a + bz)^2}r = 0. \tag{4.26 p}$$

The solution of (4.26 p) is

$$r = A(a + bz)^{m_1} + B(a + bz)^{m_2} \tag{4.27 p}$$

wherein

$$m_1 = \tfrac{1}{4}[(2 - n) + \sqrt{4 - 3n^2}]$$
$$m_2 = \tfrac{1}{4}[(2 - n) - \sqrt{4 - 3n^2}]$$

and A and B are arbitrary constants to be evaluated by the initial conditions.

In cathode-ray tubes one is normally interested in an electrostatic field confined to a relatively small region, with equipotential regions beyond its boundaries. Such a field is shown in Fig. 3.2. The axial distribution of potential of such a field conforms to the specification that $V =$ constant for $z_0 > z > z_1$ and $V = V(z)$ for $z_0 \geqslant z \geqslant z_1$. The axial distribution of potential actually

[1] Henceforth, $V(z)$ stands for the axial distribution of potential, thus replacing $V_0(z)$.

existing in cathode-ray tubes is not representable by such functions as given by Eqs. (4.22 p) and (4.25 p). However, although such expressions do not represent the axial distribution over the entire range $z_0 \leqslant z \leqslant z_1$, they may be useful in representing it over a limited range.

It is worth noting, in passing, that one may obtain almost any axial distribution of potential by means of a resistance cylinder wall, *i.e.*, a cylinder wall composed of resistance material with the resistance varying along the length of the cylinder. Thus, let $V(z)$ be the desired axial distribution and c the radius of the cylinder; then by Eq. (4.4), the potential at the cylinder wall is

$$V(z,c) = V(z) - \frac{c^2}{2^2}V''(z) + \cdots$$

To obtain the desired $V(z)$ a constant current i is sent through the variable resistance

$$R(z,c) = R(z) - \frac{c^2}{2^2}R''(z) + \cdots$$

where $R(z,c)$ is determined from the given axial distribution $V(z)$ by the relations

$$R(z) = \frac{V(z)}{i}, \qquad R''(z) = \frac{V''(z)}{i}, \text{ etc.}$$

If an analytical expression for $V(z)$ is available, then Eq. (4.17 p) may be integrated in the form of an infinite series. However, in practical cases there is seldom an analytical expression for $V(z)$ available. In these cases it is necessary to integrate Eq. (4.17 p) by some step-by-step method of approximation. In determining approximate solutions it is very valuable to have at one's disposal several different methods for the determination of the solution. Accordingly, there will now be described several methods for the approximate determination of the trajectory of a paraxial electron.

4.10. Calculation of Paraxial Trajectories from Eq. (4.17 p).— In order to obtain the trajectory from Eq. (4.17 p) it is necessary to know $V(z)$, $V'(z)$ and $V''(z)$. Figure 4.3 gives $V(z)$, $V'(z)$ and $V''(z)$ due to an electrode configuration as that shown in Fig. 3.2. As mentioned in Sec. 4.4, $V(z)$ may always be deter-

mined experimentally. $V'(z)$[1] and $V''(z)$ are then determined by numerical or graphical differentiation of $V(z)$. If the axis of z be divided into n intervals (not necessarily equal to each other) each sufficiently small so as to permit, within the accuracy desired, considering the quantities $V''(z)/4V(z)$ and $V'(z)/2V(z)$ as constants or linear functions of z throughout the interval, then Eq. (4.17 p) can be solved step by step.

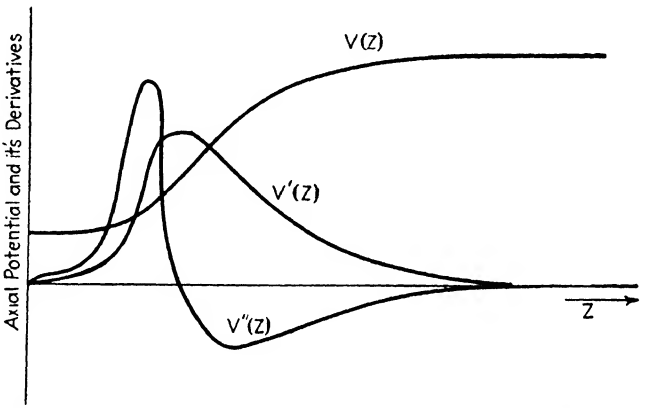


FIG. 4.3.—Axial distribution of potential and its derivatives.

Thus consider any interval, say the ith, and let

$$\frac{1}{2}\frac{V'(z)}{V(z)} = -2A \qquad \text{and} \qquad \frac{1}{4}\frac{V''(z)}{V(z)} = -B$$

where A and B are constants throughout the ith interval, then Eq. (4.17 p) becomes for this interval

$$\frac{d^2r}{dz^2} - 2A\frac{dr}{dz} - Br = 0. \tag{4.28 p}$$

The solution of this equation is

$$r = \frac{(r_i' - r_i m_2)}{2\sqrt{A^2 + B}}\epsilon^{m_1 z} - \frac{(r_i' - r_i m_1)\epsilon^{m_2 z}}{2\sqrt{A^2 + B}}. \tag{4.29 p}$$

where r_i and r_i' are the known values of r and dr/dz at the beginning of the ith interval, and $m_1 = A + \sqrt{A^2 + B}$ and

$$m_2 = A - \sqrt{A^2 + B}.$$

[1] By the use of two probes (see McArthur, *loc. cit.*) instead of one, $V'(z)$ can also be determined experimentally.

Equation (4. 29 p) permits one to determine the value of r at any point of the ith interval; the value of dr/dz at any point in the ith interval is from (4.29 p)

$$\frac{dr}{dz} = \frac{(m_1 r_i' + r_i B)\epsilon^{m_1 z}}{2\sqrt{A^2 + B}} - \frac{(m_2 r_i' + r_i B)\epsilon^{m_2 z}}{2\sqrt{A^2 + B}}. \qquad (4.30\text{ p})$$

The values of r and dr/dz at the end of the ith interval are then used as the r_{i+1} and r_{i+1}' for the next, $(i + 1)$th, interval. In this way the path of the electron through the entire electrostatic field may be determined.

Similarly, suppose that throughout any interval, say the ith,

$$\frac{1}{2}\frac{V'(z)}{V(z)} = -(a + bz) \qquad \text{and} \qquad \frac{1}{4}\frac{V''(z)}{V(z)} = -(p + qz)$$

where a, b, p and q are constants, then Eq. (4.17 p) becomes for the ith interval

$$\frac{d^2r}{dz^2} = (a + bz)\frac{dr}{dz} + (p + qz)r. \qquad (4.31\text{ p})$$

The solution of Eq. (4.31 p) may be expressed as the Taylor series

$$r(z) = r_i + r_i'(z - z_i) + \frac{r_i''}{2!}(z - z_i)^2 + \cdots + \frac{r_i^{(n)}}{n!}(z - z_i)^n + \cdots \qquad (4.32\text{ p})$$

where z_i and $r_i^{(n)}$ are the values of z and $d^n r/dz^n$ at the beginning of the ith interval. The values of z_i, r_i and r_i' are known and the values of the successive derivatives are evaluated from the differential equation itself and by successive differentiation of Eq. (4.31 p). The values of the higher derivatives are determined by the recurrent formula

$$r_i^{(n)} = (a + bz_i)r_i^{(n-1)} + [p + (n - 2)b + qz_i]r_i^{(n-2)} + (n - 2)qr_i^{(n-3)}.$$

The series (4. 32 p) is made very rapidly convergent by choosing $(z - z_i)$ sufficiently small. The value of dr/dz at any point in the ith interval is from (4.32 p)

$$\frac{dr}{dz} = r_i' + r''(z - z_i) + \frac{r'''}{2!}(z - z_i)^2 + \cdots + \frac{r^{(n+1)}}{n!}(z - z_i)^n + \cdots \qquad (4.33\text{ p})$$

The values of r and dr/dz at the end of the ith interval are then used as the $r_{(i+1)}$ and $r_{(i+1)}'$ for the next interval.

Equation (4.17 p) can be solved step by step if, in each interval, $V(z)$ is assumed to be representable by any function which makes Eq. (4.17 p) integrable. Equations (4.22 p) and (4.25 p) are examples of functions which make Eq. (4.17 p) integrable.

4.11. Calculation of Paraxial Trajectories from Eqs. (4.9 p).— The trajectory of a paraxial electron may be determined by solving Eqs. (4.9 p) step by step in a manner similar to that described in Sec. 4.10 for solving Eq. (4.17 p). Thus, let $V''(z)$ and $V'(z)$ be constant throughout the ith interval, then Eqs. (4.9 p) may be reduced to

$$\frac{d^2r}{dt^2} = K_1^2 r$$
$$\frac{d^2z}{dt^2} = K_2 \qquad (4.34\text{ p})$$

where $K_1^2 = \frac{1}{2}\frac{e}{m}V''(z)$ and $K_2 = \frac{e}{m}V'(z)$. The first equation of (4.34 p) may now be solved, obtaining

$$r = \frac{(K_1 r_i + \dot{r}_i)}{2K_1}\epsilon^{K_1 t} + \frac{(K_1 r_i - \dot{r}_i)}{2K_1}\epsilon^{-K_1 t} \qquad (4.35\text{ p})$$

wherein r_i and $\dot{r}_i$ are the values of r and dr/dt at the beginning of the ith interval. The value of dr/dt at any point in the ith interval is then

$$\frac{dr}{dt} = \frac{(K_1 r_i + \dot{r}_i)}{2}\epsilon^{K_1 t} - \frac{(K_1 r_i - \dot{r}_i)}{2}\epsilon^{-K_1 t}. \qquad (4.36\text{ p})$$

The second equation of (4.34 p) integrates immediately into

$$\frac{dz}{dt} = K_2 t + \dot{z}_i \qquad (4.37\text{ p})$$

and then into

$$z = \tfrac{1}{2}K_2 t^2 + \dot{z}_i t + z_i \qquad (4.38\text{ p})$$

where z_i and $\dot{z}_i$ are the values of z and dz/dt at the beginning of the ith interval. The time which the electron spends in any interval $\Delta z = (z - z_i)$ may then be obtained from Eq. (4.38 p). As this

requires the solution of a quadratic equation in t, it is often quicker to estimate the time by using the relation $\Delta z/v$ as a guide; in this relation v is the average speed of the electron in the interval and is calculable from the relation $v = 5.95 \times 10^7 \sqrt{V_i}$ cm./sec., where V_i is the average potential of the interval.

Another method for the determination of electron trajectories may be deduced from the equipotential line plot shown in Fig. 3.2. Thus, consider two equipotential lines sufficiently close together both in distance and in potential so that the force that would act on an electron in this interval can be taken as constant throughout this interval. The constant field is given by $\Delta V/\Delta S$ where ΔV is the difference in potential between the two equipotential lines, and ΔS is the distance between them measured along the normal to the two equipotential lines at the place under consideration. Hence, the force acting on the electron during the time it is between the two equipotential lines is $e\dfrac{\Delta V}{\Delta S}$. Decomposing this force into its two components it follows that

$$\begin{aligned} m\frac{d^2r}{dt^2} &= e\frac{\Delta V(r, z)}{\Delta r} = e\frac{\Delta V(r, z)}{\Delta S} \sin\theta \\ m\frac{d^2z}{dt^2} &= e\frac{\Delta V(r, z)}{\Delta z} = e\frac{\Delta V(r, z)}{\Delta S} \cos\theta \end{aligned} \tag{4.39}$$

where θ is the angle between the direction of the force (or Δn) and the axis. Now $\sin\theta = -\dfrac{\bar{r}}{R}$ and $\cos\theta = \dfrac{\sqrt{R^2 - \bar{r}^2}}{R}$ where R is the radius of curvature obtained by averaging the radii of curvatures of the two equipotential lines and $\bar{r}$ is the average value of r. For paraxial electrons $\cos\theta \cong 1$ and

$$\frac{\Delta V}{\Delta S} = \frac{\Delta V}{\Delta z} = \frac{dV}{dz}.$$

So that the above equations become, after integration of the second one,

$$\begin{aligned} m\frac{d^2r}{dt^2} &= -e\frac{\Delta V}{\Delta z}\frac{\bar{r}}{R} \\ \frac{m}{2}\left(\frac{dz}{dt}\right)^2 &= \frac{m}{2}v^2 = eV. \end{aligned} \tag{4.39 p}$$

Comparing Eqs. (4.9 p) and (4.39 p) it is worth noting that

$$R = \frac{2V'}{V''}. \tag{4.40 p}$$

Expression[1] (4.40 p) shows that the axial radii of curvatures of the equipotential lines of Fig. 3.2 are obtainable from the axial distribution of potential, and that the equipotential lines of Fig. 3.2 may be considered near the axis to be a series of spherical equipotential (refracting) surfaces.

As the force is assumed constant throughout the ith interval, the first equation of (4.39 p) may be written as

$$\frac{d^2r}{dt^2} = K \tag{4.41 p}$$

where K is a constant equal to $-\frac{e}{m}\frac{\Delta V}{\Delta z}\frac{\bar{r}}{R}$, $\bar{r}$ being the average value of r in the ith interval. As a first, and usually sufficient, approximation $\bar{r}$ is taken equal to r_i, the value of r at the beginning of the interval. Equation (4.41 p) integrates first into

$$\frac{dr}{dt} = Kt + \dot{r}_i \tag{4.42 p}$$

and then into

$$r = \tfrac{1}{2}Kt^2 + \dot{r}_i t + r_i. \tag{4.43 p}$$

The time which the electron spends in the ith interval is then obtained from Eq. (4.38 p) or is estimated by means of the relation $\Delta z/v$.

It is worth noting that Eqs. (4.39) are not limited to paraxial electrons, and hence with the aid of an equipotential-line plot, such as that shown in Fig. 3.2, enable one to determine the trajectory of a non-paraxial electron.

4.12. Calculation of Paraxial Trajectories Using Only V and V'. The previous methods for calculating paraxial trajectories required a knowledge of V'' or of R. The method to be outlined in this section requires a knowledge of V'/V only. This is of

[1] For a more direct derivation of Eq. (4.40 p) see: Theory of Electron Gun, I. G. Maloff and D. W. Epstein, *I.R.E. Proc.*, December, 1934, pp. 1386–1411.

considerable advantage since the quantity V'' or R is usually the least accurately known and the most difficult to obtain.

The knowledge of V'' in (4.17 p) is made unnecessary by transforming (4.17 p), which is a differential equation in r, into a differential equation in ρ, where ρ is defined by the relation

$$r = \rho V^{-\frac{1}{4}}. \tag{4.44 p}$$

Thus, differentiating and substituting (4.44 p) into (4.17 p) there results the following differential equation in ρ

$$\frac{d^2\rho}{dz^2} + \frac{3}{16}\left(\frac{V'}{V}\right)^2 \rho = 0. \tag{4.45 p}$$

Thus, the differential equation in ρ contains only V'/V and does not contain V'' or R. Equation (4.45 p) may be solved for ρ by any of the methods already outlined in the previous sections. Having determined ρ, r may be then determined by means of Eq. (4.44 p) *except where* $V = 0$.

CHAPTER 5

ELECTROSTATIC ELECTRON LENSES[1]

5.1. General.—An axially symmetric optical system is usually described in terms of paraxial imagery. The focusing action of an electrostatic field is similarly described, to a first approximation, by considering only paraxial electrons. In the last chapter there was derived the differential equation

$$\frac{d^2r}{dz^2} + \frac{V'}{2V}\frac{dr}{dz} + \frac{V''}{4V}r = 0 \qquad (4.17\text{ p})$$

which is satisfied by the trajectories of paraxial electrons in axially symmetric electrostatic fields. Equation (4.17 p) may be considered as the fundamental equation of paraxial imagery for axially symmetric electrostatic fields.

The general solution of the linear differential equation (4.17 p) may be expressed as

$$r(z) = c_1 r_1(z) + c_2 r_2(z) \qquad (5.1\text{ p})$$

where $r_1(z)$ and $r_2(z)$ are any two linearly independent solutions of (4.17 p). Equation (5.1 p) thus tells one that the trajectory of any paraxial electron is a linear combination of two independent trajectories. Hence the complete focusing action of an axially symmetric electrostatic field may be determined by calculating the trajectories of only two electrons.

5.2. Two Fundamental Trajectories.—Multiplying Eq. (4.17 p) by $\sqrt{V}$, there results the self-adjoint[2] equation:

[1] Epstein, D. W., Electron Optical System Of Two Cylinders as Applied to Cathode-Ray Tubes, *I.R.E. Proc.*, **24**, 1095 (1936).

[2] The linear second-order differential equation

$$p_0\frac{d^2u}{dz^2} + p_1\frac{du}{dz} + p_2u = 0$$

is self-adjoint if $p_1 = dp_0/dz$. See Ordinary Differential Equations, E. L. Ince, p. 215.

$$L(r) = \sqrt{V}\frac{d^2r}{dz^2} + \frac{V'}{2\sqrt{V}}\frac{dr}{dz} + \frac{V''}{4\sqrt{V}}r$$
$$= \frac{d}{dz}\left(\sqrt{V}\frac{dr}{dz}\right) + \frac{V''}{4\sqrt{V}}r = 0. \quad (5.2\text{ p})$$

Let $r_1(z)$ and $r_2(z)$ be two independent solutions of (4.17 p) representing the trajectories of two electrons, then

$$r_2L(r_1) - r_1L(r_2) = \frac{d}{dz}\left[\sqrt{V}\left(r_2\frac{dr_1}{dz} - r_1\frac{dr_2}{dz}\right)\right] = 0$$

Integrating this equation between the limits a and b, there results[1]

$$\int_a^b [r_2L(r_1) - r_1L(r_2)]dz = \left[\sqrt{V}\left(r_2\frac{dr_1}{dz} - r_1\frac{dr_2}{dz}\right)\right]_a^b = 0.$$

Substituting the limits,

$$\sqrt{V(b)}[r_2(b)r_1'(b) - r_1(b)r_2'(b)]$$
$$= \sqrt{V(a)}[r_2(a)r_1'(a) - r_1(a)r_2'(a)]. \quad (5.3\text{ p})$$

In particular, let $r_1(z)$, $r_2(z)$, $r_1'(z)$ and $r_2'(z)$ assume the following values at a and b:

$$\begin{aligned} r_1(a) &= h_1 & r_1(b) &= 0 \\ r_1'(a) &= 0 & r_1'(b) &= \tan\beta_2 \\ r_2(a) &= 0 & r_2(b) &= -h_2 \\ r_2'(a) &= \tan\beta_1 & r_2'(b) &= 0. \end{aligned} \quad (5.4\text{ p})$$

then equation (5.3 p) reduces to

$$\sqrt{V(b)}h_2\tan\beta_2 = \sqrt{V(a)}h_1\tan\beta_1. \quad (5.5\text{ p})$$

The two trajectories $r_1(z)$ and $r_2(z)$ satisfying relations (5.4 p) will be called the two *fundamental trajectories*. Figure 5.1 shows two fundamental trajectories. Any two independent trajectories may be taken as the fundamental pair; this particular pair is chosen because by means of this pair the usual optical relations are easily obtained. Thus Eq. (5.5 p) corresponds to Lagrange's law[2]

$$\mu_2h_2\tan\beta_2 = \mu_1h_1\tan\beta_1.$$

[1] This discussion is limited to electrostatic fields having finite extension, *i.e.*, $V = V(z)$ for $\alpha \leqslant z \leqslant \beta$ and $V =$ constant for $\alpha \geqslant z \geqslant \beta$ and further $a \leqslant \alpha$ and $b \geqslant \beta$.

[2] For this and other optical relations, see Houston's A Treatise on Light.

Let f_1 and f_2 be the focal lengths of the focusing system, then (see Fig. 5.1)

$$f_1 = -\frac{h_2}{\tan \beta_1} \quad \text{and} \quad f_2 = \frac{h_1}{\tan \beta_2} \tag{5.6 p}$$

by definition. Inserting (5.6 p) into (5.5 p), it results that

$$\frac{f_2}{f_1} = -\sqrt{\frac{V(b)}{V(a)}}. \tag{5.7 p}$$

Equation (5.7 p) corresponds to the well-known optical relation that the ratio of the focal lengths of a system is equal to the ratio of the indexes of refraction on the two sides of the system.

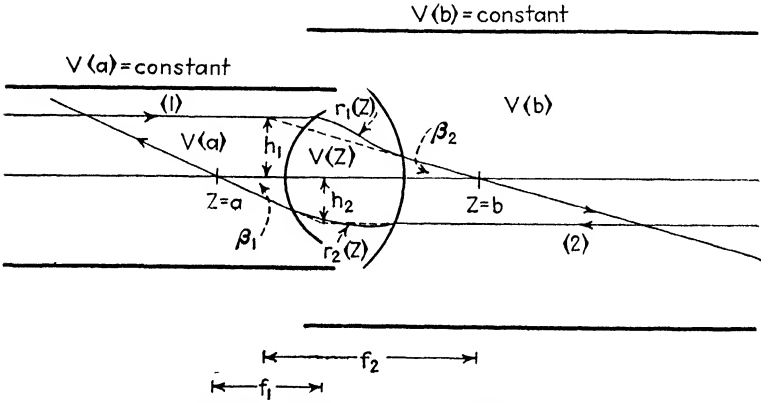


FIG. 5.1.—The fundamental trajectories.

Further let (see Fig. 5.2)

$$X_1 = \frac{h_1}{\tan \beta_1} \quad \text{and} \quad X_2 = -\frac{h_2}{\tan \beta_2} \tag{5.8 p}$$

then from (5.8 p) and (5.6 p) it follows that

$$X_1 X_2 = f_1 f_2. \tag{5.9 p}$$

X_1 is the distance between an object and the first focal point, and X_2 is the distance between the image and the second focal point, only if h_1 is the height of the object and h_2 is the height of the image. The magnification is, from (5.6 p) and (5.8 p),

$$m = \frac{h_2}{h_1} = -\frac{f_1}{X_1} = -\frac{X_2}{f_2}. \tag{5.10 p}$$

The points F_1, F_2, H_1 and H_2 shown in Fig. 5.2 constitute the set of cardinal points of the focusing system. F_1 and F_2 are the first and second focal points, and H_1 and H_2 are known as the first and second principal points, respectively.

5.3. Equivalent Lens of Axially Symmetric Field.—By determining the two fundamental trajectories $r_1(z)$ and $r_2(z)$, one determines the location of the cardinal points of the focusing system. A knowledge of the location of these cardinal points is sufficient for the determination of the paraxial focusing action of the field. It is therefore permissible to replace the electrostatic field by the

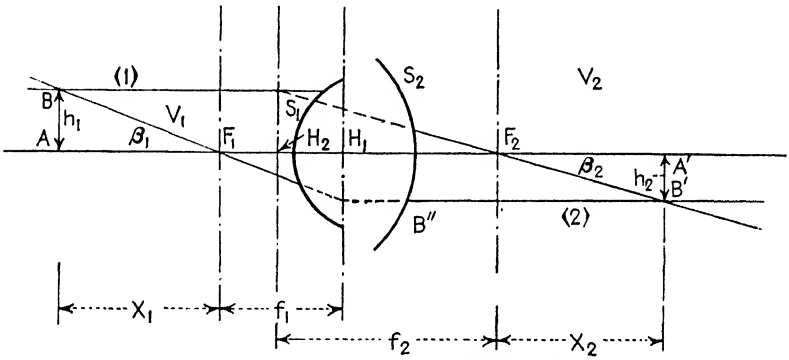


FIG. 5.2.—Location of cardinal points.

set of cardinal points. This set of cardinal points constitutes a lens, which may be called the *equivalent lens* of the field.

It is thus seen that an electrostatic field having axial symmetry will bring a paraxial electron beam to a focus. Fields possessing axial symmetry may be produced by applying various voltages to electrodes having geometric axial symmetry, such as coaxial cylinders, cones, disks with apertures, etc. Figure 5.3 gives the axial cross sections of several focusing electrode combinations, together with the distribution of potential along the axis.

5.4. Determination of Cardinal Points.—It is now appropriate to consider in some detail how to evaluate the focusing action of an axially symmetric electrostatic field, *i.e.*, how to find the location of the cardinal points of the equivalent lens. Referring to Fig. 5.2, let S_1 and S_2 be two equipotential surfaces such that the space to the left of S_1 is equipotential and is at potential V_1, and the space to the right of S_2 is equipotential and at the potential V_2. The potential in the region between S_1 and S_2 varies

continuously in some such manner as indicated in Fig. 3.2. Then the paraxial electron (1) moving parallel to the axis in the equipotential space to the left of S_1 will, after passing through the focusing system, move in a direction inclined at an angle to the axis and will pass through the axial point F_2. All paraxial electrons moving parallel to the axis in the object space[1] will pass through F_2. The point F_2 is known as the *second focal point.* The plane passing through the second focal point and

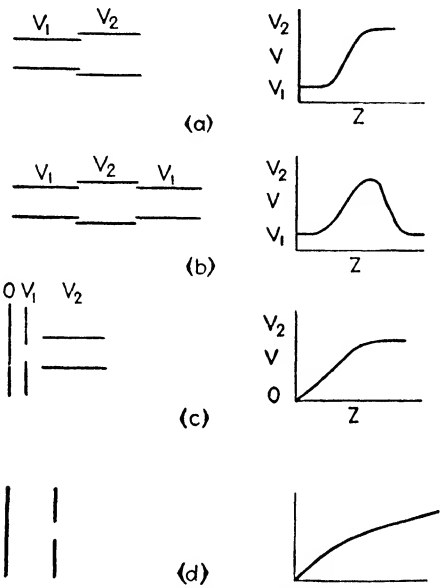


FIG. 5.3.—Types of electrostatic lenses.

perpendicular to the axis of symmetry is known as the *second focal plane.*

The plane perpendicular to the axis and passing through the point of intersection of the original and final directions of motion of electron (1) is known as the *second principal plane.* The point of intersection H_2 between the second principal plane and the axis is known as the *second principal point.* The distance H_2F_2 denoted by f_2 is known as the *second focal length.*

Similarly, the paraxial electron (2) moving parallel to the axis in the image space will, after passing through the focusing sys-

[1] The region to the left of the plane H_1 is the object space, and the region to the right of H_2 is the image space.

tem, move in a direction inclined to the axis, and will pass through the point F_1 in the object space. F_1 is known as the *first focal point*. F_1 may also be considered as that axial point in the object space from which all electrons, after passing through the focusing system, are parallel to the axis in the image space.

The plane perpendicular to the axis of symmetry and passing through the first focal point is known as the *first focal plane*. The plane perpendicular to the axis and passing through the point of intersection of the original and final directions of motion of the electron (2) is known as the *first principal plane*. The point of intersection H_1 of the first principal plane and the axis is known as the *first principal point*.

It is to be noted that in Fig. 5.2 the principal planes are crossed, that is, the object and image spaces overlap. This is a characteristic of lenses having indexes of refraction different on the two sides.

5.5. Use of Cardinal Points.—Having the cardinal points of the lens one may obtain either graphically or by means of Eqs. (5.9 p) and (5.10 p) the position and magnification of a given object. Thus in Fig. 5.2 let AB be an object from which electrons issue (to make it more concrete let the object AB be an aperture through which electrons are passing). Then a paraxial electron coming from B and moving parallel to the axis will, after passing through the lens, go in the direction F_2B'. A paraxial electron issuing from B in the direction BF_1 will, after passing through the lens, go in the direction $B''B'$; similarly, for every point of the object AB. So the inverted image $A'B'$ is obtained. The ratio $A'B'/AB$ gives the magnification. The electron image of AB becomes visible if a fluorescent screen is placed in the plane of $A'B'$.

Instead of obtaining the position and magnification of the object graphically, one can obtain them more simply with the aid of Eqs. (5.9 p) and (5.10 p). Thus, since the values of X_1, f_1 and f_2 are known, X_2, the distance between the second focal plane and the image, is calculated from Eq. (5.9 p) as

$$X_2 = \frac{f_1 f_2}{X_1}.$$

And the magnification m is, by (5.10 p),

$$m = -\frac{f_1}{X_1} = -\frac{X_2}{f_2}.$$

Similarly, if the position and size of the image are known, one can, with the aid of Eqs. (5.9 p) and (5.10 p), determine the position and size of the object.

5.6. Types of Electrostatic Lenses.—Electrostatic lenses may be classified by various distinguishing characteristics, such as number of electrodes, types of electrodes, voltages on electrodes, etc. For the purposes of cathode-ray tubes it is convenient to distinguish the following types of electrostatic lenses:

1. By a *bipotential lens* will be understood one which has different equipotential regions on its two sides, *i.e.*, $V_1 \neq V_2$ (see Fig. 5.2). The lens shown in Fig. 3.2 is a bipotential lens. The bipotential lens is further subdivided into:

a. The *direct bipotential lens,* which has the potential of the image space greater than that of the object space ($V_2 > V_1$).

b. The *inverted bipotential lens,* which has the potential of the image space less than that of the object space ($V_2 < V_1$). Figure 5.3(*a*) shows the axial distribution of potential of a simple direct bipotential lens. The bipotential lenses used in TCR tubes are almost always of this simple direct type. More complicated bipotential lenses arise from using more than two electrodes. Such complicated lenses may be further subdivided.

2. By a *unipotential lens* will be understood one which has identical equipotential regions on its two sides, *i.e.*, $V_1 = V_2$ (see Fig. 5.2). The unipotential lens may be further subdivided into any number of various types by considering the number of electrodes, type of electrodes, voltages on electrodes, etc. However, it is so rarely used in TCR tubes that no further subdivision will be made here. Figure 5.3(*b*) shows the axial distribution of potential of a three-electrode unipotential lens with the potential of the middle electrode greater than that of the end electrodes.

3. By an *immersion lens* will be understood one such as that shown in Fig. 5.3(*c*). This is the type of lens existing at the cathode of most TCR tubes, and its name is suggested by the fact that the object is immersed in the lens.

4. A fourth type of lens which is here mentioned but which is of little interest in the field of TCR tubes is the *aperture lens,* shown in Fig. 5.3(*d*). It is worth noting that the aperture lens is bounded by regions of constant field and not of constant potential. The only reason it is here mentioned is that it is the only electrostatic lens of those mentioned that can be made convergent or divergent. The unipotential, bipotential and immersion lenses are always *convergent.*

5.7. Thin Lens.—In general all lenses are "thick lenses." A knowledge of the positions of the focal and principal points is sufficient for the determination of the focusing action of a thick lens. To compute the positions of the cardinal points it is necessary to determine the two fundamental trajectories. The

calculation of the fundamental trajectories is accomplished by the integration of Eq. (4.17 p) or its equivalent, as described in Chap. 4.

A thin lens is one having negligible thickness along the axis of symmetry and is characterized by the fact that the two principal planes are assumed to coincide with, say, the center of the lens. The paraxial focusing action of a thin lens may, therefore, be completely determined as soon as the location of the lens and one of the focal lengths, say f_1, is known, for by Eq. (5.7 p) f_2 also becomes known.

Since for only a lens of zero thickness will the two principal planes coincide with the lens, a practical thin lens will only approximately have the characteristics of the ideal thin lens.

For a thin lens, the electrostatic field is confined to such a narrow range of z that the electron is in the field for such a short time that r remains sensibly unchanged during the time the electron is in the field. Thus let $r = r_0$ be the value of r when the electron is in the field, so Eq. (4.17 p) becomes

$$\frac{d^2r}{dz^2} + \frac{V'}{2V}\frac{dr}{dz} + \frac{V''}{4V}r_0 = 0. \qquad (4.17'\text{ p})$$

Let

$$\frac{\sqrt{V}}{r_0}\frac{dr}{dz} = P$$

then Eq. (4.17′ p) becomes

$$\frac{dP}{dz} + \frac{V''}{4\sqrt{V}} = 0.$$

Let the electrostatic field be confined within the narrow range $a \leq z \leq b$, then integrating between these limits

$$P_b - P_a + \int_a^b \frac{V''}{4\sqrt{V}}dz = 0.$$

Now consider an electron in the object space $z \leq a$ moving parallel to the axis at the distance r_0 from the axis, then at a, $dr/dz = 0$, and so $P_a = 0$. At b the electron issues from the lens still at the distance r_0, but with the slope $dr/dz = r'(b)$, so that

$$P_b = \frac{\sqrt{V_b}}{r_0}r'(b) = \frac{\sqrt{V(b)}}{f_2},$$

since $\frac{1}{f_2} = \frac{r'(b)}{r_0}$ by definition, and it follows that

$$\frac{1}{f_2} = \frac{1}{4\sqrt{V(b)}}\int_a^b \frac{V''}{\sqrt{V}}dz. \tag{5.11 p}$$

From Eqs. (5.11 p) and (5.7 p) it follows that

$$\frac{1}{f_1} = -\frac{1}{4\sqrt{V(a)}}\int_a^b \frac{V''}{\sqrt{V}}dz. \tag{5.12 p}$$

By partial integration it may be deduced that

$$\int_a^b \frac{V''}{\sqrt{V}}dz = \frac{1}{2}\int_a^b \frac{(V')^2}{\sqrt{V^3}}dz.$$

So Eqs. (5.11 p) and (5.12 p) become

$$\frac{1}{f_2} = \frac{1}{8\sqrt{V(b)}}\int_a^b \frac{(V')^2}{\sqrt{V^3}}dz \tag{5.13 p}$$

$$\frac{1}{f_1} = -\frac{1}{8\sqrt{V(a)}}\int_a^b \frac{(V')^2}{\sqrt{V^3}}dz. \tag{5.14 p}$$

5.8. Magnification of Lenses.—One of the problems of the TCR tube is to produce on the screen a bright spot of given dimensions. For a given position and size of object, and given position of the lens and image (screen), the size of the spot on the screen will vary directly with the magnification of the lens, assuming the aberrations of the lens to be negligible. It is, therefore, of interest to compare the magnification of the unipotential and bipotential lenses for given positions of object and image and lens.

From Eq. (5.5 p) and the definition of magnification it follows that

$$m = \frac{h_2}{h_1} = \frac{\tan \beta_1}{\tan \beta_2}\sqrt{\frac{V_1(a)}{V_2(b)}}. \tag{5.15 p}$$

From Fig. 5.2 it follows that $\tan \beta_1 = \frac{h_2 + h_1}{f_1 + X_1}$ and $\tan \beta_2 = \frac{h_2 + h_1}{f_2 + X_2}$ so that Eq. (5.15 p) may be written as

$$m = \frac{h_2}{h_1} = \frac{f_2 + X_2}{f_1 + X_1}\sqrt{\frac{V(a)}{V(b)}}. \tag{5.16 p}$$

To a first approximation, the fraction $\frac{f_2 + X_2}{f_1 + X_1}$ may be considered as a constant k which is independent of the type of lens used. Hence for this case

$$m = k\sqrt{\frac{\bar{V}(a)}{\bar{V}(b)}}. \qquad (5.16' \text{ p})$$

Equation (5.16′ p) thus says that to a first approximation the direct bipotential lens ($V_b > V_a$) will produce a smaller spot than the unipotential lens ($V_b = V_a$) or the inverted bipotential lens ($V_b < V_a$). This is one reason why the direct bipotential lens is used in TCR tubes in preference to unipotential or inverted bipotential lenses. This also shows that, everything else remaining the same, the direct bipotential electrostatic lens will produce a smaller spot than will the magnetostatic lens which is similar to the unipotential lens.

CHAPTER 6

ELECTROSTATIC LENSES OF TELEVISION CATHODE-RAY TUBES

6.1. General.—Figure 6.1 gives the cross section through the axis of a TCR tube. The electrostatic focusing system consists of an immersion lens close to the cathode and a direct bipotential lens near the end of the first anode. The operation of the immersion lens as used in TCR tubes is extremely complicated and not very well understood; the operation of the bipotential lens is much better understood and will therefore be treated first. As

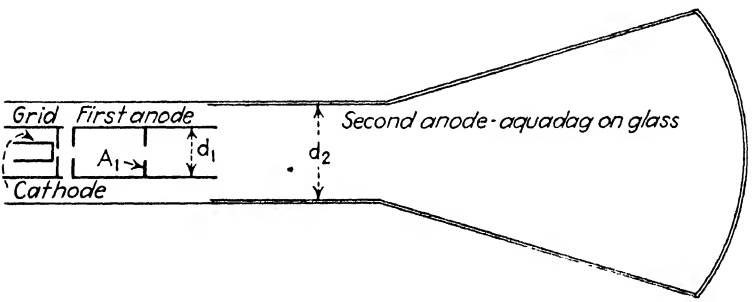


FIG. 6.1.—Cross section of a TCR tube.

far as the bipotential lens is concerned, the immersion lens may be thought of as concentrating the electrons emitted by the cathode into a small new source which serves as the object which the bipotential lens images on the fluorescent screen producing the visible spot.

Before going into a detailed study of the bipotential lens it might be well to recall briefly what has been accomplished so far. Methods have been given for determining electron trajectories through axially symmetric electrostatic fields. It has been shown that by determining the two fundamental trajectories, one is enabled to assign cardinal points to the axially symmetric electrostatic field and so replace the electrostatic field by an equivalent lens. The equivalent lens will in general be a thick lens.

It has also been shown that if X_1 and X_2 be the object and image distances measured from the first and second focal points, respectively, then

$$X_1X_2 = f_1f_2 \qquad (5.9\text{ p})$$

and the magnification m is

$$m = -\frac{f_1}{X_1} = -\frac{X_2}{f_2}. \qquad (5.10\text{ p})$$

A knowledge of the position of the cardinal points, together with the position of the object (or image), enables one to calculate the position of the image (or object) and the magnification of the object (or image).

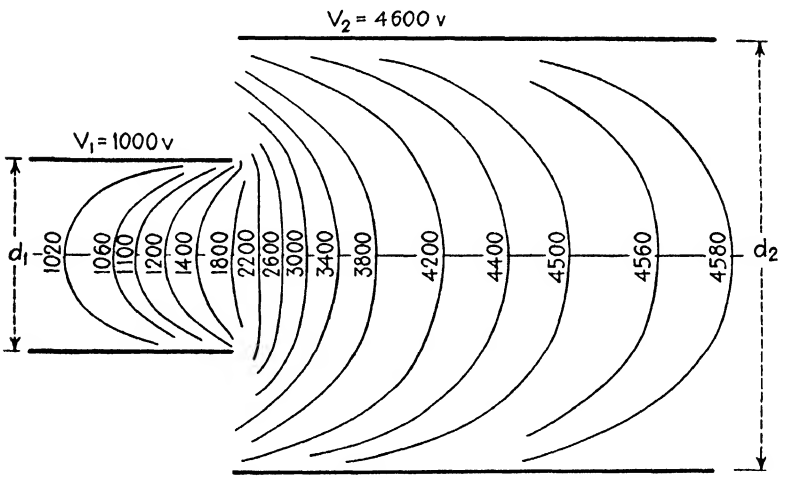


FIG. 6.2.—Equipotential line plot of bipotential lens.

6.2. Electrostatic Field of Two Cylinders.—Consider the bipotential lens formed by two metallic cylinders of diameters d_1 and d_2 and charged to potentials V_1 and V_2 relative to the cathode. Figure 6.2 represents a cross section through the axis of the equipotential surfaces of such a lens.

If one of the diameters, say d_1, is chosen as the unit of length (in both the r and z directions) then a given ratio of cylinder diameters d_2/d_1 always represents the same configuration of electrodes, *i.e.*, two cylinders, one having a diameter of unit length and the other having a diameter of d_2/d_1 units of length. This unit shall be designated as the g.d. (gun diameter).

A given ratio of diameters d_2/d_1 will, then, for given voltages V_2 and V_1 always produce the same distribution of potential. Thus Fig. 6.2 represents the potential distribution for any d_1 and d_2 for which d_1 is taken as the unit of length and d_2/d_1 is that given in Fig. 6.2. The use of the g.d. as the unit of length greatly simplifies the presentation of information. Thus the two fundamental trajectories and the accompanying cardinal points determined for given cylinder diameters d_2 and d_1 are also the fundamental trajectories for all cylinder diameters d_2' and d_1' for which $d_2'/d_1' = d_2/d_1$. The differential equation (4.17) for the trajectory of an electron remains unchanged if one replaces V by kV where k is a constant. Hence the trajectory of an electron will remain unaltered if the voltages on the electrodes are all multiplied by the same factor. The two fundamental trajectories and the accompanying cardinal points depend, therefore, not on V_1 or V_2 but on V_2/V_1. Hence, if the cardinal points are determined for given potentials V_1 and V_2, they remain the same for any V_1', V_2' so long as $V_2'/V_1' = V_2/V_1$. Hence, a given voltage ratio and a given diameter ratio uniquely determine the positions of the cardinal points of the electrostatic field if the g.d. be used as the unit of length.

From the fact that the potential function is relative, it follows that a given equipotential plot, such as that shown in Fig. 6.2, is independent of V_1 or V_2 but depends solely upon $V_2 - V_1$. Now let $V_2 - V_1 = a$. Then for any given value of a it is possible to determine V_2 and V_1 such that $V_2/V_1 = b$ where b is any preassigned number. So a given equipotential plot may be used for the determination of the fundamental trajectories for any ratio V_2/V_1.

Hence the important result that a given distribution of potential determined for two cylinders of diameters d_1 and d_2 and charged to potentials V_1 and V_2 enables one to determine the positions of the cardinal points of an electrostatic field of two cylinders of diameters d_2' and d_1', if $d_2'/d_1' = d_2/d_1$ and with any voltages on the electrodes.

6.3. Cardinal Points of Electrostatic Field of Two Cylinders.—Figure 6.3 shows the cardinal points as they are actually located inside the tube for a given voltage ratio and diameter ratio. It is convenient to locate the focal points F_1 and F_2 by their distances from the end of the gun. Hence, starting with Fig. 6.3,

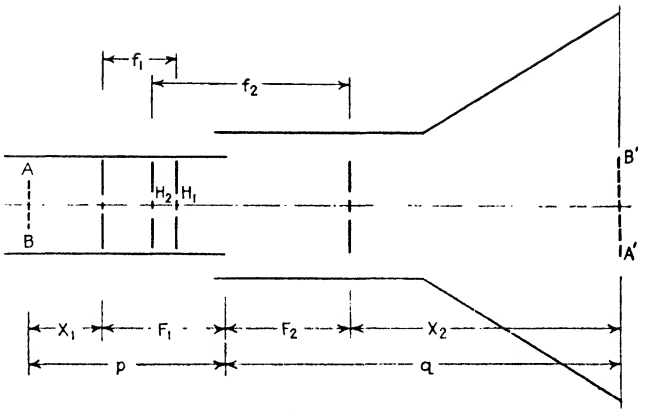


FIG. 6.3.—Positions of cardinal points.

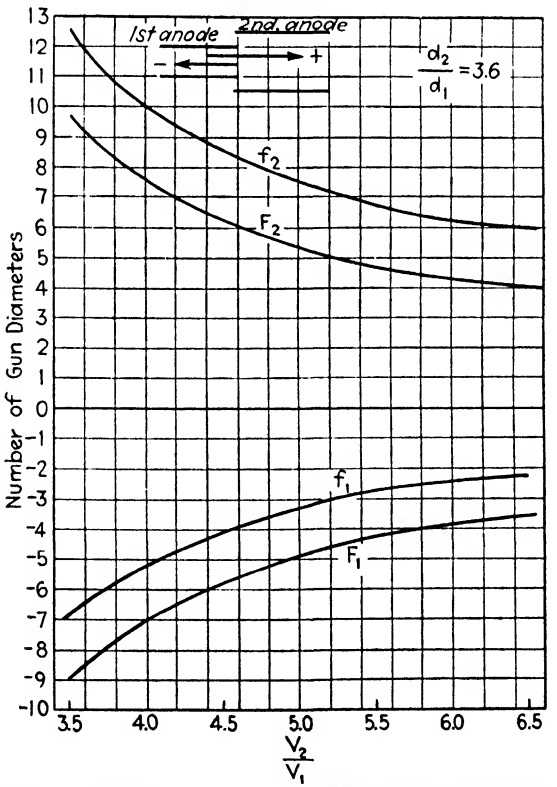


FIG. 6.4.—Variation of optical constants with voltage ratio.

F_1 and F_2 are indicated as lengths. It is to be noted that (1) the principal planes are crossed—*i.e.*, the object and image spaces overlap; (2) the focal length of the image space f_2 is greater than the focal length of the object space f_1; (3) the principal planes are located inside the first anode (for $V_1 < V_2$). These properties are inherent in a direct bipotential lens.

Figure 6.4 shows how the cardinal points vary with the voltage ratio for a diameter ratio 3.6. It is seen that the focal lengths

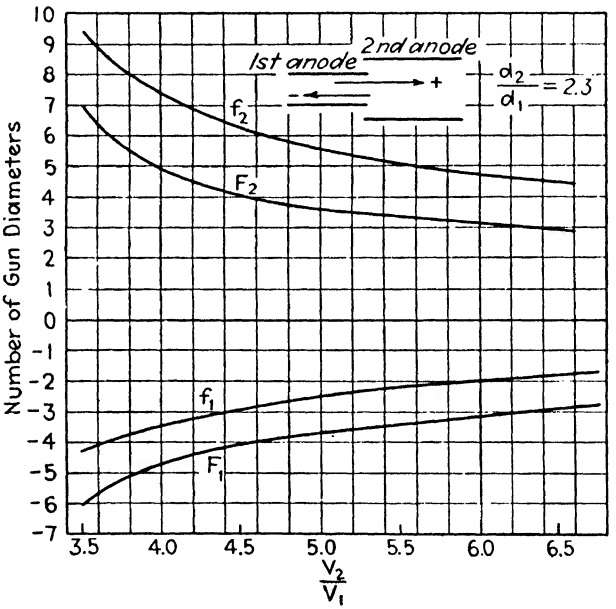


FIG. 6.5.—Variation of optical constants with voltage ratio.

decrease, *i.e.*, the power of the lens increases, as the voltage ratio is increased. Figures 6.5, 6.6 and 6.7 show the variation of the positions of the cardinal points with voltage ratio for different diameter ratios. Figure 6.8 shows how the cardinal points vary with the diameter ratio for a given voltage ratio. The locations of the cardinal points shown in these figures were determined by calculating the fundamental trajectories.

It should be clear that to a given voltage ratio and diameter ratio there corresponds but one set of cardinal points if the g.d. is chosen as the unit of length. It is worth noting that a given ratio of diameters determines the shape (curvature) and the

distance between the refractive (equipotential) surfaces (see Fig. 6.2), while the voltage ratio determines the indexes of refraction between the various surfaces.

The curves of Figs. 6.4 to 6.8 show how the cardinal points vary when V_2/V_1 and d_2/d_1 are varied. From these curves, one can determine how the cardinal points vary by changing only one of the variables V_1, V_2, d_1 and d_2 and keeping the other three constant. Thus by increasing V_2 while keeping V_1, d_1 and d_2 constant, V_2/V_1 increases and the focal lengths decrease; by

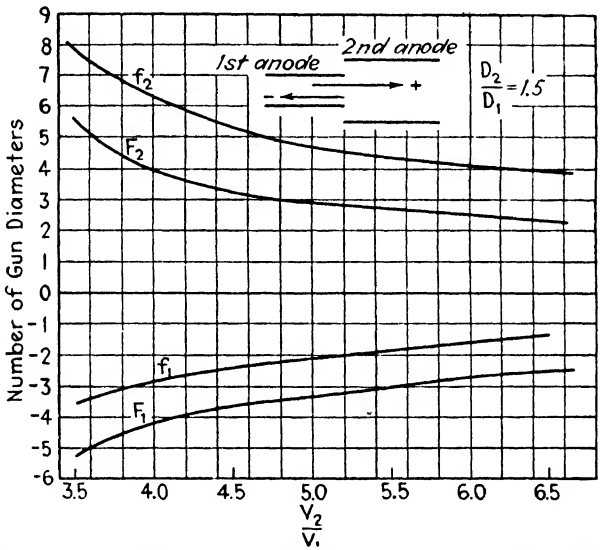


FIG. 6.6.—Variation of optical constants with voltage ratio.

increasing V_1 the focal lengths are increased. By increasing d_2, d_2/d_1 increases, and so the focal lengths are increased; and by increasing d_1, d_2/d_1 decreases and the focal lengths are decreased. It must be carefully noted, however, that by increasing d_1 the unit of length is increased.

Bipotential lenses composed of an electrode combination different from that of two cylinders have been used in TCR tubes. These electrode combinations are usually more complicated and much more difficult to construct and are seldom more suitable than the simple cylinder combinations. In some cases, however, it may be of advantage to use other than simple cylinder combinations, as they may permit one to focus the object on the screen

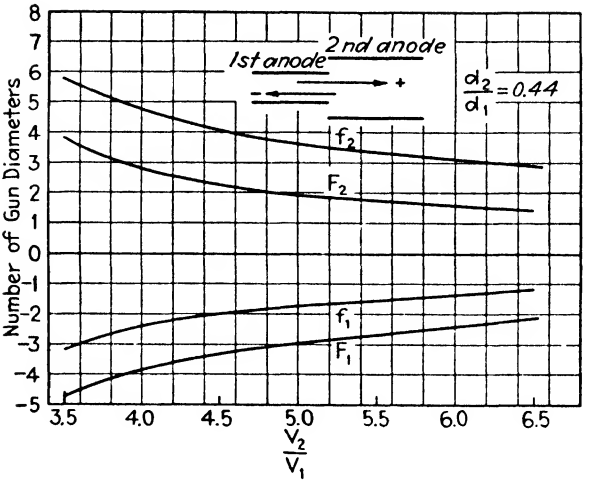


FIG. 6.7.—Variation of optical constants with voltage ratio.

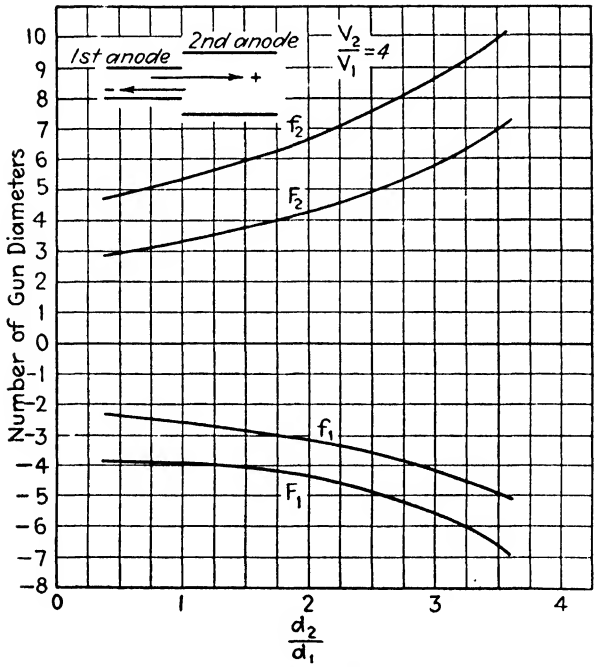


FIG. 6.8.—Variation of optical constants with diameter ratio.

at a lower voltage ratio or with less aberration. Figures 6.9 and 6.10 give the location of the cardinal points for two such combinations.

6.4. Use of Curves of Cardinal Points.—As an example to illustrate the use of the curves of Figs. 6.4 to 6.8, consider the

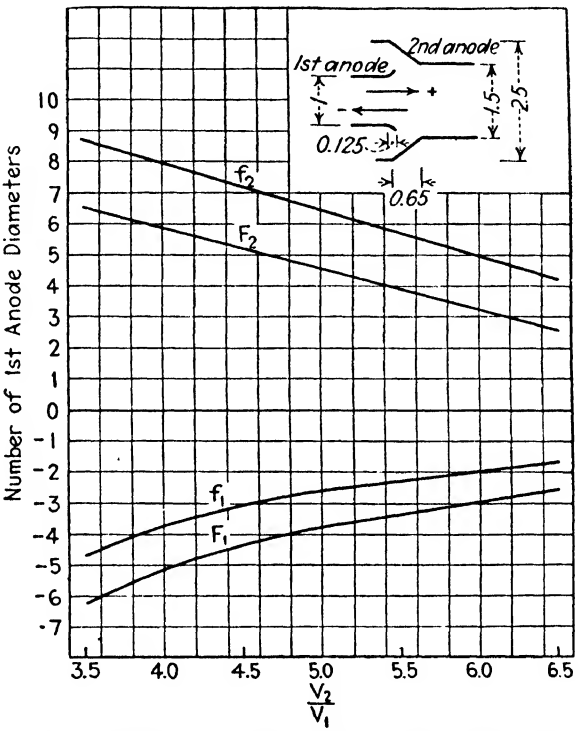


FIG. 6.9.—Variation of optical constants with voltage ratio.

cathode-ray tube shown in cross section in Fig. 6.1. Suppose it was found that to obtain the minimum spot on the screen with 1,000 volts on the first anode it is necessary to have 5,000 volts on the second anode. Further, let $d_2 = 3.5$ cm. and $d_1 = 1.5$ cm. Then from Fig. 6.5 note that for $d_2/d_1 = 2.3$ and $V_2/V_1 = 5$.

$$f_1 = -2.5 \text{ g.d.} \qquad F_1 = -3.7 \text{ g.d.}$$
$$f_2 = +5.6 \text{ g.d.} \qquad F_2 = +3.6 \text{ g.d.}$$

(It must be remembered that these cardinal points hold only for paraxial electrons. To use these cardinal points it is, therefore,

necessary to have the aperture A_1 sufficiently small to permit only paraxial electrons to enter the lens.)

Let q, the distance between the screen and the end of the gun, be 15 g.d., then

$$X_2 = q - F_2 = 15 - 3.6 = 11.4 \text{ g.d.}$$

So by Eq. (5.9 p)

$$X_1 = \frac{f_1 f_2}{X_2} = \frac{-2.5(5.6)}{11.4} = -1.2 \text{ g.d.}$$

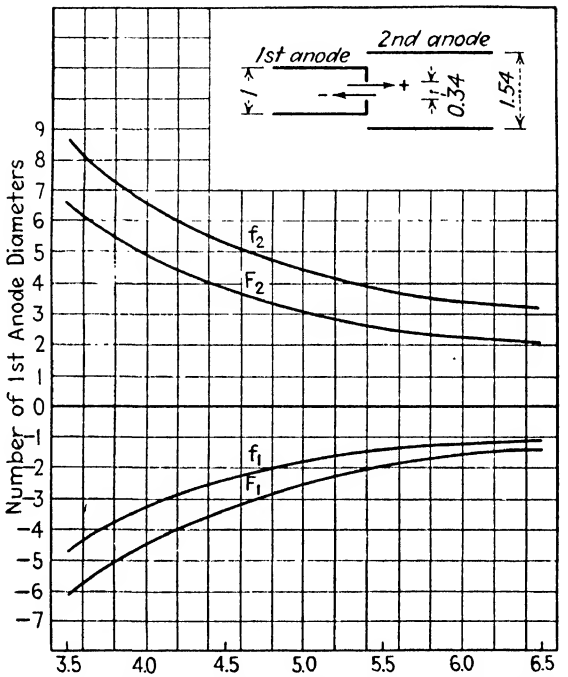


FIG. 6.10.—Variation of optical constants with voltage ratio.

and p, the distance between object and end of gun, is

$$p = X_1 + F_1 = -1.2 - 3.7 = -4.9 \text{ g.d.}$$

By means of Eq. (5.10 p) the magnification of the object is determined. Thus

$$m = -\frac{f_1}{X_1} = -\frac{2.5}{1.2} = -2.1$$

and if the size of the spot on the screen is measured to be 0.5 mm., the size of the object being imaged on the screen will be

$0.5/2.1 = 0.24$ mm. So, by knowing V_2/V_1, d_2/d_1, q and the size of the spot, one can, by means of the curves of Figs. 6.4 to 6.7 and Eqs. (5.9 p) and (5.10 p) determine the position and the size of the object that is imaged on the screen.

The knowledge of the positions of the cardinal points permits one to give quantitative answers to the following questions: What will be the effect on the spot size of a given tube if

1. The length of the first anode is varied?
2. The diameter of the first anode is varied?
3. The diameter of the second anode is varied?
4. The ratio of voltages on the two anodes is varied?
5. The distance between the screen and gun end is varied?

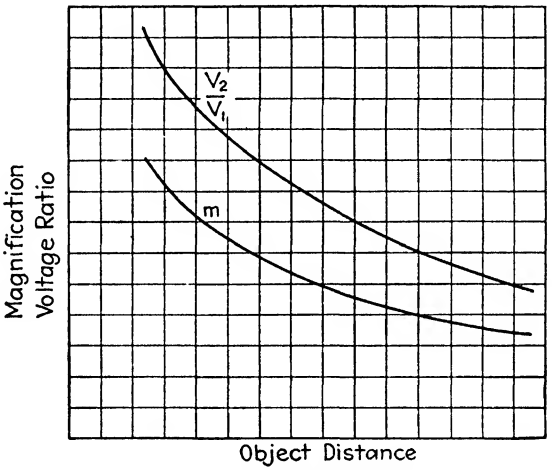


FIG. 6.11.—Magnification and voltage ratio *vs.* object distance.

It is to be noted, however, that to keep the object focused on the screen (minimum spot on the screen) it is necessary to vary at least two of these variables. Thus, if in the above example, the length of the first anode is changed, it will also be necessary to change V_2/V_1 or gun-screen distance in order to keep the minimum spot on the screen.

The calculation of the effects of changes in the various variables is therefore not straightforward but requires several trials. Thus to use the curves of Figs. 6.4 to 6.7 to calculate the effect on the spot size if the length of the first anode is changed while the remaining dimensions in the tube remain unaltered, it is first necessary to know the voltage ratio for the new object distance

at which the spot is focused. To get around this difficulty note that with a given ratio d_2/d_1 and given q, V_2/V_1 can be calculated as a function of p, and to each p (and the corresponding V_2/V_1) a magnification can be determined so that the curves shown in Fig. 6.11 may be plotted. Figure 6.11 then gives the change in spot size and voltage ratio caused by a given change in object distance. The curves of Fig. 6.11 hold for a given image distance; similar curves can be calculated for other image distances.

6.5. Thin Lens.—If the object and image are at large distances from the end of the gun, then to a fair approximation the two principal planes may be assumed to coincide and to be located at the end of the gun. Under these circumstances, the following simple formulas apply

$$\frac{\sqrt{V_1}}{p} + \frac{\sqrt{V_2}}{q} = \frac{\sqrt{V_1}}{f_1} = \frac{\sqrt{V_2}}{f_2} \qquad (6.1\text{ p})$$

and

$$m = -\frac{q}{p}\sqrt{\frac{V_1}{V_2}} \qquad (6.2\text{ p})$$

where p and q are the object and image distances measured from the end of the gun and f_1 is the first focal length, which may be calculated from the formula for the focal length of a thin lens given by Eq. (5.14 p). For this case it is not necessary to determine the fundamental trajectories.

A much better approximation is to consider the lens thin — that is, to assume that the two principal planes coincide—but to assume the thin lens to be situated between the two principal planes. Here, the above formulas still apply if p and q are measured from the assumed position of the thin lens instead of from the end of the gun. For this case it is necessary to determine one fundamental trajectory. It is seen from Figs. 6.4 to 6.7 that this thin lens is situated about 1.5 g.d. inside the first anode.

6.6. Experimental Determination of Cardinal Points.—The position of the cardinal points may be determined experimentally if the magnification is known for two given positions of object and image. Thus in Fig. 6.3, let AB, representing a fine wire mesh of known dimensions, be the object. The object is "illuminated" with electrons originating at a cathode to the left of

AB, and is imaged on a fluorescent screen. Let the distance between the mesh AB and the end of gun be p, the distance between end of gun and fluorescent screen be q and the magnification with which the mesh is imaged on the screen be m. Then from Fig. 6.3 and the relation for magnification

$$m = -\frac{f_1}{X_1} = -\frac{X_2}{f_2}$$

it follows that

$$p = X_1 + F_1 = \frac{-f_1}{m} + F_1 \qquad (6.3\text{ p})$$

$$q = X_2 + F_2 = -mf_2 + F_2. \qquad (6.4\text{ p})$$

Then if p_1, q_1 and m_1 correspond to one position of object, and p_2, q_2, and m_2 correspond to another position of the object (for the same lens, *i.e.*, the voltage ratio and diameter ratio are fixed), then it follows that

$$f_1 = m_1 m_2 \left(\frac{p_1 - p_2}{m_1 - m_2}\right) \qquad (6.5\text{ p})$$

$$f_2 = \frac{q_1 - q_2}{m_2 - m_1} \qquad (6.6\text{ p})$$

$$F_1 = \frac{m_1 p_1 - m_2 p_2}{m_1 - m_2} \qquad (6.7\text{ p})$$

$$F_2 = \frac{m_2 q_1 - m_1 q_2}{m_2 - m_1}. \qquad (6.8\text{ p})$$

Equations (6.5 p) to (6.8 p) allow one to determine the positions of all the cardinal points, provided the sets of conjugate quantities p_1, q_1, m_1 and p_2, q_2, m_2 are known. An obvious method for determining these quantities is to use a tube having two independently moving parts, the object mesh and the fluorescent screen on which to image the mesh. It is quite difficult, however, to build a tube with two independently moving parts which move over considerable distances.

It is relatively simple, however, to build a tube with one moving part, and especially so if the moving part is the object mesh. Figure 6.12 shows the cross section of a gun having a moving mesh. The mesh is welded to an aperture cup which slides inside the gun by merely tilting the tube. The position of the mesh

inside the gun is indicated by an index on the outside of the gun, which is calibrated.

To determine the two sets of quantities p_1, q_1, m_1 and p_2, q_2, m_2 by means of a tube having only a moving object, one may proceed as follows. The gun is inserted inside a glass blank with

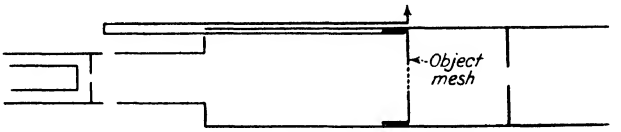


FIG. 6.12.—Gun with sliding object.

the screen at a known distance, say q_1, from the end of the gun. The voltage ratios required to focus the mesh on the screen and the magnifications of the mesh are then noted for various known positions of the mesh. The results are then plotted as shown by the curves p_1 and m_1 of Fig. 6.13. The same gun is

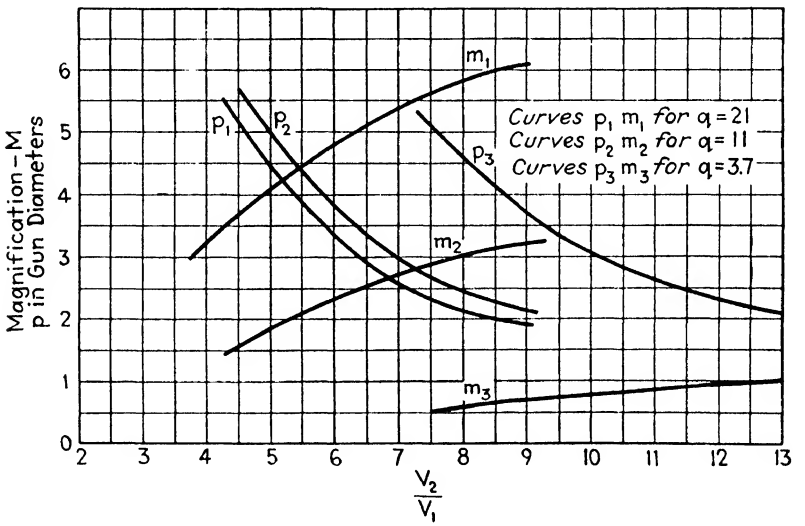


FIG. 6.13.—Magnification and object distance *vs.* voltage ratio.

then inserted into a blank with the screen at a different distance, say q_2, from the end of the gun. The measurements are repeated with this blank, and the results plotted as given by the curves m_2 and p_2 of Fig. 6.13.

The positions of the cardinal points may then be calculated by means of Eqs. (6.5 p) to (6.8 p) and the four curves p_1, m_1, p_2,

m_2 of Fig. 6.13. To do this it is necessary to choose the quantities m and p along a vertical line corresponding to a given voltage ratio.

The curves of Fig. 6.13 were determined, by the method just described, for a lens corresponding to a diameter ratio of 1:5.

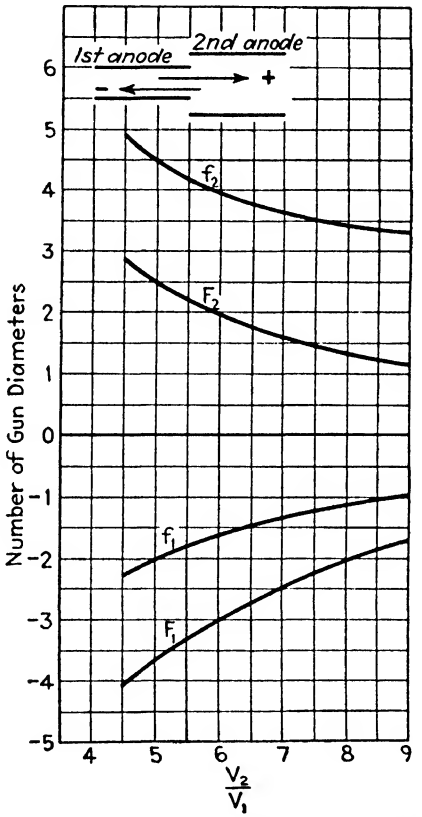


FIG. 6.14.—Optical constants *vs.* voltage ratio for $d_2/d_1 = 1.5$.

The accuracy of determining the positions of the cardinal points by means of Eqs. (6.5 p) to (6.8 p) depends greatly upon the accuracy with which the quantities $p_2 - p_1$, $q_2 - q_1$ and $m_2 - m_1$ may be determined from the curves of Fig. 6.13. A glance at the curves will show that the accuracy of the quantity $p_2 - p_1$ is rather poor, since it is the difference of two nearly equal quantities. The quantities $q_2 - q_1$ and $m_2 - m_1$ are quite accurate, however.

To avoid the use of the quantity $p_2 - p_1$ one may, instead of using Eqs. (6.5 p) to (6.8 p), proceed as follows:

Determine f_2 by means of

$$f_2 = \frac{q_1 - q_2}{m_2 - m_1}. \tag{6.6 p}$$

f_1 is then determined by the relation

$$f_1 = \frac{-1}{\sqrt{V_2/V_1}} f_2 \tag{6.9 p}$$

and the quantities F_1 and F_2 are determined by

$$F_1 = p_1 + \frac{f_1}{m_1} = p_2 + \frac{f_1}{m_2} \tag{6.10 p}$$

$$F_2 = q_1 + m_1 f_2 = q_2 + m_2 f_2 \tag{6.11 p}$$

wherein f_2 and f_1 are the values obtained by means of (6.6 p) and (6.9 p).

The focal lengths and the positions of the focal points determined as described above are plotted in Fig. 6.14. In Table 6.1, the values of f_1, f_2, F_1 and F_2, determined from the experimental data, are compared with the calculated values taken from the curves of Fig. 6.6. The agreement is within experimental error.

TABLE 6.1.

V_2/V_1	f_2		f_1		F_2		F_1	
	Exp.	Theo.	Exp.	Theo.	Exp.	Theo.	Exp.	Theo.
4.5	4.9	5.1	−2.3	−2.4	2.9	3.2	−4.1	−3.8
5.0	4.4	4.7	−2.0	−2.1	3.0	2.9	−3.7	−3.4
5.5	4.2	4.3	−1.8	−1.9	2.3	2.7	−3.3	−3.1
6.0	4.0	4.1	−1.7	−1.7	2.0	2.4	−3.0	−2.8
6.5	3.8	3.9	−1.5	−1.5	1.8	2.1	−2.7	−2.6

The data as plotted in Fig. 6.13 are useful only for the determination of the positions of the cardinal points. The data have been replotted in Figs. 6.15 and 6.16, showing how the magnification and the focusing voltage ratio vary with the position of the object for three positions of the screen (or gun-screen distances). These curves are very useful as gun-design information.

The method described above determines the positions of all the cardinal points and therefore determines the thick-lens

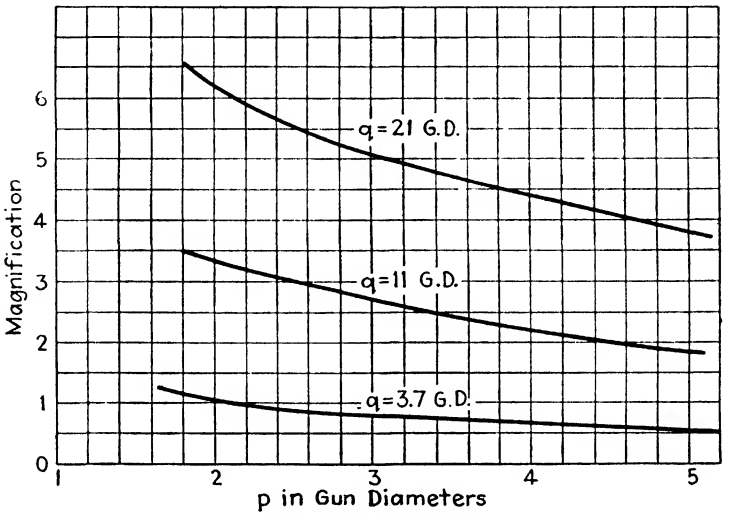


FIG. 6.15.—Magnification *vs.* object distance for $d_2/d_1 = 1.5$.

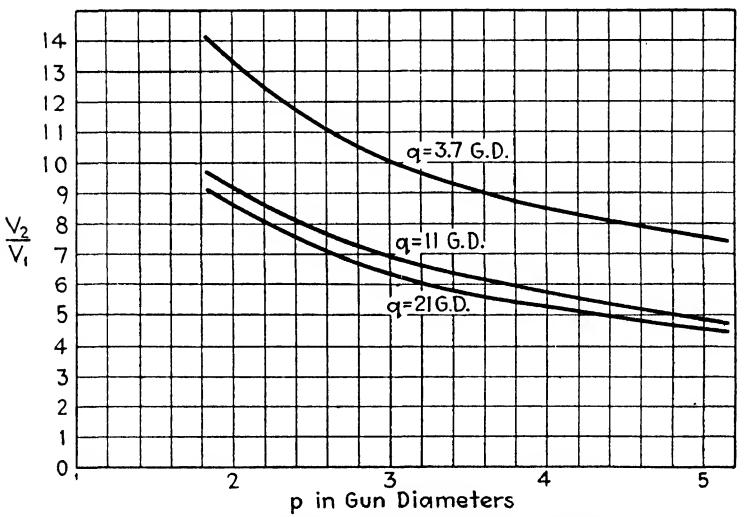


FIG. 6.16.—Voltage ratio *vs.* object distance for $d_2/d_1 = 1.5$.

equivalent of the focusing field. If the object and image are sufficiently distant from the lens, one may to a fair approximation consider the lens as thin and situated between the two principal

planes of the thick lens. The position and the focal lengths of the thin lens may be determined experimentally by simply noting the distance between object and image, the magnification and the voltage ratio at which the object is focused on the screen.

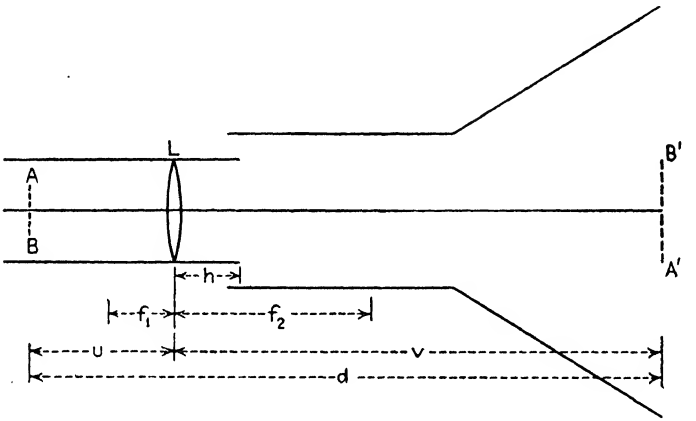


FIG. 6.17.—Position of thin lens.

Thus, in Fig. 6.17 let L be the equivalent thin lens and let u and v be the object and image distances measured from the thin lens. The following thin-lens relations then apply:

$$\frac{1}{u} + \frac{1}{v}\sqrt{\frac{V_2}{V_1}} = \frac{1}{f_1} \tag{6.12 p}$$

$$\frac{v}{u} = m\sqrt{\frac{V_2}{V_1}}. \tag{6.13 p}$$

Besides these, the following two relations also hold

$$u + v = d \tag{6.14 p}$$

$$\frac{f_2}{f_1} = -\sqrt{\frac{V_2}{V_1}} \tag{5.7 p}$$

where d is the distance between object and image.

It follows from Eqs. (6.13 p) and (6.14 p) that

$$u = \frac{d}{1 + m\sqrt{V_2/V_1}}. \tag{6.15 p}$$

As u is the distance between object and lens, Eq. (6.15 p) permits one to determine the position of the lens. Substituting (6.15 p) and (6.14 p) into (6.12 p), it results that

$$f_1 = \frac{d}{1 + m\sqrt{V_2/V_1}} \frac{m}{m+1} = u\frac{m}{m+1} \qquad (6.16\text{ p})$$

Equation (6.16 p) determines the first focal length; it is seen from Eq. (6.16 p) that for large magnifications

$$f_1 = u. \qquad (6.17\text{ p})$$

The second focal length is determined by Eq. (5.7 p).

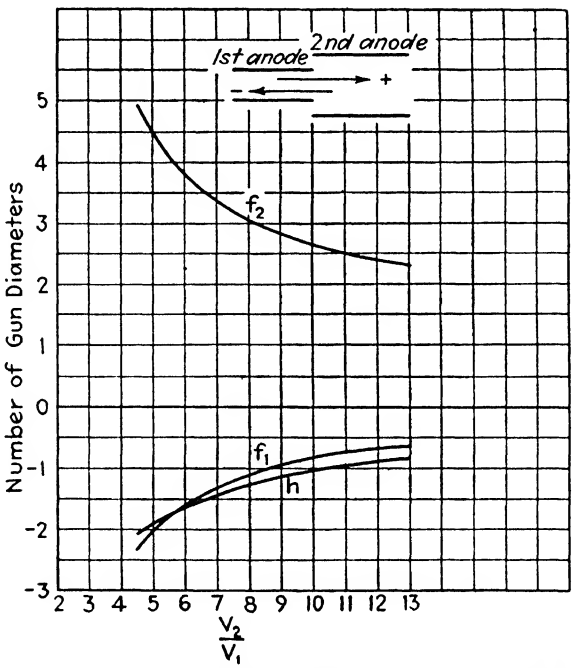


FIG. 6.18.—Position and focal lengths of thin lens *vs.* voltage ratio.

It is thus seen that to obtain the position of the thin lens and the two focal lengths f_1 and f_2 it is merely necessary to know: (1) d, the distance between the object (mesh) and the image (screen); (2) m, the magnification of the object; and (3) V_2/V_1, the voltage ratio required to focus the mesh on the screen.

Using the data of Fig. 6.13, the position and focal lengths of the thin lens were calculated by means of the above relations. Figure 6.18 gives the position and the focal lengths as functions of the voltage ratio. The position of the thin lens is given by the distance h between the lens and the end of the first anode.

Comparing Figs. 6.14 and 6.18, it is seen that the thin lens is situated between the two principal planes of the thick lens.

6.7. Immersion Lens.—The electrode system of the immersion lens of TCR tubes usually consists of a cathode, grid (sometimes also a screen grid) and first anode. Figure 6.19 shows two equipotential line plots of a simple electrode system of an immersion lens, the only difference between the two plots being in the

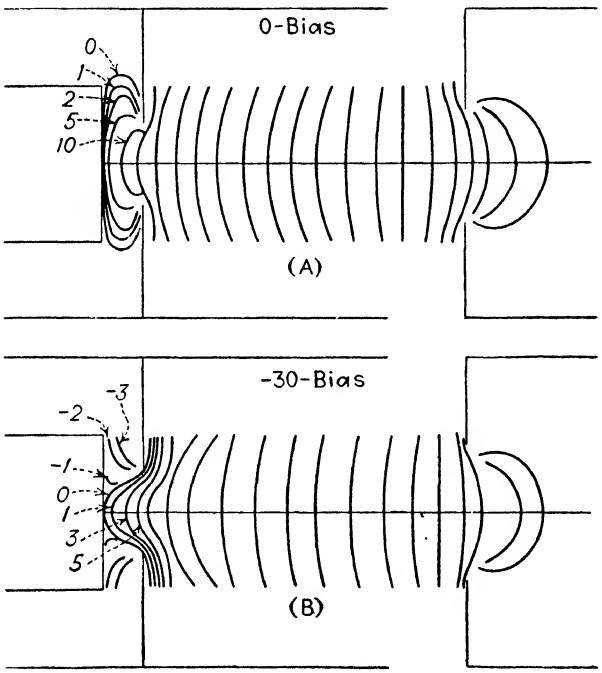


FIG. 6.19.—Equipotential lines of immersion lens.

voltage on the grid. Figure 6.19 shows how the grid voltage controls the current in the electron beam of a TCR tube. Thus Fig. 6.19(*B*) shows that a negative voltage on the grid permits only the central portion of the cathode to emit electrons, since the negative potential near the outer portions of the cathode forces the electrons emitted by these portions to return to the cathode. Further, the potential gradient near the center of the cathode is greatly reduced by the negative grid voltage, thus reducing the number of electrons emitted by the central portion of the cathode.

The position of the cardinal points of the immersion lens can be determined in the same manner as in the case of the bipotential lens; *i.e.*, by determining the two fundamental trajectories. However, there is a difficulty involved in this case which was not met in the case of the bipotential lens; *viz.*, the fact that the electrons incident on the immersion lens are not all moving with the same speed. The position of the cardinal points will, of course, depend upon the particular initial speed (this follows from the facts that the index of refraction of a point is proportional to the speed of the electron at the point and that the ratio of the focal lengths of a lens is equal to the ratio of the indexes of

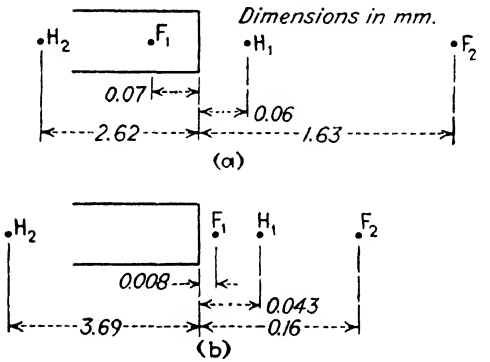


FIG. 6.20.—(*a*) cardinal points of immersion lens of Fig. 6.19(*A*). (*b*), cardinal points of immersion lens of Fig. 6.19(*B*).

refraction on the two sides of the lens). This corresponds to chromatic aberration in the case of light.

Figure 6.20 shows the positions of the cardinal points of the immersion lenses corresponding to the electrostatic fields of Fig. 6.19, and for an initial speed of the electrons of 0.2 equivalent volt. From Fig. 6.20(*b*) it is seen that the first focal length is 0.0035 cm., while the second focal length is 0.385 cm., so the ratio of the focal lengths (which is also the ratio of the indexes of refraction on the two sides of the lens) is 0.385:0.0035 = 110. Such a ratio of indexes of refraction is never met in ordinary optics. As might be expected, this large ratio produces rather unusual results.

From the positions of the cardinal points as given in Fig. 6.20(*b*) it may be deduced that the immersion lens images the cathode at 1.7 cm. in front of the cathode and with a magnifica-

tion of 4.4. Figure 6.19(B) shows that a circular area of cathode having a diameter of about 0.5 mm. is emitting at the given grid bias. Thus at a plane 1.7 cm. in front of the cathode there occurs the image of the emitting portion of the cathode, the diameter of the cathode image being 2.2 mm. If it be supposed that the bipotential lens focuses on the screen this image of the emitting portion of the cathode, it is found that the size of the minimum spot on the screen would be about 10 mm. Actually, the size of the measured minimum spot is less than 1 mm.

Whenever the position and size of the object that is imaged on the screen for a minimum spot are calculated from the constants of the bipotential lens, it is found that the object is located very

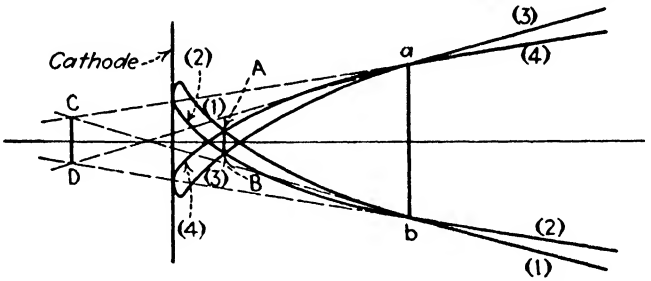


FIG. 6.21.—Beam crossover in proximity of cathode.

near the actual cathode and is considerably smaller than the emitting portion of the cathode. It is, therefore, to be concluded that, for a minimum spot, the bipotential lens does not focus the image of the cathode produced by the immersion lens. The question then naturally arises as to what it is that the bipotential lens focuses on the screen when the spot is a minimum.

To answer the above question, consider Fig. 6.21. Curve 1 of Fig. 6.21 shows the path of an electron having the initial radial velocity of +0.45 equivalent volt and zero axial velocity. Curve 2 shows the path of an electron of initial radial velocity −0.45 equivalent volt and zero axial velocity. Curves 3 and 4 represent the corresponding paths of electrons emitted by a point on the cathode below the axis. Now note that through point A pass electrons 1 and 4. Similarly, electrons of suitable velocities coming from every point of the cathode will pass through point A. Point A is, therefore, a new source of electrons; similarly with point B and with every other point of the cross section AB. The cross section AB is thus a new source of electrons and will be

called the *crossover*, for lack of a better name. Thus AB may be considered as an object to be imaged, and the position and size of the image can be found by extending rays 1 and 4 backward until they meet in C; then C is the virtual image of A. Similarly, D is the virtual image of B. It is this virtual image of the crossover that is focused on the screen when the spot is a minimum. It is to be noted that for the particular case of Fig. 6.21 the virtual image of the crossover is smaller than the emitting portion of the cathode and is located behind the cathode. Hence a much smaller spot is obtained by focusing on the screen the virtual image of the crossover instead of the image of the cathode, which is shown as *ab* on Fig. 6.21.

The reasons for the crossover—which is not ordinarily met in optics—may be understood from the following elementary considerations. The relation between the positions of the object and image in the simple case of refraction of a paraxial beam by a spherical surface may be written as

$$\frac{\mu_2}{v} - \frac{\mu_1}{u} = \frac{\mu_2 - \mu_1}{r}$$

where μ_2 and μ_1 are the indexes of refraction on the two sides of the surface, r is the radius of the surface, u is the distance between the object and refracting surface and v is the distance between the image and refracting surface.

The focal point of such a surface is defined as that point to which all the rays coming from an object at an infinite distance converge. The focal point is thus obtained by letting $\mu_1/u = 0$. It is clear, however, that μ_1/u can be zero, or negligible, with respect to the term μ_2/v either from u being very large, which is the normal case, or from μ_1 being very small. In the latter case it is to be noted that all the rays coming from an object at any distance from the surface will be focused very near the focal point.

For explanation purposes the immersion lens may be considered as consisting of two refracting surfaces, the first surface being convergent and the ratio of indexes of refraction being very large. This surface will then focus all the electrons coming from the cathode at points very close to the focal point of this surface and thus produce the crossover. The crossover occurs in front of the second refracting surface, which is divergent and, having a

much smaller ratio of indexes of refraction, produces the virtual image of the crossover.

It is interesting to follow what is being focused on the screen as the second anode voltage is varied, everything else remaining unchanged and in the normal operating condition. If the second and first anodes are at the same potential, then there is no bipotential lens and the immersion lens produces a bad inverted image of the cathode on the screen. As the voltage of the second anode is slightly increased there appears a bipotential lens, and this lens will focus a cross section of the electron beam (as seen by the bipotential lens)[1] very distant from the second lens. As the voltage on the second anode is further increased, the bipotential lens becomes stronger, and the distance between the cross section being imaged on the screen and the lens is decreased. As this distance is decreased, the diameter of the cross section being imaged decreases and the spot on the screen decreases. This continues until the image on the screen is that of the virtual image of the crossover. At this position the diameter of the cross section is a minimum, and so the spot on the screen is a minimum.[2]

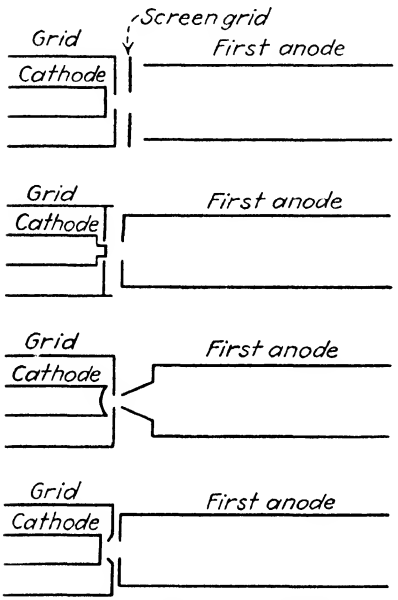


FIG. 6.22.—Electrode systems of immersion lenses used in TCR tubes.

As the second anode voltage is further increased, the focal lengths of the bipotential lens further decrease, and the spot on the screen increases since the cross section of the beam being

[1] By the electron beam as seen by the bipotential lens is meant the beam obtained by extending backwards indefinitely lines having the direction of the electrons just entering the lens.

[2] It may happen that by focusing a slightly larger cross section of the beam than the minimum, but farther removed from the bipotential lens, the spot on the screen will be smaller due to the decreased magnification. The importance of this effect depends on how rapidly the size of the cross section varies with distance.

imaged increases. The image on the screen, resulting from a further increase in second anode voltage, is that of the image of the cathode produced by the immersion focusing field. This image, however, is erect and is much better than the inverted one produced when the second anode voltage is nearly that of the first anode.

The discussion so far has dealt specifically with the immersion lenses corresponding to the electrostatic fields and electrode system shown in Fig. 6.19. The above discussion applies, to a large extent, also to the other forms of immersion lenses used in TCR tubes. The electrode systems of several such immersion lenses are shown in Fig. 6.22. As described above, the immersion lens is used for forming a crossover and its image, the latter being focused on the screen by the bipotential lens as the minimum spot. It is also possible to use an immersion lens for focusing on a small aperture most of the electrons emitted by the cathode. This aperture is then focused on the screen by the bipotential lens as the minimum spot. The size of the aperture has to be about 0.1 mm. to produce a spot sufficiently small for television purposes.

CHAPTER 7

DEFECTS OF ELECTRON-FOCUSING SYSTEM OF TCR TUBES

7.1. Types of Defects.—By defects of electron-focusing systems is meant the departures of focusing action of these systems from the relatively simple theory that has been presented thus far. The defects may be roughly classified as:

1. Those associated with construction.
2. Those associated with aberrations.
3. Those associated with space charge or mutual repulsion between the electrons in the beam.

The defects of the first class may be caused by:

a. Cylinders and apertures not being coaxial.

b. Cylinders and apertures being out of round.

c. Cathode emitting plane and grid plane not being coplanar, etc.

These and other similar defects may all be overcome by a more careful construction of the tube. In what follows it will be assumed that the tube is well constructed.

7.2. Aberrations.—In optics it is customary to speak of first, third, fifth, etc., order of imagery. Paraxial imagery is synonymous with first-order imagery. Third-order imagery is usually considered as first-order imagery upon which are superimposed third-order monochromatic aberrations. Similarly, with the fifth and higher orders of imagery. The number and complexity increase rather rapidly with the order of imagery, so that the usual treatment of the aberrations limits itself to the aberrations of the third order.

Similarly, one may speak of first and higher orders of imagery in electron optics. Thus one may speak of the various orders of imagery according to the number of terms that is used in the expansion for the potential (Eq. 4.4) and its derivatives (Eqs. 4.5 and 4.6). For first-order or paraxial imagery, terms containing higher powers of r and dr/dz than the first are neglected (see Eqs. 4.4 p and 4.9 p). For third-order imagery, all terms con-

taining higher powers of r and dr/dz than the third are neglected. Hence for third-order imagery, Eqs. (4.4), (4.5) and (4.6) reduce to

$$V(r,z) = V_0(z) - \frac{r^2}{2^2}V_0''(z) \tag{7.1}$$

$$\frac{\partial V(r,z)}{\partial r} = -\frac{r}{2}V_0''(z) + \frac{r^3}{2^2 4}V_0^{(4)}(z) \tag{7.2}$$

$$\frac{\partial V(r,z)}{\partial z} = V_0'(z) - \frac{r^2}{2^2}V_0(z). \tag{7.3}$$

By expanding $1/V(r,z)$ into a power series in r and neglecting terms containing powers of r greater than the third, there results

$$\frac{1}{V(r,z)} = \frac{1}{V - \dfrac{r^2}{2^2}V''} = \frac{1}{V} + \frac{r^2}{2^2}\frac{V''}{V^2} \tag{7.4}$$

where, as before, V stands for $V_0(z)$. Inserting Eqs. (7.2), (7.3) and (7.4) into Eq. (4.17), it results that the trajectory in the meridian plane of an electron of third-order imagery is satisfied by the differential equation

$$\frac{d^2r}{dz^2} + \frac{V'}{2V}\frac{dr}{dz}\left[1 + r^2\left(\frac{V''}{4V} - \frac{V^{(3)}}{4V'}\right) + \left(\frac{dr}{dz}\right)^2\right] + \frac{V''r}{4V}\left[1 + r^2\left(\frac{V''}{4V} - \frac{V^{(4)}}{8V''}\right) + \left(\frac{dr}{dz}\right)^2\right] = 0. \tag{7.5}$$

If the electron trajectory is not in a meridian plane (plane containing the axis of the system) then Eq. (7.5) is not sufficient for the determination of the trajectory. The calculation of skew trajectories (trajectories resulting from the fact that the electron has a component of velocity not in the meridian plane) requires a three-dimensional analysis and is normally too complicated for practical purposes. Fortunately, oblique trajectories play a relatively unimportant part in the operation of TCR tubes.

There are five monochromatic aberrations resulting from third-order imagery. Although these aberrations are largely interdependent, it is customary to speak of them as independent defects. These aberrations are enumerated below and some of their characteristics given:

1. *Spherical aberration* signifies that all electrons coming from an object point on the axis do not recombine to form a point

image. As this aberration is of great importance in the bipotential lens, a whole section will be devoted to it later.

2. *Coma* is of no importance if the object point lies on the axis, but becomes important for object points at short distances from the axis. The image of an object point off the axis is principally formed by the skew electrons, which may form rather unexpected patterns. The patterns usually resemble a comet, suggesting the name *coma* for this type of aberration. Coma increases as the distance of an image point from the axis and as the square of the aperture.

3. *Astigmatism* is characterized by the fact that the electrons originating from a single object point at a considerable distance off the axis are focused to two lines at right angles and at different distances from the object. The two lines are called the *sagittal* and *tangential foci*. At the sagittal focus, the cross section of the electron bundle is stretched out in the direction towards the axis, whereas at the tangential focus the stretching is in a direction at right angles to the plane containing the axis and object point. The astigmatic difference in focus increases as the square of the distance of an image point from the axis but is independent of the aperture. The length of the focal lines increases as the first power of the aperture and as the second power of the distance of the image point from the axis.

4. *Curvature of the field* causes the image of a plane object to be formed upon a curved surface rather than on a flat surface.

5. *Distortion* causes a displacement of the image point from the axis. If the inner parts of the object are imaged on a larger scale than the outer, it is known as *positive* or *barrel distortion*. If the outer parts are imaged on a larger scale than the inner, it is known as *negative* or *pincushion distortion*. The amount of distortion is independent of the size of the aperture but increases as the cube of the distance of the image point from the axis.

Of the five monochromatic aberrations enumerated above, it is only the first three that affect the definition of the image points; the last two affect only the position of the image point.

All the aberrations just described are present in the immersion lens. However, as the immersion lens in TCR tubes is not used to focus the cathode (as it is used in the electron microscope), but rather to form the crossover, the effect of the aberrations is not so serious. The aberrations do, however, increase the size

of the crossover. The determination of the exact amount of increase of the size of the crossover and its axial extension due to the aberrations is a difficult task which has, as yet, not been completed.

The size of the object that is focused by the bipotential lens of a TCR tube is usually about 0.1 mm. If the gun is well lined-up, one may consider this small object to be wholly on the axis. As the only type of aberration displayed by a lens in imaging an axial object point is spherical aberration, one may assume that the bipotential lens of well-constructed TCR tubes displays spherical aberration only. If the gun is not well lined up, the small object will lie off the axis and electrons coming from this object will strike the bipotential lens obliquely. In this case, the bipotential lens will also display coma and astigmatism in imaging the object. Coma and astigmatism may be present even though no spherical aberration is present. The spherical aberration of the bipotential lens will be treated in the next section.

Until now it has been assumed that all the electrons are moving with the same velocity. If they are not, further distortions arise. The distortion caused by the velocity distribution of the electrons is known as *chromatic aberration.* It is sometimes necessary to distinguish between two types of chromatic aberration. The displacement of an image along the axis because of a change in the initial velocity of the electrons is known as *axial chromatism,* while the variation in size of the image is known as *lateral chromatism.*

The chromatic aberration of the bipotential lens resulting from the Maxwellian distribution of velocities of the electrons emitted from the cathode is negligible. It may happen, however, that secondary electrons, resulting from the impact of the primary beam striking some part of the gun (aperture, wall, etc.), are also focused by the second lens. In this case there will be considerable chromatic aberration in the second field. As will be shown later, this chromatic aberration is easily avoided.

The chromatic aberration of the immersion lens is considerable and is due to the Maxwellian distribution of the speeds of the electrons emitted from the cathode. An idea of the chromatic aberration of this lens may be had from Fig. 7.1, which shows the variation of the first focal length f_1 with the initial speed of the

electrons. Figure 7.1 was determined for the electrostatic field of Fig. 6.19(*B*). An interesting fact is that the effect of the initial speeds on chromatic aberration is to alter the cardinal points of the object space only; the cardinal points of the image space, *viz.*, F_2 and H_2, are not appreciably altered.

Figure 7.2 shows the variation of the position of the cathode image with the initial speed of electrons.

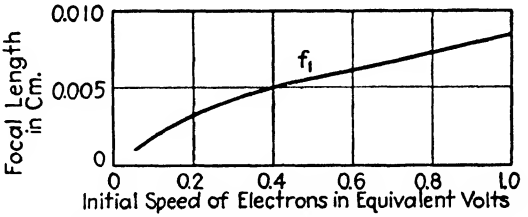


FIG. 7.1.—Variation of first focal length with initial speed of electrons.

7.3. Spherical Aberration of Bipotential Lens.—A discussion of the first-order or paraxial effects of the bipotential lens is an approximate description of the imagery, which is very useful for most purposes but is not sufficiently complete to serve as a basis for the final design of a TCR tube. The assumptions underlying paraxial imagery are true only if the aperture of the

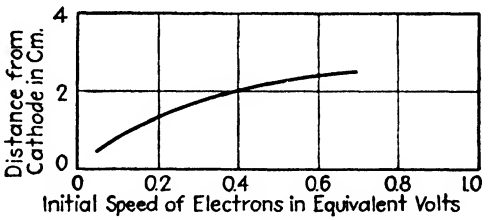


FIG. 7.2.—Variation of position of cathode image with initial speed of electrons.

lens and the size of the object are very small. To obtain the required current in the beam, it is necessary to have quite a large aperture. Therefore, the imagery actually existing in cathode-ray tubes departs from paraxial. For a final design it is therefore necessary to know the aberrations of the bipotential lens. As mentioned in the preceding section, the only aberration that need be considered in a well constructed TCR tube is spherical aberration.

The task of theoretically determining the spherical aberration of the bipotential lens is extremely difficult, and in the end is

not so useful as the experimental determination. Further, in the experimental determination of the spherical aberration it is not necessary to limit oneself to third-order imagery.

In the presence of spherical aberration, all electrons coming from an object point on the axis do not recombine at one point on the axis—as paraxial theory predicts—but rather intersect the axis at various distances as shown in Fig. 7.3. In Fig. 7.3, I represents the paraxial image of O; IL_1 and IL_2 are defined as the longitudinal spherical aberrations; and IT_1, IT_2 are defined as the transverse spherical aberrations for the various apertures.

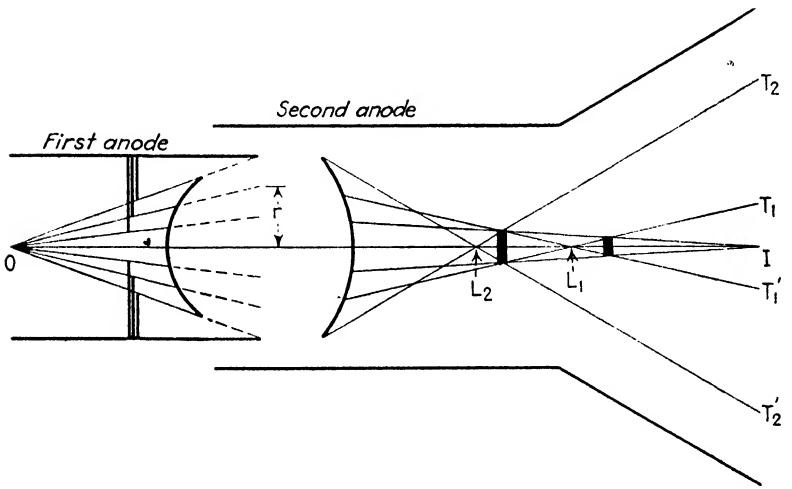


FIG. 7.3.—Imaging in the presence of spherical aberration.

The longitudinal spherical aberration is said to be positive, if, as in Fig. 7.3, L_1 and L_2 are to the left of I, and negative if to the right of I. The transverse spherical aberration is said to be positive for electrons T_1 and T_2 and negative for T_1' and T_2'. The heavy vertical lines in the image space represent the disks of least confusion.

Let r be the distance between any electron and the axis at the end of the gun (see Fig. 7.3), and let L be the longitudinal spherical aberration for the electron of height r, then

$$L = a_2r^2 + a_4r^4 + a_6r^6 + \cdots \tag{7.6}$$

That L is a function of only the even powers of r follows from the fact that L is the same for plus or minus r; *i.e.*, L is the same for

an electron above or below the axis. Similarly, if T represents the transverse spherical aberration, then

$$T = a_1 r + a_3 r^3 + a_5 r^5 + \cdots \tag{7.7}$$

since a change in the sign of r changes the sign of T only.

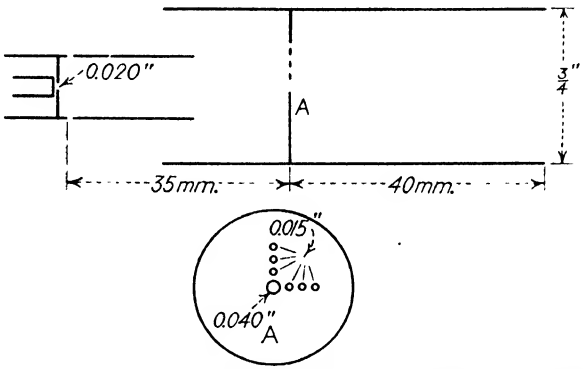


FIG. 7.4.—Gun useful in determination of spherical aberration.

It is to be noted that the effect of spherical aberration is perfectly symmetrical; *i.e.*, the disks of least confusion, which represent the smallest spots possible with the given amount of aberration, are perfectly symmetrical about the axis.

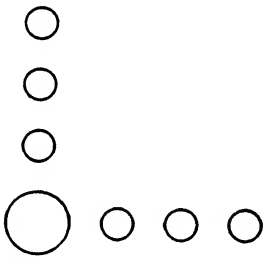


FIG. 7.5.—(*a*), appearance of unfocused spots.

FIG. 7.5.—(*b*), appearance of spots for paraxial focus.

Experimentally, it is simpler to determine the transverse spherical aberration, and Fig. 7.4 shows the type of gun useful in the determination of that type of aberration. Figure 7.5(*a*) represents the appearance of the spots on the screen when $V_2 = V_1$. If there were no spherical aberration, all the spots would, on focusing, unite to form the small paraxial-image spot.

Figure 7.5(*b*) shows the appearance of the screen when the electrons issuing from the central 0.040-in. aperture are focused. In Fig. 7.5(*b*) it is seen that, owing to spherical aberration, the non-paraxial electrons are already over-focused when the axial electrons (those passing through the central 0.040-in. aperture) are just focused. The transverse spherical aberration is thus directly obtained from Fig. 7.5(*b*).

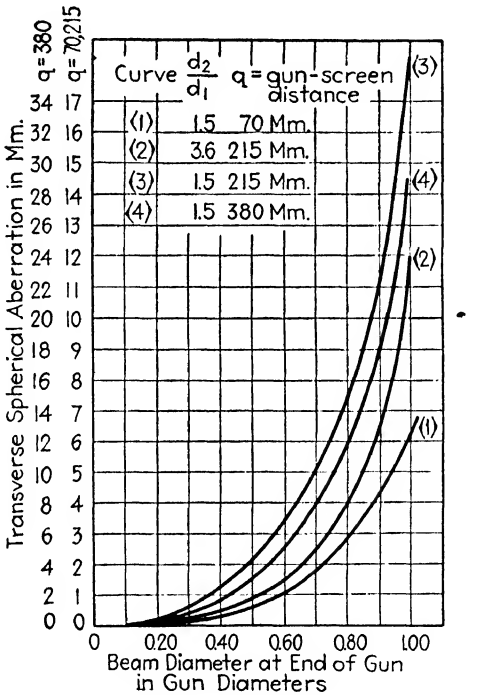


FIG. 7.6.—Transverse spherical aberration *vs.* beam diameter.

Figure 7.6 shows how the transverse spherical aberration for various image distances and diameter ratios varies with the width of the beam at the end of the gun. The curves of Fig. 7.6 may be fairly well represented by the first two terms of Eq. (7.7). The width of the beam at the end of the gun is obtained from Fig. 7.5(*a*) and from the given positions of aperture and screen.

As shown in Fig. 7.3, the plane containing the disk of least confusion occurs nearer the lens than the position of the image as predicted by paraxial theory. If the focusing voltage ratio

is set for the paraxial image, then the disk of least confusion will occur between the end of the gun and the screen; in other words, *i.e.*, the spot on the screen is over-focused. To focus the disk of least confusion on the screen, it will be necessary to lower the voltage ratio. Figure 7.7 shows the variation of the voltage ratio required to focus the disk of least confusion for various

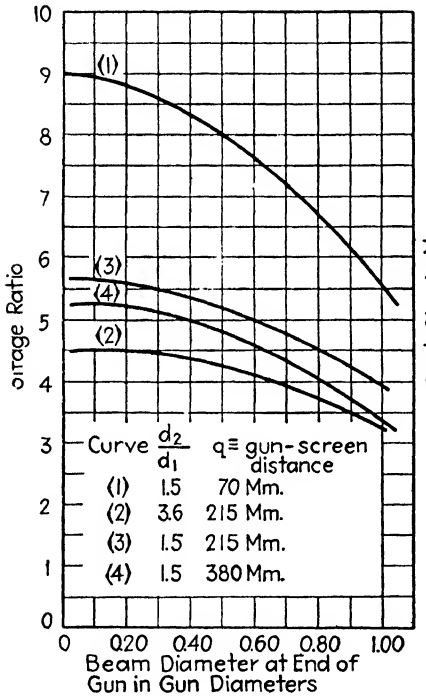


FIG. 7.7.—Focusing voltage ratio *vs.* beam diameter.

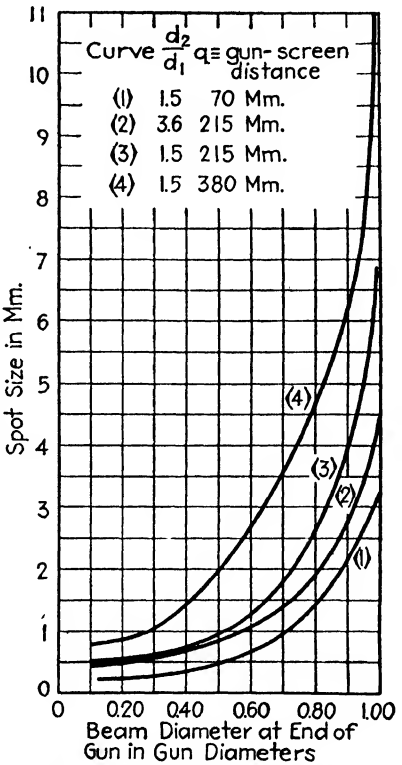


FIG. 7.8.—Minimum spot size *vs.* beam diameter.

beam widths. The voltage ratio required to focus the disk of least confusion and the size of the disk were determined by simultaneously focusing to a minimum spot the beams passing through the central 0.040-in. aperture and two or more of the other apertures, considering the center of the central spot (spot formed by beam through 0.040-in. aperture) as the center of the disk of least confusion.

The size of the disk of least confusion represents the minimum spot obtainable. Figure 7.8 shows the variation of the size of the disk of least confusion with the ratio of the beam diameter at the end of the gun to gun diameter.

Figure 7.9 shows the same curves as given in Fig. 7.8, plotted in terms of the paraxial image size. Figure 7.9 shows the very interesting result that the disk of least confusion, measured in terms of the paraxial size of the image, is independent of the image distance, *i.e.*, of the type of tube.

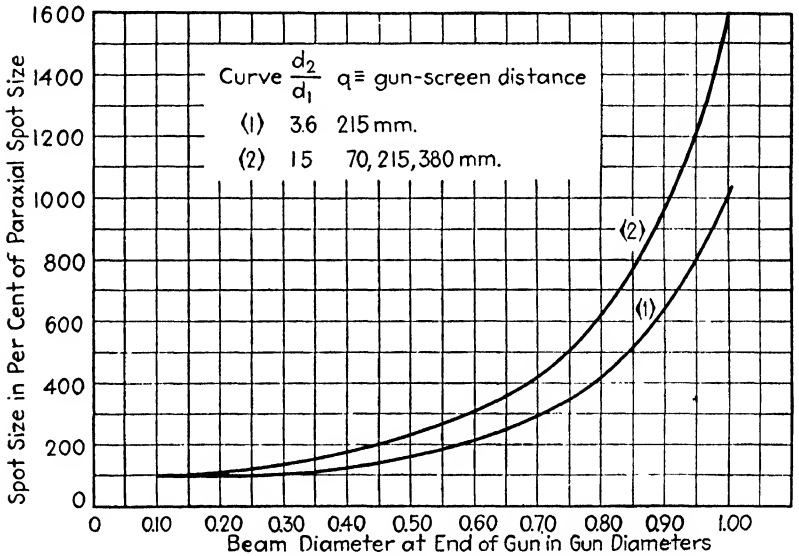


FIG. 7.9.—Minimum spot size *vs.* beam diameter.

7.4. Space Charge in Region Near Cathode.—The focusing action of electronic lenses is complicated by the fact that the density of the electrons in the beam may become sufficiently large to appreciably alter the focusing properties of the lens. This is especially true in the region near the cathode where the electron velocities are low.

If the space charge is negligible, the path of an electron can be determined from an equipotential-line plot, such as shown in Fig. 6.19. If, however, the space charge is not negligible, the potential distribution in space is no longer given by Fig. 6.19 but becomes a modified distribution. To find this modified distribution, it is necessary to find the potential distribution in

space due to the space charge and the charges induced on the electrodes by the space charge, and add this to the distribution shown in Fig. 6.19. The path of an electron through this modified field is then the actual path followed by an electron in the presence of the space charge.

The exact calculation of the potential distribution in space due to the space charge would, of course, require that the potential at a point due to each electron in the region be found and added, this being done for all the points for which the potential is required. It can be shown that for the electron concentrations normally encountered it is not necessary to consider the individual electrons in order to calculate the potential at a point due to them, but that it is sufficiently accurate to consider a uniform distribution of charge, the volume density of charge ρ being a statistical average of the charge density existing in a small, but not too small, volume element.

It is unnecessary to consider any collision between electrons in the calculation of the effects of space charge. The greatest density of electrons occurs near the cathode, and for the normal currents drawn in TCR tubes there are about 2×10^{11} electrons/c.c. For this purpose the electrons may be considered as cubes with sides measuring 4×10^{-13} cm. The volume occupied by the electrons for 1 c.c. of space is thus

$$2 \times 10^{11} \times 64 \times 10^{-39} = 1.28 \times 10^{-26}.$$

The mean free path, then, is of the order of 10^7 *miles*. In other words, on the average, an electron will travel 10^7 miles before it encounters another.

Although the exact calculation of the effect of space charge in the cathode region is very nearly impossible, an approximate solution of the problem may be obtained by the following procedure:

1. Neglecting space charge, the shape of the beam near the cathode is obtained from the envelope of electron paths, or the shape of the beam may be determined experimentally. Figure 7.10 shows the shape of the beam as determined by the envelope of electron paths through the field of Fig. 6.19(*b*).

2. From the equipotential-line plot and the initial velocities of the electrons, the velocity of any electron at any point is calculated. Knowing the velocity of the electrons, the shape of

the beam and the current in the beam, the density of the charge at any point can then be calculated. Figure 7.11 gives the charge density along the axis for the beam shown in Fig. 7.10.

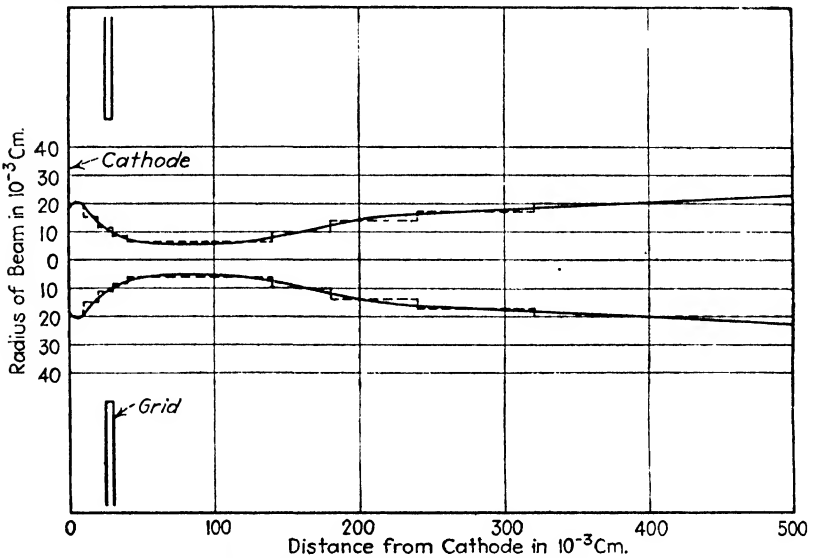


FIG. 7.10.—Envelope of electron beam.

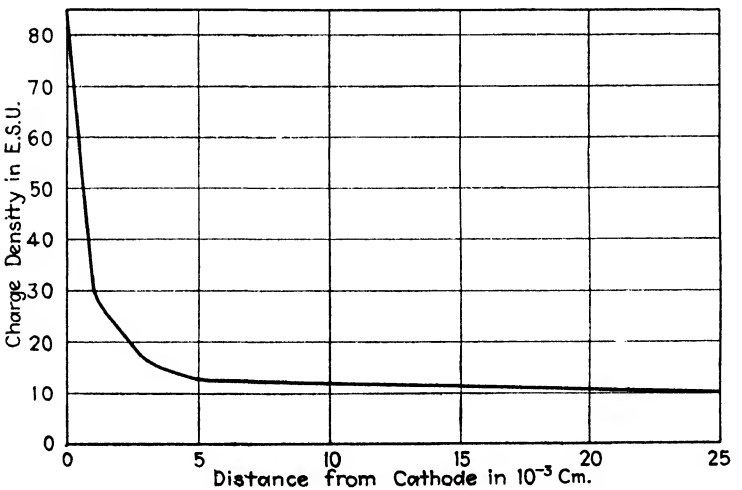


FIG. 7.11.—Charge density along axis of electron beam.

3. From the charge density and the shape of the beam, the potential at any point due to the beam is determined. Curve 1,

Fig. 7.12, gives the calculated potential along the axis due to the beam of Fig. 7.10.

4. The charges induced on the electrodes are then calculated, and the potential distribution in space due to these induced charges is obtained. Curve 2, Fig. 7.12, gives the potential along the axis due to the charges induced on the electrodes by the space charge.

5. The difference between the potential distributions obtained in 3 and 4 gives the distribution that is to be subtracted from the Laplacian distribution to obtain the field as modified by the

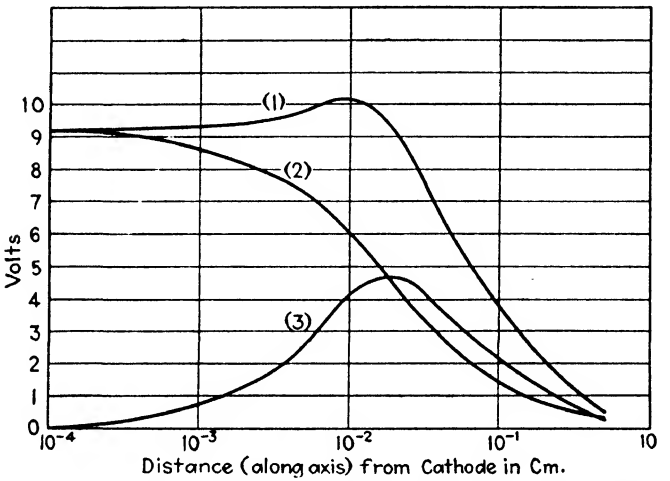


FIG. 7.12.—Curve (1), potential along the beam axis due to beam itself. Curve (2), potential along the axis due to induced charges. Curve (3), difference between curves (1) and (2) (half scale).

space charge. Curve 3, Fig. 7.12, is the potential along the axis obtained as the difference between curves 1 and 2.

The solid lines of Fig. 7.13 show the potential distribution without space charge, *i.e.*, the distribution of Fig. 6.19(*b*), and the dotted lines give the distribution after the space charge has been taken into account.

A better approximation may be obtained by repeating the above procedure, starting with the modified distribution (as shown with dotted lines in Fig. 7.13) instead of the original distribution (as shown by solid lines in Fig. 7.13).

Some of the problems met in the above procedure will now be discussed. In calculating the density of charge near the cathode,

one is confronted with the problem of the potential minimum. The value of the potential minimum is very easily obtained from the ratio of the total beam current (first and second anode current) to the total emission current and the normal velocity distribution of the emitted electrons. Thus, if one knows that half of the current emitted is returned to the cathode, then it is merely necessary to find that negative voltage which is able to stop 50 per cent of the emitted electrons. This voltage can be obtained from Eq. (2.7), since what is sought is that value of V_1 for which all electrons emitted with velocities between $V_2 = \infty$

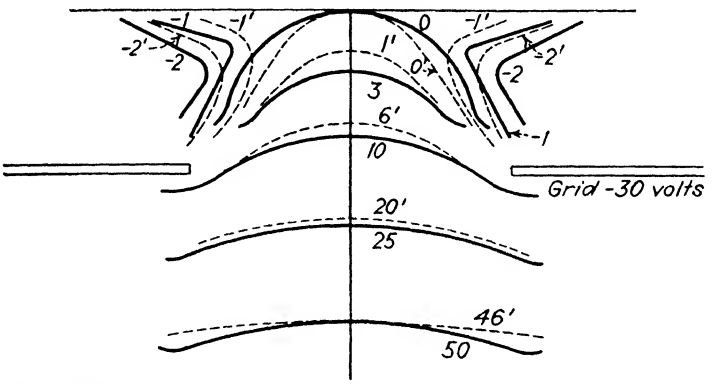


FIG. 7.13.—Effect of space charge on potential distribution of immersion lens.

and $V_1 = V_{min}$ will constitute 50 per cent of the total of emitted electrons. Thus Eq. (2.7) reduces to

$$\frac{\Delta N}{N} = \frac{I_{beam}}{I_{emitted}} = 0.5 = \epsilon^{-\frac{1.16 \times 10^4 V_{min}}{T}}. \tag{7.8}$$

The value of V_{min} can be obtained from Table 2.1. Thus for a cathode at 1200°A., about -0.07 volt will return 50 per cent of the emitted electrons.

As an approximation, the position of V_{min} can be calculated by considering the cathode and anode as infinite planes. This yields the result that the potential minimum is located less than 0.001 cm. in front of the cathode. Such a thin disk of charge existing between the potential minimum and the cathode will produce a negligible potential in space. This is true (in spite of the fact that the density of charge of this disk is greatest) because

the disk is so close to the cathode that the charges induced on the cathode by the disk almost annihilate its effects.

Thus in Fig. 7.14, curve 1 gives the potential distribution along the axis due to a disk with a charge 10^{-3} cm. thick, having a radius of 20×10^{-3} cm. and located next to the cathode. Curve 1′ gives the potential distribution due to the charges induced on the cathode by this disk. It is seen that curves 1 and 1′ nearly coincide. Their difference is therefore very small.

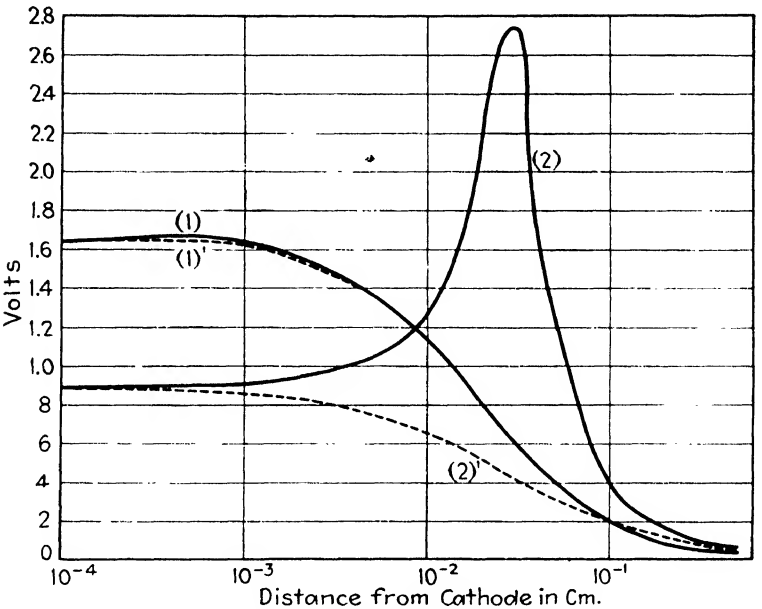


FIG. 7.14.—Curve (1), potential distribution along axis due to a disk of charge located near cathode. Curve (2), potential distribution along axis due to charges induced on cathode by disk. Curve (2)′, potential produced by the charges induced on cathode by the cylinder of charge producing curve (2).

The existence of a potential minimum does not destroy the Maxwellian distribution of the electrons emitted from the cathode. To show that this is so, it is merely necessary to rewrite Eq. 2.2 in terms of voltage instead of velocity, so that

$$dN_V = \frac{Ne}{300kT} \epsilon^{-\frac{eV}{300kT}} dV. \tag{7.9}$$

Now if a potential minimum is present

$$V = V_1 + V_{\min} \qquad \text{and} \qquad dV = dV_1$$

where V_1 corresponds to the velocity of the electrons after passing through the potential minimum V_{min}. An electron having a velocity less than that corresponding to V_{min} will not pass through. Equation (7.9) can therefore be rewritten as

$$dN_{V^1} = \frac{Ne}{300kT} \epsilon^{-\frac{eV_{min}}{300kT}} \epsilon^{-\frac{eV_1}{300kT}} dV_1$$

or

$$dN_{V_1} = \frac{N'e}{300kT} \epsilon^{-\frac{eV_1}{300kT}} dV_1. \tag{7.10}$$

Equation (7.10) is of exactly the same form in V_1 as Eq. (7.9) is in V. Therefore the velocity distribution of the electrons after passing through the potential minimum is exactly the same as that of the electrons emitted by the cathode.

However, it is seen from Eqs. (7.9) and (7.10) that

$$N' = N\epsilon^{-\frac{eV_{min}}{300kT}}$$

i.e., the potential minimum has decreased the number of electrons getting through.

Thus it may be assumed that no potential minimum exists, that the cathode emission is less (the emission corresponding to current in the beam), but that all the electrons emitted are drawn into the beam.

To calculate the density of charge near the cathode, one may, therefore, start with the distribution given in Table 2.2. Only the normal velocity distribution need be taken into consideration, for the radial component will not contribute to the charge density.

Consider a cathode which has an emission of I abamp./sq. cm., then the charge density ρ in e.m.u. is given by

$$\rho = \frac{I}{v}$$

where v is the velocity of the electrons in centimeters per second. As the electrons near the cathode are not all moving with the same velocity, ρ will be given by

$$\rho = \frac{I_1}{v_1} + \frac{I_2}{v_2} + \frac{I_3}{v_3} + \cdots \frac{I_n}{v_n} \tag{7.11}$$

where I_n is that fraction of the current whose electrons are moving at the velocity v_n. This equation may be rewritten as

$$\rho = \rho_1 + \rho_2 + \rho_3 + \cdots \rho_n \tag{7.12}$$

where ρ_n is the contribution to the space charge by the electrons whose initial velocity is v_n.

Having the shape of the beam and the charge density along the beam (the density along a cross section of the beam may usually be assumed constant), one may calculate the potential in space by considering only small portions of the beam at a time. The small portions being so chosen that in each portion: (1) the beam is either a disk or cylinder; (2) the density along the axis is either constant or varies linearly; and (3) the density normal to the axis is either constant throughout or the disk or cylinder of charge can be obtained from the superposition of several disks or cylinders of various radii and charge densities.

The equations found most useful in the calculation of the potential distribution due to the space charge are:

1. The potential along the axis due to a cylinder of charge of constant density ρ, length $2l$ and radius a is given by

$$V = \pi\rho\Big[(l-x)\sqrt{(x-l)^2+a^2} + (l+x)\sqrt{(x+l)^2+a^2} - 4lx + a^2 \log \frac{\sqrt{(x-l)^2+a^2} + (l-x)}{\sqrt{(x+l)^2+a^2} - (l+x)}\Big] \tag{7.13}$$

where x is measured from the center of the cylinder.

2. The potential at any point (r,θ) due to a disk of total charge Q and radius a is

$$V = \frac{2Q}{a}\Big[\frac{1}{2}\frac{a}{r} - \frac{1\cdot 1}{2\cdot 4}\Big(\frac{a}{r}\Big)^3 P_2(\cos\theta) + \frac{1\cdot 1\cdot 3}{2\cdot 4\cdot 6}\Big(\frac{a}{r}\Big)^5 P_4(\cos\theta) - \cdots\Big] \qquad r > a \tag{7.14}$$

$$V = \frac{2Q}{a}\Big[1 - \frac{r}{a}P_1(\cos\theta) + \frac{1}{2}\Big(\frac{r}{a}\Big)^2 P_2(\cos\theta) - \frac{1\cdot 1}{2\cdot 4}\Big(\frac{r}{a}\Big)^4 P_4(\cos\theta) + \cdots\Big] \qquad r < a, \qquad \theta < \frac{\pi}{2}$$

where the P's are Legendre's polynomials.

3. The potential at any point (r,θ) due to a circular ring of small cross section of total charge Q and radius a is

$$V = \frac{Q}{a}\left[\frac{a}{r} - \frac{1}{2}\left(\frac{a}{r}\right)^3 P_2(\cos\theta) + \frac{1\cdot 3}{2\cdot 4}\left(\frac{a}{r}\right)^5 P_4(\cos\theta) - \cdots\right] \qquad r > a$$

$$= \frac{Q}{a}\left[1 - \frac{1}{2}\left(\frac{r}{a}\right)^2 P_2(\cos\theta) + \frac{1\cdot 3}{2\cdot 4}\left(\frac{r}{a}\right)^4 P_4(\cos\theta) - \cdots\right] \qquad r < a. \quad (7.15)$$

4. The potential at any point due to a line of charge of total charge Q and length l is

$$V = \frac{Q}{l}\log\frac{r + r' + l}{r + r' - l} \quad (7.16)$$

where r and r' are the distances from the point under consideration to the two ends of the line of charge.

5. The potential along the axis of a cylinder of charge of length $2l$, radius a, and density ρ which varies along the axis is

$$V_x = 2\pi\int_{-l}^{+l}\rho[\sqrt{(x - z)^2 + a^2} - |\, x - z\,|]dz \quad (7.17)$$

where x is the point at which the potential is desired, measured from the center of the cylinder and along the axis of the cylinder; z, also measured from the center of the cylinder and along the axis, is the variable of integration, ρ is the density of charge as a function of z only.

The potential distribution in space that is found in 3 above is that due to a distribution of charge approximately duplicating the beam and suspended in space with no other charges in the vicinity. According to the distribution in 3 it is found, for example, that for the case of Fig. 6.19(*b*) the potential at the center of the cathode is −9.25 volts, which means that the space charge has reduced the potential of that point of the cathode by 9.25 volts. Furthermore, the potential of different points of the cathode is reduced by different amounts. However, the potential of the cathode is kept constant by the battery. The battery accomplishes this by removing the required number of electrons from the cathode and leaving the cathode sufficiently positively charged so as to just neutralize the effect of the electrons in the space charge on the surface of the cathode. Exactly the same is true of the grid and other electrodes.

In determining the charges induced on the electrodes by the charges in the beam, it is therefore necessary to find that charge distribution on the electrodes which is just sufficient to neutralize the effect of the space charge on the electrodes. The only electrodes that are sufficiently close to the beam so that the induced charges are not negligible are the cathode and grid.

The distribution of charge on the cathode may be assumed to be the superposition of various disks of charge of various radii and surface-charge densities, the surface-charge density of each disk being constant and all disks having the same axis; similarly with

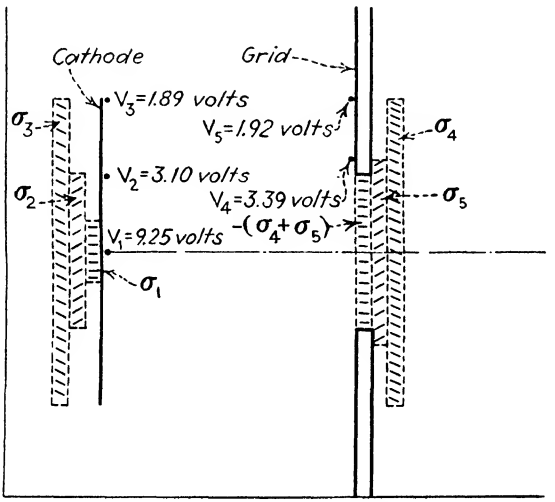


FIG. 7.15.—Approximate distribution of charge induced on cathode and grid by beam.

the grid. Thus Fig. 7.15 pictures the positive induced charges on the cathode to consist of three disks of charge of radii 20×10^{-3} cm., 50×10^{-3} cm. and 100×10^{-3} cm. and of surface-charge densities σ_1, σ_2 and σ_3, respectively. The induced charges on the grid are similarly assumed to consist of three disks of charge of radii 50×10^{-3} cm., 60×10^{-3} and 100×10^{-3} and of surface-charge densities $-(\sigma_4 + \sigma_5)$, σ_4 and σ_5, respectively. It is seen that the charge density of the aperture of the grid is $-(\sigma_4 + \sigma_5) + \sigma_4 + \sigma_5 = 0$.

Having calculated the surface-charge densities, the voltage distribution in space due to these disks of induced charges is then calculated. Curve 2 of Fig. 7.12 gives the potential distribution

along the axis due to these disks of induced charges. A good approximation to curve 2 is obtained by assuming the cathode to be an infinite plane and calculate the potential due to the image of the beam in the plane of the cathode. As the potential in space due to each cylinder in the beam has already been calculated, it is but necessary to change the sign of the potential to obtain the potential due to the image of the cylinder. The image is, however, located behind the cathode so that care has to be taken in adding up the potentials due to the various image cylinders.

Curve 1′ of Fig. 7.14 shows the potential along the axis due to the charges induced on the cathode by the disk of charge producing curve 1. Curve 2′ is the potential produced by the charges induced on the cathode by the cylinder of charge producing curve 2.

From Fig. 7.13 the effects of the space charge may be enumerated as:

1. Decreases the cathode area of emission.
2. Decreases the potential gradient along the axis in the region between cathode and grid.
3. Increases the curvature of the equipotential surfaces in the region between cathode and grid.
4. Slightly decreases the curvature of the equipotential surfaces in the region between grid and first anode.

The calculation of the paths of paraxial electrons through the dotted equipotential-line plot of Fig. 7.13 shows that the effect of space charge is to slightly increase the size of the object (crossover) that is focused on the screen, the increase in size calculated to be 15 per cent. Space charge increases the focal lengths of the immersion lens; thus the space charge increased the first and second focal lengths of Fig. 6.20(*b*) to 0.014 cm. and 1.5 cm., respectively.

7.5. Space Charge in Region of First and Second Anodes.—Whereas the different rays of a beam of light do not affect each other, the various electrons in a beam do influence each other. Two electrons moving in the same direction repel each other on account of their like charges and attract each other because of their magnetic fields. In TCR tubes the velocity of the electrons is sufficiently low so that the attraction due to their magnetic fields is negligible in comparison with the repulsion due to their

electrostatic fields. Like the aberrations discussed in the last section, the mutual electron repulsion acts to increase the minimum spot obtainable in TCR tubes. Unlike these aberrations, however, it acts throughout the length of the beam and depends upon the intensity of the beam.

Figure 7.16 illustrates an object being imaged by a perfect lens. The object of this section is to give means for determining the effect of the mutual repulsion on the electron trajectories indicated in Fig. 7.16. For the present it will be assumed that the beam consists of two portions of cones—one in the object space, with the velocity of the electrons corresponding to the

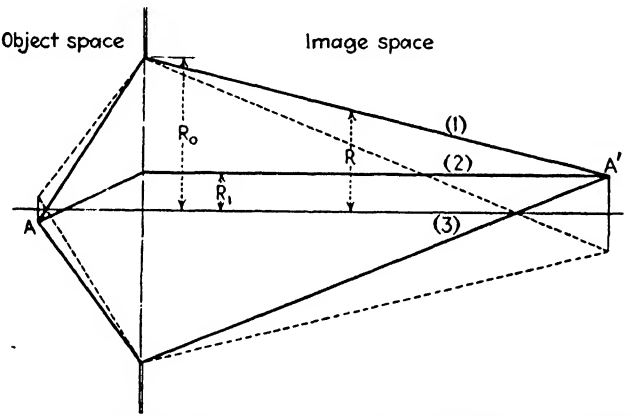


FIG. 7.16.—Paraxial electron focusing with negligible space charge.

first-anode voltage, and the other in the image space with the velocity of the electrons corresponding to the second-anode voltage. Hence the object is to find the trajectory of an electron moving inside or just outside of a frustum of cone of charge such as that formed by the beam in the object or image space. In the discussion which follows, it is assumed that the charge density is uniform along a cross section and that the electrons are moving so fast that any axial force that may be exerted by the space charge on an electron is negligible.

From the fact that the radial electrostatic force at a point inside a long cylinder of charge is independent of the radius of the cylinder, depending merely on the charge density and the distance of the point from the axis, it follows that to a sufficient degree of approximation the radial electrostatic force inside or

just outside the frustum of cone is the same as that inside a cylinder of charge, the density of which varies along the axis. The problem therefore reduces to that of finding the trajectory of an electron moving *inside* a cylinder of charge, the density of which varies along the axis.

To find the variation of density along the axis, note that the charge per unit length, Q, is

$$Q = \frac{I}{v}. \tag{7.18}$$

Since I, the current in the beam, and v, the velocity of the electrons, are constants, Q is constant. The charge density at any point z, where the radius of the cone (see Fig. 7.16) is R, is

$$\rho = \frac{Q}{\pi R^2}. \tag{7.19}$$

But from Fig. 7.16

$$R = R_0 - \frac{(R_0 - R_1)}{l} z = R_0 \left[1 - \left(1 - \frac{R_1}{R_0} \right) \frac{z}{l} \right].$$

So ρ as a function of z is given by

$$\rho = \frac{Q}{\pi R_0^2 \left[1 - \left(1 - \frac{R_1}{R_0} \right) \frac{z}{l} \right]^2}. \tag{7.20}$$

Under the assumptions made, Poisson's equation becomes, after a single integration,

$$\frac{dV}{dr} = -2\pi\rho r. \tag{7.21}$$

The equation of motion of an electron inside the cylinder thus becomes, after replacing z by vt,

$$\frac{d^2r}{dt^2} = \frac{\alpha r}{(1 - \beta t)^2} \tag{7.22}$$

where

$$\alpha = 2 \frac{e}{m} \frac{I}{v} \frac{C^2}{R_0^2} \tag{7.23}$$

and

$$\beta = \left(1 - \frac{R_1}{R_0} \right) \frac{v}{l} \tag{7.24}$$

e/m and I being expressed in e.m.u.; $C = 3 \times 10^{10}$.

If one lets

$$y = \log (1 - \beta t)$$

then Eq. (7.22) becomes, upon transformation,

$$\frac{d^2r}{dy^2} - \frac{dr}{dy} - kr = 0 \tag{7.25}$$

where $k = \alpha/\beta^2 =$ a constant.

The solution of (7.22) is therefore

$$r = \frac{1}{m_1 - m_2}\left[\left(m_1 r_0 + \frac{\dot{r}_0}{\beta}\right)(1 - \beta t)^{m_2} - \left(m_2 r_0 + \frac{\dot{r}_0}{\beta}\right)(1 - \beta t)^{m_1}\right] \tag{7.26}$$

where $m_1 = \dfrac{1 + \sqrt{1 + 4k}}{2}$ $\quad m_2 = \dfrac{1 - \sqrt{1 + 4k}}{2}$

and $\quad r_0 = r \quad$ and $\quad \dot{r}_0 = \dfrac{dr}{dt} \quad$ when $\quad t = 0.$

From Eq. (7.26) one can obtain the value of r at any time t. The value of dr/dt at any time t is obtained from

$$\frac{dr}{dt} = \frac{\beta}{m_1 - m_2}\left[m_1\left(m_2 r_0 + \frac{\dot{r}_0}{\beta}\right)(1 - \beta t)^{-m_2} - m_2\left(m_1 r_0 + \frac{\dot{r}_0}{\beta}\right)(1 - \beta t)^{-m_1}\right]. \tag{7.27}$$

To determine A' (Fig. 7.16), the image of A, it is necessary to know the path of several rays. Rays 1, 2 and 3 of Fig. 7.16 are suitable. Consider a tube with very little current so that little space charge exists, then all three rays will intersect each other at A'. Now as the beam current is increased the space charge will increase and will deflect rays 1 and 2 above A' and ray 3 below A'. To reunite them as much as possible, it will be necessary to increase the voltage ratio E_{p_2}/E_{p_1}. When the voltage ratio is altered the values of r_0 for rays 1 and 3 and the value of r_0 for ray 2 are altered. For a given change in E_{p_2}/E_{p_1} one can determine the changes in r_0 and $\dot{r}_0$ from the constants of the lens.

Knowing $\dot{r}_0$ and r_0 one can find, by means of Eq. (7.26), the value of r at the screen for each of the rays. The value of E_{p_2}/E_{p_1} is then varied until the value of r at the screen for the rays 1 and

2 is the same. This gives a good approximation of the minimum spot on the screen. Table 7.1 shows the calculated value of r at the screen for various values of E_{p^2}/E_{p^1}. It is seen that under the conditions indicated in the Table 7.1, the space charge of the image space increased the spot size from 0.060 to 0.068 cm.

TABLE 7.1

Ray	E_{p^2}/E_{p^1}					
	4:20		4:24		4:25	
	r_0, cm.	r, cm.	r_0, cm.	r, cm.	r_0, cm.	r, cm.
1	0.200	0.052	0.200	0.034	0.200	0.032
2	0.030	0.036	0.028	0.034	0.028	0.034
3	−0.200	0.014	−0.200	0.033	0 200	0.036

Diameter ratio = 2.28 Gun-screen distance = 21 cm
Initial beam diameter = 4 mm. I = 100 μa. V = 2,400 volts
E_{p^2}/E_{p^1} = 4.2 and r = 0.030 for all rays when space charge is negligible

The assumptions underlying Eq. (7.26) are valid only if the effect of space charge is not too large. If the space-charge effect is large, then the envelope of the beam departs sufficiently from that of the cone shown in Fig. 7.16 so that Eq. (7.26) yields wrong results.

For large space-charge effects it is necessary to use some method of successive approximations. Thus a beam envelope may be estimated from Eq. (7.26). This envelope is decomposed into a few sections, each section being considered as a cone or cylinder of charge. The paths of rays 1, 2 and 3 are then obtained so that rays 1 and 2 intersect at the screen. The path of ray 1 will usually not coincide with the beam envelope chosen, meaning that the wrong envelope has been assumed. A different beam envelope is then assumed, and the process repeated until the path of ray 1 and the beam envelope nearly coincide.

The solid lines of Fig. 7.17 give the trajectories of three electrons in the presence of space charge. The dotted lines give the trajectories of the two outer electrons when the space charge is negligible. The focusing-voltage ratio required to focus the electrons in the absence of space charge is 4.6, while that required in the presence of space charge is about 5.2.

It is seen from Fig. 7.17 that the presence of space charge has reduced the quality of the image. In general, the smaller the object the worse the quality of the image. Under the conditions indicated in Fig. 7.17 it was found that for object sizes less than 0.0005 cm. in radius very few electrons cross the axis (there is practically no real image formed).

The solid line of Fig. 7.18 gives the variation in the size of the image (size of spot on screen) and the size of object, under the

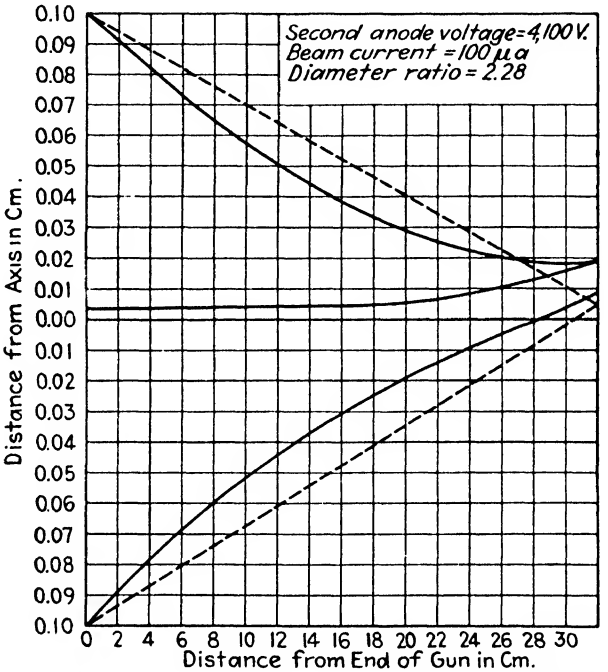


FIG. 7.17.—Effect of space charge on paraxial electron trajectories in image space.

condition indicated. The variation of the size of the image with that of the object, with no space charge present, is depicted by the dotted line of Fig. 7.18. The interesting thing to notice from Fig. 7.18 is that if the space charge be appreciable, the image size does not decrease proportionately with the object size, but less slowly. Thus in the particular case shown in Fig. 7.18, nothing is gained by decreasing the object size below 0.0028 cm. in radius (corresponding to a radius of 0.01 cm. of the image size without space charge).

The spreading of the beam in the object space due to the mutual electron repulsion produces a virtual object larger in size and closer to the end of the gun than the actual object. The size and position of this virtual object may be determined with the aid of Eqs. (7.26) and (7.27). For a given stopping

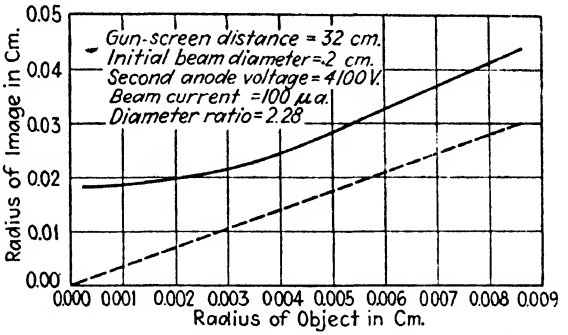


FIG. 7.18.—Effect of space charge in image space on magnification.

aperture in the first anode, the spreading of the beam also decreases the amount of second-anode current. Hence the mutual electron repulsion in the object space causes an increase in spot size with a decrease in current.

In conclusion, it is worth noting that the effect of the mutual electron repulsion is decreased by: (1) an increase in beam diameter, (2) a decrease in beam length, (3) an increase in electron velocity, and (4) a decrease in beam current.

CHAPTER 8

MAGNETOSTATIC FOCUSING[1]

8.1. General.—Magnetostatic focusing in cathode-ray tubes is accomplished by means of axially symmetric magnetostatic fields created by passing direct current through accurately wound

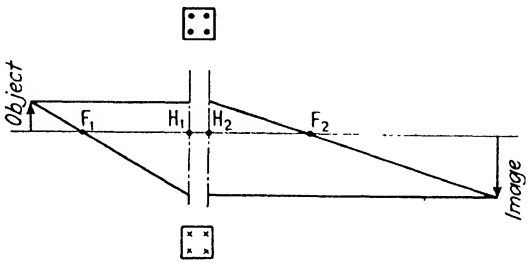


FIG. 8.1.—Focusing with a short coil.

coils having axial symmetry. Magnetostatic focusing is usually accomplished by either long or short coils. Figure 8.1 shows a magnetostatic focusing system using a short coil. Long coils extend over the entire length of the beam and are therefore not used in TCR tubes. Farnsworth uses a long coil in his image-

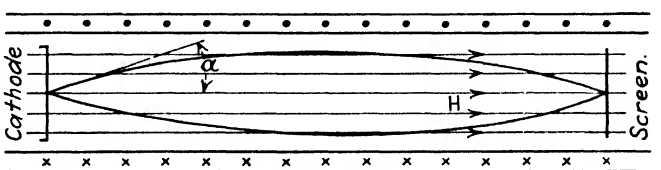


FIG. 8.2.—Focusing with a long coil.

pickup tube. The long coil produces a magnetostatic field extending over the entire length of the tube.

8.2. Focusing Action of Long Coil.—The theory of the focusing action of the long coil is relatively simple, and may serve as an introduction to the more difficult problem of focusing by means of a short coil. In Fig. 8.2 let the dots and crosses represent

[1] BUSH, H., *Ann. d. Phys.*, **81**, 976 (1926); *Archiv f. Elekt.*, **18**, 583 (1927).

the current in the coil, and let the horizontal lines represent the constant magnetic field due to this current. Now consider an electron moving with the velocity v making the angle α with the magnetic field, then the velocity component $v \cos \alpha$ along the magnetic field is not affected by the field, but the component $v \sin \alpha$ perpendicular to the field is affected by the field as though the parallel component were not present.

If the parallel component $v \cos \alpha$ were not present, the magnetic field would cause the electron to follow a circular path, the plane of the circle being perpendicular to the magnetic field. In the presence of the parallel component, the path of the electron is a helix, but the projection of the path upon a plane perpendicular to the magnetic field is a circle. The radius of this circle is

$$R = \frac{v \sin \alpha}{\frac{e}{m} H} \tag{8.1}$$

where H is the strength of the magnetic field. The time taken by the electron in describing the circle is

$$T = \frac{2\pi R}{v \sin \alpha} = \frac{2\pi m v \sin \alpha}{eHv \sin \alpha} = \frac{2\pi}{\frac{e}{m} H}. \tag{8.2}$$

Hence, the time is independent of v, α or R, and depends merely on H, so that for a given H (a given current in the coil) the time taken for *any* electron to describe the circle is the same.

Now consider the case of a point source from which electrons issue with the constant speed v and make small angles with the lines of magnetic force. As α is very small, $\cos \alpha = 1$, and the component of velocity along the lines of force $v \cos \alpha$ is constant and equal to v. Hence there results the case of all the electrons issuing from a point having the same constant speed v along the lines of force and different small speeds $v \sin \alpha$ normal to the lines of force. As was seen above, the time taken by the various electrons in describing their various circles is the same for all of them, so at a distance

$$l = Tv = \frac{2\pi}{\frac{e}{m} H} v \tag{8.3}$$

from the source, the electrons will meet again. Hence l is the distance between the object and the image. It is seen that l is directly proportional to v and inversely proportional to H. It is further to be noted that the image has unit magnification and is upright.

If the angle α is not sufficiently small so as to let $\cos \alpha = 1$, then

$$l = Tv \cos \alpha \doteq \frac{2\pi v}{\frac{e}{m}H}\left(1 - \frac{\alpha^2}{2}\right). \tag{8.4}$$

In this case the lens displays longitudinal spherical aberration and all the electrons emitted from one point do not again meet in one image point.

8.3. Axially Symmetric Magnetostatic Fields.—Magnetostatic fields due to permanent magnets may be treated exactly the same as electrostatic fields; *i.e.*, they may be uniquely described by means of a scalar potential. Magnetostatic fields due to currents are not uniquely described by a scalar potential, but are uniquely described by a *vector* potential usually denoted by $\bar{A}$ and defined by the vector relations

$$\begin{aligned} \bar{H} &= \text{curl } \bar{A} \\ 0 &= \text{div } \bar{A} \end{aligned} \tag{8.5}$$

where $\bar{H}$ is the strength of the magnetic field. The components of H in cylindrical coordinates are then

$$\begin{aligned} H_r &= \text{curl}_r \bar{A} = \frac{1}{r}\left(\frac{\partial A_z}{\partial \theta} - \frac{\partial (rA_\theta)}{\partial z}\right) \\ H_z &= \text{curl}_z \bar{A} = \frac{1}{r}\left(\frac{\partial (rA_\theta)}{\partial r} - \frac{\partial A_r}{\partial \theta}\right) \\ H_\theta &= \text{curl}_\theta \bar{A} = \left(\frac{\partial A_r}{\partial z} - \frac{\partial A_z}{\partial r}\right). \end{aligned} \tag{8.6}$$

In case of axial symmetry

$$A_r = A_z = \frac{\partial A_\theta}{\partial \theta} = 0$$

and Eqs. (8.6) reduce to

$$H_r = -\frac{1}{r}\frac{\partial(rA_\theta)}{\partial z}$$
$$H_z = \frac{1}{r}\frac{\partial(rA_\theta)}{\partial r}$$
$$H_\theta = 0. \tag{8.7}$$

In space free from currents

$$\text{curl } \bar{H} = \text{curl curl } \bar{A} = 0 \tag{8.8}$$

Inserting (8.7) into (8.8), it results that A_θ satisfies the differential equation

$$\frac{\partial^2 A_\theta}{\partial z^2} + \frac{\partial^2 A_\theta}{\partial r^2} + \frac{\partial}{\partial r}\left(\frac{A_\theta}{r}\right) = 0. \tag{8.9}$$

The *similarity* but *not identity* of Eq. (8.9) with Laplace's equation should be carefully noted.

Since A_θ is an odd function of r it may be expanded into the series

$$A_\theta = rf_1(z) + r^3 f_3(z) + r^5 f_5(z) + \cdots \tag{8.10}$$

Differentiating this series, inserting it into (8.9) and equating the coefficients of equal powers of r to zero,

$$f_3 = -\frac{f_1''}{2 \cdot 4}, \qquad f_5 = \frac{f_1^{(4)}}{2 \cdot 4^2 \cdot 6}, \cdots$$

and so Eq. (8.10) becomes

$$A_\theta(r,z) = rf_1(z) - \frac{r^3}{2 \cdot 4} f_1''(z) + \frac{r^5 f_1^{(4)}}{2 \cdot 4^2 \cdot 6}(z) \cdots \tag{8.11}$$

From Eqs. (8.7) and (8.11) it thus follows that

$$H_z(r,z) = \frac{1}{r}\frac{\partial(rA_\theta)}{\partial r} = 2f_1(z) - \frac{r^2}{2} f_1''(z) + \frac{r^4}{2 \cdot 4^2} f^{(4)}(z) \cdots$$

On the axis

$$H_z(o,z) = 2f_1(z)$$

so $f_1(z)$ is $\frac{1}{2}$ the z component of the magnetic field on the z axis.

Hence Eq. (8.11) may be rewritten as

$$A_\theta(r,z) = \frac{r}{2}H(z) - \frac{r^3}{2^2 \cdot 4}H''(z) + \frac{r^5}{2^2 \cdot 4^2 6}H^{(4)}(z) \cdots \quad (8.12)$$

where $H(z)$ is the strength of the magnetic field along the z axis.

8.4. Motion of Electron in Axially Symmetric Magnetostatic and Electrostatic Fields.—At a given point an electron of velocity $\bar{v}$ moving through a magnetic field of strength $\bar{H}$ experiences a force which is proportional to the product of the velocity of the electron, the strength of the field and the sine of the angle between the directions of the velocity and field at the point, and is directed along the normal to the plane determined by $\bar{v}$ and $\bar{H}$. This is expressed by the vector equation

$$\bar{F} = m\bar{a} = -e[\bar{v}\times\bar{H}] \quad (8.13)$$

where $\bar{F}$ is the force, $\bar{a}$ is the acceleration and $[\bar{v}\times\bar{H}]$ is the vector product of $\bar{v}$ and $\bar{H}$. If there is an electrostatic field superimposed upon the magnetostatic field, the above equation becomes

$$\bar{F} = m\bar{a} = -e(-\text{ grad } V + (\bar{v}\times H)). \quad (8.14)$$

In the case of axial symmetry this equation becomes, when broken into its three components in cylindrical coordinates

$$a_r = \ddot{r} - r\dot{\theta}^2 = -\frac{e}{m}\left[\dot{\theta}\frac{\partial}{\partial r}(rA_\theta) - \frac{\partial V}{\partial r}\right] \quad (8.15\text{ a})$$

$$a_z = \ddot{z} = -\frac{e}{m}\left[\dot{\theta}\frac{\partial}{\partial z}(rA_\theta) - \frac{\partial V}{\partial z}\right] \quad (8.15\text{ b})$$

$$a_\theta = \frac{1}{r}\frac{d}{dt}(r^2\dot{\theta}) = -\frac{e}{m}\left[-\frac{1}{r}\frac{d}{dt}(rA_\theta)\right]. \quad (8.15\text{ c})$$

Equation (8.15 c) integrates into

$$r^2\dot{\theta} = \frac{e}{m}rA_\theta + C$$

or

$$r\dot{\theta} = \frac{e}{m}A_\theta \quad (8.16)$$

the constant C being assumed zero since A_θ is zero on the axis of an axially symmetric field (see Eq. 8.12). Inserting Eq. (8.16) into Eqs. (8.15 a) and (8.15 b) there results, after reduction,

the following equations of motion

$$\ddot{r} = \frac{e}{m}\frac{\partial}{\partial r}\left[V - \frac{1}{2}\frac{e}{m}A_\theta^2\right]$$
$$\ddot{z} = \frac{e}{m}\frac{\partial}{\partial z}\left[V - \frac{1}{2}\frac{e}{m}A_\theta^2\right]. \qquad (8.17)$$

Inserting the expansions for V (Eq. 4.4) and A_θ into Eqs. (8.17) it follows that

$$\ddot{r} = -\frac{e}{m}\left[\frac{r}{2}\left(V'' + \frac{1}{2}\frac{e}{m}H^2\right) - \frac{r^3}{2^2 4}\left(V^{(4)} + 2\frac{e}{m}HH''\right) + \cdots\right]$$
$$\ddot{z} = \frac{e}{m}\left[V' - \frac{r^2}{2^2}\left(V''' + \frac{e}{m}HH'\right) + \cdots\right] \qquad (8.18)$$

Equations (8.18) are the equations of motion of an electron in axially symmetric electrostatic and magnetostatic fields. In these equations V, V', V'' and V''' represent the electrostatic potential along the axis and its first, second and third derivatives with respect to z, respectively. H, H' and H'' represent the z component of the magnetic field along the axis and its first and second derivatives, respectively.

8.5. Trajectory of Paraxial Electron.—In the case of a paraxial electron the equations of motion (8.18) reduce to

$$\ddot{r} = -\frac{e}{m}\frac{r}{2}\left(V'' + \frac{1}{2}\frac{e}{m}H^2\right).$$
$$\ddot{z} = \frac{e}{m}V'. \qquad (8.18\text{ p})$$

It follows from Eqs. (8.18 p) that the differential equation for the trajectory of a paraxial electron is

$$\frac{d^2r}{dz^2} + \frac{V'}{2V}\frac{dr}{dz} + \frac{1}{4V}\left(V'' + \frac{e}{2m}H^2\right)r = 0. \qquad (8.19\text{ p})$$

As in the pure electrostatic case, the differential equation (8.19 p) for the trajectory of a paraxial electron is linear and of the second order. Hence if any two linearly independent solutions $r_1(z)$ and $r_2(z)$, corresponding to two electron trajectories, are known, then any other trajectory may be determined by the relations

$$r(z) = C_1 r_1(z) \qquad \text{and} \qquad C_2 r_2(z)$$

where C_1 and C_2 are suitable constants. Hence the focusing action of a magnetostatic field superimposed upon the electrostatic field may be determined by obtaining the two fundamental trajectories by solving Eq. (8.19 p).

8.6. Rotation of Image.—In obtaining Eqs. (8.17) from (8.15), $r\dot{\theta}$ was eliminated by means of (8.16). So Eq. (8.19 p), which was deduced from (8.17), really represents the equation for the trace of the paraxial electron trajectory on the $r - z$ plane through the electron. The actual trajectory of a paraxial electron is of the form of a helix. As a result magnetic focusing rotates the image. From Eqs. (8.12) and (8.16) it follows that

$$\dot{\theta} = \dot{z}\frac{d\theta}{dz} = \frac{e}{m}\frac{A_\theta}{r} = \frac{e}{m}\left(\frac{H}{2} - \frac{r^2}{2^2 \cdot 4}H'' + \cdots\right). \quad (8.20)$$

For a paraxial electron

$$\frac{d\theta}{dz} = \frac{e}{m}\frac{1}{\sqrt{2\frac{e}{m}V}}\frac{H}{2} = \frac{1}{2}\sqrt{\frac{e}{2m}}\frac{H}{\sqrt{V}}$$

and

$$\theta = \frac{1}{2}\sqrt{\frac{e}{2m}}\int_{z_1}^{z_2}\frac{H(z)}{\sqrt{V(z)}}dz \quad (8.21 \text{ p})$$

where $z_2 - z_1$ gives the axial region over which the field $H(z)$ extends; *i.e.*, $H(z)$ is zero to the left of z_1 and to the right of z_2.

It is to be noted that for a given $H(z)$ and $V(z)$, θ is a constant (depending on z_1 and z_2 only) and so is the same for any paraxial electron. Hence the magnetostatic field rotates the image as a whole through the angle θ, but does not distort the image.

8.7. Pure Magnetostatic Focusing.—In cathode-ray tubes electrostatic and magnetostatic focusing are seldom superimposed, either one or the other type being used. In this and the remaining sections of this chapter the discussion will be limited to magnetostatic focusing only.

In this case V is constant throughout the range of the magnetostatic field and is proportional to the square of the constant speed with which the electron traverses the focusing magnetostatic field. Setting $V = V_0$ (a constant) in Eq. (8.19 p), it follows that the differential equation for the trace of a paraxial trajectory

on the r-z plane through the electron in an axially symmetric magnetostatic field is

$$\frac{d^2r}{dz^2} + \frac{e}{8m}\frac{H^2}{V_0}r = 0. \tag{8.22 p}$$

The positions of the cardinal points of the magnetostatic field are then determined by solving (8.22 p) for the two fundamental trajectories. The angle θ through which the image is rotated is, from Eq. (8.21 p),

$$\theta = \frac{1}{2}\sqrt{\frac{e}{2mV}}\int_{z_1}^{z_2} H(z)dz. \tag{8.23 p}$$

8.8. Thin Magnetostatic Lens.—In the special case of a thin lens, $H(z)$ is different from zero for only a narrow region, so that one may let $r = r_0$ (a constant) throughout this region. Let

$$P = \frac{1}{r_0}\frac{dr}{dz}$$

then Eq. (8.22 p) reduces to

$$\frac{dP}{dz} + \frac{e}{8m}\frac{H^2}{V_0} = 0 \tag{8.24 p}$$

so

$$P_b - P_a = -\frac{e}{8mV_0}\int_a^b H^2dz.$$

By reasoning identical with that used for the pure electrostatic thin lens it may be deduced that

$$\frac{1}{f_2} = \frac{e}{8mV_0}\int_a^b H^2dz = \frac{0.022}{V(\text{volts})}\int_a^b H^2(\text{gauss})dz. \tag{8.25 p}$$

As the index of refraction is the same on both sides of the lens, the ratio of the focal lengths is -1, and so

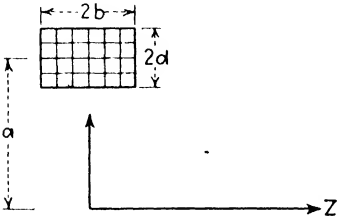
$$\frac{1}{f_1} = -\frac{e}{8mV_0}\int_a^b H^2dz. \tag{8.26 p}$$

The interval $b - a$ is the narrow range over which the magnetostatic field extends.

It is to be noted from Eq. (8.25 p) that a magnetostatic lens is always convergent and independent of the sign of $H(z)$; *i.e.*,

the lens is convergent for current flowing in either direction of the focusing coil.

From Eq. (8.23 p) it follows that a thin magnetostatic lens will rotate the image through an angle θ given by



$$\theta = \frac{1}{2}\sqrt{\frac{e}{2mV}}\int_a^b H dz = \frac{0.15}{\sqrt{V(\text{volts})}}\int_{ab} H(\text{gauss}) dz. \quad (8.27\text{ p})$$

Equation (8.27 p) shows that the rotation of the image is reversed by reversing the direction of the current through the focusing coil.

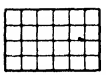


FIG. 8.3.—Multiple-layer coil of rectangular cross section.

From Eqs. (8.25 p) and (8.27 p) it is seen that the characteristics of a magnetostatic thin lens are known as soon as H, the z component of the magnetostatic field along the axis, is known. If there is no iron near the coil, H may usually be calculated. Thus the field

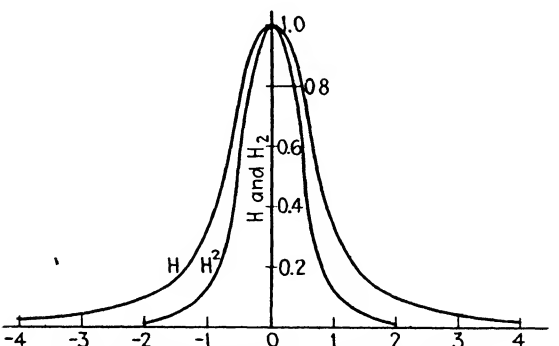


FIG. 8.4.—H and H^2 of a circular turn of wire.

on the axis of a circular turn of wire is

$$H = \frac{2\pi R^2 i}{(Z^2 + R^2)^{\frac{3}{2}}}$$

where Z is the distance along the axis of a point considered from the center of the circle and R is the radius of the circle. The

field on the axis of the square coil shown in Fig. 8.3 is

$$H = \frac{\pi ni}{2bd}\left((z+b)\log_\epsilon \frac{(a+d)+\sqrt{(z+b)^2+(a+d)^2}}{(a-d)+\sqrt{(z+b)^2+(a-d)^2}} - (z-b)\log_\epsilon \frac{(a+d)+\sqrt{(z-b)^2+(a+d)^2}}{(a-d)+\sqrt{(z-b)^2+(a-d)^2}}\right)$$

where n is the number of turns and i is the current. In order to confine the magnetic field to a narrow range, the coil is partially surrounded with iron, and in such cases H is determined by measurement.

Figure 8.4 shows how the quantities H and H^2 vary with the distance along the axis. The area between the H curve and the z axis gives the amount of image rotation, and the area between H^2 curve and the z axis gives the power $1/f$ of the lens. In determining the areas it is sufficiently accurate to take for the limits a, b as the two points for which H is 10 per cent of the maximum. Figure 8.5 shows how, in general, the focal length and the angle of image rotation vary with the current through the coil.

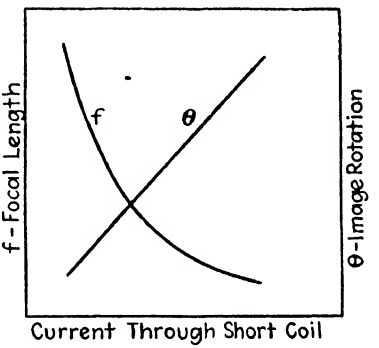


FIG. 8.5.—General form of variation of focal length and image rotation with current through a short coil.

It is sometimes desirable to have no image rotation. This may be accomplished by placing two short coils next to each other in such a way that the current will pass through them in opposite directions. By this scheme it is possible to arrange so that

$$\int_{-a}^{+a} H dz \doteq 0$$

and

$$\int_{-a}^{+a} H^2 dz > 0.$$

PART II
TELEVISION CATHODE-RAY TUBE

CHAPTER 9

THE ELECTRON GUN

9.1. General.—The function of the TCR tube is to reproduce a television image on a luminescent screen by scanning it with a focused and modulated electron beam. From the performance standpoint the three essential elements of TCR tubes are the luminescent screen, the scanning or deflecting arrangement and the arrangement for generation, control and focusing of the electron beam. The latter arrangement is commonly called an *electron gun*, and this chapter is devoted to its treatment. The luminescent screens and the deflecting systems are treated in the two chapters immediately following.

A typical electron gun is shown in Fig. 9.1. It has an axially symmetric structure and comprises a cathode, a grid, a first anode and a second anode. The second anode, while physically not always assembled on the press, is a part of the electron gun, is an essential member of the arrangement and should always be considered as a part of it.

In the terminology of electron optics as developed in Part I, this electron gun consists of an electron source, an immersion lens and a direct bipotential lens. The individual functions of the optical components of an electron gun are as follows: The immersion lens concentrates the electrons emitted by the cathode into a small cross section or a crossover and forms a virtual image of it. The bipotential lens focuses the virtual image of the crossover on the luminescent screen.

9.2. Electron Gun Requirements.—In order that a TCR tube may operate satisfactorily in a given television receiver, it must have an electron gun capable of the following functions:

At the available second-anode voltage E_{p2} it must generate not less than a certain amount of beam current I_{p2}, with a width of the scanning line λ not greater than a certain value determined by the resolution of the system. The beam cut-off bias value and modulation characteristics of the TCR tube must be such

that the video output of the receiver is ample to vary the beam current from zero to the maximum usable value. The maximum usable beam current corresponds to a value at which the line width is just beginning to limit the resolution of the system. The first-anode current I_{p1} should never exceed a value at which the regulation of its source begins to affect the focusing of the beam.

The task is to design an electron gun that satisfies the above requirements as closely as possible within the given limits. The

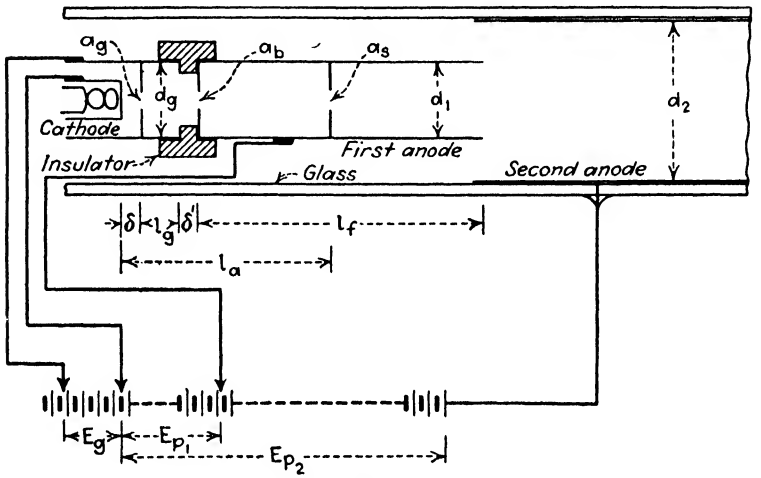


FIG. 9.1.—The electron gun.

extent of this task may be realized by considering the particular type of gun shown in Fig. 9.1 and noting the number of variables that may be necessary to consider for an optimum design. The independent variables, *i.e.*, those things which one may vary at will without necessarily changing any of the others, are:

1. Cathode-grid spacing ..... δ.
2. Diameter of grid aperture ..... a_g.
3. Diameter of grid skirt ..... d_g.
4. Length of grid skirt ..... l_g.
5. Grid-first anode spacing ..... δ'.
6. Diameter of first-anode backing aperture ..... a_b.
7. First-anode length ..... l_f.
8. Diameter of first-anode stopping aperture ..... a_s.
9. Distance between stopping aperture and cathode .... l_a.
10. Diameter of first anode ..... d_1.
11. Diameter ratio of second to first anode ..... D.

12. Distance between first-anode end and screen ..... l_s.
13. Voltage on first anode ..... E_{p1}
14. Voltage ratio ..... E.
15. Grid voltage ..... E_g

The dependent variables, each of which is a function of most of the independent variables, are:

1. Cut-off voltage ..... E_{cf}.
2. Total current ..... $I_t = I_{p1} + I_{p2}$.
3. Beam current ..... I_{p2}.
4. Width of scanning line (resulting from deflection of the spot at normal scanning speed and normal repetition rate) λ.

It is to be noted that each of the dependent variables is a function of independent variables only, and when one dependent variable varies with another dependent variable it is only through the fact that they are functions of the same independent variables.

An ideal method of solving the problem would be to determine the exact and complete relationship between each dependent and its independent variables, and accordingly adjust the variables to satisfy the imposed conditions. The determination of the relationships could be either analytical or experimental, but in either case an experimental verification and cross-checks of the data obtained would be necessary. The number of independent variables, however, precludes any such straightforward method of accumulating data and solving the given problem. Practically, the problem is approximately solved with the aid of a limited number of simple relations of dependence.

9.3. Experimental Determination of Gun Data.—A theoretical determination of the relationships existing between each of the dependents and its independent variables is usually feasible for only a few of the variables. Thus, for paraxial electrons, relationships between λ and l_f, l_s, D, E are deducible from the theoretically determined curves of Figs. 6.4 to 6.10. For nonparaxial beams, however, it is necessary to use the experimentally determined curves of Figs. 7.8 and 7.9, in conjunction with those of Figs. 6.4 to 6.10. Relationships between λ and its other variables, as well as those between E_{cf}, I_{p1}, I_{p2} and their variables, usually have to be determined experimentally. The determination of these relationships very often becomes an extremely difficult task.

As an example, consider an experimental determination of the relationship between each of the dependent variables and the grid-skirt length, l_g. Apparently, this may be accomplished by constructing a series of tubes differing only in skirt length. However, besides being expensive and time consuming, this method has several fundamental limitations which apply quite generally. For the results to be meaningful, it is necessary to keep the other independent variables at suitably fixed values throughout the experiment. Thus it is usually desirable to have the cathode-grid spacing about 0.010 in. and it is extremely difficult to repeat spacings within 0.002 in. A difference in grid-cathode spacing of 0.002 in. is sufficient to considerably modify the result of a change in grid-skirt length. Another

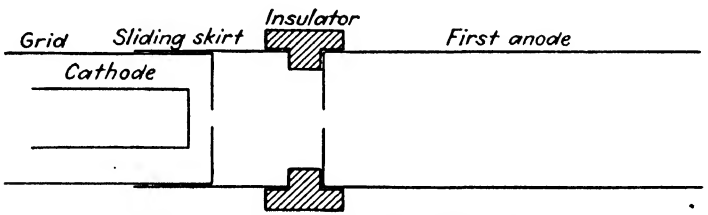


FIG. 9.2.—Gun with sliding grid skirt.

difficulty is the fact that oxide-coated cathodes are not usually reproducible, so the determination of the effect of a change in grid-skirt length on current is fairly unreliable.

These difficulties may be overcome by performing the experiments on a single tube having a gun with a sliding grid skirt as shown in Fig. 9.2. The length of the grid skirt is varied by simply tilting the tube and is indicated on the outside of a calibrated gun.

This example illustrates the usefulness of the sliding member in experimenting on electron guns (see Fig. 6.12). If the variations are small, as they would be in experiments on the determination of the effects of changes in cathode-grid spacings, sliding parts are not practical, and other schemes have to be devised to accomplish the desired end. The effect of changes in apertures may be determined with the aid of a weighted disk that may rotate about an axis not coaxial with that of the gun, and as the gun (tube) is rotated the disk remains fixed and apertures of various sizes, punched in the disk, are brought into positions

coaxial with the gun. A very useful, although expensive, apparatus is a large demountable tube which permits the change

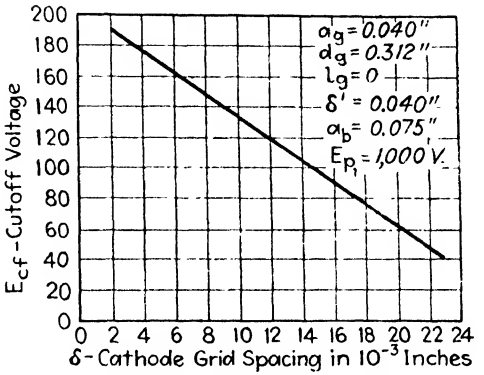


FIG. 9.3.—Variation of E_{cf} with δ.

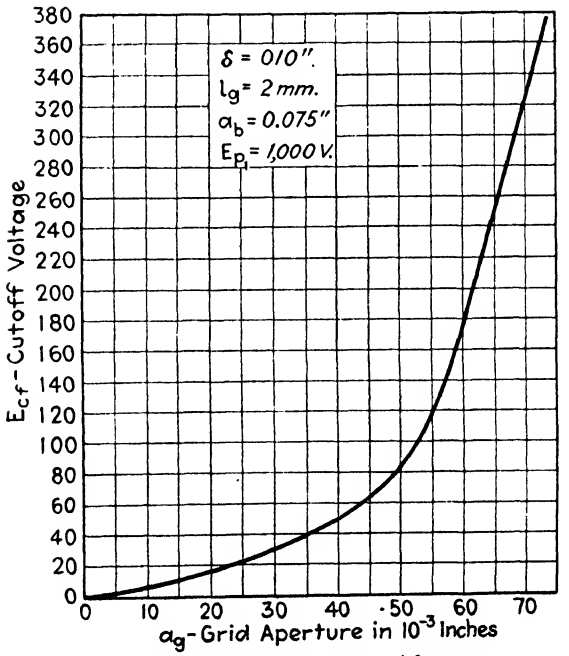


FIG. 9.4.—Variation of E_{cf} with a_g.

and accurate adjustment of the various parts of the gun. The sections following present some of the experimentally determined

relations between the dependent and independent variables of the gun of Fig. 9.1.

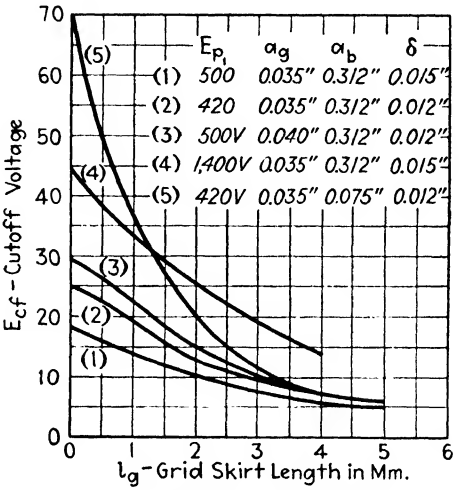


FIG. 9.5.—Variation of E_{cf} with l_g.

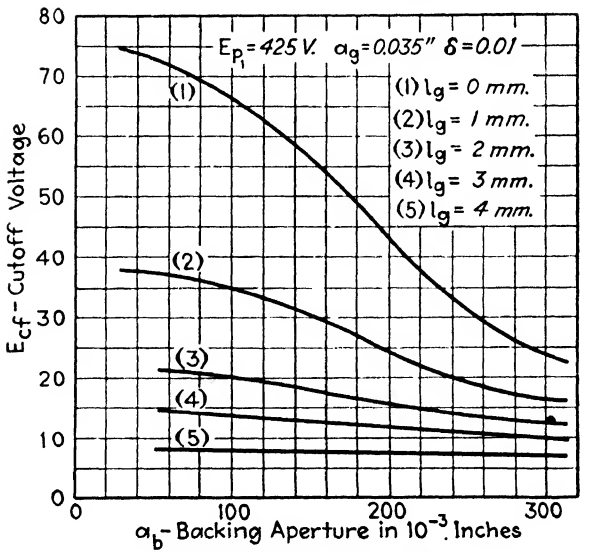


FIG. 9.6.—Variation of E_{cf} with a_b.

9.4. Cut-off Voltage.—The cut-off voltage E_{cf} is a function of the seven independent variables: δ, a_g, d_g, l_g, δ', a_b and E_{p1}. Figure 9.3 shows that, under the conditions indicated, a practi-

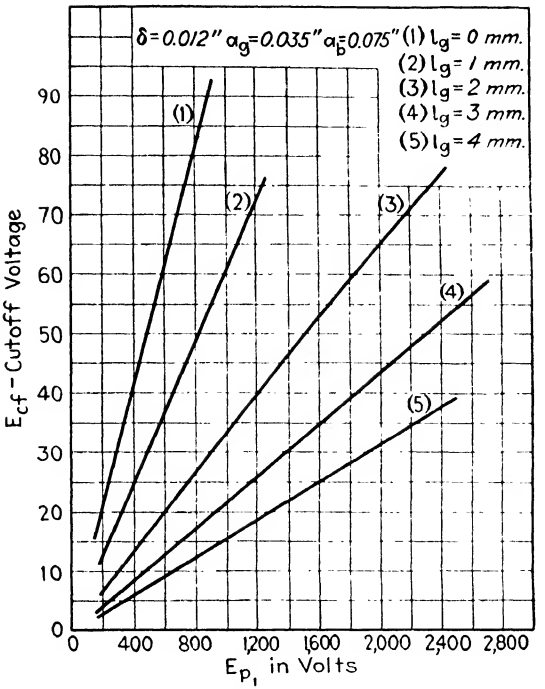


FIG. 9.7.—Variation of E_{cf} with E_{p1}.

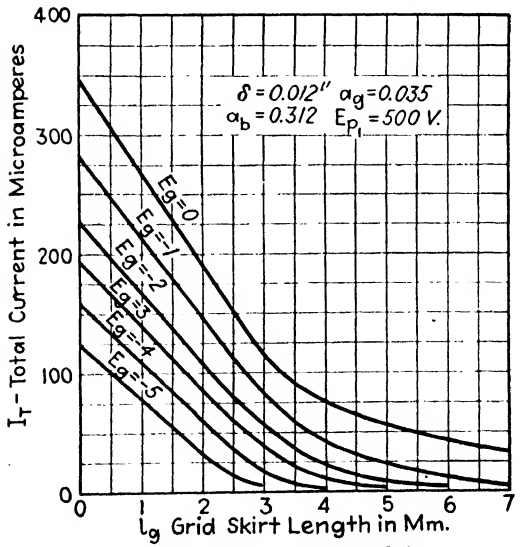


FIG. 9.8.—Variation of I_T with l_g.

cally linear relation exists between E_{cf} and δ in the range shown. The curve shown in Fig. 9.4 depicts the effects of the size of the grid aperture on E_{cf}; it is seen that for the larger grid apertures E_{cf} increases rapidly with aperture size. Relations between E_{cf} and grid-skirt length for several different conditions are shown in Fig. 9.5. These curves show that the variation of E_{cf} with l_g is greatest for short skirt lengths and small backing apertures.

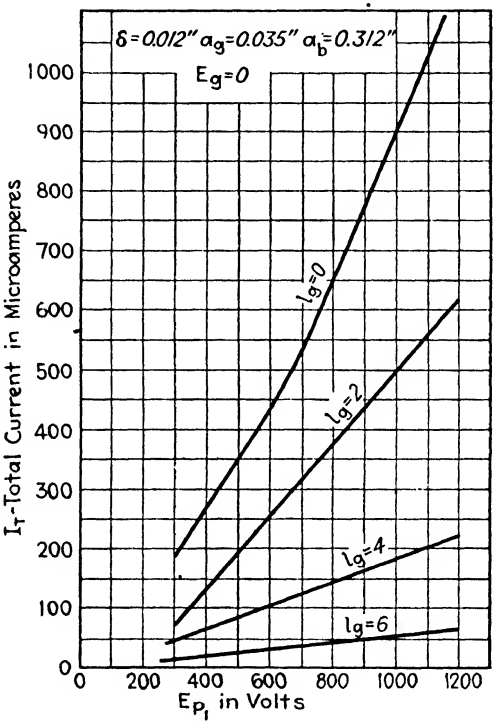


FIG. 9.9.—Variation of I_T with E_{p1}.

The variation of E_{cf} with the size of the backing aperture is depicted by the curves of Fig. 9.6. It is worth noting that the effect of a_b on E_{cf} is negligible for the longer grid-skirt lengths. As expected, the variation of E_{cf} with E_{p1} is depicted by a straight line through the origin. Figure 9.7 shows the linear relations between E_{cf} and E_{p1} for five values of l_g. In dimensioning a gun to obtain a value of E_{cf} as dictated by receiver considerations, it is necessary to consider not only the curves of Figs. 9.3 to 9.7 but also the fact that the beam current, first-anode current,

and line width are also functions of the same independent variables as is the cut-off voltage.

9.5. Total Current.—The total current (current drawn by first and second anodes) is the current in the beam formed by the immersion lens (see Fig. 9.11). It is a function of the eight independent variables: δ, a_g, d_g, l_g, δ', a_b, E_{p1} and E_g. The total current is, with the exception of E_g, a function of the same vari-

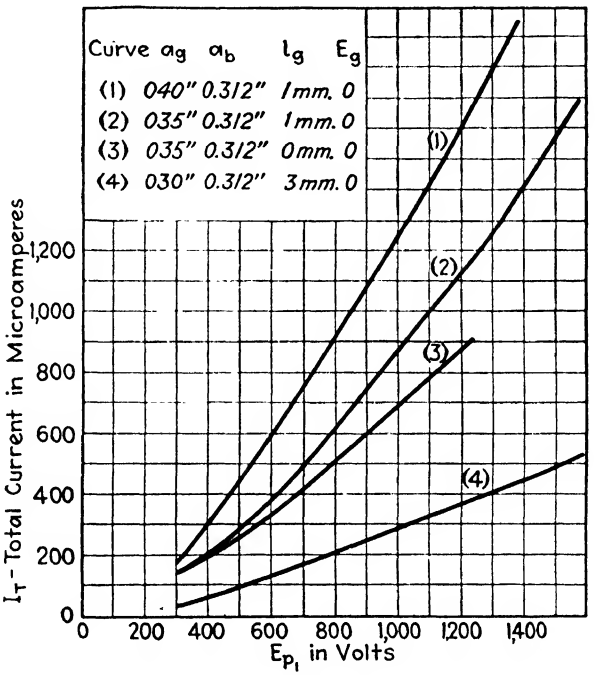


FIG. 9.10.—Variation of I_T with E_{p1} for various tubes.

ables as the cut-off voltage. In general, any change in the dimensioning of a given type of gun which tends to increase the cut-off voltage likewise increases the total current. The curves of Fig. 9.8 depict the relation between I_T and l_g for various values of E_g. The curves of Fig. 9.9 show the variation of I_T with E_{p1} for various l_g. And the curves of Fig. 9.10 give the variation of I_T with E_{p1} for various tubes. It is seen that to a fair approximation I_T increases linearly with E_{p1}.

9.6. Beam Current.—The useful current in a TCR tube is that portion of the total current which constitutes the beam that

strikes the screen to form the luminescent spot. From Fig. 9.11 it is seen that only that portion of the total current which passes through the stopping aperture is the useful current. Although both the total and the useful currents form beams, it has become

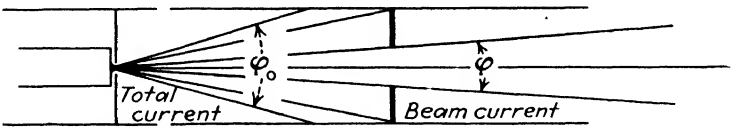


FIG. 9.11.—Solid angles subtended at cathode by total and beam currents.

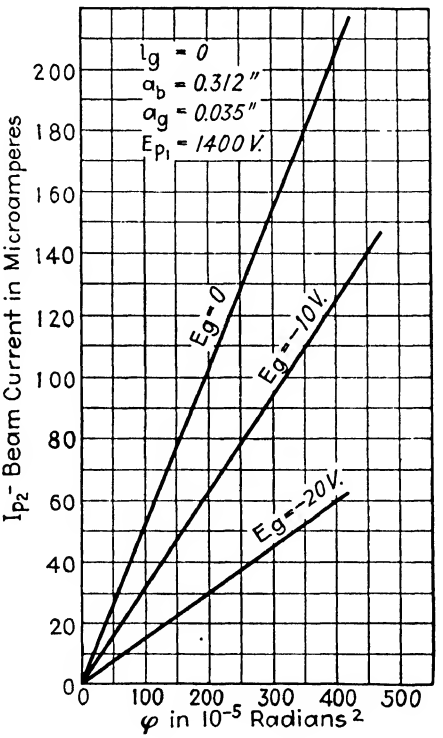


FIG. 9.12.—Variation of I_{p2} with φ for three values of E_g.

customary to reserve the name "beam current" for the useful current. If it be assumed that the current density along any cross section of the total current is constant; then to a good approximation the beam current I_{p2} is given by the relation

$$I_{p2} = I_T \frac{\varphi}{\varphi_0} (\varphi < \varphi_0) \tag{9.1}$$

where φ is the solid angle subtended at the point of intersection between cathode and axis by the stopping aperture, and φ_0 is the solid angle subtended at the same point by the total current

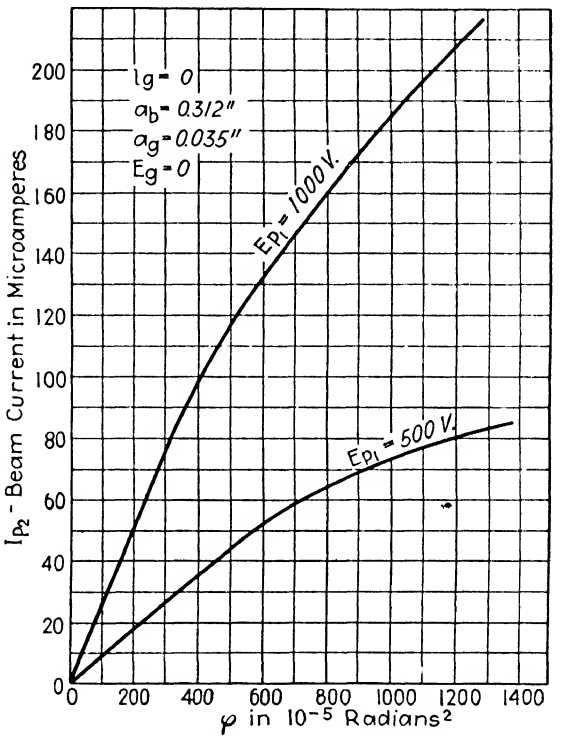


FIG. 9.13.—Variation of I_{p2} with φ for two values of E_{p1}

(see Fig. 9.11). For the small solid angles met in TCR tubes φ, in radians, is given by

$$\varphi = \pi\left(\frac{a_s}{l_a}\right)^2. \tag{9.2}$$

Figure 9.12 shows curves giving I_{p2} as a function of φ for three values of grid voltage. It is seen that the function is linear, indicating that for the range of solid angles given in Fig. 9.12 Eq. (9.1) is valid. However, the curves of Fig. 9.13, giving I_{p2} as a function of φ for two values of first-anode voltage, indicate that Eq. (9.1) is not valid when extended to larger solid angles than those given in Fig. 9.12.

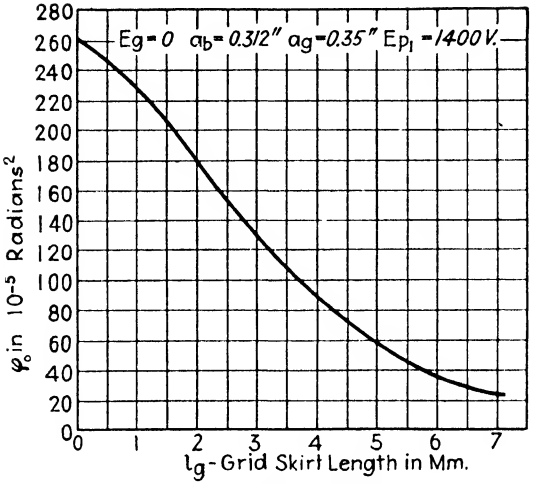


FIG. 9.14.—Variation of φ_0 with l_g.

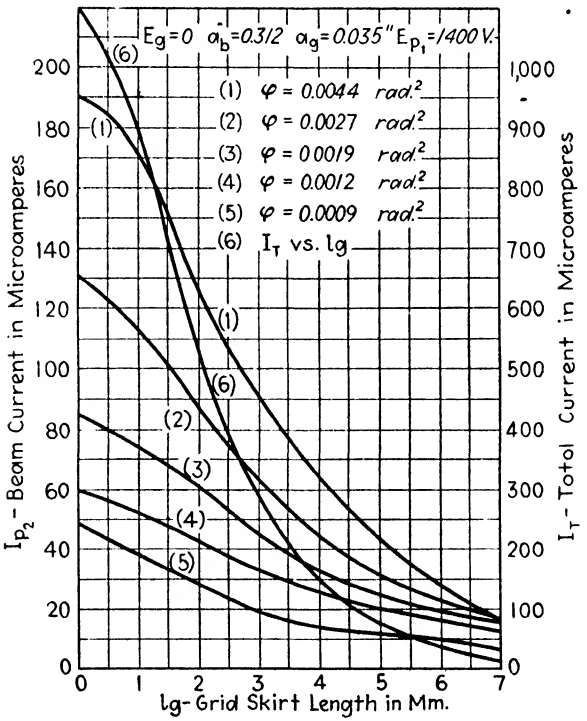


FIG. 9.15.—Variation of I_{p2} with l_g.

From Eqs. (9.1) and (9.2) it follows that φ_0 in radians is given by

$$\varphi_0 = \pi\left(\frac{a_s}{l_a}\right)^2 \frac{I_T}{I_{p2}}. \tag{9.3}$$

It is to be noted that φ_0 is not the actual solid angle subtended by the total current. It is the solid angle which the total current

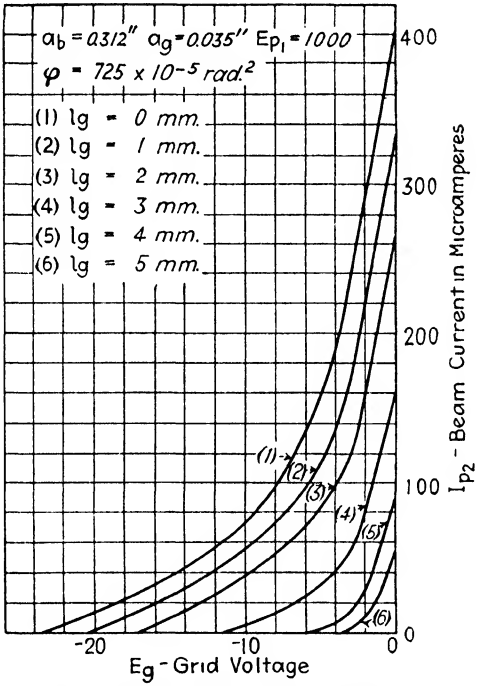


FIG. 9.16.—Variation of I_{p_2} with E_g

would subtend if the current density in the total-current beam were constant throughout any cross section. Figure 9.14 shows how φ_0, calculated with the aid of Eq. (9.3), varies with the length of the grid skirt. Although the actual solid angle subtended by the total current is considerably larger (see Fig. 9.13) than φ_0, the variation of actual solid angle with l_g is fairly well depicted by Fig. 9.14.

Curves 1 to 5 of Fig. 9.15 show the variation of beam current with grid-skirt length for five values of φ. Curve 6 of Fig. 9.15

gives the total current plotted to a scale five times that of curves 1 to 5. The curves of Fig. 9.16 give the beam current as a function of grid voltage for six values of grid-skirt length.

9.7. First-anode Current.—First-anode current, which is the difference between the total and the beam currents, is an important variable in a TCR tube. In general, power-supply considerations, as well as efficient utilization of the cathode, require the first-anode current to be small. Moreover, as the average

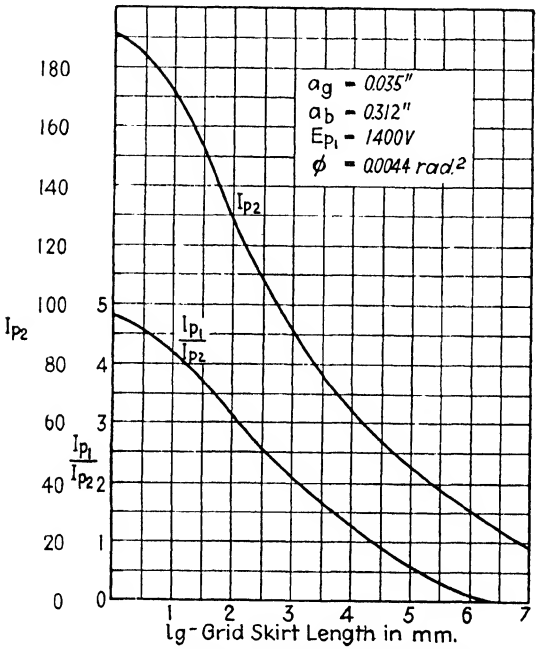


FIG. 9.17.—Variation of I_{p2} and I_{p1}/I_{p2} with l_g.

illumination of the reproduced picture varies between a minimum and a maximum, the average first-anode current also varies between the minimum or zero value corresponding to an all-black picture and a certain maximum which corresponds to an all-highlight picture. The instantaneous variations of the first-anode current (and similarly of the second-anode current) do not affect the voltage delivered by the power-supply unit because of the smoothing effect of the filter condenser. However, the slow variations of the mean current, which flows through the reactor of

the filter, cause corresponding variations in the supply voltage. These variations cause fluctuations of the focusing-voltage ratio E_{p2}/E_{p1}, which will be correct only for a limited range of average brightness of the reproduced picture.

For a given size and position of stopping aperture (given φ), the first-anode current may be reduced by increasing the length of the grid skirt. Figure 9.17 gives a curve showing the ratio of first-anode current to beam current as a function of skirt length for a given φ and E_{p1}. On the same figure is also shown the variation of beam current and skirt length for the same φ and E_{p1}. It is seen that although an increased skirt length decreases the first-anode current, it also decreases the beam current. To improve the efficiency of a given gun, *i.e.*, to decrease first-anode current without a decrease in the beam current, one may resort to any of the following means: Use a long grid skirt combined with a substantial increase in the first-anode voltage; increase the diameter of the stopping aperture; reduce the length of the first anode; etc. However, one must not forget that any of the above changes modifies a number of other operating parameters of the tube, and these changes should be taken into account.

9.8. Size of Luminescent Spot.—As was shown in Part I, the small luminescent spot produced on the screen by applying suitable potentials to the first and second anodes depicts the electronic image (usually with aberrations) of an object (crossover, aperture, etc.) situated at or close to the cathode. The size of the luminescent spot depends upon (1) the size of the object being imaged, (2) the magnification of this object by the bipotential lens imaging it, (3) the aberrations of the bipotential lens and (4) the mutual electron repulsion existing in the electron beam. If the current required in the beam is small (as in the case of the iconoscope), one may deal exclusively with paraxial electrons. Then the size of the luminescent spot will depend merely on factors (1) and (2). In this case, the size of the luminescent spot can be reduced by decreasing the size of the object or the magnification, or both. The size of the object may be reduced by decreasing the grid aperture in the case of imaging a crossover, or the size of the aperture in the case of imaging an aperture. The magnification can be reduced by increasing the length of the first anode or by increasing proportionately both the length and diameter of the first anode. Curves such as those given in

Chap. 6 will then be sufficient to enable one to calculate dimensions of the first and second anodes in order to obtain a desired magnification.

However, one of the most important requirements of the gun is the production of a small luminescent spot of high intensity. The intensity of the luminescent spot is approximately proportional to the product of the second-anode voltage and beam current, *i.e.*, to the beam wattage. Suppose that in a given case it is desirable to have a 3-watt beam in the highlights. This means that a tube operating with 6,000 volts on the second anode requires a maximum beam current of 500 μa. As far as the gun is concerned, a simpler and more efficient[1] manner of obtaining a small spot with a 3-watt beam is by operating the tube with 24,000 volts on the second anode and with a maximum beam current of about 125 μa. However, such a solution may be of no practical value as the cost of the power-supply unit increases roughly as a cube of the voltage, and deflecting difficulties also rise rapidly with the increase of the second-anode voltage.

In contemplating a reduction in object size, it is necessary to consider the effect of this reduction on the total current and the solid angle subtended by this total current. The size of the object must not be reduced beyond the point where a further reduction would decrease the maximum beam current below 500 μa. Moreover, in the case of a very intense spot, curves such as those given in Chap. 6 are no longer sufficient to enable one to dimension the first and second anodes. It now becomes necessary to consider the effects of spherical aberration and mutual electron repulsion which have been treated in Chap. 7.

It is worth noting that an increase in the length of the first anode does not necessarily decrease the size of a spot with a given beam current. An excessive increase in the length of the first anode may result in an increase in the size of a spot with a given beam current. This contingency is brought out by the following example:

Consider the case of a gun having a first anode 3 g.d.[2] long, a beam width of 0.35 g.d. at the end of the first anode and

[1] As shown in Chap. 11 luminescent screens display saturation for high current densities but not for high voltages.

[2] The abbreviation "g.d." means "gun diameter" and has been defined in Sec. 6.2 as the diameter of the first anode.

requiring a voltage ratio of 8 for focusing. Suppose the length of the first anode increases to 7 g.d., then the (paraxial) magnification decreases to 58 per cent of what it was with the 3-g.d. first anode. However, the beam width at the end of the first anode now becomes about 0.82 g.d., so by curve 2 of Fig. 7.9 the spot has been increased to 58 per cent $\times$ 4.3 = 250 per cent of what it was with the 3-g.d. first anode. However, due to the longer first anode and the spherical aberration, the focusing-voltage ratio will have decreased to about 4, thus doubling the first-anode voltage (for a fixed second-anode voltage). Hence the current increases about $2\frac{1}{2}$ times.

Reducing φ (the solid angle subtended at the cathode by the stopping aperture) to about 40 per cent of its original value will reduce the beam to its original value. The reduction of φ to 40 per cent will cause a decrease of beam width to 63.5 per cent of 0.82 g.d. = 0.52 g.d. According to curve 2 of Fig. 7.9, the spot size has thus been increased to 58 per cent $\times$ 1.7 = 99 per cent. Hence for constant beam current, no decrease in spot size has been accomplished by increasing the length of the first anode. However, the first-anode current has been increased by a factor of at least 3, and thus the cathode has a harder service to perform. The increased length of first anode has also made it possible for other aberrations to appear on account of the magnification of slight misalignments, and the spot most likely will be of odd shape and effectively of greater size. Further increase in the first-anode length is certain to cause a spot increase for the same current in the beam.

The paraxial magnification may be decreased by proportionately increasing the length and diameter of the first anode, since this causes a reduction in image distance measured in g.d. Moreover, this permits keeping φ constant so that no change in current need accompany the decrease in magnification if the second-anode diameter is so modified as to keep the voltage ratio constant. However, deflection difficulties, both as regards the amount of deflection and the distortion arising from deflection of very wide beams, often make the use of large first anodes a disadvantage. In contemplating any means for reducing the size of the luminescent spot, one must also consider the effect on this reduction on (1) cut-off voltage, (2) total current, (3) beam current and (4) deflection difficulties.

The determination of φ is normally controlled by the maximum beam width permissible on account of the spherical aberration of the bipotential lens. A given φ may be obtained with various positions and sizes of the stopping aperture. However, the location of the stopping aperture should be in a perfectly unipotential region.

The reason for such condition is that even a very weak field existing at the aperture is sufficient to direct into the strong focusing field of the bipotential lens a considerable number of the slow secondary electrons which result from the impact of electrons striking the aperture disk and wall of the first anode. The bipotential lens, which acts as an extremely strong lens for the slow secondaries, sprays these electrons onto the luminescent screen. These electrons strike the screen with a velocity corresponding to the difference in potential between the second and first anodes. The illumination of the screen resulting from the impact of these electrons is known as *haze*. This haze greatly reduces the contrast of a television picture and is to be avoided. Practically all the haze disappears if the stopping aperture is placed about 2 g.d. from the end of the first anode.

9.9. Width of Scanning Line.—The apparent width of the lines appearing on the scanning pattern and resulting from the fast deflection of the luminous spot depends upon: (1) the dimension of the spot perpendicular to the direction of scanning, (2) the shape of the spot, (3) the distribution of luminous intensity over the spot, (4) the background illumination and (5) the speed with which the spot moves along the scanning line.

The measured width of the scanning line is always less than the dimension of the stationary spot measured perpendicularly to the direction of scanning. Thus if the undeflected spot is an ellipse with a major axis of a mm. and a minor axis of b mm., the width of the scanning line will be considerably less than a mm. with the scanning along the minor axis and less than b mm. with the scanning along the major axis. If the spot is circular, the line width is considerably less than the diameter of the spot. For a picture with equal horizontal and vertical detail it is necessary that the spot have equal dimensions along and normal to the direction of scanning. The spot of a well designed and constructed tube is circular in shape, and hence the discussion of the remainder of this section will be limited to circular spots.

The measured width of the scanning line decreases as the speed of the spot is increased. To see this, consider the case of a spot of diameter $2r$, and of uniform intensity, deflected into a scanning pattern. Although the spot is actually present at any point of the pattern for only a very short time during each frame, one may,[1] for light-distribution considerations, think of the spot as permanently drawn out into a line. The length of this hypothetical line equals the product of spot speed along the scanning line by the time of one frame. Let n be the number of lines per frame, m the number of frames per second, ρ the ratio of line return time to line scanning time, and l the length of one scanning line; then the speed of the spot along the scanning line is

$$v_s = mn(1 + \rho)l \qquad (9.4)$$

and therefore the length of the hypothetical line is

$$l_s = n(1 + \rho)l. \qquad (9.5)$$

Thus the speed of the spot along the scanning line for a 343-line, 30-frame picture for which $\rho = \frac{1}{16}$, and $l = 20$ cm. is 2.26×10^5 cm./sec., and the length of the hypothetical line is 7.55×10^3 cm.

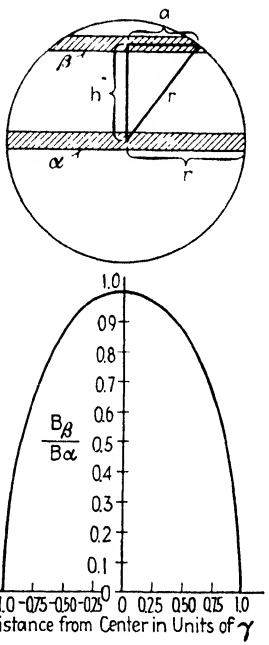


FIG. 9.18.—Distribution of intensity along width of scanning line of a deflected circular spot of uniform intensity.

Both strips α and β of the spot shown in Fig. 9.18 may be considered as stretched into strips of length l_s. The ratio of brightness of the center of the scanning line (resulting from the deflection of strip α) to that of the spot is $2r/l_s$, whereas the ratio of brightness of a point near the edge of the line (resulting from the deflection of strip β) to that of the spot is $2a/l_s$. Thus for a 1-mm. spot in the above example, the ratio of the brightness of the center of the line to that of the spot is

[1] By Talbot's law, which states that $B_{ave.} = \frac{1}{T}\int_0^T B dt$.

$$\frac{1}{7.55 \times 10^3} = 1.32 \times 10^{-5}.$$

If B_β be the brightness of the deflected strip β, and B_α that of the deflected strip α, then (see Fig. 9.18)

$$\frac{B_\beta}{B_\alpha} = \frac{a}{r} = \sqrt{1 - \frac{h^2}{r^2}}. \tag{9.6}$$

The curve on Fig. 9.18 shows the variation in brightness of the scanning line along a normal to direction of scanning. It was

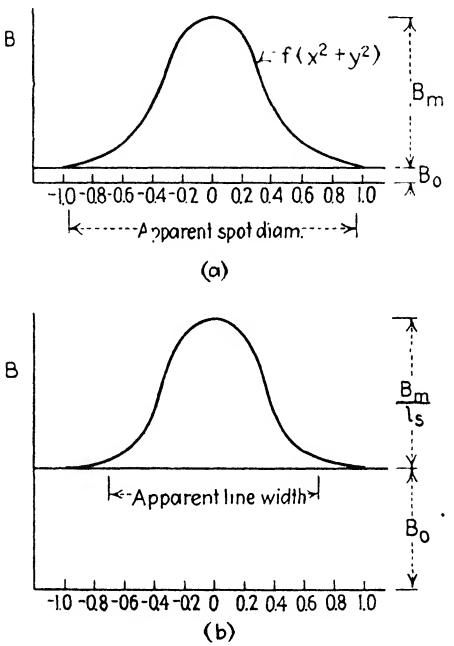


FIG. 9.19.—Distribution of intensity; (*a*) along diameter of stationary spot, (*b*) along width of scanning line.

calculated from Eq. (9.6), which presumes the deflection of a circular spot of uniform intensity. In practical cases, the screen of a TCR tube has always a background illumination resulting from multiple reflections within the tube itself as well as from room illumination. As a difference in brightness of about 5 per cent is required for clear distinction, the line width measured against the illuminated background will appear smaller than the diameter of the undeflected spot.

If, as is usually the case, the intensity distribution of the spot is not uniform but is symmetrical and given by some such curve as shown in Fig. 9.19(*a*), then there is a further reduction in the measured width of the scanning line. The distribution of brightness along the width of the line resulting from the deflection of a circular spot having the intensity distribution $f(x^2 + y^2)$ is

$$\frac{B_\beta}{B_\alpha} = \frac{\int_0^{\sqrt{r^2-y^2}} f(x^2 + y^2)dx}{\int_0^r f(x^2 + y^2)dx} \tag{9.7}$$

Figure 9.19(*b*) shows such a distribution of brightness superimposed upon the background illumination B_0. It is seen that due to the necessary difference in brightness of about 5 per cent for distinction, the measured line width is considerably less than the diameter of the stationary spot.

In general, the greater the aberration of the immersion and the bipotential lenses, the more non-uniform is the intensity distribution of the spot and the greater is the ratio of the undeflected-spot diameter to the scanning-line width. In TCR tubes of low aberration this ratio is about 1.5, while in tubes of high aberration ratios of 10 have been observed. On first thought it might appear that this high ratio is desirable. Further consideration will disclose, however, that normally such a high ratio is to be avoided as it reduces the maximum contrast obtainable by increasing the background illumination.

The measured line width for best focus in a given tube is practically independent of first-anode voltage, backing-aperture size and grid-skirt length.[1] As noted in Chap. 6 the action of the grid aperture is two-fold: it controls the number of electrons leaving the cathode and, together with the cathode and first anode, determines the immersion lens. By applying a more negative voltage on the grid, the cathode emitting area is reduced and the constants of the immersion lens are altered. The result of this is that the object imaged by the bipotential lens is reduced in size and moved farther from the bipotential lens. As a result, the line width decreases with increased grid bias. Further, as

[1] The line width does vary with grid-skirt length in those tubes where an increase in skirt length decreases the beam width and reduces the effect of aberration.

increased grid bias increases the object distance, the voltage ratio necessary for focusing will decrease. Figure 9.20 gives curves showing the variation of line width and voltage ratio with grid bias. The variations shown in Fig. 9.20 are often undesirable and may be avoided by focusing an aperture (<0.1 mm.) instead of the virtual image of the crossover.

9.10. Screen Grid.—Some constructors of TCR tubes make use of electron guns with an additional electrode between the grid and the first anode. This electrode is usually called *screen grid.* The electron gun, an equipotential-line plot of which is shown on

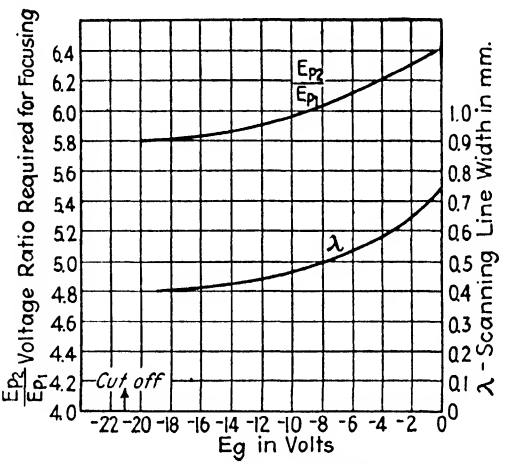


FIG. 9.20.—Variation of λ and E_{p2}/E_{p1} with E_g.

Fig. 4.1, has such an electrode. As in Fig. 4.1 the screen grid usually consists of an aperture disk placed close to the control grid, and it is usually operated at a potential lower than that of the first anode. The potential of the screen grid shown in Fig. 4.1 is 100 volts, and in operation of the TCR tube draws no current.

A characteristic of the screen grid is that it makes the total current drawn from the cathode almost independent of the first-anode voltage. This is especially true for low ratios of first-anode voltage to the screen-grid voltage (E_{p1}/E_s). Figure 9.21 gives two curves showing the variation of total cathode current with ratio E_{p1}/E_s. The effect of the ratio E_{p1}/E_s upon the ratio of beam current to first-anode current (I_{p2}/I_{p1}) is shown by the curve of Fig. 9.22.

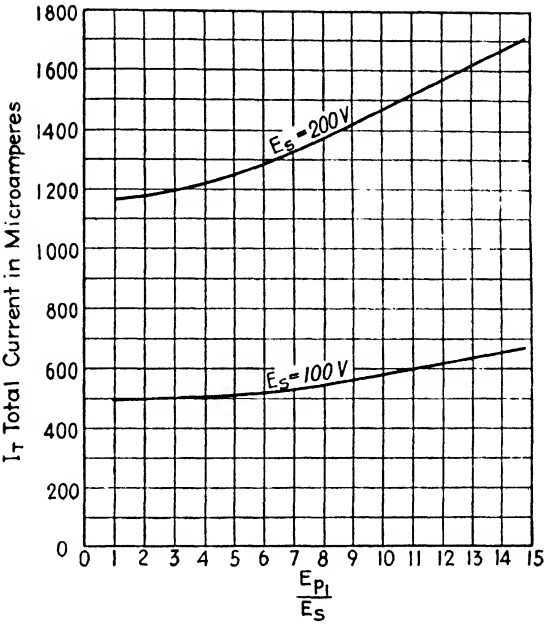


FIG. 9.21.—Variation of I_t with E_{p1}/E_s.

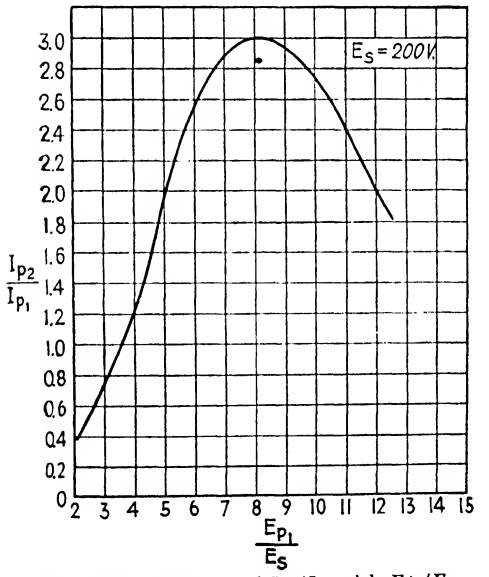


FIG. 9.22.—Variation of I_{p2}/I_{p1} with E_{p1}/E_s.

9.11. Cathode.—One of the most vital parts of the electron gun is the oxide-coated cathode used to generate the electron beam. If there were no limit to the number of electrons that may be drawn from a cathode, then the task of producing a small intense spot would be very simple. It would merely require the use of an extremely small cathode area. According to many authorities, for a reasonable cathode life not much more than 2 ma./sq. mm. may be drawn from the oxide-coated surface. Even such current densities require very careful preparation of the cathode.

The heater of present TCR tubes consists of several inches of tungsten wire less than 0.010 in. in diameter wound in a double spiral so as to neutralize its magnetic field. The winding of this double spiral is accomplished essentially by inserting the middle of the tungsten wire, which is stretched into a straight line, into a notch cut in the end of a cylindrical rod, and then imparting to the rod the motion of a screw. This noninductive heater winding is then cleaned and dipped into a suspension of aluminum oxide (Al_2O_3) in amyl acetate with a small percentage of nitrocellulose as binder. The coated heater is then dried in air and fired in hydrogen or vacuum at about 2000°A. for several minutes. After firing, the aluminum oxide insulation forms a fairly solid and uniform layer about 0.02 in. thick around the tungsten wire.

The cathode consists of a cylindrical nickel sleeve about 0.120 in. in diameter and 0.5 in. long, with a snugly fitting nickel cap. The emitting coating is sprayed on this cap by means of an ordinary spray gun. The spraying material usually consists of about 60 per cent barium carbonate ($BaCO_3$) and 40 per cent strontium carbonate ($SrCO_3$) suspended in amyl acetate with a small amount of nitrocellulose as binder. In spraying the cathode, great care has to be taken that the compressed air used in spraying is clean; the presence in the air of a very small amount of an impurity, such as oil, is sufficient to ruin a cathode. Cleanliness of all parts is important. It is advisable to hydrogen- or vacuum-clean the cathode before spraying. The amount of emitting material usually sprayed on is about 10 mg./sq. cm. of cathode surface. Although heavier coatings usually give longer life, their activation is much more critical and difficult.

The heater coated as described above is then inserted into the cathode sleeve, touching it at many points under light pressure.

In this manner a fairly good thermal connection and electrical insulation (about 10^7 ohms) is provided between the heater wire and cathode sleeve operating in vacuum. The heater normally operates at 2.5 volts, 2.2 amp. and at about 500°C. above the temperature of the cathode.

Having assembled the gun on the press and sealed the press into the glass blank, the tube is sealed to the pumps for evacuation (see Chap. 14). After evacuation and outgassing of the glass and electrodes, the cathode is ready for activation. The cathode is activated with the tube still sealed to the pumps.

There is no one definite way for activation. An oxide-coated cathode may be activated in several apparently different ways and yield satisfactory results. However, to obtain a satisfactory cathode the activation process has to be carried out differently according to the thickness of the coating, the manner in which the coating is applied, etc. The activation of the cathode is usually controlled by noting the magnitude of the current drawn from the cathode. The following process of activation is suitable for coatings deposited as described above:

1. The heater voltage is slowly increased to about 2 volts and kept there for a few minutes for outgassing.

2. With about 50 volts positive on all the gun elements the heater voltage is slowly increased until the heater current is about 3 amp. The heater current is kept at 3 amp. until the emission current (current drawn by gun elements) ceases increasing.

3. The filament current is then lowered to 2.2 amp. and the voltage on the gun elements is removed for a minute.

4. Emission readings are then taken with 5 and 10 volts on all the elements. Satisfactory cathode yields (for close grid-cathode spacings) are as below:

Voltage	Emission
5	>3 ma.
10	>7 ma.

Figure 9.23 shows the simple circuit arrangement used for cathode activation.

9.12. Similitude Relations.—The two similitude relations about to be stated are general aids of considerable importance for the design and understanding of TCR tubes. The first

relation states that if all the dimensions of the electrodes of a tube are multiplied by a constant factor, then the trajectories of the electrons in the tube are also multiplied by the same factor. The second states that the trajectory of an electron remains unchanged by multiplying the voltages on all the electrodes by any constant factor. These two relations are valid for any electrostatic field and are not limited to the case of axially symmetric fields. Thus these relations hold in the case of a tube containing deflecting plates. In the case of axially symmetric fields, these relations follow, as already noted in Sec. 4.7, from Eq. (4.17).

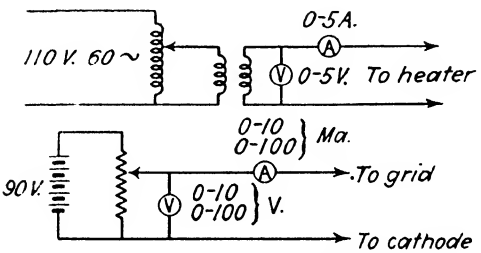


FIG. 9.23.—Circuit for cathode activation.

These two relations have already been used in the case of the bipotential lens where it was seen that the focusing action of the lens depends only on the *ratio* of cylinder diameters and the *ratio* of voltages on the cylinders. If the g.d. be chosen as the unit of length, then the truth of the first relation is obvious. Thus the trajectories of electrons measured in g.d. will remain unaltered if the dimensions of the tube measured in g.d. remain unchanged, and no change in the dimensions of a tube measured in g.d. is effected by multiplying all the dimensions measured in, say, centimeters by a constant factor. Hence, measured in g.d., the trajectories of the electrons in a tube are unchanged by multiplying the dimensions of all the parts of a tube measured in centimeters by any constant factor.

As an example illustrating the usefulness of the similitude relations, consider the following case:

Given a tube employing electric deflection and having a luminescent screen 10 in. in diameter, suppose that the tube is capable of reproducing a picture of satisfactory detail and brightness. It is desired to design a tube having a screen 5 in.

in diameter, utilizing electric deflection and capable of reproducing a picture one-quarter as large in area but of equal detail and brightness as the picture on the 10-in. tube. Equal detail and brightness will be obtained if the spot size and beam power of the 5-in. tube are respectively one-half and one-quarter those of the 10-in. tube.

A solution to this problem obtained with the aid of the two similitude relations is merely to multiply the dimensions of all the parts of the 10-in. tube by a factor of $\frac{1}{2}$, and to multiply the voltages on the electrodes by a factor of $\frac{1}{\sqrt{2}}$. That this is really a solution may be seen as follows: Consider first that every dimension of the 10-in. tube has been halved but that the voltages have not been altered. Then, measured in g.d., the size of the luminous spot and the amount of deflection are unchanged. But, measured in centimeters, both the size of the spot and the amount of deflection have been halved. Although the emitting area of the cathode has been reduced to one-quarter, the current will only be halved since the voltage gradient at the cathode has been doubled. If the voltages on all the electrodes are now decreased by a factor $1/\sqrt{2}$, then, assuming a linear relation between current and first-anode voltage (see Fig. 9.9), the beam power will have been reduced to one-quarter.

CHAPTER 10

DEFLECTION OF ELECTRON BEAMS

10.1. General.—There are three kinds of systems for deflection of electron beams used in television:[1] first, by means of two magnetic fields at right angles to each other; second, with two electric fields at right angles to each other; and third, with an electric and with a magnetic field parallel to each other. A deflecting system comprises the arrangements for generating synchronized electrical impulses of saw-tooth shape and also the arrangements of electrodes and poles for the generation of electric and magnetic fields for the deflections *per se*. The electrical network is called the *driving circuit for all-magnetic deflecting system, driving circuit for all-electric deflecting system* and *driving circuit for combined electric-magnetic deflecting system*, respectively. The arrangement of poles and coils for magnetic deflection is called a *magnetic deflecting yoke* and is invariably an arrangement external to the glass envelope of the TCR tube. The arrangement for producing electric fields for deflecting the electron beam is called *deflecting plates* and is invariably included inside the TCR tube envelope. A combined magnetic-electric system calls for a magnetic deflecting yoke and also for a pair of deflecting plates.

Four factors are important when considering a particular arrangement for deflecting or scanning. First, the system must not require more than a reasonable amount of power for a full-size pattern; second, the luminous spot must maintain its size and shape when deflected to the edges of the pattern; third, the pattern must not deviate from its normal rectangular shape; and fourth, the system must be capable of giving a high enough ratio of the picture to return sweep. The properties corresponding to these requirements are:

1. Deflection sensitivity.
2. Freedom from defocusing of the luminous spot.

[1] MALOFF, I. G., *Proc. Radio Club of Amer.*, **12** (1935).

3. Freedom from distortion of the pattern.
4. High enough overall frequency response.

The above requirements apply to any system of deflection, but the mechanics of magnetic and electrostatic deflection differ greatly.

10.2. Fundamental Relations of Magnetic Deflection.—Some of the relations which will be now given were already treated in Part I. They will be repeated only for the convenience of the reader.

While the magnetic field which deflects the beam is formed in a medium the permeability of which is unity (vacuum or air), the magnetic yoke often contains iron for the purpose of confining the field and reducing reluctances of return paths, thereby reducing the total energy stored in the magnetic field for a given deflecting effect.

A magnetic field is usually described by its field intensity $\bar{H}$. The force on an electron moving in such a field with a velocity $\bar{v}_0$ is

$$(\bar{H} \times \bar{v}_0 e)$$

which is related to the mass m of the electron and the resultant acceleration $\bar{a}$, as follows:

$$m\bar{a} = (e\bar{H} \times \bar{v}_0)$$

and the acceleration becomes

$$\bar{a} = \left(\bar{H} \times \bar{v}_0 \frac{e}{m}\right).$$

This becomes in scalar notation:

$$a = \frac{e}{m} H v_0 \sin \theta$$

where θ is the angle between the direction of the velocity of the electron and the magnetic-field intensity. The direction of a is at right angles to the plane through H and v_0.

Since the direction of the acceleration is at right angles to v_0 the electron cannot gain speed while going through a magnetic field; it can change only its direction. The kinetic energy of an electron moving in a magnetic field is a constant quantity, and therefore the radius of curvature R of the orbit can be calculated

from the law of conservation of energy

$$\frac{mv^2}{R} = Hev \sin \theta$$

$$R = \frac{mv}{eH \sin \theta}. \tag{10.1}$$

When $\sin \theta$ is equal to unity the latter expression becomes

$$R = \frac{mv}{eH} \tag{10.2}$$

where m, e, and v maintain their original values throughout the phenomena.

10.3. Computation of Magnetic Deflection.—Consider an electron (or a narrow electron beam) entering a layer of uniform magnetic field of intensity H. In Fig. 10.1 a is the path of an incident electron and SS are the intersections of plane boundaries of the magnetic field with the plane of the paper. The direction of H is vertical and perpendicular to the plane of the paper. Let R be the radius of a circle which an electron of the beam will begin to describe while under the action of the magnetic field. Then

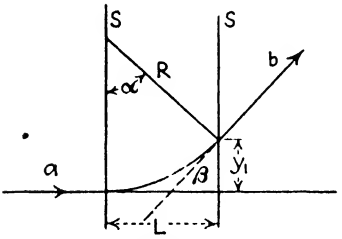


FIG. 10.1.—Electron trajectory in magnetic deflecting field.

$$R = \frac{mv}{eH} = \frac{m\sqrt{2V\frac{e}{m}}}{eH} = \frac{\sqrt{2V\frac{m}{e}}}{H} = \frac{k}{H}$$

where V is the potential difference through which the electron has fallen and k is a constant. If the angle between the incident and the refracted beam is β, then by inspection of Fig. 10.1 one may write

$$\sin \beta = \sin \alpha = \frac{L}{R} = \frac{LH}{\sqrt{2V\frac{m}{e}}} = \frac{LH}{k}$$

and

$$y_1 = R(1 - \cos \alpha) = R(1 - \sqrt{1 - \sin^2 \alpha})$$

$$= R\left(1 - \sqrt{1 - \frac{L^2}{R^2}}\right) = R - \sqrt{R^2 - L^2}. \tag{10.3}$$

Now referring to Fig. 10.2, if W is the total deflection and D is the distance from the point where the beam enters the uniform

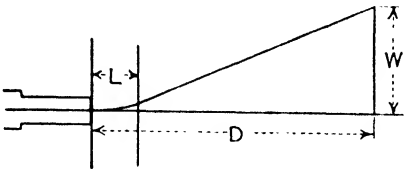


FIG. 10.2.—Magnetic deflection of electron beam.

magnetic field to the screen, then

$$W = y_1 + (D - L)\tan\alpha$$

and $$\tan\alpha = \frac{\sin\alpha}{\cos\alpha} = \frac{\frac{L}{R}}{\sqrt{1 - \frac{L^2}{R^2}}} = \frac{L}{\sqrt{R^2 - L^2}}.$$

Therefore

$$W = R - \sqrt{R^2 - L^2} + (D - L)\frac{L}{\sqrt{R^2 - L^2}}$$

$$= R + \frac{DL - R^2}{\sqrt{R^2 - L^2}}. \tag{10.4}$$

Unfortunately, W in this relation is not a simple function of either R, L or D. However, if L^2 is small compared to R^2 this relation becomes

$$W = R + \frac{DL - R^2}{R} = D\frac{L}{R}. \tag{10.5}$$

Since at present an L/D ratio much larger than about .2 is seldom used, it will be well worth while to compute W for several values of R, L and D by the exact Eq. (10.4) and also by the approximate Eq. (10.5) to see if the approximation just given is justifiable. The results of such computation are given in Table 10.1.

TABLE 10.1

D	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
L	1	1	1	2	2	2	5	5	5	10	10	10	20	20	20
R	10	5	2.5	20	10	5	50	25	12.5	100	50	25	200	100	50
W	10	20.3	43.4	9.95	20.2	43.2	9.9	19.9	42.7	9.5	19.4	41.4	9	18.4	39.1
DL/R	10	20	40	10	20	40	10	20	40	10	20	40	10	20	40

Another useful approximation for deflection angles used in television and oscillographic work is

$$\alpha \cong \tan \alpha \cong \frac{W}{D} \cong \frac{L}{R}. \tag{10.6}$$

It is interesting to know the variation of current through the coils of a deflecting yoke as a function of the length of the lamination stack, for constant values of W, D and the total inductance of the circuit. To keep the inductance of the yoke constant we must satisfy the relation

$$LN^2 = 1 \qquad \text{or} \qquad N = \frac{1}{\sqrt{L}}$$

The current I through the inductance is proportional to the H desired and to the reciprocal of the number of turns N

$$I = \frac{H}{N} = H\sqrt{L}.$$

But

$$H = \frac{W}{DL} \text{ approx.}$$

Therefore

$$I = \frac{W}{D\sqrt{L}}. \tag{10.7}$$

In other words, for a magnetic yoke of increasing length L, with the inductance kept at a constant value by a corresponding reduction in the number of turns, the current required for a given deflection is proportional to the square root of the reciprocal of the length of the magnetic yoke, to a fair degree of approximation. The end effects of the magnetic field (not taken into account in these computations) will tend to decrease this effect slightly.

The power required for a given deflection and the energy stored in the magnetic field are inversely proportional to the length of the deflecting yoke.

The statement above means that if a given beam can be deflected under certain conditions by means of two power tubes, doubling the length of the deflecting yoke will require the use of only one tube to accomplish the same result. In case the length of a deflecting circuit is increased beyond 20 per cent of the

gun-screen distance, the approximate formula which we have just derived does not hold. A computation by an exact formula for a case where W and D are kept constant while L and R are varied, and where the inductance of the yoke is kept constant by a corresponding adjustment in the number of turns, gives the results shown in Table 10.2. Table 10.2 indicates that there is

TABLE 10.2

D	100	100	100	100	100	100	100	100	100	100	100	100	100	∞
L	100	90	80	70	60	50	40	35	30	20	10	5	2	1
W	20	20	20	20	20	20	20	20	20	20	20	20	20	20
I[1]	0.0385	0.0380	0.0354	0.0356	0.0358	0.0368	0.0386	0.0401	0.0417	0.0485	0.0650	0.0800	0.1400	0.1970
I[1] %	19.5	19.3	18.0	18.1	18.2	18.7	19.6	20.3	21.2	24.6	33.0	40.6	71.0	100.0

[1] Current for constant inductance and deflection.

no reason to increase L more than 80 per cent of the gun-screen distance and that 30 per cent for all practical purposes is as good as 80 per cent, since the accuracy of the computation is impaired by neglecting the edge effects.

In the computations given on the preceding pages, no attention was paid to units and proportionality factors. The relation for R is

$$R = \frac{\sqrt{2V\frac{m}{e}}}{H}.$$

This relation calls for either of the absolute systems of units, e.m.u. or e.s.u. To convert it into practical units it is necessary to express all the quantities in either of the systems, say e.m.u., and use the conversion factors

$$V_{\text{e.m.u.}} = 10^8 V_{\text{volts}}$$
$$H_{\text{e.m.u.}} = H_{\text{gauss}}$$
$$\frac{e}{m} = 1.77 \times 10^7 \text{ e.m.u./g.}$$

Therefore

$$R_{\text{em.}} = \frac{\sqrt{1.13 \times 10^{-7} V_{\text{e.m.u.}}}}{H_{\text{e.m.u.}}} = \frac{\sqrt{1.13 \times 10^{-7} \times 10^8 V_{\text{volts}}}}{H_{\text{gauss}}}$$
$$= 3.36\frac{\sqrt{V_{\text{volts}}}}{H_{\text{gauss}}}[1] \qquad (10.8)$$

[1] In air the field intensity H is equal to the flux density B and can be expressed in either gauss or gilberts per centimeter.

and of course

$$H_{\text{gauss}} = 3.36\frac{\sqrt{V_{\text{volts}}}}{R_{\text{cm.}}}. \tag{10.9}$$

The measure of sensitivity of a particular deflecting yoke is the amount of total energy stored in the magnetic field for a given full deflection or a given angle between extreme deflections of the luminous spot from one edge of the pattern to the opposite. This quantity is

$$E = \frac{L_0 I^2}{2}. \tag{10.10}$$

Here E is in joules (watt-seconds), L_0 is in henries and I is in amperes. If the picture sweep is repeated n times per second,

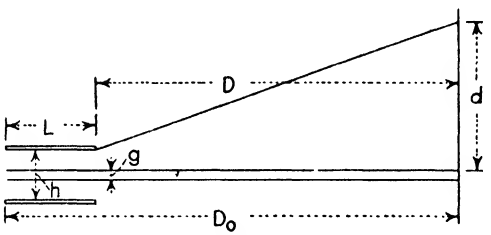


FIG. 10.3.—Electrical deflection of electron beam.

and after each picture sweep this energy is dissipated, then the output tube should be capable of delivering nE watts to the yoke.

The reason for taking current needed to sweep the spot from one edge of the tube to the opposite one for computation of the total energy stored in the magnetic field is because the normal deflecting driving circuit is connected to the yoke through a large condenser. So far as deflecting energy is concerned, this is equivalent to shifting the gun in the tube so that the spot rests at one edge of the tube with no current through the yoke.

10.4. Computation of Electric Deflection. *a. With Parallel Deflecting Plates.*—There are two shapes of electrostatic deflecting plates in wide use today, parallel plates and curved plates. This section will deal with the case of parallel plates and with the calculation of optimum length and separation of parallel plates.

In the following computation of the electrostatic deflection, an assumption is made that there is no stray field of any kind present; neither is there an edge effect around the deflecting plates. If the plates are parallel and the plate separation is h, then the gradient of the electrostatic field between the plates is V/h, where V is the potential difference between the plates.

Taking a practical problem where the total path D_0 from the beginning of the plates to the luminous screen is fixed, and letting d be the desired maximum deflection, one may compute the optimum length L_0 of the deflecting plates and the optimum separation h_0 (see Fig. 10.3).

The intensity of the electrostatic field is

$$\frac{V}{h}.$$

Also
$$D = D_0 - L.$$

The transverse or radial acceleration of an electron due to this electrostatic field is

$$a = \frac{V}{h}\frac{e}{m}.$$

The time that an electron is under influence of the field is:

$$t = \frac{L}{v_0}$$

where v_0 is the axial velocity of the electron. The final radial velocity of the electron as it leaves the field is

$$at = \frac{V}{h}\frac{e}{m}\frac{L}{v_0}.$$

The radial deviation or deflection of the electron as it leaves the field is

$$S_r = \frac{at^2}{2} = \frac{1}{2}\frac{V}{h}\cdot\frac{e}{m}\cdot\frac{L^2}{v_0^2}.$$

However

$$v_0^2 = 2E\frac{e}{m}$$

where E is the voltage through which the electron has fallen before entering the deflecting field.

Therefore

$$S_r = \frac{1}{4}\frac{V}{E}\frac{L^2}{h}.$$

The final slope of the trajectory as the electron leaves the field is

$$\frac{v_r}{v_0} = \frac{1}{2}\frac{V}{E}\frac{L}{h}.$$

Hence the deflection after the electron left the field is

$$S_r' = \frac{1}{2}\frac{V}{E}\frac{L}{h}(D_0 - L).$$

The total deflection at the screen then is

$$d = \frac{1}{4}\frac{V}{E}\frac{L^2}{h} + \frac{1}{2}\frac{V}{E}\frac{L}{h}(D_0 - L)$$

or

$$d = \frac{1}{2}\frac{V}{E}\frac{L}{h}\left(D_0 - \frac{L}{2}\right) \tag{10.11}$$

where both E and V may be expressed in practical volts and the rest of the quantities in centimeters. For a given maximum deflection d_0 the field must be

$$\frac{V}{h} = 2d\frac{E}{L}\frac{1}{D_0 - \frac{L}{2}}. \tag{10.12}$$

It is desired to find the smallest possible value of h, which for the beam of finite diameter g becomes

$$h = \frac{1}{2}\frac{V}{h}\frac{L^2}{E} + g. \tag{10.13}$$

If h is restricted to this value, after substituting (10.12) into (10.13), it follows that

$$h = \frac{1}{2}2d\frac{E}{L}\frac{1}{D_0 - \frac{L}{2}}\frac{L^2}{E} + g$$

or

$$h = \frac{Ld}{D_0 - \frac{L}{2}} + g \tag{10.14}$$

which is a new function of L and can be substituted into the final equation of the field required (10.12). Then

$$\frac{V}{\dfrac{Ld}{D_0 - \dfrac{L}{2}} + g} = \frac{2dE}{L} \frac{1}{D_0 - \dfrac{L}{2}}.$$

Transposing one gets

$$\frac{V}{E} = 2\left[\left(\frac{d}{D_0 - \dfrac{L}{2}}\right)^2 + \frac{g}{L} \frac{(d)}{\left(D_0 - \dfrac{L}{2}\right)}\right]. \tag{10.15}$$

Differentiating (10.15) with respect to L and equating the derivative of the right-hand member to zero one gets the expression for optimum L

$$\frac{d^2}{\left(D_0 - \dfrac{L_0}{2}\right)^3} - gd \frac{D_0 - L_0}{L_0^2 \left(D_0 - \dfrac{L_0}{2}\right)^2} = 0.$$

After suitable transformations the optimum equation can be brought to a form

$$(2d - g)L_0^2 + 3gD_0L_0 - 2gD_0^2 = 0$$

the solution of which is:

$$L_0 = \frac{-3gD_0 + \sqrt{(3gD_0)^2 + 4(2d - g)(2gD_0^2)}}{4d - 2g} \tag{10.16}$$

since only the positive root is of interest. Here L_0 is the optimum value of L.

The optimum separation h_0 is computed from Eq. (10.14) as follows:

$$h_0 = \frac{L_0 d}{D_0 - \dfrac{L_0}{2}} + g. \tag{10.17}$$

Now a specific problem, that of a typical 9-in. TCR tube, will be considered. Here $d = 10$ cm., $D_0 = 30$ cm., and $g = 0.5$ cm. By means of relations (10.16) and (10.17) one arrives at the optimum length of plates $L_0 = 5.9$ cm. and the optimum separa-

tion $h_0 = 2.68$ cm. The corresponding minimum ratio of the voltage needed for the full deflection is found to be equal to 0.338 [from Eq. (10.15)].

In Fig. 10.4 the variation of V/E for a 9-in. TCR tube with L varying from zero to $L = D_0$ is shown graphically. The case is restricted to the best possible separation for each length of the plates.

It may be here noted that the sensitivity of an electrostatic tube can be very conveniently expressed in terms of the ratio of the voltage on the plates needed for full deflection to the accelerating or the second-anode voltage. *This value* is independent

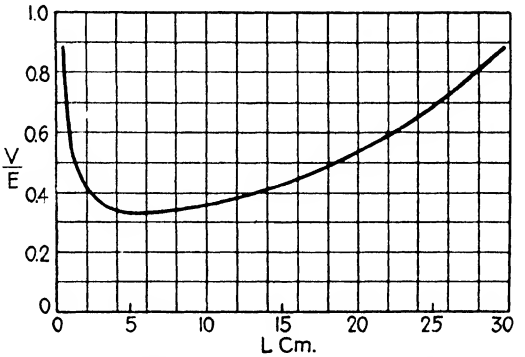


FIG. 10.4.—Variation of deflection sensitivity with length of deflecting plates.

of the actual magnitudes of the voltages and *depends only on the geometry of the tube.*

10.5. Computation of Electric Deflection. *b. With Curved Plates.*—The best sensitivity, however, cannot be obtained with straight parallel deflecting plates. It can be obtained if at any point along the path the plates are brought together as closely as possible without touching the beam. In the following calculations the assumption is made (which is quite accurate especially where the field is most concentrated) that the electrostatic potential gradient anywhere along the axis L of the tube is equal to V/h, where V is the potential difference between the plates and h is the separation of plates measured along a line perpendicular to L.

Then if $h = f(L)$, the transverse or radial acceleration is

$$\frac{e}{m}\frac{V}{f(L)}$$

and this radial acceleration is equal by its definition to the second derivative of the radial path (or transverse path) with respect to time. This can be expressed as follows:

$$\frac{d^2S_r}{dt^2} = \frac{V}{f(L)}\frac{e}{m} \tag{10.18}$$

but $t = \frac{L}{v_0}$, also $v_0^2 = 2E\frac{e}{m}$, where v_0 is the longitudinal velocity of the electron and E is the accelerating voltage, or the potential through which the electron has fallen. Substituting these relations into (10.18),

$$\frac{d^2S_r}{dL^2} = \frac{1}{2}\frac{V}{E}\frac{1}{f(L)}. \tag{10.19}$$

But we want the plates to be flush with the outer boundary of the beam throughout the path. Therefore, the plate separation must be, at any point along the axis L, equal to twice the radial path plus the diameter of the beam (by diameter the actual diameter plus the allowances is meant). So

$$h = f(L) = 2S_r + g$$

where g is the diameter of the beam. Therefore

$$\frac{d^2S_r}{dL^2} = \frac{1}{2}\frac{V}{E}\frac{1}{2S_r + g}. \tag{10.20}$$

Let $S = 2S_r + g,$

then $dS = 2dS_r$

$d^2S = 2d^2S_r$

$d^2S_r = \frac{1}{2}d^2S.$

We may rewrite (10.20) in this form

$$\frac{d^2S}{dL^2} = \frac{V}{E}\frac{1}{S}$$

which is a differential equation of a standard form. Integrating it once,

$$\frac{dS}{dL} = \sqrt{2\frac{V}{E}\log S + C}$$

but $S = 2S_r + g$ and $\frac{dS}{dL} = 2\frac{dS_r}{dL}$.

Therefore

$$\frac{dS_r}{dL} = \sqrt{\frac{1}{2}\frac{V}{E}\log(2S_r + g) + C}.$$

We know that for $S_r = 0$, $dS_r/dL = 0$. Therefore

$$\sqrt{\frac{1}{2}\frac{V}{E}\log g + C} = 0 \qquad \text{and} \qquad C = -\frac{1}{2}\frac{V}{E}\log g.$$

Consequently

$$\frac{dS_r}{dL} = \sqrt{\frac{1}{2}\frac{V}{E}}\sqrt{\log\left(\frac{2S_r + g}{g}\right)}. \tag{10.21}$$

Rewriting (10.21),

$$dL = \frac{dS_r}{\sqrt{\frac{1}{2}\frac{V}{E}}\sqrt{\log\frac{2S_r + g}{g}}}.$$

Integrating once again,

$$L = \sqrt{2\frac{E}{V}}\int_0^{S_r} \frac{dS_r}{\sqrt{\log\frac{2S_r + g}{g}}}.$$

Let $\frac{2S_r + g}{g} = R$, then $S_r = \frac{g}{2}(R - 1)$ and $dS_r = \frac{g}{2}dR$.

One may then write

$$L = \frac{g}{2}\sqrt{2\frac{E}{V}}\int_1^{\frac{2S_r+g}{g}} \frac{dR}{\sqrt{\log R}}. \tag{10.22}$$

The definite integral of the expression can be evaluated graphically. A table at the end of this section gives several values of this integral.

The best possible value of voltage ratio V/E is computed from relation (10.22) as follows:

$$\frac{V}{E} = \frac{1}{2}\left(\frac{g}{L}\int_1^{\frac{2S_r+g}{g}} \frac{dR}{\sqrt{\log R}}\right)^2 \tag{10.23}$$

where for L and S_r, the total path and the desired deflection are substituted. Once the best value of V/E is determined it is

substituted into (10.22), and then the relation between S_r and L can be tabulated. When S_r is plotted against L the graph gives the optimum shape of the deflecting plate.

For the average 9-in. static TCR tube, the beam can be taken as 0.5 cm. in diameter. The normal path is 30 cm., and the deflection is 10 cm.

From (10.23),

$$\frac{V}{E} = \frac{1}{2}\left(\frac{0.5}{30}\int_1^{41} \frac{dR}{\sqrt{\log R}}\right)^2$$

because

$$\frac{2S_r + g}{g} = 41.$$

Consequently

$$\frac{V}{E} = 0.094.$$

For the intermediate values of L

$$\frac{g}{2}\sqrt{2\frac{E}{V}} = \frac{L}{\int_1^{41} \frac{dR}{\sqrt{\log R}}} = \frac{30}{26} = 1.15.$$

The relation between L and S_r may now be tabulated from the following relation:

$$L = 1.15\int_1^{\frac{2S_r+0.5}{0.5}} \frac{dR}{\sqrt{\log R}}.$$

TABLE 10.3.—TABULATION OF L AS FUNCTION OF S_r

S_r	10	8	6	5	4	3	2	1.5
L	30.0	25.0	20.2	17.5	14.7	12.1	8.87	7.37
S_r	1.0	0.8	0.6	0.4	0.2	0.1	0.05	0.025
L	5.64	4.95	4.03	3.23	2.15	1.46	1.01	0.68

Figure 10.5 shows the optimum shape of deflecting plates for a 9-in. kinescope, computed with the enumerated assumptions.

10.6. Defects of the Scanning Pattern.—There are two main forms of defects of the scanning patterns on the screens of cathode-ray tubes. The first is defocusing of the luminous spot, and the second is distortion of scanning pattern.

By defocusing of the luminous spot is meant an increase in the size of the spot when deflected. By distortion of the scanning pattern is meant the deviation of the pattern from its normal rectangular shape.

The degree to which the above defects may be present in a particular deflecting system is determined primarily by the shapes and types of the deflecting fields. There are two more common defects caused more or less by the deflecting circuit as a whole. They are non-uniform distribution of the scanning pattern, and cross-talk between the vertical and horizontal circuits. For the

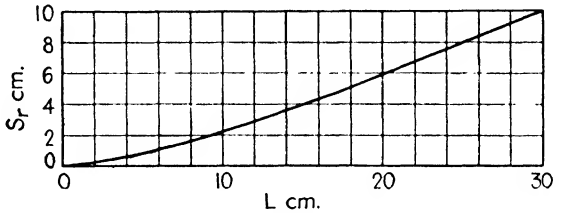


FIG. 10.5.—Optimum shape of deflecting plates.

first of these, the wave shape of the driving circuit and the frequency response of the arrangement are responsible. For the second, either the coupling (of any kind) between the driving circuits or the coupling between the magnetic fields of the yoke may be the cause. All three systems of deflection, *viz.*, the all-magnetic, the all-electric and the combined electric-magnetic, are subject to the enumerated defects. The general character of defects of the magnetic and static systems differs so greatly, however, as to require independent treatment.

10.7. Defocusing of the Luminous Spot by Magnetic Deflecting Field.—The defocusing or the increase of the size of the luminous spot when deflected by means of a magnetic field is due to two factors: First, for a given non-uniform magnetic field it is a direct function of the beam diameter as it goes through the deflecting field, with slow electrons, if present, accentuating the effect. Second, for a given cathode-ray tube it is a function of non-uniformity of the field in the direction of deflection.

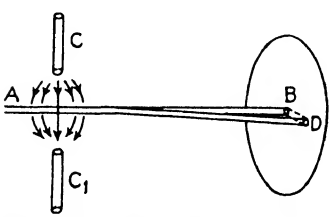


FIG. 10.6.—Defocusing magnetic deflecting field.

The above statement will be better understood by looking into the details of the mechanism of defocusing.

Referring to Fig. 10.6, consider an electron beam *A* of a circular cross section and with electrons moving parallel to each other.

Such a beam when undeflected will produce a luminous spot B on the screen. This spot will be of a circular shape also. Now let this beam be deflected to one side of the screen by means of a magnetic field produced by electromagnets (or permanent magnets) C, C_1. Following the corkscrew rule, the beam will be so deflected that the spot will shift to D. The magnetic field produced by two coaxial round bar magnets will be of a shape such as shown in Fig. 10.7, giving the distribution of lines of flux through a plane cutting the field along the plane of deflection. The spaces between adjacent circles on this figure enclose equal amounts of lines of force. (In air and vacuum, the magnetic-flux density is numerically equal to the magnetic-field intensity.)

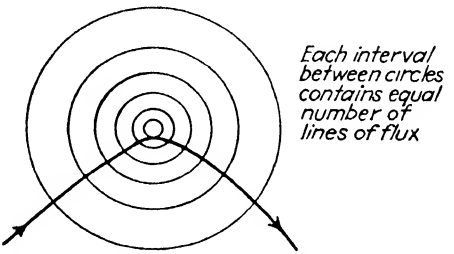


FIG. 10.7.—Magnetic field of two bar magnets.

Now suppose this cylindrical electron beam has an initial direction towards a point on the axis of the magnetic field and this point is equidistant from the magnets C and C_1 in Fig. 10.6. Such a beam will be deflected by the field in a manner shown in Fig. 10.7 and will miss the axis of the magnetic field. The side of the beam which will pass the closer to the axis will be acted upon by a denser field and therefore will be deflected more. The opposite side will be deflected less. The spot, by the time it is deflected, will be compressed along the direction of deflection and will tend to take the shape of an oval or ellipse, with its major axis perpendicular to the direction of deflection.

The field shown in Figs. 10.6 and 10.7 possesses a barrel-shape form when its cross section through the axis of symmetry is observed. This means that at any distance from the axis, the lines of force possess a definite curvature. It can be shown mathematically that any non-uniform magnetic field possesses also a certain curvature, which is a function of non-uniformity. (This holds true for any magnetic field in a continuous medium of constant permeability.)

Fig. 10.8 shows a beam of a cylindrical shape being deflected away from the center of a barrel-shaped field. Away from the plane of symmetry of the field, the curvature of the field results in a component of the field parallel to the plane of symmetry. These components, however, have opposite directions on the opposite sides of the plane of symmetry.

In the case shown in Fig. 10.8 (which is the same case as shown in Fig. 10.7), the upper and lower parts of the beam will be stretched away from the plane of symmetry in opposite directions, changing the shape of the luminous spot from a circle to that of an ellipse with a major axis perpendicular to the direction of deflection. Therefore, the non-uniformity of the field and the curvature of the field both act to change the luminous spot into an ellipse with the major axis perpendicular to the direction

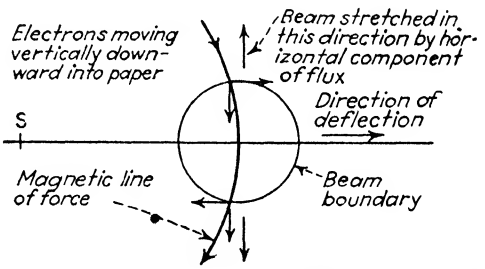


FIG. 10.8.—Beam defocusing by magnetic deflecting field.

of deflection. But this will hold only if the deflection is away from the region of the field where the field is more concentrated.

When a cylindrical beam is pulled into a field towards the region where it is more concentrated, the beam is stretched into an ellipse with its major axis parallel to the direction of deflection.

One may look at the effects of a non-uniform field from another angle. It affects a cylindrical beam as a divergent cylindrical lens. For deflection towards weaker regions of the field, the axis of this lens is parallel to the plane in which the direction of deflection lies. For deflection towards stronger regions of the field, the axis is perpendicular to this plane, and the larger the beam diameter the larger is the effect of a given field.

So far only cylindrical beams have been considered. In practice there are always converging beams, which are either focused, under-focused or over-focused. It can be shown, by reasoning similar to that given in the preceding paragraphs,

that if a field stretches an over-focused beam in a particular direction, a readjusting of the focusing field to give an under-focused condition will stretch the spot in a direction perpendicular to the former.

One could go on indefinitely trying to evaluate these and other effects of non-uniformities and curvatures of deflecting fields. It seems, however, entirely useless to do so, since it is known that these two factors are to be reduced to a minimum and that, as will be shown later, fields are known which, when deflecting normal conventional beams, give negligible defocusing and therefore possess sufficient uniformity for practical purposes.

10.8. Distortion of the Scanning Pattern by an All-magnetic Deflecting System.—By distortion of the scanning pattern is meant the deviation of the pattern from its normal rectangular shape. When all the four corners are pulled away farther than they should be, the pattern is called *pincushioned*, and when these corners are not pulled away enough the pattern is called *barrel-shaped*.

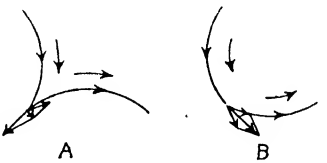


FIG. 10.9.—Distortion of scanning pattern by magnetic deflecting field.

Distortion of the scanning pattern, just as defocusing, is caused by non-uniformity and curvature of the deflecting field. A combination of two magnetic deflecting fields, each of which is of barrel-shape distribution, causes a pincushion pattern. A combination of two pincushion fields produces a barrel-shape pattern. The reason for these effects can be better understood by considering Fig. 10.9.

Fig. 10.9(*a*) shows how the components of two pincushion fields add together and give a comparatively small resultant for corner deflection and a barrel-shape pattern. Similarly, the components of two barrel-shape fields add together as shown in Fig. 10.9(*b*) and give a comparatively large resultant and pincushion pattern.

As in the case of defocusing, it is not necessary to evaluate these effects any further since it is known that deflecting fields should possess sufficient uniformity to reduce the distortion of the scanning pattern to a negligible amount. Furthermore, as will be shown later, fields are known which have the required degree of uniformity.

10.9. Defocusing of Luminous Spot by Electric Deflecting Field.—Elementary explanations of the defocusing of the luminous spot by an electric deflecting field are apt to be misleading. Such explanations assume no edge effect of the field of deflecting plates. While the edge effect does not alter much the results of computations of deflection, it is very important in estimating defocusing effect.

The old deflecting technique was to connect one of the deflecting plates to the second anode and to swing the other plate

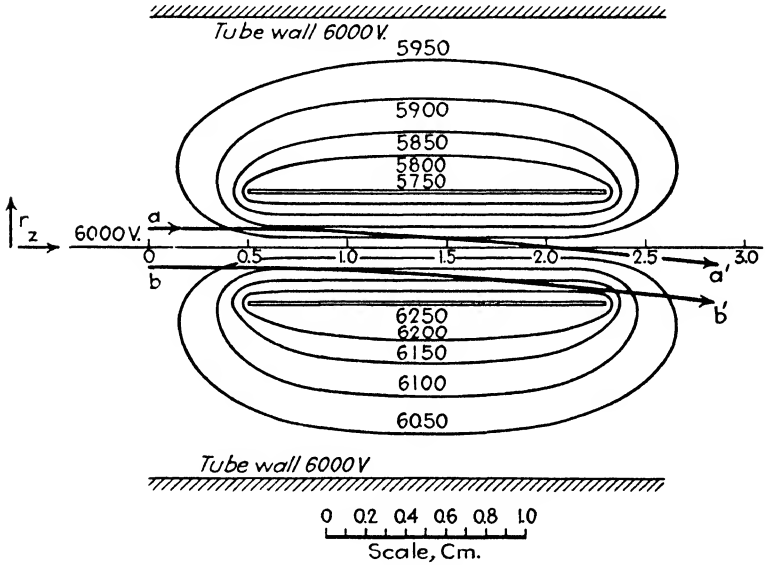


FIG. 10.10.—Symmetrical electric deflecting field.

above and below the second-anode potential. This method of unbalanced deflecting field gave very large defocusing and has been superseded by the balanced-field method. In the latter each plate is connected to the second anode through a high resistance (one or more megohms) and for deflection purposes one plate is lowered in potential by exactly the same amount as the other plate is raised. Fig. 10.10 shows an experimentally taken equipotential plot of a balanced or symmetrical deflecting field. Fig. 10.11 shows a roughly extrapolated plot of an unbalanced deflecting field.

Trajectories of two electrons initially parallel to the axis of the tube through the deflecting field shown in Fig. 10.10 were

calculated. These trajectories are shown in the figure as aa' and bb'. A plot similar to that shown on the figure, except for greater detail, was used in the actual calculations. The results obtained are as follows: from $z = 0$ to $z = 0.5$ cm. the r coordinate for a electron changed from 0.1 to 0.099 cm. while its z component of velocity diminished to 5,925 equivalent volts. The r coordinate for b electron, equal to -0.1 at $z = 0$, became -0.101 at $z = 0.5$ cm., while its z component of velocity became

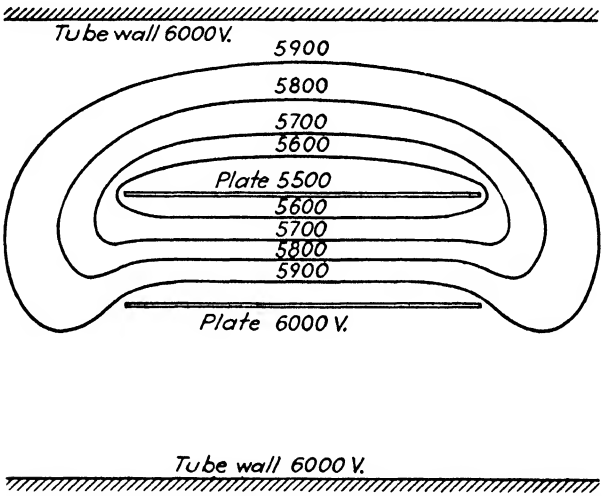


FIG. 10.11.—Electric deflecting field with one plate tied to second anode.

6,075 equivalent volts. The r component of velocity at $z = 0.5$ cm. comes out as -0.272×10^7 cm./sec. for both electrons. A conclusion follows that the edge effect on the entrance side of deflecting plates results in slowing down the upper electron and in speeding up the lower electron, with no change in separation between the two.

From $z = 0.5$ up to $z = 2.3$, the r coordinates of the electrons a and b change to -0.034 and -0.231 cm., respectively; the total velocities become 6,025 and 6,175 equivalent volts, respectively; while tangents of the angles of trajectories with the axis are 0.1414 and 0.138, respectively. This difference in angle, while small at the first glance, results in a 1-mm. increase in spot size for an initially focused electron beam 30 cm. long, for a total deflection of 8 deg.

Elementary explanations of electric defocusing stop at this point and predict an amount of defocusing which is many times greater than the amount observed in well constructed tubes.

It is true that the slower electron *a* is deflected more than the faster electron *b*, and the beam would appear over-focused. However, a further calculation shows that the edge effect of the electric field on the exit side of deflecting plates is such that the lower, faster electron is deflected more than the upper, slower electron. In the particular case illustrated in Fig. 10.10, the edge effect compensates the focusing action caused by the speed of the electrons to within the accuracy of the graph and calculations.

The amount of defocusing of the luminous spot by a balanced electric deflecting field is a residual of two large effects, the effect of the difference in speeds of electrons on the boundaries of the beam and the edge effect of the exit side of deflecting plates, the latter counteracting the former. In balanced deflecting fields, the defocusing effect is smaller than either of the two mentioned large effects.

The angle of deflection of 8 deg. in the example given is rather small, as in practical TCR tubes deflecting angles as high as 30 deg. away from the axis are not unusual.

In an unbalanced deflecting field such as shown in Fig. 10.11, the edge effect of the entrance side of the plates is of the order of the edge effect of the exit side, and is of the same sign. Since the speed effect is always present, the resultant defocusing by an unbalanced field is always large, and such field finds little use in television.

CHAPTER 11

LUMINESCENT SCREENS FOR TCR TUBES

11.1. Luminescence, Fluorescence and Phosphorescence. – Any emission of light not directly ascribable to incandescence and therefore occurring at low temperatures is called *luminescence*. In order to emit light, a substance must receive or consume energy in some form. This energy, which usually comes from without, is called the energy of excitation. There are a great many forms of luminescence differing one from another in forms of excitation, such as photoluminescence, cathodo-luminescence, chemi-luminescence, etc. In cathode-ray tubes the only type of luminescence utilized is cathodo-luminescence, or luminescence with electron bombardment as excitation.

The luminescence of a substance during the excitation is called *fluorescence*, while the luminescence of a substance after the excitation has been removed is called *phosphorescence*.

11.2. Phosphors.—Solid luminescent materials possessing properties of fluorescence and phosphorescence are commonly called *phosphors*. The material of the luminescent screens in TCR tubes is therefore a phosphor. The theory of luminescence of phosphors is very involved and as yet has not been firmly established in all its details.[1] The detailed theory of luminescence, however, is of more use to a person making and developing phosphors than it is to a maker and user of TCR tubes. In this treatise this theory will be avoided as much as possible without impairing the object of the work.

There are a great many phosphors known. Some of them were first discovered as natural minerals while the others were developed in laboratories. There are organic and inorganic phosphors. Since the phosphor in TCR tubes is operated in high vacuum and since in preparation of TCR tubes the phosphor

[1] NICHOLS, E. L., H. L. HOWES and D. T. WILBER, Cathodo-Luminescence and Luminescence of Incandescent Solids (a publication of Carnegie Inst. of Washington) (1928); also R. TOMASCHEK, Optik und Elektronik fester und flüssiger Stoffe, *Die Physik*, **2**, 33 (1934).

has to withstand temperatures of the order of 500°C., the organic phosphors are of no interest here.

The inorganic phosphors are more or less complicated chemical compounds of metals, such as zinc, cadmium, calcium and a few others, on one side, and oxygen, sulfur, silica, tungsten, etc., on the other side. A presence of small amounts of an impurity, usually metallic, is essential to the performance of some phosphors. This impurity is called activator, and some of the chemically, or rather spectroscopically, pure materials are practically non-luminescent. Of the great number of phosphors known, only two groups are commonly used for screens in TCR tubes. They are silicate phosphors and sulfide phosphors. The predominant metal (but not the only metal) in either of the two groups is zinc. Of silicate phosphors the most widely used is willemite. Chemically, willemite is zinc orthosilicate (Zn_2SiO_4) with small traces of manganese or other metals present. It is very popular with American makers of TCR tubes.[1] Of the sulfide phosphors, zinc sulfide (ZnS) is the best known. In Europe zinc sulfide, either silver- or copper-activated, and zinc-cadmium sulfide[2] have gained far greater popularity than willemite.

11.3. Efficiency of Luminescent Screen.—The subject of light efficiencies of luminescent screens involves a concept of brightness of surfaces which is rather complicated. There is a general confusion of units and terms used by physicists and engineers. The authors for their own use have adopted definitions and units used by illuminating engineers in the United States. The definitions and terms as applied to the subject of brightness of surfaces will be discussed briefly.

The unit of luminous intensity is the International Candle agreed upon in 1909 by the National Standardizing Laboratories of France, England and the United States. It is the same as the Pentane Candle, Bougie Candle, and American Candle.[3] Candlepower is luminous intensity expressed in candles.

[1] PERKINS, T. B., and H. W. KAUFMANN, Luminescent Materials for Cathode Ray Tubes, *I.R.E. Proc.*, **23**, p. 1324 (1935).

[2] LEVY, L., and D. W. WEST, Fluorescent Screens for Cathode Ray Tubes for Television and Other Purposes, *J. Inst. Elec. Eng.* (London), **79**, 11 (1936).

[3] In Germany, Hefner Candle is widely used. Hefner Candle is equal to 0.9 International Candle.

A point source of light having a luminous intensity of one candle generates $4\pi = 12.6$ lumens of luminous flux. In other words it will produce a flux density of one lumen per unit solid angle in all directions. A source of light may give this value of luminous-flux density along one direction and, say, half of that along some other. In such case its intensity is said to be one candlepower along the former and half of that along the latter.

When some source, other than the point source, produces a luminous-flux density of n lumens per unit solid angle in a particular direction, it is said to have an intensity of n candlepower along that particular direction. When a uniform flat surface produces along a direction normal to itself a luminous-flux density of n lumens per unit solid angle, its *normal candlepower* is said to be n candles.

To a flat surface either emitting or reflecting light applies a law which is called the *cosine law of emission*. Its application implies that the light is emitted in a perfectly diffuse manner. In reality there is no perfect diffusing surface, so the law never strictly applies, but it may be said that all flat sources approximately obey this law. The law states that for a flat surface emitting or reflecting light, the luminous intensity in any direction will be its normal candlepower times the cosine of the angle that particular direction makes with the normal to the surface.

From the cosine law, by a simple integration of the flux over a solid angle of 2π it results that a flat surface having a normal candlepower of n candles generates a total of $\pi \times n$ lumens of luminous flux.

The brightness of a surface viewed from any direction is the ratio of either the luminous-flux density or the luminous intensity emitted (or reflected) in lumens or candlepower in that direction to the area of the surface projected on a plane perpendicular to the direction from which it is viewed.

For a perfectly diffusing surface, the brightness of it is independent of the angle from which it is viewed, but, of course the candlepower (total) varies according to the cosine law.

When talking of the candlepower of a particular luminescent screen one usually means the normal candlepower as measured by any suitable photometer. If the luminescence is of other color than white, the photometer should be equipped with suitable filters. n candlepower of a monochromatic source of light

means that its illumination is such that a white source of n candlepower appears as bright as the monochromatic source.

Under normal conditions, the candlepower (of a given luminescent screen) per watt of the beam power input varies some, but not very much, with the values of current and voltage used in television. For accurate comparisons, however, the curves of screen efficiency *vs.* beam current and voltage are required.

It can be said that screen efficiency of the better luminescent screens of TCR tubes is of the order of 1 and 2 candlepower per watt when viewed through the glass bulb from the side opposite to the one from which the electron bombardment takes place. When viewed from the gun side, the efficiency is higher by at least 50 per cent. A simple calculation then shows that the total light generated by a good TCR-tube screen is between 8 and 16 lumens per watt of beam power.

For light of wave length of maximum visibility the theoretical equivalent of one watt is 621 lumens. Quoting S. Dushman,[1] "For every color there exists a certain ideal efficiency which depends upon sensitivity of the eye for that particular color, and for white light this maximum theoretical efficiency is about 225 lumens per watt." The same authority gives a value for efficiency of 12 lumens per watt for the ordinary 60-watt tungsten lamp. The conclusion follows that luminescent materials producing nearly white light at 16 lumens per watt are (1) more efficient than incandescent electric lights, and (2) are operating at 7 per cent efficiency in energy conversion.

The most efficient artificial source of light is the sodium-vapor lamp yielding up to 60 lumens per watt of nearly monochromatic light at 12 per cent efficiency.

11.4. Willemite.—The name "willemite" was given to the natural material by its discoverer, A. Levy, in honor of King Willem I of the Netherlands. The natural willemite is found in the United States near Franklin Furnace, N. J., and is suitable for use in TCR tubes. The variation in the impurity contents, however, renders the natural mineral rather undependable. In the first decade of this century, W. S. Andrews, of the General Electric Co., succeeded in producing synthetic willemite with manganese as activator. Most of the modern TCR tubes made in this country use synthetic willemite for luminescent screens.

[1] Dushman, S., Atoms as a Source of Light, *Elec. Eng.*, **52**, 173 (1933).

Manganese is still considered as the most suitable activator for it. The colors of fluorescence and of phosphorescence of the best known types of willemite are green and yellow.

Synthetic willemite is prepared by mixing suitable proportions of zinc oxide, silica, activator and flux and fusing it in a furnace

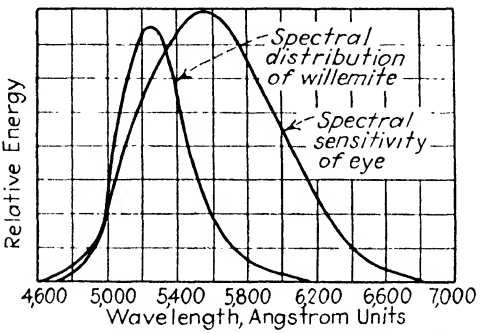


FIG. 11.1.—Spectral characteristic of green willemite.

into a uniform substance. After cooling, the resultant silicate is ground, sieved and put into a solution, which is applied to the inside surface of the TCR tube. The willemite thus obtained is a glass-like inorganic substance capable of withstanding, without decomposing, all the temperatures that a TCR tube has to

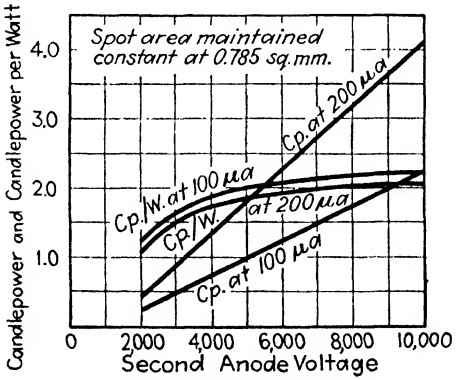


FIG. 11.2.—Luminescent efficiency of green willemite as a function of voltage.

undergo to insure a long life. These temperatures reach as high as 500°C. in vacuum and 550°C. in air to secure an optimum outgassing and annealing of glass. The two commonly used methods of application to the walls of a tube are spraying (with a suitable

binder) and settling from a solution and slowly decanting. The spectral distribution of the light emitted by the green willemite under electron bombardment is shown in Fig. 11.1.

The overall efficiency of a willemite screen in a TCR tube varies between 1 and $2\frac{1}{2}$ candlepower per watt, depending on the following factors: size of willemite particles, thickness of the luminescent screen, second-anode voltage, degree of saturation of willemite, type and concentration of activator, etc.

Figure 11.2 shows the candlepower per watt as a function of beam voltage at two values of current for a typical green willemite screen, manganese activated, sieved through 400 mesh and

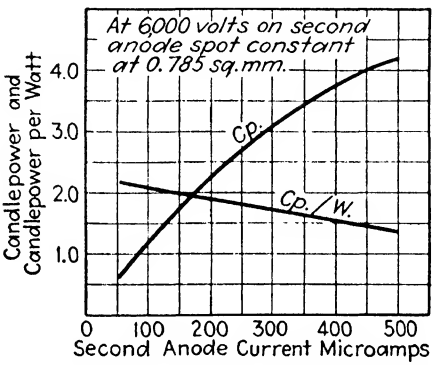


FIG. 11.3.—Luminescent efficiency of green willemite as a function of current.

applied to a Pyrex bulb in quantity of 0.7 mg./sq. cm. Figure 11.3 shows the candlepower as function of beam current at a constant value of beam voltage for the same screen.

11.5. Zinc Sulfide.—Zinc sulfides, while very popular with TCR tube makers in Europe, have only lately reached any noticeable prominence in the United States. There is a great number of trade names used in connection with this material, and there exists a great number of combinations of it with other sulfides. The activator in case of zinc sulfides is either silver or copper, and its quantity is so minute that preparation of these sulfides usually begins with purification of a salt of zinc to remove the excess of copper and other heavy metals. After purification, the sulfide is mixed with fluxes and fused in a neutral atmosphere to prevent oxidation. The resultant crystalline preparation is ground, sieved and applied to the inside surface of the glass bulb. The glass surface is first coated with either

sodium or potassium silicate to make it adhesive, and the phosphor is blown on the surface. The excess powder then is shaken off. The processing of TCR tubes having zinc sulfide screens is somewhat more complicated than that of those having willemite screens. This is due to the danger of oxidation of the sulfide when baked in air. The maximum baking temperature for zinc sulfide screens seems to be 450°C.

The predominant color of luminescence of zinc oxide is blue. By varying the ingredients of this phosphor, it is possible to obtain screens having almost white luminescence, with very small traces of either pink or blue color.[1]

The efficiency and persistence characteristics of sulfide phosphors vary just as much as their spectral characteristics. For television purposes, however, there are available a number of very satisfactory sulfide phosphors. In general it may be said that screen efficiency of good zinc sulfide is very little different from that of willemite; usually somewhat lower than the latter at low current densities and higher at high current densities. In other words, it can be said that willemite shows signs of saturation at lower current densities than does zinc sulfide.

11.6. Behavior of Willemite Screen under Electron Bombardment.—In 1903 Lenard[2] formulated a relation which later became known as *Lenard's equation* for light output from luminescent materials under electron bombardment. It is of the form

$$CP = AI(V - V_0)$$

where CP is the total candlepower output, A is a constant characteristic of the phosphor (luminescent material), I is the bombarding-beam current, V is the total potential through which the electrons have fallen and V_0 is a certain constant value of potential also characteristic of the phosphor, and which is often called *dead voltage*. Observations on a large quantity of television cathode-ray tubes with willemite screens, and several experiments especially made for determination of light output of

[1] See Levy and West (*loc. cit.*) for description of white zinc-cadmium sulfide. Unfortunately, these writers worked with poor willemite samples, and the performance values they give for willemite are four to five times lower than those obtained with regular synthetic willemites made in the United States.

[2] LENARD, P., *Ann. d. Physik*, **12**, 462 (1903).

willemite screens as a function of electron-beam current and voltage, have led to the conclusion that Lenard's equation holds[1] for willemite screens in television cathode-ray tubes under normal operating conditions. The only stipulation needed is that the current density of the luminous scanning spot should be kept constant.

In this section the discussion will be limited to the bombardment-voltage range between 2,000 and 10,000 volts, as occurs in normal TCR tubes under normal operating conditions.[2] By a normal cathode-ray tube is meant a tube made in a typical way, without any special precautions for cleanliness, with the usual getter and with a final vacuum of the order of 10^{-6} mm. of mercury. By normal operating conditions is meant conditions of scanning, *i.e.*, bombarding the area of the screen progressively by a beam of the cross section of the order of 1 sq. mm. with scanning-line frequency of the order of 10,000 lines per second. The willemite used in all the observations described was manganese-activated synthetic willemite ($2ZnO \cdot SiO_2$:Mn). Most of the willemite was prepared by H. W. Leverenz of the RCA Manufacturing Co., Camden, N. J., and the methods of preparation of the screens were similar to those described in one of his papers.[3]

In general, the behavior of a luminescent screen in a TCR tube is determined by the following factors: first, the velocity of the electrons in the beam; second, the secondary-emission characteristics of the luminescent material (phosphor) itself; third, the secondary-emission characteristics and the potential of the base

[1] Recently announced and published experiments of Nottingham, *Phys. Rev.*, **51**, 591 (1937), and **51**, 1008, (1937), led him to believe that the term within the parentheses of the Lenard equation should be squared to give the true dependence. He found a large potential difference between the willemite screen and the second anode, and after substituting for the electron accelerating voltage the voltage difference between the willemite screen and cathode which he experimentally determined, he arrived at the mentioned conclusion. Assuming that Nottingham's measurements be free from error in method, his conclusions do not seem to apply to the case of willemite screens in television cathode-ray tubes operating under normal operating conditions.

[2] T. S. Brown studied willemite performance under continuous bombardment at potentials between 200 and 1,500 volts [*J.O.S.A.*, **27**, 186 (1937)].

[3] LEVERENZ, H. W., Problems Concerning the Production of Cathode Ray Tube Screens, *J.O.S.A.*, **27**, 25 (1937).

on which the phosphor is deposited; fourth, the conductivity of the phosphor; and fifth, the electron-beam current.

The effect of secondary emission on the equilibrium potential of the screen has been explained in Sec. 2.7 of Part I by way of analogy with the secondary emission from an insulated metal electrode. Such analogy is very rough because the secondary-emission characteristics of dielectrics are radically different from those of metals and conductors. The literature is meager on the subject of the secondary emission from dielectrics.[1]

The purpose of this discussion is to describe a number of typical experiments, interpretation of which leads to the following conclusions:

1. The surface potential of dielectrics, such as Pyrex glass and willemite, under mentioned conditions of bombardment follows very closely the potential equivalent to the electron velocity.

2. Pyrex glass by itself acquires a potential by 5 to 15 volts below that of the beam under the same bombardment.

3. The potential of a thin (of the order of one milligram per square centimeter) willemite screen under bombardment depends largely upon the base on which it is deposited. Willemite on a metal base slides farther away from the beam potential than willemite on glass.

4. The current density of the scanning spot appears to be the factor of largest importance in determining the potential of the willemite screen in each particular case.

5. The transverse conductivity of thin willemite screen under electron bombardment is ample for maintenance of both sides of the screen at approximately equal potentials.

6. The light output is a linear function of beam voltage with beam current density kept constant.

Experiment 1.—Figure 11.4 shows total light output *vs.* second-anode voltage at a constant beam current and constant spot size for a special tube with willemite screen. A diagrammatic sketch of the tube is shown at top of Fig. 11.4. A willemite screen of order of 0.7 mg./sq. cm. was sprayed on platinized mica. The platinum coating was connected to a separate electrode. The measurements of light were made by means of a carefully calibrated Macbeth illuminometer with calibrated green filter.

[1] KOLLATH, R. The Secondary Emission of Solid Bodies, *Phys. Zeits.*, **38**, 202–224 (1937).

Each point on the curves is an average of at least nine observations. Curve *a* of Fig. 11.4 shows that when the platinum coating is connected to the second anode, the light output obeys the Lenard rule between 2,500 and 10,000 volts, assuming that the surface of willemite is at the potential of the platinum coating. Curve *b* shows that when the platinum coating is left floating, the light output appears not to obey the Lenard rule, dropping from it by an equivalent of 1,400 out of 10,000 volts and by 600 out of 8,000 volts.

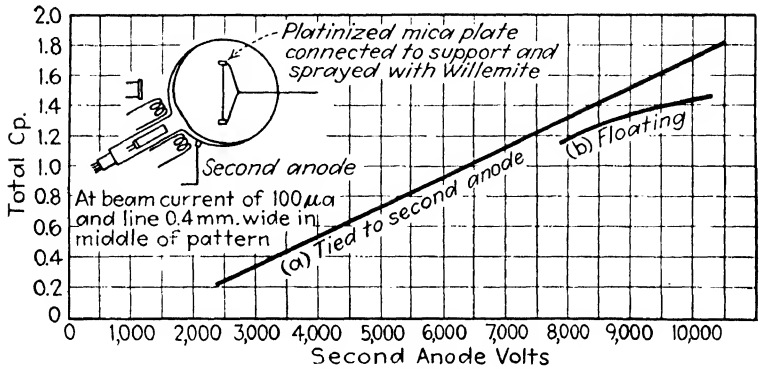


FIG. 11.4.—Light output *vs.* second-anode voltage for willemite on a platinized mica screen.

Experiment 2.—In Fig. 11.2 curves of light output *vs.* second-anode voltage for a normal television cathode-ray tube are shown. In this tube the willemite is deposited on clean glass in a layer of approximately 0.7 milligram per square centimeter. The inside walls of the tube are covered with aquadag and form the second anode. The willemite on the glass screen is floating, of course. The curves in Fig. 11.2 show that the light output from a conventional TCR-tube screen is a linear function of the second-anode voltage much in the same manner as in the case of the willemite screen on platinized mica kept at the second-anode potential. Although the willemite on the glass screen is floating, it shows no indication of any drop such as is shown by curve *b* of Fig. 11.4.

Special care should be exercised in making measurements for obtaining curves such as in Fig. 11.2: the spot size or the scanning-line width should be maintained constant. In case the line

width decreases with the voltage increasing, the curve of light *vs.* second-anode voltage will bend down from the straight-line relation. Figure 11.3 shows light output of the same tube *vs.* beam current at a constant spot size and a constant voltage. It explains the need for the just-mentioned precaution by demonstrating that a willemite screen saturates with current.

Experiment 3.—The cathode-ray tube used in this experiment was built for contrast study and contains a screen made of six different sections side by side. The only section used in this experiment is the lower left section made of platinized mica sprayed with willemite. The willemite coating was partially rubbed off in the center of the section, as the photograph in Fig. 11.5 shows. The rubbed-off portion retained a minute quantity of willemite which was sufficient to show the location of an electron spot. A stationary beam was deflected by direct current so that it struck the middle of the rubbed-off portion of the section. The platinum coating was connected to the second anode, and the position of the light spot noted. Then the platinum coating was allowed to float to its own equilibrium potential. The light spot moved towards the edge of the screen a distance of about 2 mm. This deflection is due to an electric field set up between the metallized plate and the second-anode coating, both seen in Fig. 11.5. The stationary spot was then moved to a place covered with a normal coating of willemite. The difference between the position of the light spot with metallized mica floating and connected to the second anode was the same as for the rubbed-off portion. The conclusion may be deduced that, under an electron bombardment, a thin willemite coating has enough transverse conductivity to acquire the potential of the base. This conclusion is statement 5 given in the beginning of the section.

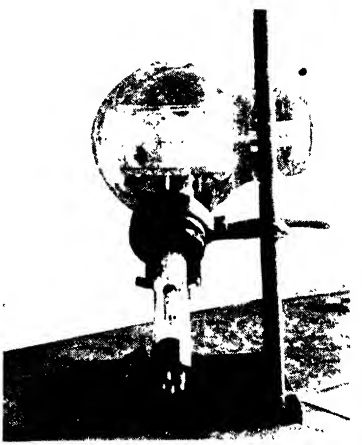


FIG. 11.5.—Experimental tube with sectional screen.

Another conclusion follows: that curve *a* in Fig. 11.4 has for the abscissa the true potential of the willemite surface, and that the

Lenard rule holds in the region of 2,500 to 10,000 volts. The latter conclusion is the earlier statement 6.

One more conclusion may be deduced from the above experiments; *viz.*, that willemite on Pyrex glass also obeys the Lenard rule, and besides, it does not build up a large potential difference with respect to the second anode.

Experiment 4.—Since the screen in the tube used in Experiment 1 seemed to acquire a negative potential with respect to the second anode, this potential was measured by means of an

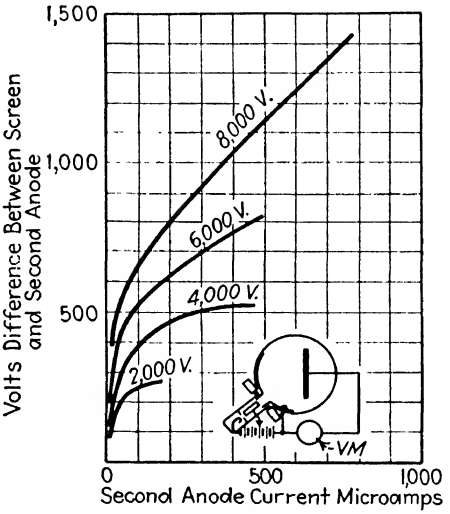


FIG. 11.6.—Secondary-emission characteristics of willemite on a platinum screen.

electrostatic voltmeter of very low leakage. The circuit diagram of the measuring setup is shown in Fig. 11.6. The measurements were made with beam currents varying from very small values up to the maximum that the cathode would deliver at zero grid bias. The spot was adjusted to the same size at each measurement. The curves of Fig. 11.6 show the data obtained for several values of second-anode voltage. They indicate that the difference of potentials between the willemite-on-platinum screen and the second anode varies continuously with varying beam current and voltage. In this experiment the luminous spot had an area of approximately 0.25 sq. mm. Data given in Fig. 11.6 should be used with caution, since even in the same tube the values shown would be much lower (50 and 75 per cent) should the spot be doubled in size by adjusting the first-anode voltage.

Experiment 5.—In another cathode-ray tube shown in Fig. 11.7, the willemite screen was deposited in a normal manner on the inside wall of a Pyrex bulb. The second-anode coating was

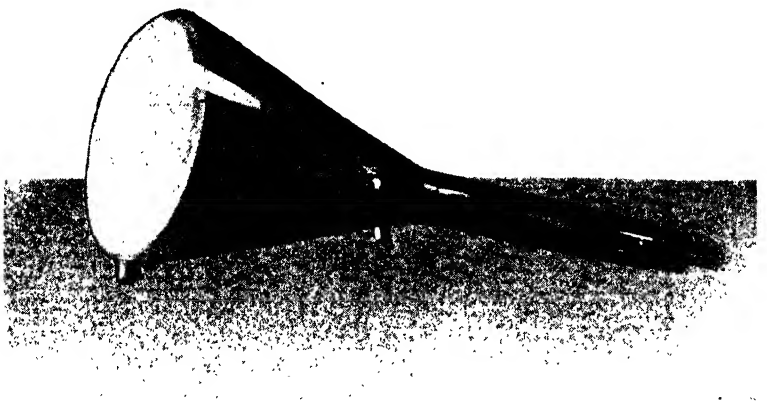


FIG. 11.7.—Experimental tube for determination of willemite potential.

deposited only in the immediate vicinity of the first anode. An additional lead-in wire was connected to a narrow silver coating directly under the edge of the willemite screen. About nine-tenths of the willemite screen was on glass and only about one-tenth of it was on silver. A scanning pattern of about one-fifth the area of the willemite screen was shifted around the screen by direct current through the deflecting coils, while an electrostatic voltmeter read the potential difference between the silvered area and the second anode. With the scanning pattern covering the willemite-over-silver portion, the reading of the voltmeter was the same as when only the willemite on glass was bombarded, indicating that the small silvered portion under willemite had little or no effect on the potential of the screen. Figure 11.8 shows the data

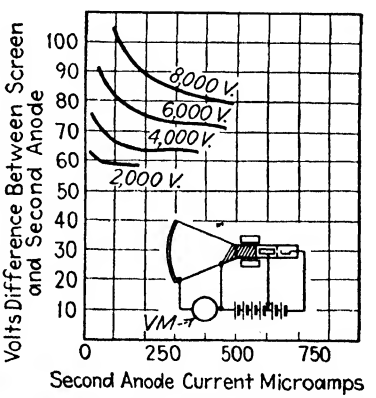


FIG. 11.8.—Secondary-emission characteristics of willemite on Pyrex glass screen.

obtained. The maximum voltage read was about 100 volts at 8,000 volts on the second anode, a little more than 1 per cent. With the accuracy of light measurements not better than 2 per cent, the charge built by the willemite-on-glass screen would not be apparent on the curves shown in Fig. 11.2. It may be of interest to mention that when, in the last experiment, the scanning pattern was shifted completely onto glass, the probe electrode showed a potential difference with respect to second anode of 5 to 15 volts.[1]

11.7. Contrast in Reproduced Picture.—In general, vision is caused by two factors, a difference in color and a difference in brightness of objects. In television reproduction, however, there is no difference in color, so the difference in brightness is the only factor which makes television reception possible. Brightness contrast, or simply contrast, is the term used to designate the difference in brightness. Contrast is usually given as a number resulting from the division of the brightness of the brighter object by the brightness of the darker object and is a plain numeric. Illuminating engineers say that two adjacent surfaces of the same color will appear as one if the difference in their brightnesses is less than 1 per cent.

The maximum degree of contrast obtainable in the received picture with a given television system depends mostly on the type of TCR tube used. This of course holds true only when the video-signal strength is sufficient and the receiving system is capable of providing sufficient amplification. Room illumination, of course, also affects the contrast.

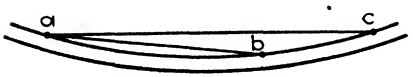


FIG. 11.9.—Effect of the curvature of a luminescent screen on contrast.

Several factors limit the maximum contrast obtainable in television reproduction. The importance of the individual factors varies with different types of tubes. In approximate order of their importance, these factors are as follows:

1. *Curvature of Luminescent Screen.*—Referring to Fig. 11.9, showing a cross section of a screen of a conventional TCR tube, if a point *a* corresponds to a dark portion of a generally bright picture, then points *b* and *c* will directly illuminate point *a*. The

[1] The latter values are in close agreement with results described by A. V. Afanasieva, *et al.*, *J. Tech. Phys.*, **6**, 149 (1936) (in Russian).

degree of this illumination is directly dependable on the curvature of the screen and the brightness of the rest of the picture. Naturally, the contrast will be the best for a bright spot on a black field and the poorest for a dark spot on a bright field.

2. *Halation.*—If a point a in Fig. 11.10 is brightly illuminated, the rays from it will go through the clear-glass wall of the tube. They will be refracted according to the well known laws of optics. The refraction will take place for all rays falling on the glass-air boundary at angles ranging from normal to that surface up to the so-called critical angle. For the critical angle and angles smaller than that, total reflection will take place.

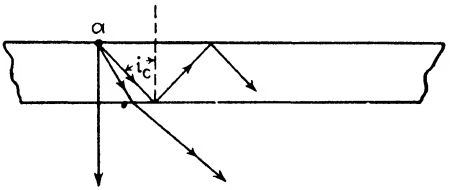


FIG. 11.10.—Total reflection as the cause of halation.

Total reflection takes place when a ray of light passes from a medium of index of refraction μ_1 into a medium of index of refraction μ_2, provided that μ_1 is larger than μ_2 and the sine of the angle of incidence i is equal to or larger than μ_2/μ_1.

In the example of Fig. 11.10, total reflection will take place when

$$\sin i \geq \frac{\mu_2}{\mu_1} \geq 0.666 \tag{11.1}$$

or for an angle of incidence

$$i \geq 42^\circ \text{ approx.}$$

Considering that $N\pi$ lumens of light are emitted by an elementary area of a perfect diffusing surface of willemite on glass, the index of refraction of which is equal to 1.5, the number of lumens I_1 that go through the glass is

$$I_1 = N\int_{\theta=0}^{\theta=2\pi} d\theta \int_{\varphi=48^\circ}^{\varphi=90^\circ} \cos\varphi \sin\varphi d\varphi = 0.448\pi N \text{ lumen.} \tag{11.2}$$

The number of lumens I_2 reflected back into the glass is

$$I_2 = N\int_{\theta=0}^{\theta=2\pi} d\theta \int_{\varphi=0}^{\varphi=48^\circ} \cos\varphi \sin\varphi d\varphi = 0.552\pi N \text{ lumens.} \tag{11.3}$$

In Eqs. (11.2) and (11.3), θ and φ are the usual angle variables

of a spherical system of cordinates. The fact that 55 per cent of the light emitted is thrown back at the willemite screen does not mean that all of it will be contributing to a loss of contrast. Willemite is a glass-like substance with an index of refraction of about 1.7. Therefore, wherever willemite is in optical contact with the glass, the totally reflected ray will emerge from the glass and illuminate the willemite screen. Naturally, not all the area of the glass is in optical contact with willemite, and in the areas between the crystals of willemite the light will be totally reflected again and will propagate inside of the glass away from the illuminated point a without impairing the contrast. The totally reflected rays will not reach the luminescent screen in the immediate vicinity of the original point a. The nearest place to the point a where these rays will fall upon willemite is along the circumference of a circle of radius r determined by the relation

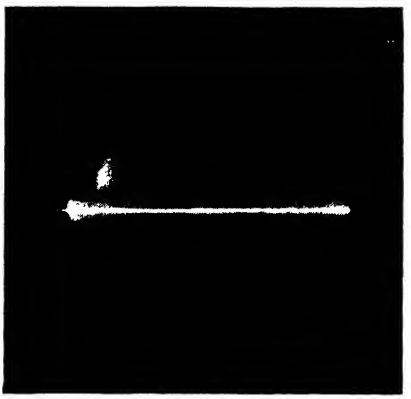


FIG. 11.11.—Photograph of halation.

$$r = 2h \tan i_c \qquad (11.4)$$

where h is the thickness of the glass and i_c is the critical angle of the combination. For most glasses used in TCR tubes, μ is around 1.5 and $\tan i_c = \tan 42° = 0.9$. For most of the practical cases, therefore, an approximative relation will hold:

$$r = 1.8h. \qquad (11.5)$$

The halation circle is the brightest at the radius given by Eq. (11.5), is dark inside and is gradually diminishing in intensity at larger values of the radius.

The effects of the halation on the reproduced picture are, first, that it reduces the general contrast of the picture by scattering light over the whole surface of it, and, second, that every bright object of the reproduced picture appears surrounded by a bright line at a distance r. In Fig. 11.11 is shown a photograph of a bright line surrounded by halation as reproduced on the end of a conventional TCR tube.

Equations (11.4) and (11.5) hold accurately for flat glass. For glass with a certain curvature, such as in conventional TCR tubes, the value of r is somewhat larger, depending on the degree of curvature. The fact that 55 per cent of the light emitted is lost should not be overestimated in importance. It cannot be used anyhow, since an observation of a picture from an angle of 45° and more away from normal is inconvenient.

3. *Stray Electrons.*—Stray electrons striking a luminescent screen with sufficient velocity will cause a loss of contrast. The sources of stray electrons are as follows: first, cold emission from the electron gun; second, secondary emission from the electron gun; and third, reflected secondary electrons from the willemite itself.

Cold emission from the electron gun results from insufficient polishing and rounding off of sharp edges of the gun and should not be present in a well-designed and well-made tube.

Secondary emission from the electron gun is caused by the electrons striking the masking aperture of the gun. If the focusing field penetrates as far as the stopping aperture, then the random moving secondaries will be drawn by the field into the region of the second anode, and eventually they may strike the screen, causing a loss of contrast. This effect is never present in a well-designed tube.

The third source of stray electrons, *i.e.*, the reflected and secondary electrons from the willemite itself, is practically negligible. The number of reflected (high-velocity) electrons is so small that their effect is hardly detectable. The velocity of the majority of the secondaries, on the other hand, is also so small that they have not enough energy to excite the luminescent screen.

4. *Direct Reflection of Light.*—Reflections from other parts of the tube are another factor contributing to the loss of contrast. The glass bulb itself reflects only about 10 per cent or less of light falling on it. A more serious offender is the conductive coating of the second anode which, even when covered with carbon or graphite, reflects a considerable amount of light.

The best conventional tube of the type having a screen on the inside surface of its bulb suffers from all the enumerated sources of loss of contrast except the stray electrons. There are other types of TCR tube which eliminate to some degree some or most of these undesirable effects. These types will be described in Chap. 12.

CHAPTER 12

CLASSIFICATIONS, RATING AND CHARACTERISTICS OF TCR TUBES

12.1. Classifications of TCR Tubes.—TCR tubes vary greatly in form, size, method of use, etc. Regarding the manner in which the reproduced picture is viewed, they may be divided into two groups: direct viewing tubes; projection tubes. According to whether the luminescent screen is deposited in a translucent layer on the inside of the glass envelope or is placed in

Fig. 12.1. -Nine-inch TCR tube for all-electric deflection.

the form of a thick layer on a plate, which in turn is mounted inside of the envelope, TCR tubes may be classified either as tubes with translucent screens or as tubes with reflective screens. When referring to a particular type of tube, the maximum diameter of the glass envelope is usually mentioned, *e.g.*, "3-in. projection tube."

Since all modern tubes are of the high-vacuum type, their classification into "hard" and "soft" tubes may be considered obsolete. As a rule, tubes are designed for a particular method of deflection of the beam, and depending on this they are divided into three groups: (1) tubes for all-magnetic deflection; (2) tubes for all-electric deflection; (3) tubes for combined electric-mag-

netic deflection. Another classification specifies the method of focusing of the electron beam, either electrostatic or magnetostatic.

Figure 12.1 shows a 9-in., direct-viewing tube, electrostatically focused, made for all-electric deflection and having a translucent

FIG. 12.2.—Three-inch projection tube for all magnetic deflection.

screen. Figure 12.2 shows a 3-in. projection tube electrostatically focused, made for all-magnetic deflection and having a translucent screen.

Figure 12.3 shows a 5-in., direct-viewing TCR tube with a translucent screen, made for all-magnetic deflection and magnetostatic focusing; while Fig. 12.4 shows a 12-in., direct-viewing

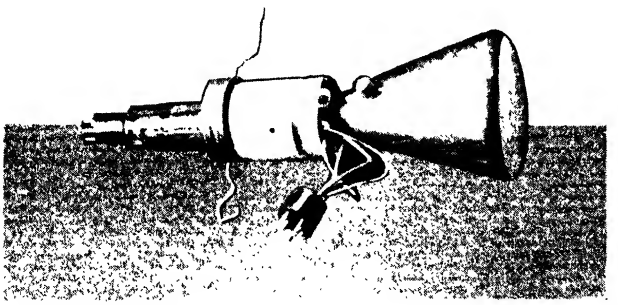


FIG. 12.3.—Five-inch TCR tube with magnetostatic focusing and all-magnetic deflection.

TCR tube with a reflective screen, electrostatically focused and made for all-magnetic deflection.

In Fig. 12.5 the construction details of TCR tubes made for all-magnetic, all-electric and combined electric-magnetic deflection are shown by *a*, *b* and *c*, respectively.

Previous to this section, the discussion has been centered on the conventional type of TCR tube, *i.e.*, the direct-viewing tube with a translucent screen. The projection tube and the tube with

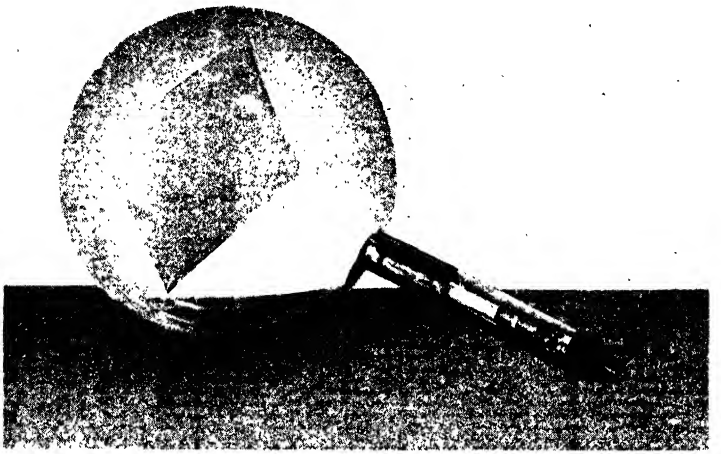


FIG. 12.4.—Twelve-inch direct-viewing TCR tube with electrostatic focusing and all-magnetic deflection.

reflective screen have just been mentioned for the first time and require explanation.

12.2. TCR Tubes with Reflective Screens.—In direct-viewing tubes with translucent luminescent screens, the electron bombardment takes place on one side of the screen, while the lumi-

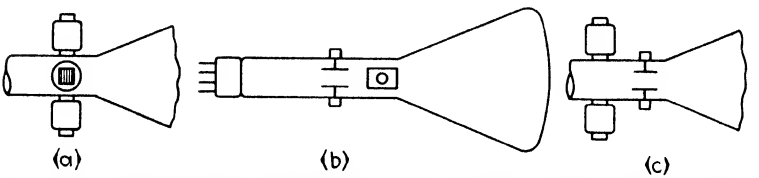


FIG. 12.5.—Details of neck construction of TCR tubes. (*a*) all-magnetic deflection; (*b*) all-electric deflection; (*c*) combined electric-magnetic deflection.

nescence is observed from the opposite side. In Sec. 11.3 it has been mentioned that when a luminescent screen is viewed from the bombardment or the gun side, the screen brightness is higher by at least 50 per cent. TCR tubes with reflective screens have been developed to take advantage of the higher screen brightness

on the gun side. Figure 12.4 in the preceding section shows a photograph of such a tube. In the tube shown, the luminescent screen is deposited in a thick nontranslucent layer on a thin glass sheet 6 by 8 in. in size. The phosphor also may be deposited on a metal plate, with a resultant loss in potential of the screen, as has been shown in Sec. 11.7 unless the metal plate is connected to the second anode.

The advantages of TCR tubes with reflective screens over those with translucent screens are as follows:

1. An increase of 50 to 100 per cent in screen efficiency (candlepower per watt) due to direction of viewing.
2. An increase in maximum contrast obtainable.

The increase in contrast results from two factors: first, the absence of the curvature of the luminescent screen; and second, the absence of halation. While the maximum contrast obtainable with conventional tubes is of the order of 20 to 1, values as high as 100 to 1 are obtainable with reflective screens.

The direct-viewing tubes, as well as the projection tubes, may be made with reflective screens, with corresponding gains in luminescent screen efficiency and maximum contrast obtainable. A very serious and an inherent disadvantage prevents, however, a widespread use of such tubes: the electron gun has to be placed at an angle of about 60° to the luminescent screen in order not to interfere with unobstructed viewing of the television image. Therefore, in order to obtain a regular rectangular scanning pattern on the screen, one must modulate the horizontal deflection in synchronism with the vertical. An arrangement for such a function is called a *keystone correction arrangement* and in its simplest form requires several tubes and circuit elements plus several additional controls on the receiving set. Another disadvantage of the arrangement is the unequal distance between the top and the bottom of the screen and the electron gun, making it difficult to get evenly focused scanning lines over the entire pattern.

12.3. Projection TCR Tubes.—As the name itself implies, a projection TCR tube is used in television systems where the image is projected by means of a suitable optical (not electron optical) lens system on a screen. A typical 3-in. projection TCR tube is shown in Fig. 12.2. Figure 12.6 shows the same tube and its lens in a typical mounting. Projection TCR tubes are often

operated at second-anode potentials as high as 40,000 volts and require special precautions in their construction for preventing excessive cold emission. As will be shown in Sec. 12.4, projection television systems are characterized by a large loss of light. To counteract this loss, much more power has to be supplied by the electron beam. Since the luminescent screen area of a projection tube is necessarily small, the current density of the luminous spot is very high, and the phosphor is being operated far beyond the region of saturation shown in Fig. 11.3. As Fig. 11.2 indicates, the willemite phosphor shows no voltage saturation at all up to 10,000 volts. It has been found that this

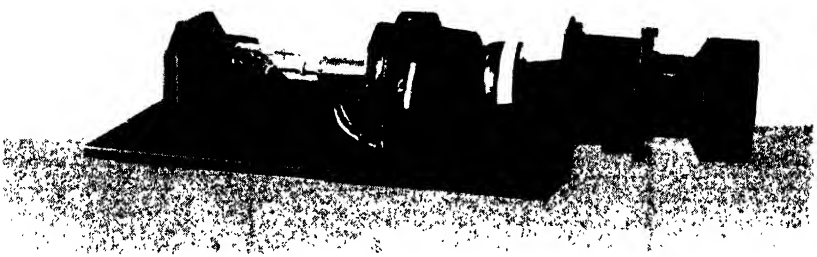


FIG. 12.6.—Typical mounting for projection TCR tube.

holds true up to at least 20,000 volts, which is the cause for the high operating potentials of projection TCR tubes.

12.4. Light Efficiency of Projection Television System.—As in any projection arrangement, in the system under consideration the question of light efficiency is of utmost importance. The unfortunate fact about the luminescent screen in a cathode-ray tube is that it emits light, not as a bundle of parallel rays but very nearly according to Lambert's law or the cosine law of diffuse emission. This diffuse emission of light accounts for the greatest loss of light in any television projection system. Figure 12.7 shows schematically the most essential details of such a system.

Referring to Fig. 12.7, consider the lens of a free aperture a and a focal length f. The speed of the lens or the so-called "F number" is a/f and is usually denoted by S. If the object

distance is denoted by U, it follows from elementary optics that:

$$a = \frac{f}{S'} \qquad d = f + U \qquad \text{and} \qquad U = \frac{f}{m}$$

where d is the total distance from the cathode-ray tube screen to the lens and m is the linear magnification. It also follows that

$$d = \frac{f}{m}(m + 1).$$

If the luminous intensity of the tube screen in the direction normal to the screen is I_0 candlepower, the lens will receive an amount of light flux equal to the product of luminous intensity

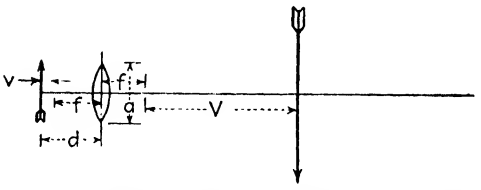


FIG. 12.7.—Optical system of projection TCR tube.

by the ratio of the lens area to the squared distance between the lens and the tube screen. Denoting the light flux received by the lens by F_L it follows that[1]

$$F_L = I_0 \frac{\pi a^2}{4} \frac{1}{d^2}.$$

Substituting the previously obtained values of a and d,

$$F_L = I_0 \frac{\pi}{4S^2} \frac{m^2}{(m + 1)^2}.$$

Since an average good high-speed lens system has at least four separate lenses with a loss of light of, say, 5 per cent per glass surface, and since there is a certain amount of dispersion of the light by the body of the glass, only about 50 per cent of the light flux F_L will pass through the lens into the image space. Denoting the light flux striking the projection screen by F_i and taking into account the lens transmission, it follows that

[1] A small correction for cosine law distribution of light on the surface of the lens from points on the object has been neglected.

$$F_i = \frac{F_L}{2} = I_0 \frac{\pi}{8S^2} \frac{m^2}{(m+1)^2}.$$

If the amount of light flux F_i strikes a perfectly diffusing, 100 per cent reflecting screen, the luminous intensity of the screen I_i' becomes

$$I_i' = I_0 \frac{1}{8S^2} \frac{m^2}{(m+1)^2}$$

because 1 lumen of light flux from a perfectly diffusing, 100 per cent reflecting screen gives a normal luminous intensity of $1/\pi$ candlepower. However, there are many kinds of reflective and translucent screens which possess directive characteristics, and while reducing the angle from which the image should be viewed they concentrate the light within that angle. The best screen that the author has seen is the translucent one, which is 4.8 times better than a perfect diffusing screen. On the average, the improvement is around four to one. This factor is greatly in favor of the projection system, so that for a directive screen

$$I_i = I_0 \frac{1}{2S^2} \frac{m^2}{(m+1)^2}. \tag{12.1}$$

If the area of the object screen is A, then the area of the image screen is Am^2; and if B_o is the normal brightness of the object while B_i is the normal brightness of the image, then

$$B_i = \frac{I_i}{Am^2} = B_o \frac{1}{2S^2} \frac{1}{(m+1)^2}. \tag{12.2}$$

Equation (12.1) shows the loss of light caused by the projection. The right-hand term of this equation contains three factors. The first factor is the candlepower of the cathode-ray tube screen. The second factor may be called the *factor of lens speed*, and its meaning is that in projection of television images the effect of lens speed is a reduction of the candlepower of the projection screen with respect to that of the tube by twice the square of the lens speed. For the best lens suitable for such projection, the speed or the F number is one and one-half, which means that the best lens reduces the original amount of light to 1/4.5. The last factor of the right-hand term of Eq. (12.1) may be called the *factor of magnification*, and its value as a function of

m is shown in Fig. 12.8. For large or infinite magnification, its value is unity, and for one to one magnification its value is 0.25.

12.5. Rating of TCR Tubes.—Rating of TCR tubes should be based on the same rational considerations which apply in the case of almost any electrical apparatus intended for exacting service. Extended breakdown and life tests should be the final proofs of an intended rating. A great deal of care should be used in choosing test procedures to insure test conditions which correspond to actual service. While a reasonable degree of conservativeness is expected in any rating, too much in formulating

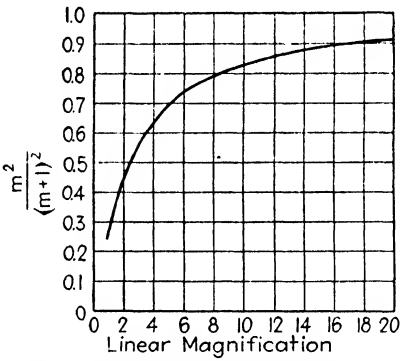


FIG. 12.8.—Factor of magnification of Eq. (12.1).

a rating for a particular device is just as unfair to the device as extreme optimism.

The rating of an electrical device should be consistent; *e.g.*, if the insulation of a deflecting plate from the second anode is such that, with full permissible deflecting voltage a full deflection cannot be obtained at second-anode voltages higher than, say, 5,000 volts, the maximum safe second-anode voltage should not be rated higher than 5,000 volts.

A differentiation should always be made between the rating and the performance characteristics. A rating is a brief statement describing what service a given device is intended for and giving either normal operating or limiting safe values of the principal operating parameters of the device. The rating seldom refers to just one device; usually it is intended for a type of apparatus. The meaning of the term "performance characteristics" is rather self-evident, covering a more or less complete

set of data on the performance of a device or a type of device. A typical rating sheet for a TCR tube is given below:

Name:—Television Cathode-ray Tube. Made by XXX Co.
Type:—No. 1,000, direct-viewing, with 9-in. translucent screen, for all-magnetic deflection.

Rating:—		
	Heater voltage	2.5 volts, a.c. or d.c.
	Heater current	2.0 amp., a.c. or d.c.
	Second-anode voltage	7,000 volts (max.)
	First-anode voltage	2,000 volts (approx.)
	Cutoff grid voltage	−50 volts
	Line width at zero bias	1.0 mm. for 343 lines, 30 frames
	Beam current at zero bias	over 500 μa (1 ma. max.)
	First-anode current at zero bias	under 1,000 μa
	Average mutual conductance	10 micromhos
	Gun-screen distance	30 cm.
	Deflecting yoke length	2.5 in. (max.)
	Deflecting yoke diameter	1.5 in. (min.)
	Maximum picture size	6 in. by 8 in.
	Overall length	23 in., plus/minus ¼ in.
	Maximum diameter	9⅛ in.
	Base	medium 5-pin
	Average screen efficiency	1.8 cp. per watt at 7,000 volts, 100 μa
	Screen	Synthetic willemite, green
	Heater	Insulated from cathode

In the case of a tube made for all-electrostatic deflection, the rating sheet would contain the data on the deflection sensitivity of the particular system of plates.

12.6. Performance Characteristics of TCR Tubes.—Complete performance characteristics of a device should be such as to enable one to judge whether or not this particular device is capable of delivering a certain performance under a chosen set of conditions. When performance characteristics have been taken on a number of devices of the same type, and average performance characteristics have been deduced, they should enable one to judge whether this type of device is capable of doing certain duty under a particular set of conditions. When the performance characteristics are supplemented by breakdown characteristics and by rational life tests, a rational safe rating can be deduced, and a prediction of performance of this type of device over a period of time can be made with reasonable certainty. Strictly speaking, life-test results and breakdown characteristics

are integral parts of performance characteristics, but obviously they cannot be taken when one intends to use the particular device which is submitted for tests. In such a case the breakdown and life values should be at least estimated by considering the design and construction specifications of the device. The TCR tube is not an exception to the above rules, and the matter of the rating and of performance characteristics is one of major importance to its user.

So far as its current-voltage characteristics are concerned, a TCR tube is primarily a multi-electrode thermionic tube. The

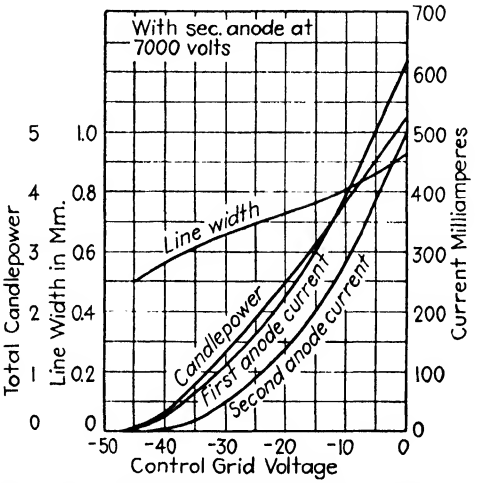


FIG. 12.9.—Operating characteristics of TCR tube.

purposes of various electrodes are different, but the general shapes of the current-voltage characteristics are very similar. There are two most generally used arrangements of elements in TCR tubes. The first type, or the four-electrode tube, has the cathode, the control grid, the first or the auxiliary anode and the second anode. The second type, or the five-electrode unit, has the same cathode, the control grid, the screen grid, the first anode and the second anode. This arrangement is usually maintained in both the electrostatically and electromagnetically focused tubes. In the latter, the first and the second anodes are often connected to the same source of potential.

In the preceding section a rating sheet of a certain TCR tube was given as an illustration. Figure 12.9 shows, among other

data, a "beam current *vs.* control-grid voltage" characteristic of the same tube at maximum rated second-anode voltage and at the rated voltage on the heater. The general appearance of the characteristics is that of a static characteristic of an ordinary vacuum tube. Strictly speaking, the characteristics of TCR tubes cannot be taken statically because a focused spot, when stationary, damages the luminescent screen. As a rule, therefore, the characteristics are taken when the luminescent spot is scanning the regular pattern on the screen in a normal fashion. The potential on the first anode is adjusted for the best focus and is kept at that value if it is independent of the control-grid voltage. It seldom varies much, but if it does vary, it should be also plotted on the same sheet. The important current, which does the work, is the second-anode current I_{p2}, or the beam current, but the undesirable one, which flows to the first anode, is just as important to know. I_{p2}, when multiplied by the corresponding second-anode voltage E_{p2}, gives the power dissipated by the beam in striking the screen, in watts. The total light produced by the screen is proportional to the beam power, and the coefficient of proportionality, called the *screen efficiency*, is expressed in candlepower per watt. Therefore, the beam power multiplied by the screen efficiency and divided by the picture area will give the brightness of the screen in candlepower per, say, square foot at a given set of conditions. The value of screen efficiency, however, while not far from being a constant,[1] varies sufficiently to warrant plotting it on the same curve sheet with the I_{p1}, I_{p2} and E_{p1}. It may be given in brightness units or in total candlepower for the maximum available area scanned by a given beam. The latter value is preferable, as it gives an indication at first glance as to how well the luminescent material is utilized.

Another characteristic also belongs on this already crowded sheet. It is the line-width characteristic. In the preceding chapters it has been explained that the area of the cathode from which the beam of electrons is drawn is dependent on the value of the control-grid potential. The diameter of the beam cross-over or the object to be focused on the screen is a direct function of the cathode emitting area. Therefore, the diameter of the

[1] The efficiency of a translucent screen depends on its thickness as well as on the intrinsic efficiency of the phosphor.

focused spot and also the width of the scanning line on the screen is a function of the control-grid voltage. In practice, the actual width of the scanning line in a scanning pattern is usually measured indirectly. The amplitude of the vertical deflection is increased so as to give a spacing between the lines equal to, say, one-tenth of the line's length. The focusing voltage is adjusted for the best spot, and the width of the scanning line is measured by means of a microscope equipped with a micrometer scale. Naturally, the whole system is synchronized to have the scanning pattern stand still. By the adjustment for the best spot is meant an adjustment of focusing voltages for as round and small a spot as is obtainable.

CHAPTER 13

ACCESSORIES

13.1. Saw-tooth Scanning.—It has been previously mentioned that Nipkow's scanning scheme calls for straight-line scanning, *i.e.*, for scanning with the luminous spot moving along straight parallel lines and with uniform speed. With magnetic deflection this requires a saw-tooth wave of current through the coils of the deflecting yoke, while with electric deflection it calls for a saw-tooth wave of voltage on the deflecting plates. Since a saw-tooth wave of current through a pure inductance produces a rectangular impulse of voltage across it, and *vice versa*, the devices for generation of either are similar in principle if not in all details. The requirements for vertical and horizontal deflecting arrangements are somewhat different on account of the difference in frequencies.

In general, a driving circuit of a deflecting system consists of an impulse generator, a discharge tube and an output tube. There are, however, circuits where all three functions are performed by a single tube. Also, there are driving circuits where the impulse generator is omitted and the synchronizing impulses are used directly instead. The latter circuit finds little use outside of monitors, because any interruption in synchronizing stops the operation of both deflecting circuits, with resulting damage to the luminescent screen by the stationary electron beam. For the above reason it has become a standard practice to use local impulse generators in television receivers for both vertical and horizontal driving circuits.

There are two customary means of impulse generation in general television practice, relaxation oscillators and blocking oscillators. There is also a circuit utilizing one tube in a complete driving circuit, which is not, strictly speaking, an impulse generator and is in a class by itself.

A relaxation oscillator generates self-sustained electric oscillations, which are characterized by periods of comparative calm or relaxation during which the energy is stored in a reactive

element, followed by periods of quick "trip-over" during which this element discharges quickly. Relaxation oscillators in their extreme forms give rise to voltage and current waves of rectangular and saw-tooth shape, with slopes either constant or changing abruptly. The best known forms of relaxation oscillator circuits are as follows:

1. An electric arc working into a series resonant circuit.
2. A glow tube working into a series resonant circuit.
3. A dynatron working into a parallel resonant circuit.
4. A combination of two vacuum tubes working into a series resonant circuit (Bloch's multivibrator).
5. A combination of two vacuum tubes, one working into a parallel resonant circuit, while the other reverses the phase (Kirschstein's multivibrator).
6. A four-electrode tube working into a series resonant circuit (van der Pol's form).

All of the above forms produce relaxation oscillations only when the L/C ratio of the resonant circuit is abnormal, either too high or too low.

The blocking oscillator is a highly damped, sine-wave oscillator with grid-leak resistance sufficiently high not to give the grid condenser time to discharge during that part of the cycle when the grid is negative. It blocks after one impulse of grid current and remains idle until the charge on the grid condenser leaks off.

13.2. Relaxation Oscillators.—Relaxation oscillation is a form of electric oscillation in systems consisting of reactive elements and devices possessing negative resistance. A symmetrical relaxation oscillator generates a wave of voltage characterized by a quick "trip-over" from extreme negative value to an extreme positive value, followed by a period of "relaxation" during which the value of voltage decreases slowly until a certain critical value is reached and the phenomena repeat themselves in the opposite sense. The presence of the trip-over and relaxation periods in the performance of an oscillator brings it into the relaxation-oscillator class.

Figure 13.1 shows an example of an output wave of a typical symmetrical relaxation oscillator. In Germany the term "kipp," meaning trip-over, is used to designate the same kind of phenomena. It is descriptive of the performance referring to the steep part of the curve.

In an unsymmetrical relaxation oscillator, the positive and the negative parts of the wave are, as the name implies, unequal in magnitude and duration, as shown in Fig. 13.2. In extreme cases one may practically disappear.

B. van der Pol[1] defines a relaxation oscillator as one in which the period is more nearly determined by the product RC of the

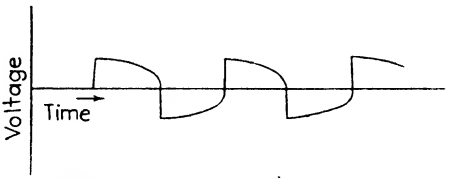


FIG. 13.1.—Voltage output of a symmetrical relaxation oscillator.

oscillating circuit than by its LC. This occurs when the L/C ratio is abnormally small. From his general case it appears that the same effect is encountered when this ratio is abnormally large. An additional requirement may be formulated according to van der Pol's point of view: that for maintenance of self-sustained relaxation oscillations the presence of two reactive components is essential, notwithstanding how small one of the two may be. Friedlander,[2] in a treatise lacking somewhat in rigor, maintains

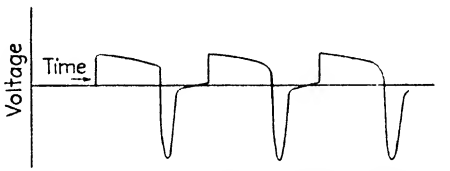


FIG. 13.2.—Voltage output of unsymmetrical relaxation oscillator.

that pure relaxation oscillation is of an entirely new kind requiring only one energy-storing component—either inductance or capacity. Kirschstein,[3] more recently treats of a case where one of the two reactive elements of the oscillating circuit gradually approaches zero. In this way he obtains a limiting case, which he calls a *pure relaxation oscillator.*

[1] VAN DER POL, B., On Relaxation Oscillations, *Phil. Mag.*, November, 1926.

[2] FRIEDLANDER, E., Relaxation Oscillations in Electron Tubes, *Arch. f. Elektrot.*, **16**, 273 (1926).

[3] KIRSCHSTEIN, F., New Methods of Graphical Treatment of Relaxation Oscillations, *Arch. f. Elektrot.*, **24**, 731 (1930).

The only contradiction between the above interpretations is between one of Friedlander's and one of van der Pol's, the former claiming that one reactive element is sufficient, while the latter asserts that at least traces of the second element are essential to carry the system through the equilibrium (dead) points. These contradictions were somewhat minimized by Kirschstein, who showed that a limiting case of van der Pol's solution gives a result similar to Friedlander's.

This controversy, however, does not affect the practical results in any way, as in practice there is always a residual and measurable capacity in any inductance and a residual and measurable inductance in any capacity. Therefore it follows that Friedlander dealt with a mathematical fiction unobtainable in practice. The definition of relaxation oscillator as given in Sec. 13.1 does not enter into controversy and therefore appears more suitable.

It may be stressed here that blocking oscillator is not a relaxation oscillator. It delivers one complete cycle of sine-wave shape and then blocks itself until either the charge on the grid blocking condenser leaks off or a synchronizing impulse forces it to go through the same sine-wave cycle again.

Of the six forms of the relaxation oscillators given in Sec. 13.1, the first two find no application in the modern television systems because of a considerable and erratic deionization period. The dynatron and the multivibrator have been in wide use for some time, and only recently began to yield first place in popularity to the blocking oscillator. The dynatron and the multivibrator essentially are either two-reactive- or one-reactive-component devices, depending on the point of view, and their free unsynchronized frequencies depend on the cathode emission and the applied voltages. In other words, they are emission-limited devices. The blocking oscillator is a three-reactive-component oscillator. Its free frequency depends mostly on the values of these reactive components, and it is not an emission-limited device.

13.3. Dynatron Impulse Generator.—The dynatron was defined by its inventor[1] as a "tube of high vacuum, possessing negative electric resistance." The dynatron oscillator is a vacuum-tube oscillator utilizing this negative resistance for

[1] HULL, A. W., The Dynatron, *I.R.E. Proc.*, **6**, 9 (1918).

generation of sustained oscillations. The cause of the negative plate resistance is the secondary electron emission from the plate. In Fig. 13.3, a schematic diagram of a dynatron oscillator is shown. The losses in the oscillating circuit are shown as a parallel resistance. In iron-cored oscillating transformers, such as used in deflecting driving circuits, the losses are represented more faithfully by parallel resistance than by series resistance. Also, it can be shown that in iron-cored transformers, an equivalent capacity lumped across the winding represents the performance of the distributed capacity with a very high degree of accuracy.

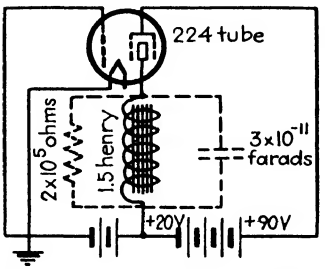


FIG. 13.3.—Dynatron impulse generator.

In Fig. 13.4 an oscillogram of the alternating-current voltage on the plate of a dynatron impulse generator is shown. The circuit with its equivalent constants is shown in Fig. 13.3. Constants as shown are typical for an impulse generator driving a horizontal deflecting circuit of a TCR tube. The tube shown is of the 224 type which operates nicely as a dynatron. The dynatron characteristics of a 224 type tube are shown in Fig. 13.5.

FIG. 13.4.—Dynatron output voltage.

The oscillation process consists principally of the storing of electrical energy in reactive elements of the oscillating circuit and in dissipating it in its losses and in the losses of the tube supplying the energy. The measures of this stored energy are the condenser voltage e_c and the inductance current i_L. In Fig. 13.6(a) these two quantities, together with plate current, are shown as functions of time. In Fig. 13.6(b), the energy stored in either of the reactive elements and the total energy stored in the oscillating circuit are shown also as functions of time. These curves are given for a

relaxation circuit slightly different from the one in Fig. 13.3, for the purpose of exaggerating some parts of the curves for ease of explanation.

From these two sets of graphs, the operation of the dynatron relaxation oscillator may be traced step by step.

The first steep positive slope of the voltage curve occurring between time $t = 0$ and $t = 3$ microseconds [Fig. 13.6(a)] is caused by the plate current sliding down the characteristic curve from point c (Fig. 13.5) through d, e and reaching as far as f. This current very quickly charges the tank condenser, which is only of 100 $\mu\mu$f capacity, and it is really the condenser voltage that makes the current slide faster along the characteristic. The current through the inductance is of high positive value and cannot change fast because of the very high inductance value of 1 henry.

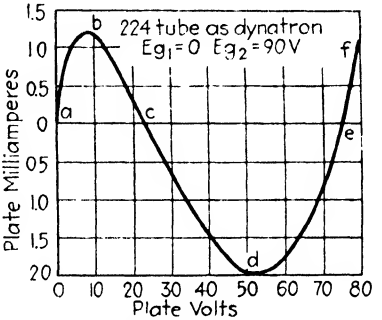


FIG. 13.5.—Dynatron characteristics of 224 tube.

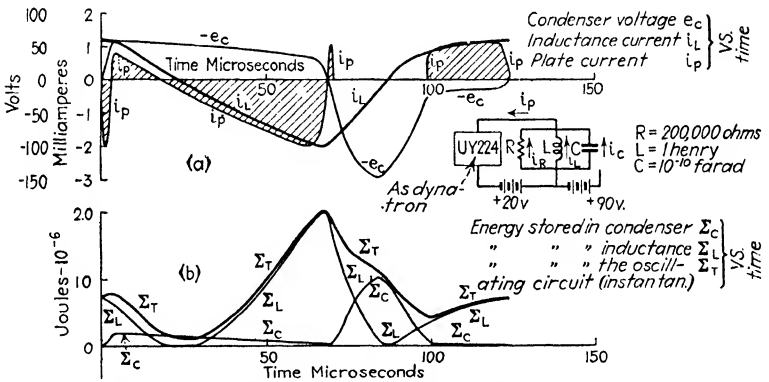


FIG. 13.6 —Detailed performance of a dynatron oscillator.

As soon as the plate current again becomes positive, the losses in the system become very high, the condenser current reverses, and the plate current slides again along the characteristics (Fig. 13.5), but in the opposite direction. This slide takes a long time because the current through the inductance has to

decrease from a large positive value to a higher negative value. It almost follows the plate current. This period is really the period of relaxation during which the energy is slowly stored in the main reactive element of the circuit. The energy curves in Fig. 13.6(*b*) show a maximum of total energy stored at the time when the plate current through the inductance is at the highest negative value.

At $t = 65$ microseconds (approximately), the plate current slope becomes positive again. The increase of the current through the inductance slows up. The plate current trips positive. The inductance current starts decreasing. A large voltage is induced on the transformer terminals. It is equal to e_c and with the sign reversed is equal to the alternating-current component of the voltage on the plate. It forces the plate to the negative side, far beyond the cut-off of the plate current. The vacuum tube is not functioning at that time; the oscillating circuit goes through a half cycle of a damped sine-wave oscillation. The period of this oscillation is determined by the constants L and C, and the damping by R.

While the inductance current is changing with an almost constant positive rate, the condenser voltage pushes the plate far below the ground potential and then decreases in the absolute value until the cut-off point of the plate current is again reached. At this point, the energy [see Fig. 13.6(*b*)] is at one of the minimums, and the circuit has difficulty in overcoming the part of the characteristic with positive resistance. Another portion of the voltage curve with an almost horizontal slope occurs. Once it is overcome and the tube is again on the negative part of its characteristic, the cycle repeats again. Were it not for the damping resistance R, the stored energy at the end of the sine-wave half cycle would be the same as at its beginning, the circuit would encounter no difficulties in overcoming this small amount of positive resistance, and the cycle would begin over again without the period of "hesitation."

In any steady state or sustained oscillation process, the additional energy stored during one complete period is equal to the energy dissipated during this complete period. But the ratio of the total energy circulating in the circuit may be many times larger than the amount dissipated and, therefore, added during the period. In a well-designed sine-wave master oscillator,

this ratio may be as high as 100:1, while in an oscillator working directly into an antenna, it is usually not over 12:1.

As an impulse generator in television deflecting circuits, the dynatron oscillator is used in two different ways. The first way is to use the long positive flat portion for synchronizing and

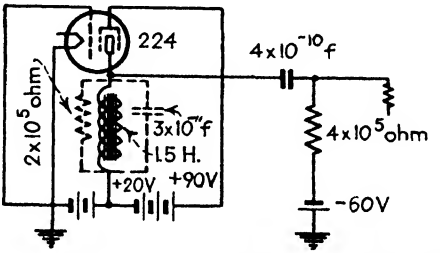


FIG. 13.7.—Dynatron impulse generator with peaking circuit.

the negative kick as the impulse. The second is to use the negative hesitation period for synchronizing and the steep positive slope for generation of positive impulse. The former is typical for a vertical deflecting circuit while the latter is quite common for a horizontal deflecting circuit.

FIG. 13.8.—Dynatron output after peaking.

The technique of synchronizing of impulse generators is described in Sec. 13.6 of this chapter. Data given in Fig. 13.6 were calculated by the so-called "isocline" method of successive approximations. Both van der Pol and Kirschstein utilized isoclines in their analyses of relaxation oscillators. The actual method used for calculations, results of which are given in this book, is described in detail in Sec. 13.4 in connection with a solution of multivibrator performance.[1]

In the case when it is desired to utilize the steep positive slope of the plate-voltage wave for generation of impulse, the steep

[1] The isocline method is suitable for calculation of the performance of oscillators with not more than two reactive elements.

wave front is applied to a resistance and capacity in series. The voltage across the resistor is proportional to the current through the combination, while this current is the charging current of the condenser, *i.e.*, an exponential function. By connecting a resistor and a condenser (peaking circuit) in the output of a dynatron, as shown in Fig. 13.7, a positive impulse, the form of which is shown on Fig. 13.8, is obtained.

13.4. Multivibrator and Isocline Method.—In Fig. 13.9 a typical circuit diagram of a non-symmetrical multivibrator is shown. If Z contains two reactive components in parallel, a differential Eq. (13.1) of second order may be used to describe its performance:

$$a\frac{d^2i}{dt^2} + b\frac{di}{dt} + ci = 0. \quad (13.1)$$

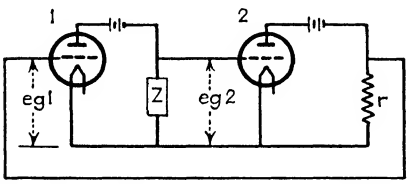


FIG. 13.9.—Non-symmetrical multivibrator.

The arrangement will give sustained electrical oscillations if the coefficient b is variable, or at least its sign is variable. In the majority of cases, the relations between the variables are such that they can be best expressed graphically. A general solution of such a case is unknown, to say the least. A particular solution may be obtained either by a method of successive approximations or by graphical means. Among the latter, the Isocline method[1] seems to be one of the most convenient and useful.

Isocline, as the word itself implies, is a line joining points of equal inclination or slope. In the method about to be described, it joins points of equal slope or derivative of a function. This method is a graphical method of solution of two simultaneous nonlinear differential equations. It is particularly suitable for the solution of performance of oscillators with two reactive elements. A typical problem of this kind is the multivibrator problem.

The voltage-current characteristic of any two-terminal device will be designated as the *terminal characteristic*.[2] The device may be just a tube, or a tube including batteries and circuit elements, or any combination of tubes, or anything with two

[1] MALOFF, I. G., Solution of Vacuum Tube Problems by the Isocline Method, *Broadcast News* (RCA Publication), February, 1934.

[2] Kirschstein's definition.

terminals to which, if one applies a voltage, a current flows. The voltage-current characteristic as obtained from these two terminals is what will be called the terminal characteristic.

Figure 13.10 shows an oscillating circuit connected across the device D. To obtain its characteristic, apply a set of values of voltages e (or e_c) across the device and obtain the current entering or leaving the device. By plotting e as ordinates and i as abscissae, the terminal characteristic is obtained. The terminal characteristic may be obtained experimentally or calculated. Figure 13.9 shows a device composed of two coupled tubes or the well known form of unsymmetric multivibrator or relaxation oscillator.

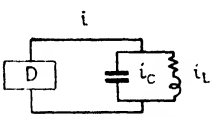


FIG. 13.10.—Parallel oscillating circuit.

If the oscillating circuit is removed and in its place a battery of variable voltage is substituted, with a voltmeter across it and an ammeter in series with it, the experimental terminal characteristic of the arrangement may be determined.

If the tubes are of 227 type, with 50 volts, direct current, on the plates and with a range of +20 and −50 volts across the oscillating circuit, the terminal characteristic for the arrangement looks as shown in Fig. 13.11. The same characteristic could be calculated from the static characteristics of the tube. It may be mentioned that the grid current is a very important factor and under no consideration can be neglected, *i.e.*, if the knowledge of the exact magnitude and shape of the output wave is desired.

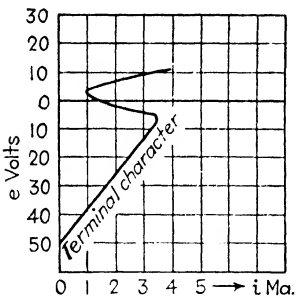


FIG. 13.11.—Terminal characteristic of non-symmetrical multivibrator.

The problem to solve may be stated as below.

For a multivibrator as shown on Fig. 13.9, compute the exact current and voltage output as functions of time. It is desired to know the exact shape and magnitude of current and voltage output of the multivibrator. Let the constants be R equal to 1,000 ohms, L equal to 1 henry and C equal to 10^{-8} farad.

The fundamental relations for the circuit are

$$i = i_L + i_c$$
$$i = f(e_c)$$

$$e_c + i_L R + L\frac{di_L}{dt} = 0$$

$$i_c = -C\frac{de_c}{dt}.$$

The sign in the last equation is taken negative because the assumed direction of i_c is such as to charge the condenser in the opposite sense of e_c. (If the reverse sense were assumed, the isoclines on the following pages would be rotated 180 deg. about the e_c axis.)

The four relations above can be reduced to the following two:

$$i_L = f(e_c) + C\frac{de_c}{dt} \tag{13.2}$$

and

$$e_c = -i_L R - L\frac{di_L}{dt}. \tag{13.3}$$

Equations (13.2) and (13.3) show relations between three variables—e_c, i_L and t. Since the process of oscillation depends chiefly upon the changes in the energy content of the system, and since e_c and i_L are measures of the energy stored in the condenser and inductance, respectively, the elimination of the parameter t from Eqs. (13.2) and (13.3) would tell the story of the manner in which these changes take place. Since the expression for $f(e_c)$ is given graphically, the relation between e_c and i_L is most easily expressed graphically.[1]

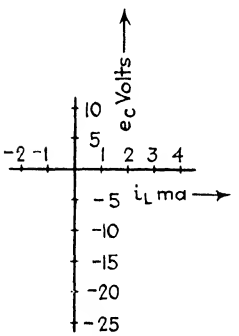


FIG. 13.12.—Rectangular coordinates of isocline method.

The isoclines offer a means of accomplishing this. The procedure is as follows:

Set up a system of rectangular coordinates with i_L as abscissae and e_c as ordinates, as in Fig. 13.12. The next step is to trace two sets of isoclines, or lines of equal derivatives, on this set of coordinates.

First, assign to $C\frac{de_c}{dt}$ (equal to $-i_c$) a number of values like 0, 1, 2, etc., and plot the equation

$$i_L = f(e_c) + C\frac{de_c}{dt}$$

[1] Even if it were expressed analytically, the graphical solution may be the only way out, as, for it, the complexity of $f(e_c)$ is not an obstacle.

for each of these values (see Fig. 13.13). For $C\frac{de_c}{dt} = 0$ the isocline is merely a graph of the terminal characteristics as given in Fig. 13.11.

Each of these curves is an isocline, as it joins all the points in the e_c-i_L plane for which the derivative $C\frac{de_c}{dt}$ is equal to a given constant. By means of this family of isoclines the value of the derivative $C\frac{de_c}{dt}$,[1] for any combination of values e_c and i_L, is read directly from the diagram.

By assigning several constant values to $L\frac{di_L}{dt}$ in Eq. (13.3), and by plotting them on the same set of coordinates, a second set of isoclines is obtained. Each of these isoclines joins all the points in the e_c-i_L plane for which the derivative $L\frac{di_L}{dt}$ is equal to a given constant By means of the second family of isoclines the value of derivative $L\frac{di_L}{dt}$,[2] for any combination of values e_c and i_L is read directly from the diagram.

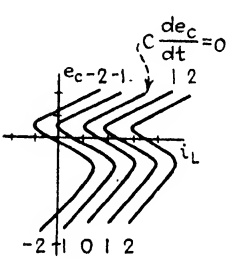


FIG. 13.13.—Isocline plot.

These two families of isoclines can be superimposed. The resultant construction represents a graphical solution of Eqs. (13.2) and (13.3), as for any combination of e_c and i_L it gives directly the values of $C\frac{de_c}{dt}$ and $L\frac{di_L}{dt}$, which satisfy the original Eqs. (13.2) and (13.3).

Figure 13.14 shows a complete problem of a relaxation oscillator worked out by the isocline method. Referring to this plot, take the point $e_c = -14$ volts, $i_L = +4$ ma.; the two isoclines $C\frac{de_c}{dt} = +1 \times 10^{-3}$ and $L\frac{di_L}{dt} = +10$ pass through this point. Putting these values into Eqs. (13.2) and (13.3), the following identities are obtained:

$$4 \times 10^{-3} = 3 \times 10^{-3} + 10^{-3}$$
$$-14 = -4 \times 10^{-3} \times 10^{+3} - 10.$$

[1] It is really the value of derivative multiplied by a constant C.

[2] Multiplied by a constant L.

It shows that Eqs. (13.2) and (13.3) are satisfied at this point as well as at any other point of the plane. The reason for this is that the families of isoclines were drawn from Eqs. (13.2) and (13.3).

To get a curve which is a particular solution of Eqs. (13.2) and (13.3), it is necessary to assign a boundary condition such as $e_c = 0$, $i_L = 1.6$ at time $t = 0$. Next, choose an arbitrary interval of time Δt, small enough to assume that derivatives are constant for its duration. To get an accurate solution, Δt

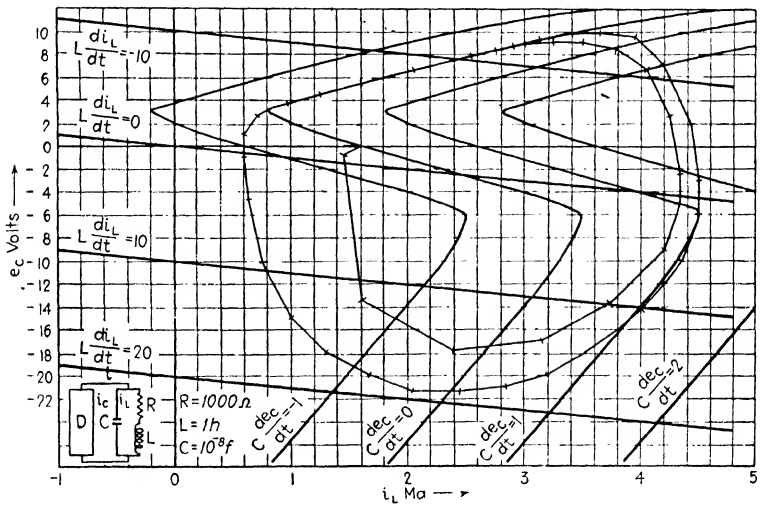


FIG. 13.14.—Isocline solution of a multivibrator.

must be small in comparison with the period of oscillation. In this example, Δt is chosen as 100 microseconds and 50 microseconds.

By drawing the isoclines as outlined above, every point of the plane e_c-i_L becomes a solution of Eqs. (13.2) and (13.3), and so there is no further use for these equations. From now on the computation proceeds from the two approximate relations

$$\Delta i_L = \frac{1}{L}\left(L\frac{di_L}{dt}\right)\Delta t \tag{13.4}$$

$$\Delta e_c = \frac{1}{C}\left(C\frac{de_c}{dt}\right)\Delta t. \tag{13.5}$$

Starting then at the point $i_L = 1.6$, $e_c = 0$, it is seen that the two isoclines $L\frac{di_L}{dt} = -1.6$ and $C\frac{de_c}{dt} = 0$ pass through this point.

As a first approximation it is assumed that $L\frac{di_L}{dt}$ stays at -1.6 and $C\frac{de_c}{dt}$ stays at 0 during the interval Δt. Inserting these values in (13.4) and (13.5), the change in i_L and e_c is obtained as

$$\Delta i_L = -1.6 \times 10^{-4}$$
$$\Delta e_c = 10^8 \times 0 \times 10^{-4} = 0$$

So, if $L\frac{di_L}{dt}$ remained at -1.6 and $C\frac{de_c}{dt}$ remained at 0 for 10^{-4} seconds, i_L and e_c at the end of this interval would be 1.44 ma. and 0 volts, respectively. Drawing a straight line between the point $e_c = 0$, $i_L = 1.6$ ma.; and point $i_L = 1.44$ ma., $e_c = 0$, the segment representing the first approximation to the exact solution is obtained.

As a second approximation, find isoclines $C\frac{de_c}{dt} = C_1$ and $L\frac{di_L}{dt} = C_2$ which pass through the midpoint of the segment drawn as a first approximation, and assume that these values of $L\frac{di_L}{dt}$ and $C\frac{de_c}{dt}$ remain constant during the interval Δt. To find the isocline passing through the midpoint of the segment, interpolate along the e_c axis for $L\frac{di_L}{dt}$ and along the i_L axis for $C\frac{de_c}{dt}$. Thus, at the midpoint ($i_L = 1.52$ ma., $e_c = 0$) $C\frac{de_c}{dt} = -0.08 \times 10^{-3}$ and $L\frac{di_L}{dt} = -1.5$, and assuming that $L\frac{di_L}{dt}$ remains -1.5 volts for the interval 10^{-4} sec. and $C\frac{de_c}{dt}$ remains -0.08×10^{-8} for the interval 10^{-4} sec., there results from Eqs. (13.4) and (13.5)

$$\Delta i_L = 1 \times -1.5 \times 10^{-4} = -0.15 \text{ ma.}$$
$$\Delta e_c = 10^8(-0.08 \times 10^{-3}) \times 10^{-4} = -0.8 \text{ volt.}$$

So at the end of 10^{-4} sec., i_L and e_c have the values 1.45×10^{-3} amp. and -0.8 volt, respectively. Drawing a straight line

between the points $e_c = 0$, $i_L = 1.6$ ma and $i_L = 1.45 \times 10^{-3}$ amp., $e_c = -0.8$ volt, the segment representing the second approximation to the exact solution is obtained.

To obtain a second segment, repeat the process; *i.e.*, find the values of $L\frac{di_L}{dt}$ and $C\frac{de_c}{dt}$ as given by the isoclines at the beginning of the second segment (end of first segment), and use these values in (13.4) and (13.5). This gives the first approximation to the second segment. To obtain the second approximation, find the values of $C\frac{de_c}{dt}$ and $L\frac{di_L}{dt}$ as given by the isoclines at the midpoint of the first approximation of the second segment, and use these values of $L\frac{di_L}{dt}$ and $C\frac{de_c}{dt}$ in (13.4) and (13.5).

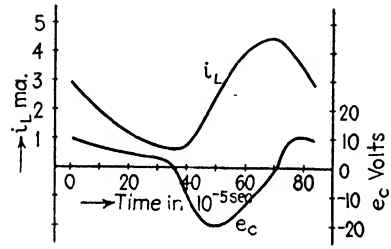


FIG. 13.15.—Solution of Fig. 13.14 plotted as function of time.

By the repeated application of (13.4) and (13.5) in the manner described above, the spiraling curve shown in Fig. 13.14 was obtained. From this spiraling curve, the curves of e_c and i_L are readily constructed as functions of time because each chosen time interval corresponds to a particular pair of Δe_c and Δi_L. Figure 13.15 shows e_c and i_L plotted as functions of time.

In general, given a circuit as that shown in Fig. 13.10, it follows from Kirchhoff's first and second laws that

$$i_L = f(e_c) + C\frac{de_c}{dt} \tag{13.6}$$

$$e_c = -i_L R - L\frac{di_L}{dt} \tag{13.7}$$

where $i = f(e_c)$ is the terminal characteristic of the device.

To solve these two simultaneous nonlinear differential equations by the isocline method, proceed as follows: First, determine in the $e_c = i_L$ plane the two families of isoclines $C\frac{de_c}{dt} = \text{constant}$ and $L\frac{di_L}{dt} = \text{constant}$, *viz.*, the curves

$$i_L = f(e_c) + \text{constant}$$
$$e_c = -i_L R - \text{constant}$$

which are obtainable by parallel displacements of the curves $i_L = f(e_c)$ along the i_L axis and $e_c = -i_L R$ along the e_c axis.

Second, assign a boundary condition; *i.e.*, decide upon a point in the plane at which to begin the integration.

Third, choose an interval of time Δt. This Δt does not have to be the same interval throughout the problem.

Fourth, calculate segments by the repeated application of the equations

$$\Delta i_L = -\frac{1}{L} L \frac{di_L}{dt} \Delta t \tag{13.4}$$

$$\Delta e_c = -\frac{1}{C} C \frac{de_c}{dt} \Delta t. \tag{13.5}$$

For a first approximation, the values of $L\frac{di_L}{dt}$ and $C\frac{de_c}{dt}$ are evaluated at the beginning of the segment; for a second approximation, $L\frac{di_L}{dt}$ and $C\frac{de_c}{dt}$ are evaluated at the midpoint of the segment determined as a first approximation.

In this manner, a curve in the e_c-i_L plane is determined which represents the process of oscillation. The closed portion of the curve depicts the steady state, while the remaining portion shows the transient state.

The isocline method may be applied to all two-terminal devices such as arcs, glow tubes, multivibrators, dynatrons, relaxation circuits, etc. For a three-terminal device, however, it cannot be applied without modifications. For a regular three-electrode tube oscillator with a grid-blocking condenser, it cannot be applied without great difficulties. The constants chosen for the numerical example just described are more nearly those of a sine-wave oscillator than of a relaxation oscillator, but were so chosen only to illustrate the method's applicability to the solution of either case.

13.5. Blocking Oscillator.—The blocking oscillator generally can be defined as a quasi-aperiodic, vacuum-tube, oscillation generator utilizing *three* reactive (one inductive and two capacitive) elements and two resistance elements in such a way that an impulse is generated at regular intervals of time. Only one impulse of grid current occurs in this interval and it is sufficient to block the grid. For a given tube, the duration of the interval is controlled by the product RC (grid-leak resistance by capaci-

tance of grid-blocking condenser). The magnitude of the impulse is determined by the inductance and capacitance of the tank circuit, and to a very marked degree by the damping resistance of the tank circuit. If this damping is reduced, the blocking oscillator becomes an ordinary sine-wave generator oscillating intermittently.

In the original form, the blocking oscillator made use of the type 227 tube with anywhere from 50 to 180 volts on the plate. The oscillating circuit, *per se*, consisted of an iron-cored transformer with windings of high distributed capacity. The grid-blocking condenser and grid leak were so proportioned as to give a desired number of impulses per second. In Fig. 13.16, a circuit diagram of such an oscillator is shown. When the values of the constants are as indicated, the oscillator will deliver approximately 4,320 impulses per second. Since the natural frequency of the oscillating transformer is 16 kc., its distributed capacity can be represented by 333×10^{-6} mf. lumped across the plate winding of 0.3 henry. Since the transformer has an iron core and nearly 100 per cent coupling, the equivalent capacity across the whole transformer can be easily evaluated. It was found, however, that the capacity across the plate winding is more convenient for a mathematical analysis.[1]

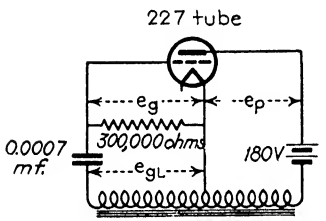


FIG. 13.16.—Blocking oscillator (natural frequency of transformer: 16 kc.).

For convenience of explanation, an example with somewhat different constants has been calculated. The oscillator performance may be divided into six unequal periods. Referring to Fig. 13.17, showing the results of calculation, the time interval after which the impulse repeats itself is approximately 970 microseconds, which would give a scanning frequency of 1,030 lines per second.

Period A.—The first period is defined as one during which the grid potential changes from the cut-off value (of plate current)

[1] For theoretical and experimental justification of lumping the distributed capacities of iron-core transformers, see Performance and Properties of Telephonic Frequency Intervalve Transformers, by D. W. Dye, *Exp. Wireless and Wireless Eng.*, September, 1924.

to a value equal to zero. The plate current begins to flow at the start of this interval, and reaches a value of the order of half of the maximum at the end of it. At first most of plate current flows through the tank condenser and, as the charge on this condenser accumulates, a voltage is impressed on the plate winding of the transformer. This voltage is equal to the difference between the impressed plate voltage E_p and the voltage e_p on the plate, and is consumed by the difference of two induced voltages—voltage induced by varying the current in the plate winding and voltage induced by varying the current in the grid winding. The latter and its variations are quite small (of the

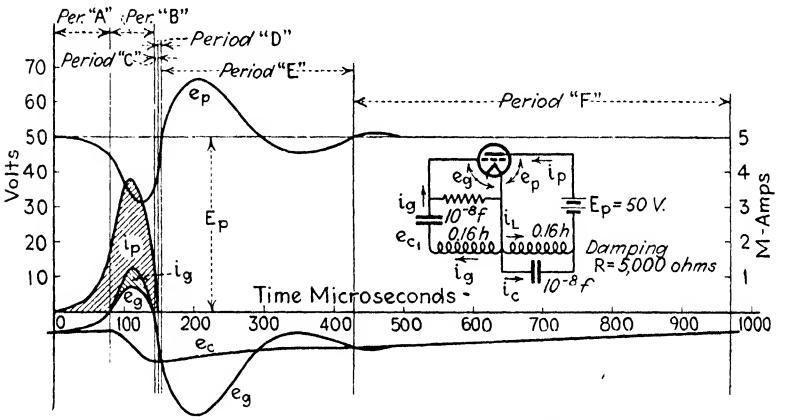


FIG. 13.17.—Detailed performance of blocking oscillator.

order of a few tenths of a volt) while the former is appreciable (of the order of several volts). The plate current flows because the grid voltage is above the cut-off potential and is getting higher. The grid is getting higher in voltage because the potential induced by the current in the plate winding is larger than that in the grid winding, and its phase is reversed.

Period B.—The second period is defined as one during which both the grid and the plate currents flow. It begins when the grid current starts to flow, and it ends when the grid current stops. The grid-blocking condenser takes a negative charge corresponding to several volts. The plate current first increases, reaches a maximum, and then decreases to a value of the order of its value at the beginning of the interval. The instantaneous plate voltage e_p reaches its minimum and begins to increase

shortly before the end of the interval. The current through the tank condenser reaches a maximum early in the interval, drops to zero, and attains a high negative value at the end. The voltage induced by the varying current in the grid winding is first positive and later negative. The voltage induced by the varying current in the plate winding is positive throughout the cycle. It is this induced voltage that holds the grid potential positive throughout the period. It is necessarily larger in the first part of the period as the voltage induced by the grid current is positive. In the second part of the period, the voltage induced by the grid current is negative, *i.e.*, it helps the voltage induced by the i_L to keep the grid voltage above the ground potential. Since the grid-blocking condenser takes on a negative charge during the interval, the current i_L reaches a value of the order of maximum plate current at the end of the interval.

Period C.—During the third period, only the plate current flows and reaches zero value at its end, at which instant the currents through the tank inductance and condenser are necessarily equal to each other.

Periods D and E.—The fourth and fifth periods are in reality one, and are separated only for simplification in computation. During these two periods the energy stored in the reactive elements of the circuit is dissipated. The damping in the circuit could best be represented by a resistance in shunt with the tank circuit. It is not shown in the diagrams.

Period F.—During the sixth period, with all the disturbances having died out, the grid condenser discharges through the grid leak, and the grid potential reaches the cut-off value for plate current, at which instant Period *A* begins anew.

The method used in computation consists of setting up differential equations for the circuit and solving them step by step by successive approximations. Referring to Fig. 13.17, these equations are as follows:

$$E_p = e_p + L\frac{di_L}{dt} - L\frac{di_g}{dt} \tag{13.8}$$

$$E_p = e_p + \frac{1}{C_0}\int i_c dt \tag{13.9}$$

$$i_p = i_L + i_c \tag{13.10}$$

$$e_g + L\frac{di_g}{dt} + \frac{1}{C_1}\int i_g dt - L\frac{di_L}{dt} = 0 \tag{13.11}$$

$$i_p = f_1(e_g, e_p) \text{ graphical} \tag{13.12}$$
$$i_g = f_2(e_g, e_p) \text{ graphical.} \tag{13.13}$$

For all negative values Eq. (13.13) becomes

$$i_g = \frac{e_g}{R_g}. \tag{13.14}$$

For positive values of e_g the grid-leak resistance R is too large compared with grid impedance, and therefore negligible.

The set of six simultaneous equations can be transformed into another set of six, one for each variable. Everyone of these new equations will be of a form known as a nonlinear differential equation of the third order. A general solution of such a form is unknown. Graphical methods such as isoclines become extremely involved when applied to third-order equations. A brute-force method of successive approximations was used for calculation of the blocking oscillator given in this section.

All the known numerical values were substituted into the original equations (13.8) to (13.14), while the curves of the nonlinear functions such as in Eqs. (13.12) and (13.13) were kept on hand. Next, certain starting conditions were picked, such as the instant of turning on the plate voltage. From then on the procedure is as follows:

For the plate voltage equal to the value of plate-supply voltage and the plate current corresponding to the zero grid voltage, compute the changes in all variables for a chosen small interval of time Δt. The best way is to assume a certain change in one variable and see what happens to all the others. If the resultant values are not consistent with the functions as given by the curves of plate and grid currents, the first chosen assumed change will have to be modified and tried again. With some experience not more than two approximations are needed to compute a point, and a complete performance of an oscillator can be completed in approximately one working day, with accuracy limited only by the time one wants to spend on a particular calculation.

A circuit of constants very nearly like the one which was computed was wired up and studied by means of a cathode-ray oscillograph. Since a 60-cycle "house" current was used to spread the image, the frequency had to be adjusted to a multiple of it to get a standing image on the screen. Eight hundred and forty cycles per second synchronized itself very well, and for this

reason the pictures of current and voltages of 840 impulses per second were taken. The wave shapes do not change appreciably with the impulse frequency. The results of the oscillographic study are shown on Fig. 13.18. The computed results in Fig.

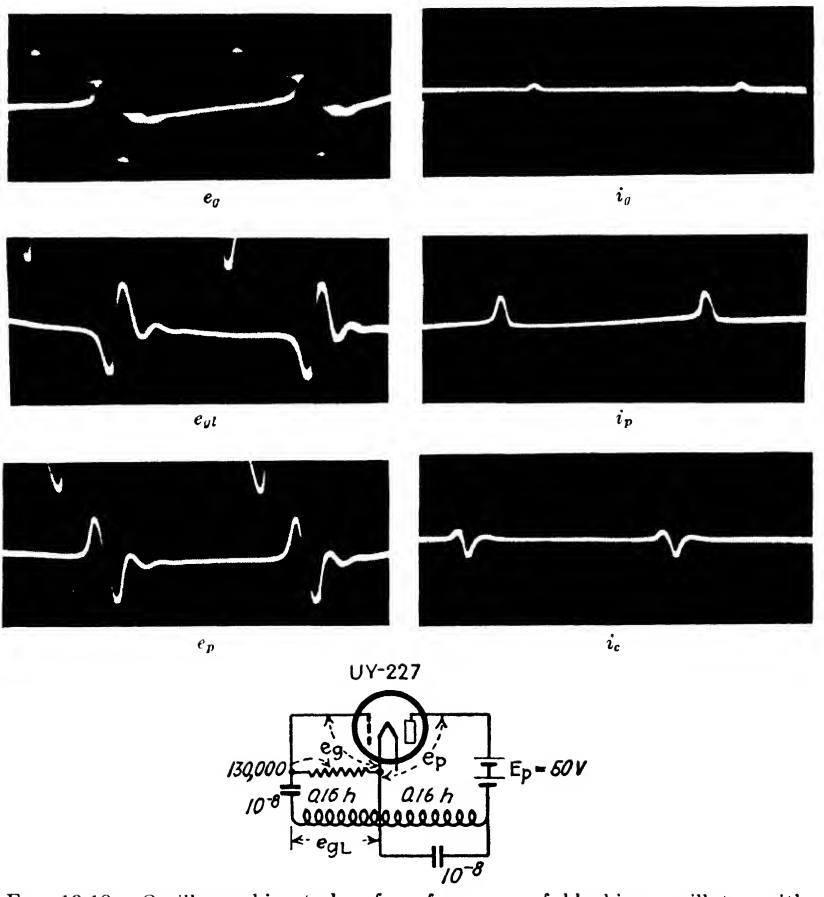


FIG. 13.18.—Oscillographic study of performance of blocking oscillator with constants shown.

13.17 are the same as the experimental results in every essential detail.

13.6. Synchronization of Impulse Generators.—Consider an impulse generator running at a given frequency. Suppose an external impulse of a different frequency is applied to the generator. The frequency at which the generator will run

following the application of the external impulse is, in general, different from that at which the generator ran before the application of the impulse. The frequency at which the generator ran before the external impulse was applied is called the *free frequency* of the generator; the frequency at which the generator ran after the impulse had been applied is called the *forced frequency* of the generator.

If the forced frequency of a generator is equal to the frequency or to any multiple or submultiple of the frequency of the external impulse, the generator is said to be synchronized by the external impulse.

The case of particular interest is when the forced frequency of the generator is the same as that of the impulse. This phenomenon occurs if the free frequency of the generator is in the neighborhood of the frequency of the impulse. Two cases are to be distinguished: (1) the free frequency of the generator is greater than the frequency of the impulse; (2) the free frequency of the generator is less than the frequency of the impulse. The mechanisms of synchronization for these two cases are radically different.

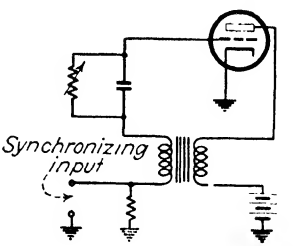


FIG. 13.19.—A method of application of synchronizing impulse to a blocking oscillator.

The above statements apply to practically all types of impulse generators; the detailed mechanisms of synchronization vary somewhat with different types. To illustrate the discussion, synchronization of a blocking oscillator will be now described in detail.

The synchronizing impulse, as the external impulse is called, is usually small compared with the maximum positive grid voltage of the oscillator. The synchronizing impulse is usually introduced into the grid circuit of the oscillator in a manner shown in Fig. 13.19.

If the free frequency of the oscillator is slightly greater than that of the synchronizing impulse, the impulse will, after several cycles, strike the oscillator during period F, and at a point where its magnitude is insufficient to drive the grid beyond the cut-off. The effect of this impulse is to reduce the rate of discharge of the blocking condenser while the impulse lasts, and in this way to

lengthen the duration of the period F. If this period is lengthened by the right amount, the forced frequency of the oscillator becomes equal to the frequency of the impulse, and the oscillator is synchronized. A somewhat exaggerated diagram of the process just described is shown in Fig. 13.20.

If the free frequency of the oscillator is slightly lower than the impulse frequency, the impulse will, after several cycles, strike

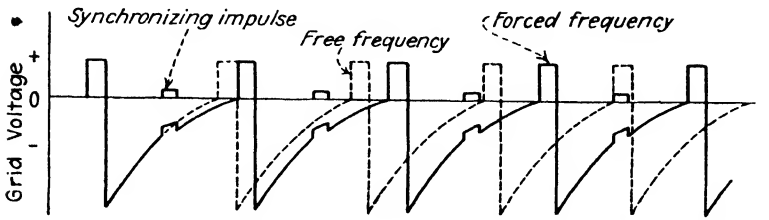


FIG. 13.20.—Synchronizing with free frequency greater than impulse frequency.

the grid at such a time as to be sufficient to drive it above cut-off. Hence, whenever the free frequency of the oscillator is slightly less than the impulse frequency, the impulse will shorten period F. If period F is shortened by the right amount, each succeeding impulse shortens it by the same amount and the forced frequency of the oscillator becomes equal to the frequency of the impulse. This is illustrated in Fig. 13.21.

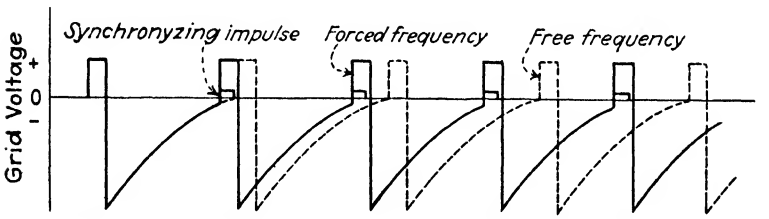


FIG. 13.21.—Synchronizing with free frequency lower than impulse frequency.

In television systems, the synchronizing impulse is applied not only to the oscillator for synchronizing, but also to the grid of the TCR tube to blank out the return lines. It has been shown that if the free frequency of the oscillator is somewhat higher than that of the impulse, the impulse will synchronize the oscillator. However, there will appear a vertical dark strip on the screen of the TCR tube, indicating the position (see Fig. 13.20) where the impulse strikes the grid of the TCR tube. This, of course, is

undesirable. This type of synchronizing is thus useless. A glance at Fig. 13.20 will show that the characteristic feature of this type of synchronizing is the fact that the impulse and grid voltage of the oscillator are not "in phase." A similar glance at Fig. 13.21 shows that here the impulse and the grid voltage are "in phase." Hence, in order to obtain synchronism with the correct phase relation (as in Fig. 13.21) *it is necessary that the free frequency of the oscillator be lower than that of the impulse.* If the oscillator runs much slower than the impulse, no synchronizing will occur.

So far the free frequency of the oscillator has been assumed constant. Actually, the free frequency of the oscillator drifts slightly. If the free frequency of the oscillator is the lowest at which it can be synchronized, a small increase in free frequency will not affect the synchronizing; but any decrease in free frequency will cause the oscillator to jump out of synchronization. If the free frequency of the oscillator is the same as the frequency of the impulse and suddenly drifts to a higher frequency, it may still be synchronized, but the black strip will appear on the screen of the kinescope. There is, therefore, a more or less narrow range of frequencies within which the oscillator must stay if there is to be the desired synchronization. Let f_0 be the lowest free frequency at which synchronization occurs, f_1 the frequency of the synchronizing impulse, and f the free frequency of the oscillator. Then synchronization without a black strip on the screen occurs if, and only if,

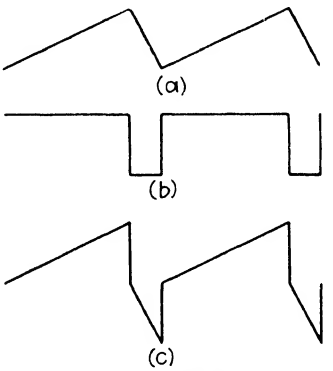


FIG. 13.22.—Voltage wave required to force a saw-tooth wave of current through (*a*) resistance, (*b*) inductance, and (*c*) resistance and inductance in series.

$$f_0 \leqslant f \leqslant f_1.$$

13.7. Discharge Tube and Peaking Circuit.—Figure 13.22 shows the voltage wave required to produce a saw-tooth wave of current through a pure resistance *a*, a pure inductance *b* and a circuit containing both resistance and inductance in series *c*.[1]

[1] HOLMES, CARLSON AND TOLSON, Experimental Television System, *I.R.E. Proc.*, **22**, (1934).

Such a voltage may be produced in a very simple circuit, such as shown in Fig. 13.23. Condenser C is charged in series with resistor r at a substantially constant rate through resistor R. The periodic positive impulses from a blocking oscillator are applied to the grid of the discharge tube, thus periodically discharging C by a definite and constant amount each time. The voltage wave across C is saw-tooth in shape, while across r it is of pure impulse form. By properly proportioning C and r it is possible to produce a voltage-wave shape containing any desired combination of saw-tooth and impulse for driving the grid circuit of the output tube of the deflecting driving circuit.

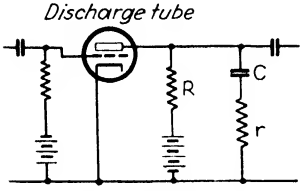


FIG. 13.23.—Discharge-tube circuit.

13.8. Output Tube.—Depending upon whether the deflecting system is for electric deflection or magnetic deflection of the electron beam, the output tube has to deliver either a saw-tooth voltage wave or a saw-tooth current wave across and through the output load, respectively. In oscillographic technique, where one deflecting plate is tied to the second anode of the cathode-ray tube, a small resistance load on the output tube is often used. The voltage across this resistance is then directly applied to the free deflecting plate, either through a coupling condenser, as shown in Fig. 13.24, or through a direct-current connection.

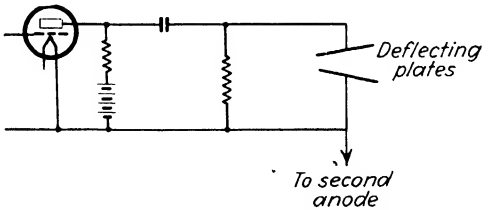


FIG. 13.24.—Circuit for unbalanced electric deflection.

In television technique the unbalanced electric deflection is unsatisfactory. Whenever the required deflecting voltage is not very large, a push-pull resistance-coupled output stage, as shown in Fig. 13.25, is quite satisfactory for obtaining balanced electric deflection. However, the usual voltage on the second anode of TCR tubes is so high (2,000 to 10,000 volts), that with normal receiving vacuum tubes (250 volts on the plate) it is

necessary to use an output transformer to obtain sufficient deflection. A diagram of such an output circuit is shown in Fig. 13.26. The transformer may work from either a single output tube or from two tubes in push-pull. The design of a transformer for horizontal deflection of TCR tubes offers a number of difficulties. Its natural frequency must be of the order of eight to ten times higher than the horizontal line frequency.

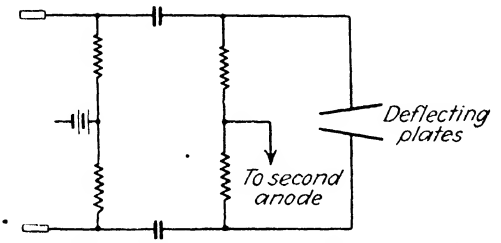


FIG. 13.25.—Circuit for balanced electric deflection.

Its inductance must be considerable, while its leakage inductance must be small and its leakage inductance-distributed capacity resonance must be critically damped by the losses of the circuit. The most important requirement, however, is that the two sides of the secondary winding must balance within the frequency range beginning with line frequency and ending with the ninth or tenth harmonic of it. A failure to satisfy the last requirement

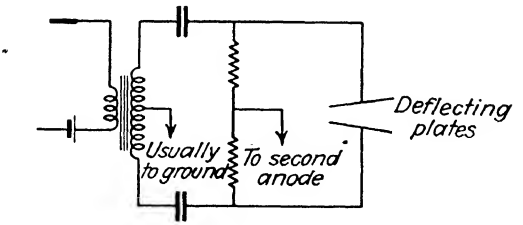


FIG. 13.26.—Circuit for balanced electric deflection utilizing output transformer.

results in appreciable defocusing. Of course, the transformer has to have a high step-up ratio to be useful. However, it is very hard to design a transformer satisfying all the above requirements with a step-up ratio higher than 5:1 or 6:1.

For magnetic deflection, the output tube works either directly into the windings of the deflecting yoke or through an output step-down transformer. Since the voltage across an inductance

with a saw-tooth current through it is of pure rectangular impulse form, the output tube in a magnetic deflecting system has to satisfy a certain set of requirements. If a given magnetic deflecting yoke has an inductance of L henries, while the current for deflecting the beam from one side of the scanning pattern to the opposite side[1] is I amperes, and the time for scanning of one line without the return is T, then the voltage across the yoke during the picture sweep is

$$E_p = \frac{LI}{T}$$

and the volt-amperes needed for the picture sweep are:

$$\text{V.A.} = \frac{LI^2}{T}.$$

A vacuum tube to operate such a circuit must have static plate-current, plate-voltage characteristics such as to permit an

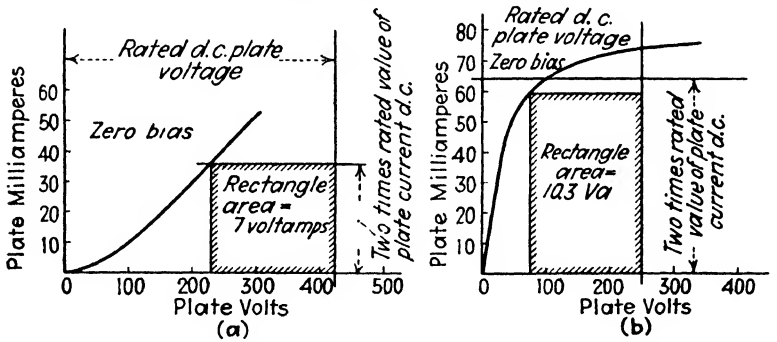


FIG. 13.27.—Computation of horizontal deflection with (*a*) 210 and (*b*) 247 type tubes.

inscription of a rectangle of LI^2/T volt-amperes in area in a space bound as follows: on the bottom by the zero-plate-current line; on the right by the line of maximum rated direct-current plate voltage; on the top by the line of twice the maximum rated direct plate current; and on the left by the plate-current, plate-voltage curve for zero grid bias. The bottom side of the rectangle should lie on the line of the zero plate current, and the right side of the rectangle must lie on the line of the maximum rated direct-current plate voltage. The point of intersection of left and top sides of the rectangle may lie anywhere on the plate-current, plate-voltage curve for zero grid bias.

[1] It is twice the value needed for deflecting the beam from the undeflected position to the edge of the pattern.

The area of the maximum rectangle, inscribed within the boundaries specified, represents the ability of a particular tube to accommodate the picture voltage of a deflecting circuit. Figure 13.27(*a* and *b*) shows these rectangles drawn on static characteristics of 210 and 247 tubes, respectively.[1]

The second requirement is that the output tube must not flash over while the return voltage is impressed between its plate and filament. The return voltage is a rectangular impulse, and its value is higher than the picture impulse by the same ratio as the ratio of durations of picture and return sweeps.

The requirements just given will be better understood after consideration of the discussion given in the next section.

13.9. Inverse Method of Calculating Vacuum-tube Performance.[2]—Any vacuum tube, irrespective of the number of the electrodes, is essentially a three-terminal device. When a tube is used either as an amplifier or as an oscillator, usually only three of its electrodes are used for connecting the external circuits. So far as the operation of either the tetrode or pentode is concerned, they may be considered as triodes of various characteristics.

The method and its variations about to be described are all based on a very definite property of triodes:

If the plate-current wave is known to be of a form

$$i_p = f_1(t)$$

and the plate-voltage wave is known to be of a form

$$e_p = f_2(t)$$

then for every value of t the grid voltage e_g has one and only one value, which can be expressed as

$$e_g = f_3(t).$$

An ordinary set of static characteristics of a vacuum tube defines the function f_3. This statement holds true for all positive values of t and also for $t = 0$, provided both f_1 and f_2 are continuous and single-valued functions of t. There are special cases in which the above relations still hold even when one of the functions

[1] Once this area has been determined, the proper number of turns, or inductance, for this particular yoke comes directly out of the dimensions of the rectangle.

[2] MALOFF, I. G., New Methods of Solution of Vacuum Tube Problems, *Broadcast News* (RCA Publication), November, 1933.

f_1 and f_2 has a number of finite discontinuities, and it becomes very handy when dealing with problems of odd wave shapes.

The term "static characteristics" is a very misleading one. The fact that these characteristics were statically taken is often considered to mean that they apply only to direct-current conditions, and that for any alternating-current problem a dynamic characteristic should be computed. The latter is not so; a dynamic characteristic of a tube shows the performance of a tube in combination with either a part or the whole circuit into which the tube is working. But the static characteristics hold for all alternating-current cases except the very high frequencies of the order of many megacycles, and it definitely states that if the instantaneous voltage on the plate is e_{p1} and at the same time the instantaneous plate current is i_{p1}, then the instantaneous grid voltage has only one value e_{g1}, and, besides, it gives the exact magnitude of that value.

For lack of a better name the method will be called the "inverse" method of calculation of amplifier performance, because in this method we assume a given output and compute the necessary input. It is just opposite to the conventional direct methods where the output for a given input is calculated.

First the inverse method as applied to a power-amplifier tube working into a given load of a given power factor will be discussed. The following is a specific problem:

With normal direct-current voltages applied, a 247 pentode is working into a purely *inductive* load of, say, 7,000 ohms and is delivering, say, 150 peak volts of fundamental frequency to this reactance. Compute the second and third harmonic content in the output wave, and the necessary grid swing.

The problem as stated above is the usual direct problem. Nevertheless, it can be solved by the inverse method if the following procedure is used:

For 150 peak volts across 7,000 ohms of inductive reactance, the peak current is 21.4 ma. and the effective volt-amperes, 1.6. Now assume that the alternating output current is of the form $i_x = 21.4 \sin \omega t$. The instantaneous plate current in the tube is

$$i_p = i_{p0} + 21.4 \sin \omega t$$

where i_{p0} is the steady rated direct plate current. The corresponding instantaneous plate voltage is

$$e_p = e_{p0} - 150 \cos \omega t$$

where e_{p0} is the steady rated direct-current voltage. Figure 13.28(a) and Fig. 13.28(b) show i_p and e_p as functions of ωt.

If the resultant i_p is plotted against the resultant e_p a perfect ellipse results. It may be plotted right on the so-called static characteristic family sheet, with the corresponding value of ωt

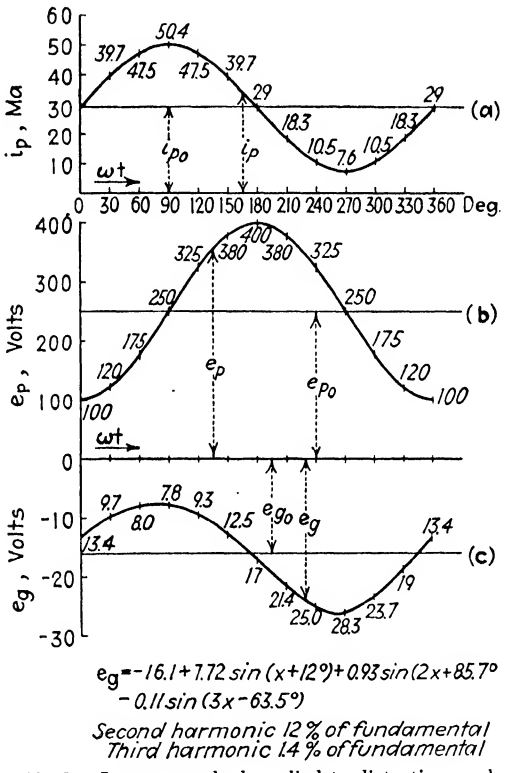


FIG. 13.28.—Inverse method applied to distortion analysis.

marked along the ellipse. (See Fig. 13.29.) Now, by remembering the fundamental principle that for any pair of instantaneous values of e_p and i_p there is only one value and a definite value of e_g, a curve may be constructed of the grid voltage e_g that will cause the tube to deliver 150 peak volts across the 7,000-ohm inductive load. By interpolating between the curves for various direct-current grid potentials, one may put next to each value of ωt

a corresponding value of e_g. Then it can be spread against time as shown in Fig. 13.28(*c*).

The next step is to analyze the grid-voltage wave by any known process. If one has forgotten the standard method, he may easily find the second- and third-harmonic content by trial. In this particular case, the standard method of harmonic analysis gives for the second harmonic 12 per cent of the fundamental and for the third, 1.4 per cent of the fundamental of the grid

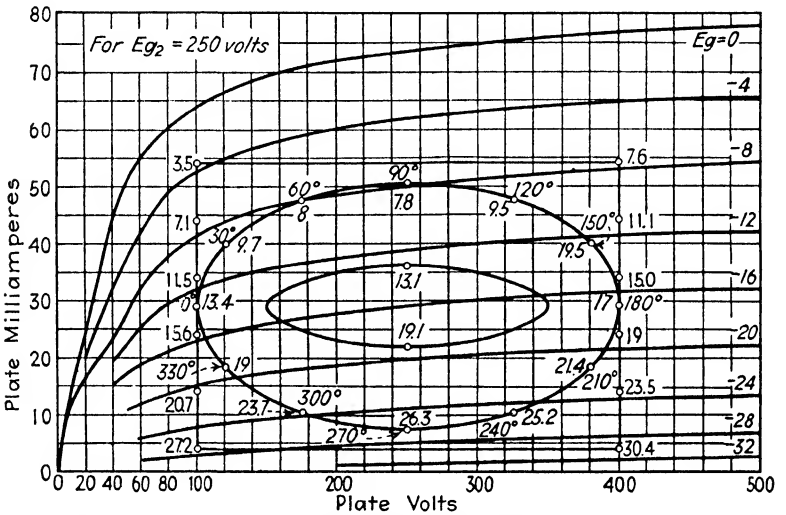


FIG. 13.29.—Average characteristics of type 247 tube with traces of solution by inverse method.

voltage. The 7,000-ohm load will become 14,000 ohms for the second harmonic, and 21,000 ohms for the third.

If now one impresses on the grid two additional waves, one of the frequency of the second harmonic of the fundamental, and the other of the frequency of the third harmonic, both 180 deg. out of phase, a pure sine-wave input on the grid and distorted current and voltage waves in the output will result.

The percentage of the harmonic in the output is estimated as follows:

Since the distortion in the amplification of the second and third harmonics will produce harmonics of the order higher than second and third of the fundamental in question, this distortion can be neglected. Choosing a convenient value for the scales

used, 100 peak volts across 14,000 ohms will require 7.14 peak ma. The corresponding ellipse will have 200 volts for major axis and 14.28 ma. for the minor. The minimum and maximum instantaneous grid voltages for such an ellipse are -13.1 volts and -19.1 volts, which gives an average peak alternating-current grid potential of 3.0 volts and the voltage-amplification factor of 100:3, or 33. Similarly the amplification for the load of 21,000 ohms is computed to be equal to 50.

The actual amplification of the fundamental is only

$$150:7.72 = 19.4,$$

and if on the grid the ratio of fundamental to first harmonic and to second harmonic is in the proportion of 100:12:1.4, then the same ratio in the output voltage is 100:20:3.6. A good theoretical reason exists, therefore, for using pentodes in push-pull to balance out the second harmonic. The peak grid swing is 7.72 volts, and the effective sine-wave voltage on the grid is 5.45 volts. The exact direct-current bias to keep the plate current at the desired value is 16.1 volts.

Following the above procedure, a number of similar problems can be solved. The method is especially applicable to problems requiring computation of a grid-voltage wave necessary to deliver either a current or voltage of odd wave shape to a reactive load.

Take a problem as follows:

It is desired to deliver a current wave of even saw-tooth shape to a pure inductive load of 3 henries. The current is to be of 500 complete periods per second, which means 1,000 sharp breaks per second. The total change wanted in the current between two successive breaks is 50 ma. It is required to find whether a 247 pentode, with normal supply voltages on its electrodes, is capable of delivering such a wave, and if so, what exact wave shape of voltage on the grid is required to produce this output.

First, plot the desired current against time. Directly under it, plot the voltage across the 3-henry inductance which constitutes the load in this case. This voltage e_x is found from the relation

$$e_x = L\frac{di}{dt}.$$

A function is continuous at a point if its limits, when it approaches the point from the right and from the left, both exist

and are equal to each other and to the value of the function at that point.

A function is said to have a finite or ordinary discontinuity at a point if its limits, when approaching this point from the right and from the left, both exist but are not equal, or not equal to the value of the function at that point.

It is apparent in this problem that while the current through the load is a continuous function of time, the voltage across it

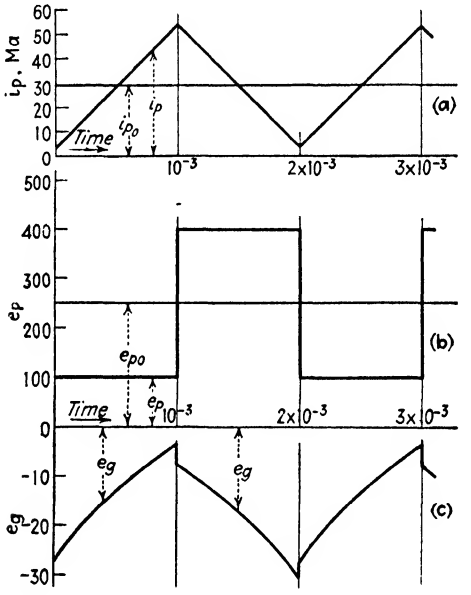


FIG. 13.30.—Inverse method applied to saw-tooth wave shape.

has finite discontinuities; also, that for the rising current the back voltage across the load is positive; and that for the decreasing current through the load the back voltage is negative.

Now one can compute the plate current and plate voltage as functions of time. They are, respectively,

$$i_p = i_{p0} + i_x$$

and

$$e_p = e_{p0} - e_x.$$

Their plots are given in Fig. 13.30(a) and 13.30(b). Now if one plots the resultant i_p against the resultant e_p, a perfect square

results. If one plots this square over the static-characteristic family of curves, an idea whether the tube can deliver the desired output may be obtained at once. From inspection of Fig. 13.29, a conclusion follows that in this particular case it can.

By going back again to the fundamental principle that for any pair of instantaneous values of e_p and i_p there is only one value, and a definite value, of e_g, one may construct the wave of the grid voltage which will make the tube deliver the current of the desired shape and magnitude to the load. Then it can be spread against time, as shown in Fig. 13.30(*c*).

The solution is complete, and all that remains to do is to connect an oscillograph to the grid of the pentode and adjust the circuits until the grid voltage of the calculated shape is obtained. The power output then will be of the desired value.

13.10. Single-tube Driving Circuit.—If a power tube of a driving circuit has an extra winding coupled to the main magnetic flux of the load, and by this means a voltage is impressed on the grid of the power tube 180° out of phase with the plate voltage, a direct generation of a saw-tooth wave of current is possible. The only requirement is that the L/C ratio of the oscillating circuit should be very large. An oscillator of such a type may be synchronized by the external impulse, although the required magnitude of the external impulse is considerably larger than with usual impulse generators.

While the above scheme appears very attractive from the standpoint of economy, it offers many practical disadvantages. A very serious one is that it is a two-reactive-element device, and therefore emission-limited. Its free frequency changes greatly with small variations in tube characteristics and with variations in applied working voltages. The only adjustment of frequency is by means of varying inductance, which is inherently inconvenient. Another important drawback of the single-tube driving circuit is that its wave shape leaves much to be desired in its approach of the perfect saw-tooth shape.

13.11. Magnetic Deflecting Yokes.—In Secs. 10.1, 10.6, 10.7 and 10.8, the requirements for a magnetic deflecting system were discussed. To satisfy these requirements, several types of yokes have been developed. A lamination of a typical iron-core yoke is similar to a stator lamination of a small motor. A

photograph of such a yoke, complete with windings and a terminal board, is shown in Fig. 13.31.

The yoke, as shown, has four windings, two for horizontal deflection and two for vertical deflection. The yoke has a deflect-

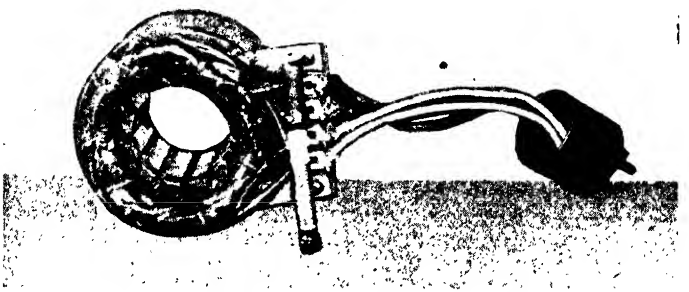


FIG. 13.31.—Improved magnetic deflecting yoke with iron core.

ing field which is uniform within 2 per cent along any line perpendicular to the undeflected beam in the direction of deflection.

Deflecting yokes may be made with air cores, or combined as shown in Fig. 13.32. The yoke shown has a pair of air-core

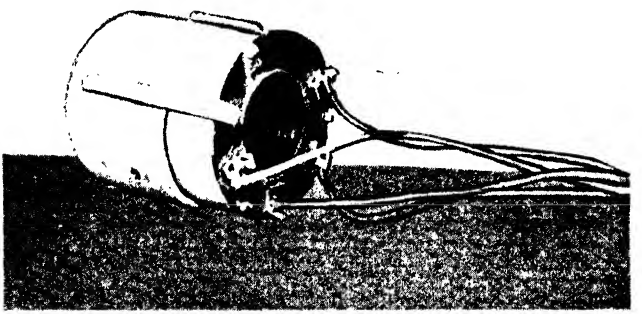


FIG. 13.32.—Improved magnetic deflecting yoke with air-core windings enclosed in iron tube.

horizontal windings next to the neck of the TCR tube. A pair of vertical windings is placed right on top of the horizontal. An iron tube is placed around the structure to reduce the reluctance of the return magnetic path of the vertical windings.

The windings are made with hollow center portions to reduce the concentration of flux in the center. The yokes, as just described, may be also made with field uniformity better than 2 per cent.

To reproduce a saw-tooth wave shape, the magnetic deflecting yoke should be capable of responding to many of the harmonics of the fundamental of the applied frequency. Other ways of obtaining the same result often have been suggested, but have not proven sufficiently advantageous to warrant treatment here.

For an infinite ratio of picture to return sweeps, a harmonic expansion is given in Byerly's "Fourier's Series" (p. 46). In a modified form it is

$$f(\omega t) = 0.636 \sin \omega t - 0.318 \sin 2\omega t + 0.212 \sin 3\omega t - 0.159 \sin 4\omega t + 0.128 \sin 5\omega t - 0.106 \sin 6\omega t + 0.091 \sin 7\omega t - 0.075 \sin 8\omega t + 0.071 \sin 9\omega t - 0.064 \sin 10\omega t + \cdots$$

The coefficients of respective terms are inversely proportional to the order of the harmonic. It follows, therefore, that to reproduce a saw-tooth of infinite return ratio, an infinite frequency response is required. In practice, however, since the time of the return is utilized for synchronizing, an infinite return ratio is not required.

For instance, in computing the coefficients of harmonic components of a 9:1 saw-tooth wave, the coefficient of 9th harmonic comes out 0.9 per cent of half of the height of the saw-tooth wave measured from peak to peak. The coefficient of 10th harmonic comes out as 0.

The development of 9:1 saw-tooth wave into its harmonic components is as follows:

$$\omega t = x$$
$$f(x) = 0\,354x \text{ for } 0 \leqslant x \leqslant 2.83$$

and

$$f(x) = 10 - 3.18x \qquad \text{for} \qquad 2.83 \leqslant x \leqslant 3.14$$
$$f(x) = a_1 \sin x + a_2 \sin 2x + a_3 \sin 3x + \cdots a_m \sin mx$$

where

$$a_m = \frac{2}{\pi}\int_0^{\pi} f(x) \sin mx dx$$

$$a_m = \frac{2}{\pi}\int_0^{162°} f(x) \sin mx + \frac{2}{\pi}\int_{162°}^{180°} f(x) \sin mx$$

$$= \frac{2}{\pi}\left[\frac{0.354}{m^2}(\sin m162° - 2.83m \cos m162°)\right.$$

$$+ \frac{10}{m}(\cos 162° - \cos 180°) + \frac{3.18}{m^2}(3.14m \cos m180°$$

$$\left. + \sin m162° - 2.83m \cos m162°)\right].$$

Table 13.1 gives values for coefficients up to the 10th harmonic for 9:1 and for an infinite ratio of return.

TABLE 13.1

m	1	2	3	4	5	6	7	8	9	10
a_m for 9:1	0.68	−0.33	0.204	−0.127	0.089	−0.059	0.037	−0.021	0.0086	0
a_m for 10:0	0.636	−0.318	0.212	−0.159	0.128	−0.106	0.091	−0.075	0.071	−0.064

Frequently in a deflecting system a serious cross-talk takes place between the horizontal and vertical circuits. Usually, the horizontal impulses find their way into the vertical circuit, and in this way produce zigzag scanning lines instead of straight lines.

It may be caused by a coupling of some sort between the driving circuits. This kind of cross-talk is usually eliminated by electrically isolating and shielding the respective circuits. Often, however, it takes place because of either electrostatic or electromagnetic coupling between the coils of the deflecting yoke. The type and degree of coupling are usually definitely connected with electric, magnetic and physical arrangements peculiar to this particular type. Therefore, it cannot be treated, in general, and has to be studied individually with every particular type of yoke. As a rule, however, the cross-talk can be eliminated by so arranging the coils on the yoke that the undesired induced voltages and currents buck out each other. Sometimes it calls for connecting horizontal coils in parallel and vertical ones in series. In other cases, both should be connected in parallel. While in some, no cross-talk is produced under any conditions.

In the preceding discussion of defects of the scanning pattern, perfectly symmetrical yokes and coils and a central position

of the beam with respect to the yoke were tacitly assumed. If however, for any reason, either the beam is not centrally located with respect to the yoke, the magnetic return legs of the yoke are not symmetrical, the coils are not symmetrically located, or the deflection coils in any given pair are of an unequal number of turns, the irregular defects of the scanning pattern result.

If the magnetic deflecting field is sufficiently uniform, however, the position of the beam with respect to the yoke is not so critical as in the case of a non-uniform field.

Any non-symmetry in the yoke ruins the uniformity of the field and immediately shows itself by producing defocusing in a part of the picture, stretching a corner or a side of the pattern, and usually producing serious cross-talk.

The symptoms of the irregular defects are such that they are easily located and eliminated by tracing defective coils and by checking the geometry of the yoke and the cathode-ray tube.

CHAPTER 14

VACUUM PRACTICE

14.1. Production of Vacuum.—The pressure in a TCR tube is about 10^{-6} mm. of mercury. There is no single pump available which is capable of producing such a low pressure. In general, the lower the pressure desired on the intake side of a vacuum pump, the smaller must be the exhaust pressure, *i.e.*, the pressure against which the pump is operating. The mercury- or oil-condensation pumps are capable of producing very high vacua, but require "fore pumps" to reduce their exhaust pressures.

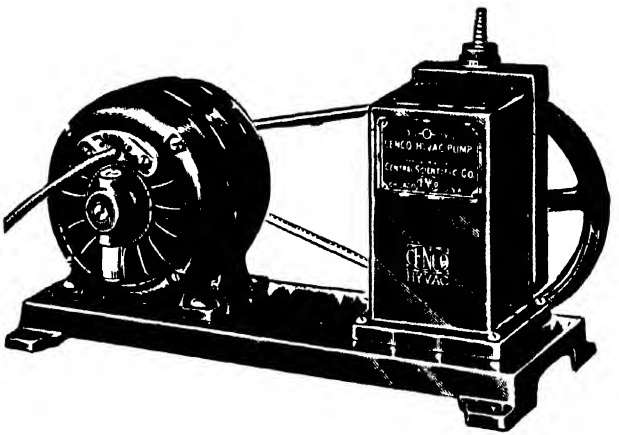


FIG. 14.1.—"Cenco-Hyvac" oil pump.

Pumps.—The fore pump most in use is a rotary oil pump. One of the many types of oil pumps is shown in Fig. 14.1. The operation of the pump may be understood from the diagrams of Fig. 14.2. The pump consists essentially of a rotor R mounted eccentrically on a shaft and rotating inside a close-fitting cylinder at a rate of 300 to 400 r.p.m. When the rotor is in position 1 of Fig. 14.2, gas from the vessel to be evacuated expands into the chamber C through the inlet tube I. As the rotor continues its rotation, chamber C increases in volume and chamber C' decreases in volume. The gas compressed in chamber C' is forced out through the output valve O. This is shown by posi-

tions 2 and 3. In position 4 the input is shut off and the chamber C occupies the entire volume. Further rotation will bring the rotor into position 1 and thus repeat the cycle. The vane V is held tightly against the rotor by spring pressure, and isolates the chambers C and C'. The pump shown in Fig. 14.1 consists of two such units, mounted on the same shaft with a small phase displacement between them, connected in series and immersed in oil. Such pumps are capable of reducing the pressure in a vessel connected to the intake to about 10^{-3} mm.

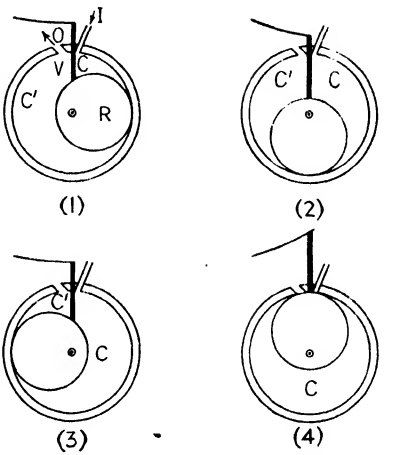


FIG. 14.2.—Operation of "Cenco" oil pump.

To produce the high vacuum needed for TCR tubes, a mercury- or oil-condensation pump is connected between the vessel to be evacuated and the fore pump. Figure 14.3 shows the cross section of a single-stage mercury-condensation pump. By heating the boiler the mercury in it is vaporized, and the mercury vapor particles traveling with considerable velocities are reflected down through jet J. The gas molecules diffuse from chamber C into the region of jet J, where they are drawn down by the fast moving mercury-vapor particles. The mercury vapor, on passing the jet, condenses on the inside wall of the water jacket W and returns to the boiler in liquid form through the U-tube. The gas molecules continue downward and are drawn away by the fore pump. It has been computed that at the

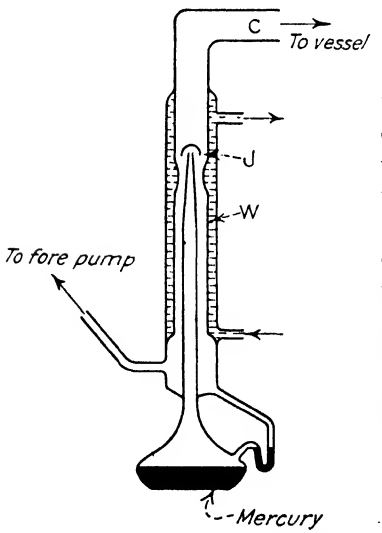


FIG. 14.3.—Single-stage mercury-condensation pump.

jet, the chance of diffusion of the gas molecules in the direction opposite to that in which the mercury vapor is passing is about 1 in 10^{20}. The mercury condensed just below the jet exerts a vapor pressure of about 10^{-3} mm. in C, so that a liquid-air trap is used to prevent the mercury from entering the tube. The back pressure against which the pump will operate depends upon the amount and velocity of the mercury vapor issuing from the jet. With sufficient heat applied to the boiler, mercury-condensation pumps have been operated against a back pressure of 1 mm. of mercury.

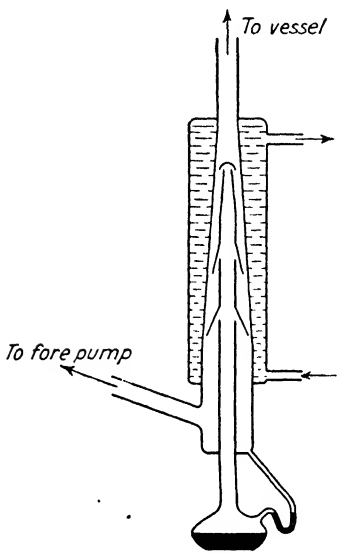


FIG. 14.4.—Triple-stage mercury-condensation pump.

Mercury-condensation pumps are built of either glass or metal and consist of two or more stages. The multistage pumps contain several jets in series, each jet creating a fore vacuum for the preceding one. The metal mercury-condensation pumps bought on the market usually require a fore vacuum of better than 1 mm. of mercury. The use of a low-vapor-tension oil instead of mercury in condensation pumps obviates the use of a liquid-air trap. However, the boiler temperature and pump construction require much more careful control and design. For this reason the general adoption of oil-condensation pumps has been slow. Figure 14.4 shows a triple-stage mercury-condensation pump.

Speed of Exhaustion.—The pressure p at any instant t in a vessel being exhausted is

$$p = p_\infty + (p_0 - p_\infty)\epsilon^{-\frac{k}{V}t} \tag{14.1}$$

where p_∞ is the lower limit of pressure attainable with the pump, p_0 is the pressure at time $t = 0$, V is the volume of the vessel and k is a constant. The rate of decrease of pressure is then

$$\frac{dp}{dt} = -\frac{k}{V}(p - p_\infty). \tag{14.2}$$

The speed of the pump has been defined as k and is thus a constant; however, the speed of exhaustion by the pump is defined as

$$S = -\frac{V}{p}\frac{dp}{dt}. \tag{14.3}$$

Inserting (14.2) into (14.3), it follows that

$$S = k\left(1 - \frac{p_\infty}{p}\right) \tag{14.4}$$

so that the speed of exhaustion decreases as the pressure p approaches p_∞, and becomes zero at $p = p_\infty$. The integration of Eq. (14.3) gives

$$S = \frac{V}{t_2 - t_1} \ln \frac{p_1}{p_2} = 2.3\frac{V}{t_2 - t_1} \log_{10} \frac{p_1}{p_2} \tag{14.5}$$

where $t_2 - t_1$ is the interval of time required for the reduction of pressure from p_1 to p_2 in a vessel of volume V. This means that a pump exhausting with a speed of S c.c./sec. will reduce the pressure of a vessel of volume S c.c. to $\frac{1}{\epsilon}$ in one second.

Connecting Tubes.—The above considerations assume that the vessel to be evacuated is attached directly to the pump. Connections between pump and vessel tend always to reduce the speed of exhaust. The connections act as an impedance. For pressures so low that the mean free path is large compared with the radius of the connecting tube, the molecular current through the tube is

$$i = (0.28 \times 10^{19})\frac{d^3}{l\sqrt{MT}}(p_2 - p_1) \tag{14.6}$$

where i is the number of molecules passing per second through the tube of length l cm. and diameter d cm., M the molecular weight of the gas, T the absolute temperature and $p_2 - p_1$ the difference in pressure, in bars, at the two ends of the tube. The impedance of such a tube is defined as the ratio $\frac{p_2 - p_1}{i}$; the impedance of a network of such tubes connected in parallel and in series obeys the same laws as a similar network of electrical impedances. Similarly, the molecular current through an aperture in a thin disk is

$$i = (0.264 \times 10^{19})\frac{A}{\sqrt{MT}}(p_2 - p_1) \qquad (14.7)$$

where A is the area of the aperture in square centimeters.

From Eq. (14.6) it is seen that the impedance of a tube to the flow of gases at low pressures varies inversely as the cube of the diameter of the tube. Small-diameter tubing should, therefore, be avoided as much as possible, particularly on the high-vacuum side. In order that a connecting tube should not greatly reduce the speed of exhaustion of very fast pumps, such as diffusion pumps, it is necessary that d^3/l be considerably greater than unity.

In practice, d^3/l is usually considerably smaller than unity, and the speed of exhaustion is usually reduced from one-half to one-tenth of what it would be if the vessel were directly connected to the mouth of the pump. Bends, constrictions and traps further reduce the speed of exhaustion.

Since there is always a pressure drop along a connecting tube, bend, constriction or trap, not only is the speed of exhaustion further reduced but the final pressure attained with a given pump is increased.

The final pressure in a system is determined by the state of equilibrium reached when the rate at which gas is removed by the pump is equal to the rate at which gas leaks into the system through connections, etc., and the rate of evolution of gas from the glass walls and metal parts in the system. With sufficient care, leaks through connections, etc., may be made small in comparison with the evolution of gas from the metal and glass parts. The evolution of gas depends upon the vapor pressure of the substances in the system and the amount of absorbed and adsorbed gases.

Table 14.1 gives approximate vapor pressures of a number of substances at various temperatures. It is practically impossible to reduce the pressure in a vessel below that of the vapor pressure present in the vessel. Thus, if there is moisture in a vessel at 20°C., the lower limit of pressure would be about 17 mm. This shows the importance of removing all moisture from a system. Similarly, if mercury is present in a system at 20°C, the lower limit of pressure would be about 10^{-3} mm. As the vapor pressure in a vessel corresponds to the temperature of the coldest

part of the vessel, it is possible to remove these vapors by a liquid-air trap (−190°C.). The vapor pressure of water vapor at −190°C. is well below 10^{-6} mm. and that of mercury well below 10^{-8} mm.

Glass and metals at ordinary temperatures and pressures contain large quantities of gases dissolved in the solid (absorbed) and condensed on the surface (adsorbed). The gases found in

TABLE 14.1.—VAPOR PRESSURE OF VARIOUS SUBSTANCES
(In millimeters of mercury)

Temp., °C.	Water	Mercury	Zinc	"Apiezon"		Copper	Nickel	Chromium
				Oil *A*	Oil *B*			
−190	10^{-6}	10^{-8}						
− 80	10^{-3}	4×10^{-6}						
− 40	0.1	7×10^{-6}						
− 20	0 8	2×10^{-5}						
0	4 6	2×10^{-4}						
20	17.5	1.2×10^{-3}		10^{-5}	10^{-6}			
50	92.5	1.2×10^{-2}						
100	760	2.7×10^{-1}						
200		17.8	6×10^{-6}					
350		680	2×10^{-2}					
500			2					
800			300			10^{-7}	10^{-8}	2×10^{-2}
1000						5×10^{-5}	5×10^{-6}	5×10^{-2}
1500						5×10^{-1}	10^{-2}	5

largest amounts are usually CO_2, H_2O, N_2 and O_2. The larger quantities of gas are usually held by the glass and consist mostly of water vapor and CO_2. (These gases may be driven out by heating the metals and glass in vacuum to relatively high temperatures). Figure 14.5 shows a simple vacuum system suitable for the evacuation of TCR tubes.

14.2. Measurement of Vacuum.—As in the case of pumps, two gauges are required in order to cover the entire range of pressures met in pumping a TCR tube. The only important gauge, however, is the one that is to cover the high-vacuum range. Although there are a great number of various types of gauges available, there are only two types that give absolute readings. They are the McLeod gauge, reading in millimeters of mercury, and Knudsen's gauge, reading in bars. All other gauges are

calibrated by comparison with one of these two. Two particularly useful gauges, due to the ease with which they are manipulated, are the thermocouple and ionization gauges; the former

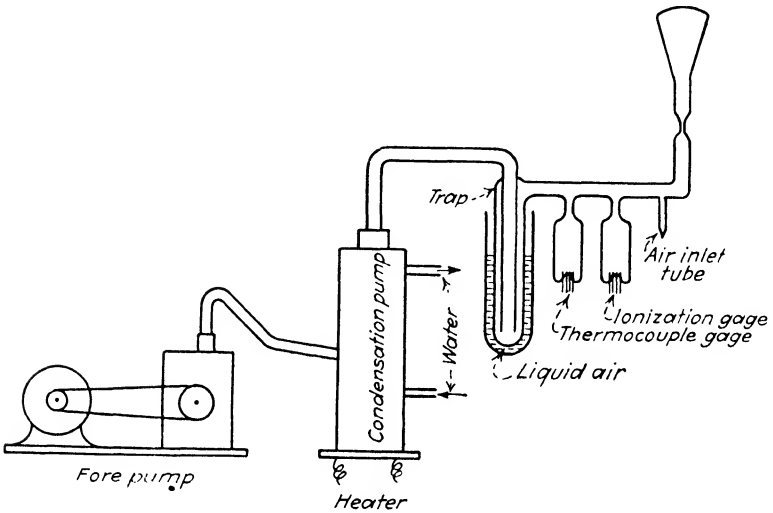


FIG. 14.5.—Simple vacuum system used in evacuating TCR tubes.

covers the low-vacuum range (10^{-1} to 10^{-3} mm.), and the latter, the high-vacuum range (10^{-3} to 10^{-7} mm.). These gauges have to be calibrated for each gas with which they are used.

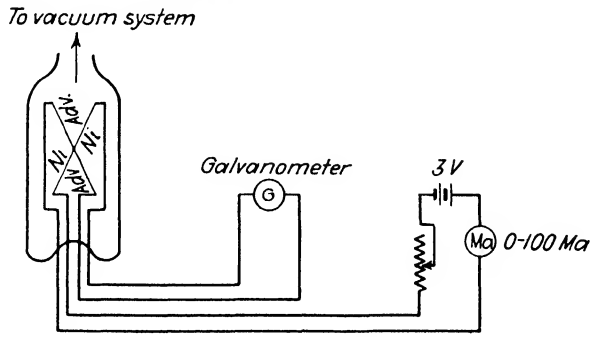


FIG. 14.6.—Connections for thermocouple gauge.

Thermocouple Gauge.—While at ordinary pressures the heat conductivity of gases is practically independent of pressure, the heat conductivity of gases at lower pressures decreases with pressure. This fact is utilized in the design of the thermocouple

gauge shown in Fig. 14.6. The thermo element is enclosed in a glass bulb which is connected to the system in which the pressure is to be measured. The thermo element may consist of a 0.002-in. nickel wire welded to a 0.002-in. advance wire, and mounted as shown in Fig. 14.6. A current passing through the junction heats it, and the thermo e.m.f. at the junction sends a current through the galvanometer. At pressures above about 0.1 mm., the e.m.f. at the junction, or the current through the galvanometer, is practically independent of the pressure, but for pressures below 0.1 mm., the heat loss by conduction decreases,

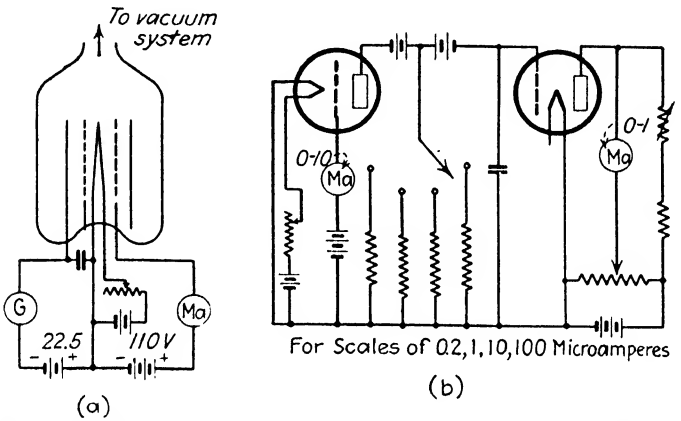


FIG. 14.7.—Ionization gauge, (*a*) connected with galvanometer, (*b*) connected with direct-current amplifier and milliammeter.

and the temperature of the junction increases, and so the e.m.f. of the junction also increases. Between 10^{-1} and 10^{-3} mm., the galvanometer current increases roughly as the logarithm of the pressure. For still lower pressures the galvanometer current reaches a maximum and stays constant as the pressure is reduced.

Ionization Gauge.—Any triode without "getter" may serve as an ionization gauge. Figure 14.7(*a*) shows the basic connections of the ionization gauge. An electron current from the filament to the grid ionizes some of the gas molecules present, and this positive ion current flows to the plate and is noted on a galvanometer. Figure 14.7(*b*) shows the galvanometer replaced by a direct-current amplifier and milliammeter, increasing the facility of the ionization gauge. For sufficiently low pressures and suitable grid and plate voltages it is found that

$$p = \frac{ki}{C} \tag{14.8}$$

where p is the gas pressure, i the positive ion current, C the electron current and k a constant for a particular gauge and type of gas. In making measurements, C is kept constant. For a triode like the 210, with the voltages shown in Fig. 14.7 and with $C = 10$ ma., Eq. (14.8) becomes for air

$$p \text{ (in mm. of Hg)} = 2.8 \times 10^{-5} i \text{ (in } \mu\text{a.)}$$

An interesting fact is that k is approximately inversely proportional to the number[1] of electrons in a molecule of the gas. Thus, keeping the pressure, electron current and plate voltage constant, it is found that the ionization current i for mercury (80 electrons) is about 5.5 times as great as for argon (18 electrons). This fact enables one to extend the use of the ionization gauge to gases other than the one with which it was calibrated.

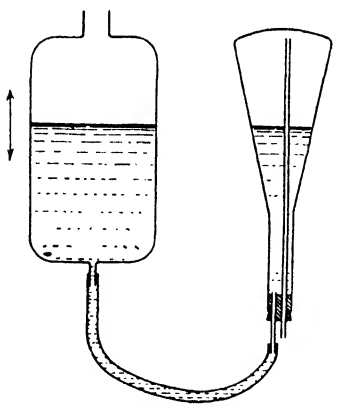


FIG. 14.8.—Method of applying aquadag coating.

14.3. Preparation of Tube.—In preparing the TCR tube for evacuation, the general rule that must be followed above all others is cleanliness. Starting with the glass (preferably Pyrex) blank, the first operation is to seal a tungsten wire into the glass, later to serve as second-anode contact. The next operation is to carefully clean the blank, first with a soapy solution such as Metzo and then with a boiling chromic acid solution, finally rinsing the blank several times with hot distilled water. Having cleaned the blank, the tungsten wire and the glass on the inside of the blank immediately surrounding the wire are painted with a silver oxide suspension which is reduced to silver by carefully heating and annealing the blank. This silver coating around the tungsten wire serves to make better contact with the second-anode coating. The blank is now ready for the application of the fluorescent material on the screen. If this material is willemite, it may be deposited upon the glass either by settling out of a water suspension or by spraying

[1] This is found not to hold for hydrogen and helium.

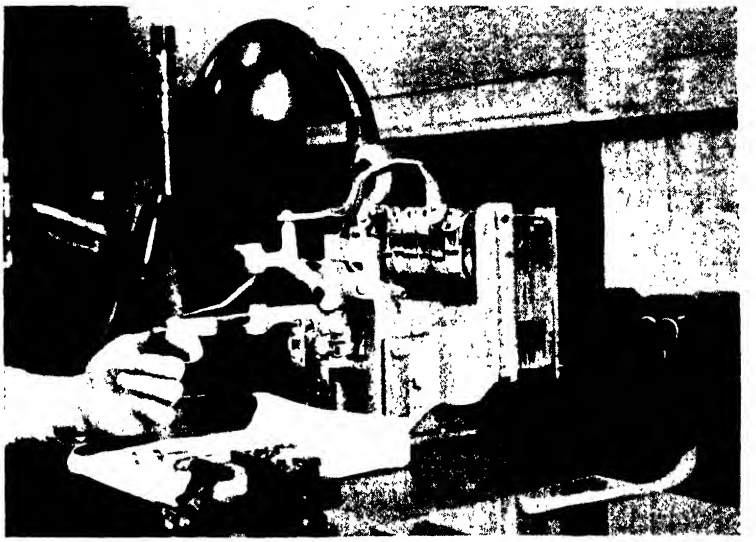


FIG. 14.9.—Photograph of a 1-kw. spot-welder.

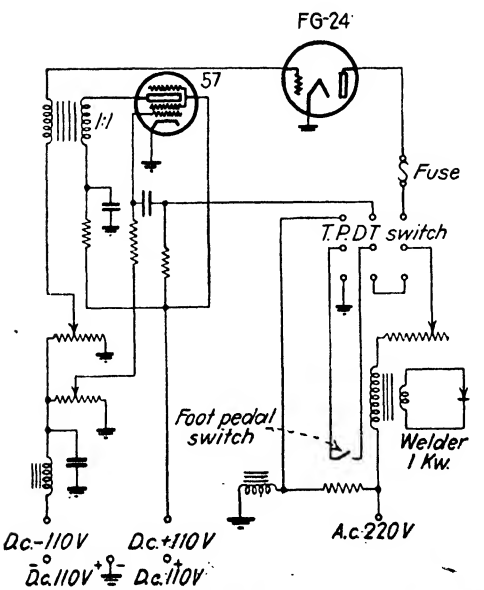


FIG. 14.10.—Circuit of a thyratron-controlled welder.

with an ordinary spray gun. The second-anode coating, usually of aquadag, is next applied. Figure 14.8 shows a convenient simple method of applying this coating.

The first-anode and grid structures are assembled by spot-welding. Figure 14.9 gives a photograph of a spot-welder. To obtain reproducibly clean and good welds, it is best that the welder be thyratron-controlled and that actual welding time be

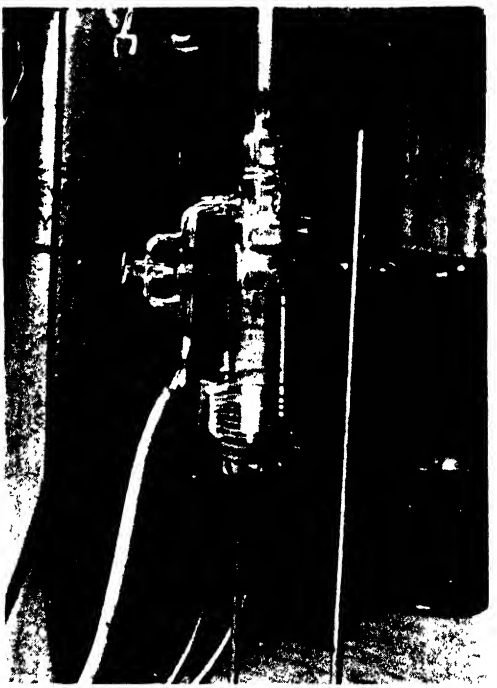


FIG. 14.11.—Parts of a gun being vacuum-cleaned.

limited to about $\frac{1}{120}$ sec. Figure 14.10 shows the circuit of a thyratron welder. The triple-pole, double-throw switch permits the use of the welder either thyratron-controlled or straight.

The assembled first anode, grid (and screen grid) and cathode are then cleaned. The cleaning is best accomplished by heating these parts in vacuum. Figure 14.11 shows the parts of a gun being vacuum-cleaned. By passing about 50 amp. of high-frequency (200–300 kc.) current through the coil, the metal parts are heated by induction. After the cathode has been thus

cleaned, it is sprayed with the suspension of calcium and barium carbonates, as described in Sec. 9.11.

Although tungsten may be sealed in Pyrex, it is quite a difficult task to make a good seal. Tungsten seals quite easily in Nonex. Accordingly, the wires of the press are sealed in Nonex, and a Pyrex flare is sealed to a special glass which in turn is sealed to the Nonex glass. The special glass has an expansion characteristic between that of Nonex and Pyrex.

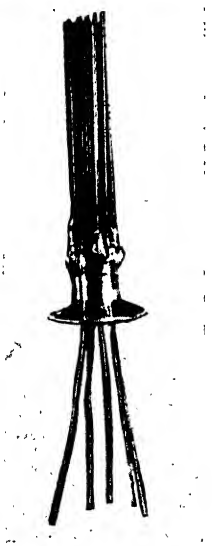


FIG. 14.12.—Press for gun assembly.

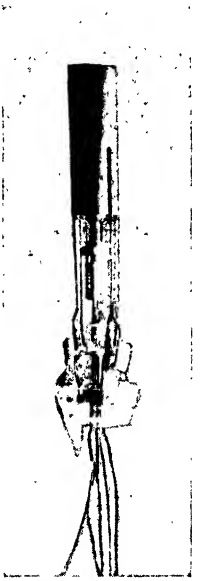


FIG. 14.13. — Gun assembled on press.

Figure 14.12 shows a press suitable for the assembly of an electron gun. Figure 14.13 shows a gun assembled on a press. Having assembled the gun on the press, the press is then sealed to the blank and this is sealed to the vacuum system. Figure 14.14 shows a TCR tube sealed to a vacuum system.

14.4. Evacuation of Tube.—The evacuation of the tube is begun by turning on the fore pump. After this pump has reduced the pressure to about 0.1 mm. or less, as read on the thermocouple gauge, the recently made seals are tested for leaks. A very convenient method of testing for pinholes in glass is with he aid of a small Tesla coil. When the high-voltage terminal

of such a coil is brought near a pinhole leak, a direct passage of a spark occurs through the leak, while a uniform glow occurs if the terminal is brought near a place where there is no leak. No glow appears if the pressure is below 10^{-3} mm. Having found no leaks, the baking oven is turned on and lowered over the tube. Figure 14.15 shows a complete vacuum system with a cylindrical oven which slides up and down along vertical guides. The temperature near the tube is measured by means of a thermocouple, shown clearly in Fig. 14.14.

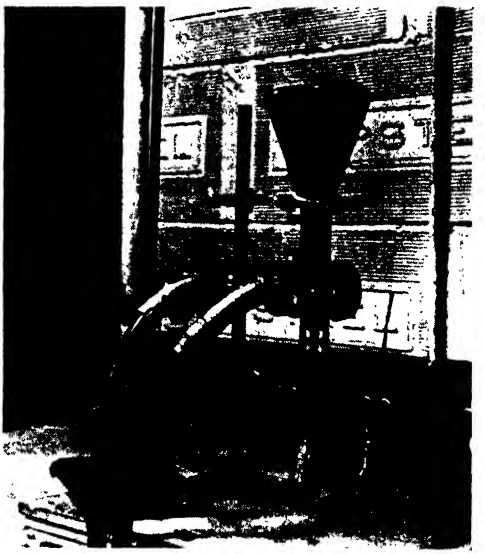


FIG. 14.14.—TCR tube sealed to vacuum system.

The purpose of the baking is to drive off most of the gas from the glass. The gas adsorbed on the surface of the glass is usually evolved at temperatures below 300°C.; however, in order to drive off the gas dissolved in the body of the glass, it is necessary to heat the glass to as high a temperature as the glass will permit. Pyrex may be heated to 475°C. without any danger of softening and collapsing. When the temperature, as indicated by the thermocouple, has reached about 350°C., the heater of the condensation pump is switched on. The reason for waiting till this temperature has been reached is to allow the great volume of gas first emitted on heating the glass to be drawn off by the

fore pump. It usually takes less than ten minutes for a mercury-condensation pump to "take hold," at which time the thermocouple-gauge galvanometer needle swings through practically its whole range. After the condensation pump has taken hold, the liquid-air bottle is slowly raised around the trap so that the liquid air does not boil too rapidly. It is to be made certain that prior to switching on the heater of the condensation pump, water has started circulating through the water jacket.

FIG. 14.15.—Vacuum system and baking oven for TCR tube.

The temperature, as measured by the thermocouple near the tube, is increased to the baking temperature (about 450°C. for Pyrex) and kept there for several hours by an automatic temperature control. When the oven has reached the maximum temperature, the ionization-gauge bulb is slowly heated with a bunsen burner to drive off the gas from the glass, and the plate of the ionization gauge is outgassed by means of a high-frequency coil, such as that shown in Fig. 14.11. The ionization gauge is switched on whenever the thermocouple gauge indicates a suffi-

ciently high vacuum ($<10^{-2}$ mm.). The length of time required for baking depends upon the condition of the glass. Sufficient baking has been given if the ionization gauge reads about 2 μa for a 10-ma. electron current.

With baking complete, the furnace is turned off and, after the system has cooled off, the gun is outgassed by heating with a high-frequency coil, as shown in Fig. 14.14. The gun, having been previously vacuum-cleaned, requires but little outgassing and should not be heated above a dull red. If the gun is heated too high, vapors from the metals evolve in large quantities and condense on the cooler portions of the system. The condensation of this vapor on insulators introduces objectionable electrical leakage. Having outgassed the gun, the cathode is activated, as described in Sec. 9.11. It is important to outgas the "getter" tab prior to activating the cathode, for if the tab is outgassed after the cathode has been activated, it is likely to be poisoned by the oxygen emitted in outgassing. The getter (usually a mixture of barium and aluminum) is flashed immediately after the activation, and the tube is then sealed off the pumps. Some time after the latter operation, the tube is tested for cut-off voltage, beam current, first-anode current and line width, as described in Sec. 12.6.

INDEX

A

B

C