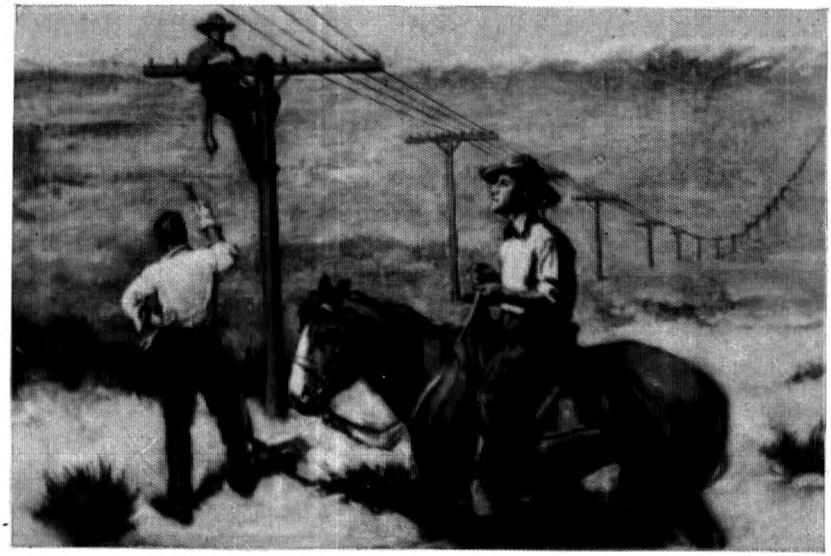


Three-color photograph transmitted over telephone line, as three separate black and white records, each corresponding to one primary color

The Bell System Technical Journal

April, 1925

The Transmission of Pictures Over Telephone Lines

By H. E. IVES and J. W. HORTON, *Bell Tel. Lab. Inc.*
R. D. PARKER and A. B. CLARK, *Amer. Tel. & Tel. Co.*

Introduction

THE problem of directly transmitting drawings, figures and photographs from one point to another by means of electricity has long attracted the attention and curiosity of scientists and engineers.[1] The broad principles of picture transmission have been recognized for many years. Their reduction to successful practice, however, required, among other things, the perfection of methods for the faithful transmission of electrical signals to long distances, and the development of special apparatus and methods which have become a part of the communication art only within the last few years. Prominent among the newer developments which have facilitated picture transmission are the photoelectric cell, the vacuum tube amplifier, electrical filters, and the use of carrier currents.

None of the systems heretofore devised have been sufficiently developed to meet the requirements of modern commercial service. The picture transmission system described in this article has been designed for practical use over long distances, employing facilities of the kind made available by the network of the Bell System.

The desirability of adding picture transmission facilities to the other communication facilities offered by the Bell System seems now to be well assured. Various engineers of the System have made suggestions and carried out fundamental studies of the possibilities for picture transmission offered by the telephone and telegraph facilities in the Bell System Plant which have aided materially in the development of the method to be described.

[1] A comprehensive account of earlier work in Picture Transmission will be found in "Telegraphic Transmission of Pictures," T. Thorne Baker, Van Nostrand, 1910, and the "Handbuch der Phototelegraphie und Telautographie," Korn and Glatzel, Leipzig, Nemnich, 1911.

The account of the picture transmission system which follows is intended to give only a general idea of the work as a whole. A number of engineers have collaborated in this work, and it is expected that later publications will describe various features of the system and its operation in greater detail.

General Scheme of Picture Transmission

Reduced to its simplest terms, the problem of transmitting a picture electrically from one point to another calls for three essential elements: The first is some means for translating the lights and shades of the picture into some characteristic of an electric current;

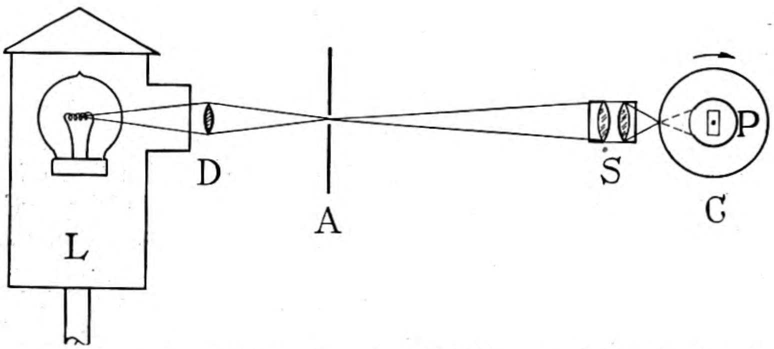


Fig. 1—Sending end optical system in section: (L) light source; (D) condensing lens; (A) diaphragm; (S) projection lens; (C) transparent picture film in cylindrical form; (P) photoelectric cell

the second is an electrical transmission channel capable of transmitting the characteristic of the electric current faithfully to the required distance; the third is a means for retranslating the electrical signal as received into lights and shades, corresponding in relative values and positions with those of the original picture.

Analyzed for purposes of electrical transmission, a picture consists of a large number of small elements, each of substantially uniform brightness. The transmission of an entire picture necessitates some method of traversing or scanning these elements. The method used in the present apparatus is to prepare the picture as a film transparency which is bent into the form of a cylinder. The cylinder is then mounted on a carriage, which is moved along its axis by means of a screw, at the same time that the film cylinder is rotated. A small spot of light thrown upon the film is thus caused to traverse the entire film area in a long spiral. The light passing into the interior of the cylinder then varies in intensity with the transmission or tone value of the picture. The optical arrangement by which a small spot of light is projected upon the photographic transparency is shown in section in Fig. 1.

The task of transforming this light of varying intensity into a variable electric current is performed by means of an alkali metal

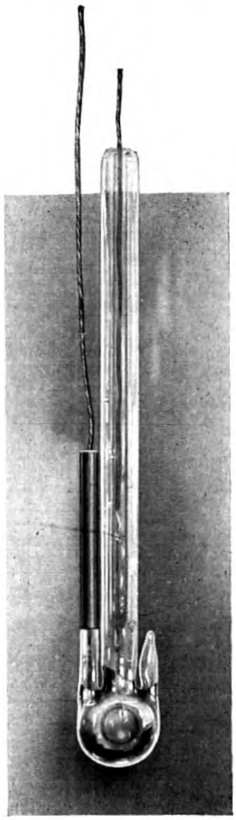


Fig. 2—Photograph of photoelectric cell of type used in picture transmission

photoelectric cell. This device, which is based on the fundamental discovery of the photoelectric effect by Hertz, was developed to a high degree of perfection by Elster and Geitel. It consists of a vacuum tube in which the cathode is an alkali metal, such as potassium. Under illumination, the alkali metal gives off electrons, so that when the two electrodes are connected through an external circuit, a current flows. This current is directly proportional to the intensity of the illumination, and the response to variations of illumination is practically instantaneous. A photograph of a photoelectric cell of the type used in the picture transmission apparatus is shown in Fig. 2. This cell is placed inside the cylinder formed by the photographic transparency which is to be transmitted, as shown in Fig. 1. As the film cylinder is rotated and advanced, the illumination of the cell and consequently the current from it registers in succession the brightness of each elementary area of the picture.

Assuming for the moment that the photoelectric current, which is a direct current of varying intensity, is of adequate strength for successful transmission, and that the transmission line is suitable for

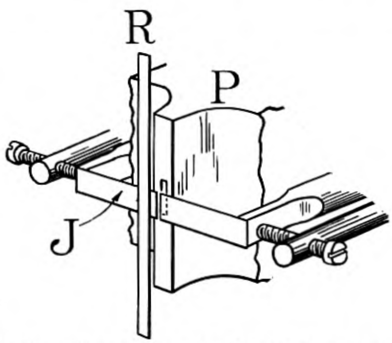


Fig. 3—Light valve details: (R) ribbon carrying picture current; (P) pole piece of magnet; (J) jaws of aperture behind ribbon

carrying direct current, we may imagine the current from the photoelectric cell to traverse a communication line to some distant point. At the distant point it is necessary to have the third element above mentioned, a device for retranslating the electric current into light and shade. This is accomplished in the present system by a device, due in its general form to Mr. E. C. Wente, termed a "light valve." This consists essentially of a narrow ribbon-like conductor lying in a magnetic field in such a position as to entirely cover a small aperture. The incoming current passes through this ribbon, which is in consequence deflected to one side by the inter-action of the current with the magnetic field, thus exposing the aperture beneath. Light passing through this aperture is thus varied in intensity. If it then falls upon a photographic sensitive film bent into cylindrical form, and rotating in exact synchronism with the film at the sending end, the film will be exposed by amounts varying in proportion to the lights and shades of the original picture. The ribbon and aperture of the light valve are shown diagrammatically in Fig. 3. Fig. 4

shows a section of the receiving end of a system of the sort postulated, with its light source, the light valve, and the receiving cylinder.

Adaptation of Scheme to Telephone Line Transmission

The simple scheme of picture transmission just outlined must be modified in order to adapt it for use on commercial electrical communication systems, which have been developed primarily for other purposes than picture transmission. Of existing electrical means of communication, which include land wire systems (telegraph and telephone), submarine cable, and radio, the wire system, as developed

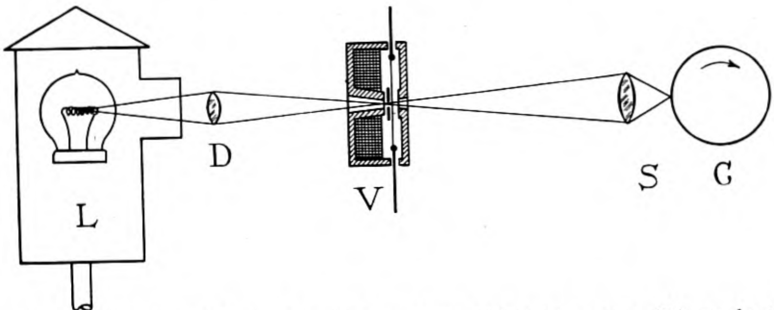


Fig. 4—Section of receiving end optical system: (L) light source; (D) condensing lens; (V) light valve; (S) projection lens; (C) sensitive film

for the telephone, offers great advantage when all factors are considered, including constancy, freedom from interference and speed. The picture transmission system has accordingly been adapted to it.

In the simple scheme of picture transmission outlined in the preceding section, the photoelectric cell gives rise to a direct current of varying amplitude. The range of frequency components in this current runs from zero up to a few hundred cycles. Commercial long distance telephone circuits are not ordinarily arranged to transmit direct or very low frequency currents, so the photoelectric currents are not directly transmitted. Moreover, these currents are very weak in comparison with ordinary telephone currents. On account of these facts, the current from the photoelectric cell is first amplified by means of vacuum tube amplifiers [2] and then is impressed upon a vacuum tube modulator jointly with a carrier current whose frequency is about 1,300 cycles per second. What is transmitted over

[2] For a very full description of the standard telephone repeater the reader is referred to "Telephone Repeaters," Gherardi and Jewett, Trans. A. I. E. E., Nov., 1919, Vol. 38, part 2, pp. 1287–1345.

Fig. 5—Portion of transmitted picture of variable width line type, enlarged

the telephone line is, then, the carrier wave [3] modulated by the photoelectric wave so that the currents, in frequency range and in amplitude, are similar to the currents corresponding to ordinary speech.

When the carrier current, modulated according to the lights and shades of the picture at the sending end, traverses the ribbon of the light valve at the receiving end, the aperture is opened and closed with each pulse of alternating current. The envelope of these pulses follows the light and shade of the picture, but the actual course of

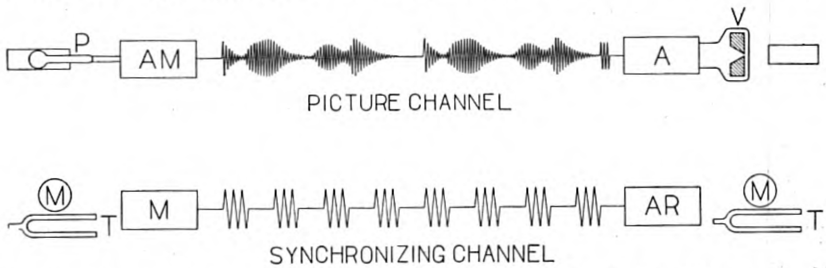


Fig. 6—Diagrammatic representation of the picture and synchronizing currents. (P) photoelectric cell; (AM) amplifier modulator; (A) amplifier; (V) light valve; (M) phonic wheel motors; (T) tuning forks; (AR) amplifier rectifier

the illumination with time shows a fine structure, of the periodicity of the carrier. This is shown by the enlarged section of a picture, Fig. 5; in this the black lines are traces of the image of the light valve aperture. Superposed on the larger variations of width, which are proportional to the light and shade of the picture, small steps will be noted (particularly where the line width varies rapidly); these are caused by the carrier pulses.

Synchronization

In order that the light and shade traced out on the receiving cylinder shall produce an accurate copy of the original picture, it is necessary that the two cylinders rotate at the same uniform rate. This, in general, demands the use of accurate timing devices. The means employed in the present apparatus consist of phonic wheels or impulse motors controlled by electrically operated tuning forks.[4] Were it

[3] A description of electrical communication by means of carrier currents will be found in "Carrier Current Telephony and Telegraphy," Colpitts and Blackwell, Trans. A. I. E. E., 1921, Vol. 40, pp. 205–300. A discussion of the relations between the several components of the signal wave employed in carrier is given in "Carrier and Sidebands in Radio Transmission," Hartley, Proc. I. R. E., Feb., 1923, Vol. 11, No. 1, pp. 34–55.

[4] A detailed description of the construction and operation of the impulse motor and its driving fork is given in "Printing Telegraph Systems," Bell Trans. A. I. E. E., 1920, Vol. 39, Part 1, pp. 167–230.

possible to have two forks at widely separated points running at exactly the same speed, the problem of synchronizing would be immediately solved. Actually this is not practical, since variations of speed with temperature and other causes prevent the two forks from operating closely enough together for this purpose. If the two cylinders are operated on separate forks, even though each end of the apparatus runs at a uniform rate, the received picture will, in general, be skewed with respect to the original. The method by which this difficulty has been overcome in the present instance is due to Mr. M. B. Long. Fundamentally the problem is solved by controlling the phonic wheel motors at each end by the same fork. For this purpose it has been found desirable to transmit to the receiving station impulses controlled by the fork at the sending end. The problem of transmitting both the fork impulses and the picture current simultaneously could be solved by the use of two separate circuits. If this were done the currents going over the two lines would be substantially as shown in Fig. 6, where the upper curve represents the modulated picture carrier for two successive revolutions of the picture cylinder, and the lower curve shows the synchronizing carrier current modulated by the fork impulses.

It would not, however, be economical to use two separate circuits for the picture and synchronizing channels, consequently the two currents are sent on the same circuit. In order to accomplish this, the picture is sent on the higher frequency carrier, approximately 1,300 cycles per second, and the synchronizing pulses are sent on the lower frequency carrier, approximately 400 cycles per second, both lying in the range of frequencies readily transmitted by any telephone circuit. These carrier frequencies are obtained from two vacuum tube oscillators.[5] The two currents are kept separate from each other by a system of electrical filters at the sending and receiving ends, so that while the current on the line consists of a mixture of two modulated frequencies, the appropriate parts of the receiving apparatus receive only one carrier frequency each.[6]

[5] The vacuum tube oscillator as a source of carrier current is described in Colpitts and Blackwell, *Loc. Cit.* A general discussion of the vacuum tube oscillator is given in the "Audion Oscillator," Heising, *Jour. A. I. E. E.* April and May, 1920. A discussion of the arrangement of the particular oscillator used with the picture transmission equipment is given in "Vacuum Tube Oscillator," Horton, *Bell System Tech. Jour.* July, 1924, Vol. 3, No. 3, pp. 508–524.

[6] The application of wave filters to multi-channel communication systems is discussed in Colpitts and Blackwell, *Loc. Cit.* More complete discussions are to be found in: "Physical Theory of Electric Wave Filters," Campbell, *Bell System Tech. Jour.* Nov., 1922, Vol. 1, No. 2, pp. 1–32.

Description of Apparatus

Mechanical Arrangements

The essential parts of the mechanism used for rotating and advancing the cylinder at the sending station, and for holding the photoelectric cell and the amplifying and modulating system are shown in the photograph, Fig. 7. At the extreme left is the phonic wheel impulse motor, which drives the lead screw through a spiral gear.

Fig. 7—Sending end apparatus showing motor, film carriage, optical system and amplifier modulator

The spiral gear ordinarily turns free of the lead screw, but may be engaged with it by a spring clutch. The lamp housing, which provides the illumination for the photoelectric cell, is in the foreground at the center of the photograph. The photoelectric cell is in a cylindrical case at the left end of the large box shown on the track and projects into the picture cylinder on which a film is in process of being clamped. The amplifier and modulator system is carried in the large box to the right, which is mounted on cushion supports to eliminate disturbances due to vibration.

The receiving end mechanism for turning and advancing the cylinder is similar to that at the sending end. The parts peculiar to the receiving end are shown in Fig. 8. They consist of the light valve, which is in the middle of the photograph, and the lens for projecting the light from it upon the cylinder. The metal cylinder

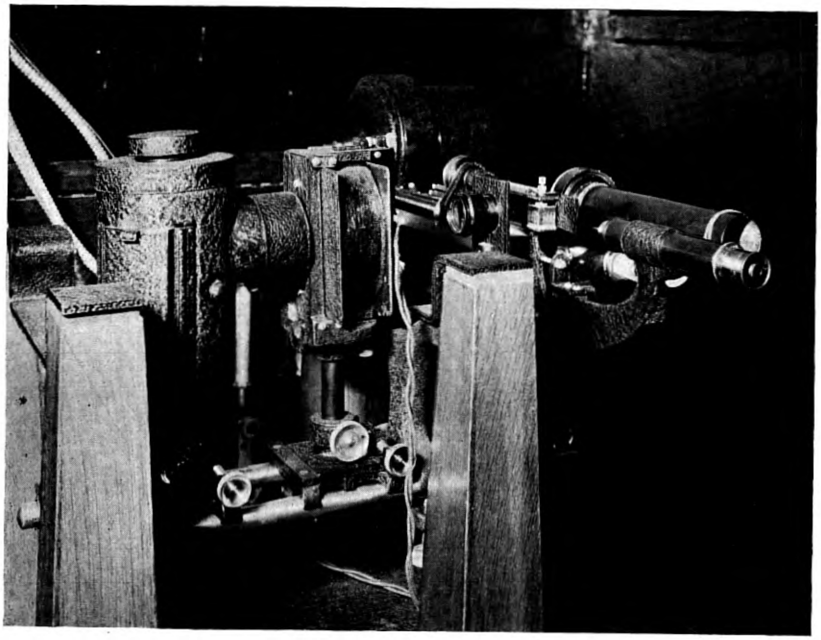


Fig. 8—View of receiving end apparatus showing light valve and observation microscope

around which the sensitive photographic film is wrapped, appears at the extreme right. The microscope and prism shown are used for inspecting the light valve aperture for adjusting purposes.

Electrical Circuits

The essential parts of the electrical circuits used are shown in the schematic diagrams, Figs. 9 and 10, in which the various elements which have been described previously are shown in their relations to each other.

Certain portions of the electrical circuits deserve somewhat detailed treatment. One of these is the amplifier-modulator system for the picture channel, the other is the filter system employed for separating the picture and synchronizing channels.

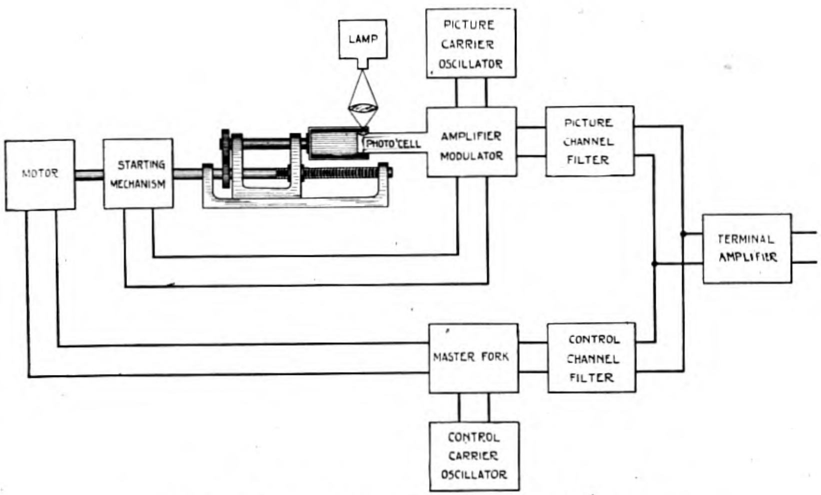


Fig. 9—Schematic diagram of sending end apparatus

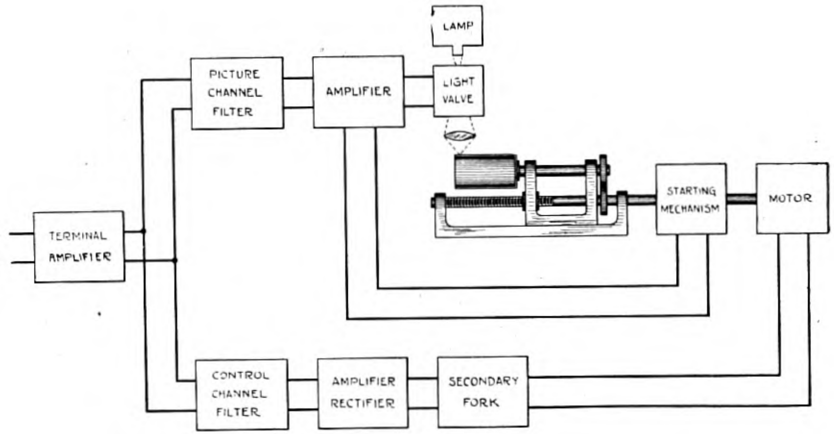


Fig. 10—Schematic diagram of receiving end apparatus

In Fig. 11 is shown (at the top) a diagram of the direct current amplifier and the modulator used for the picture channel, together with diagrams (at the bottom) showing the electrical characteristics of each element of the system. Starting at the extreme left is the

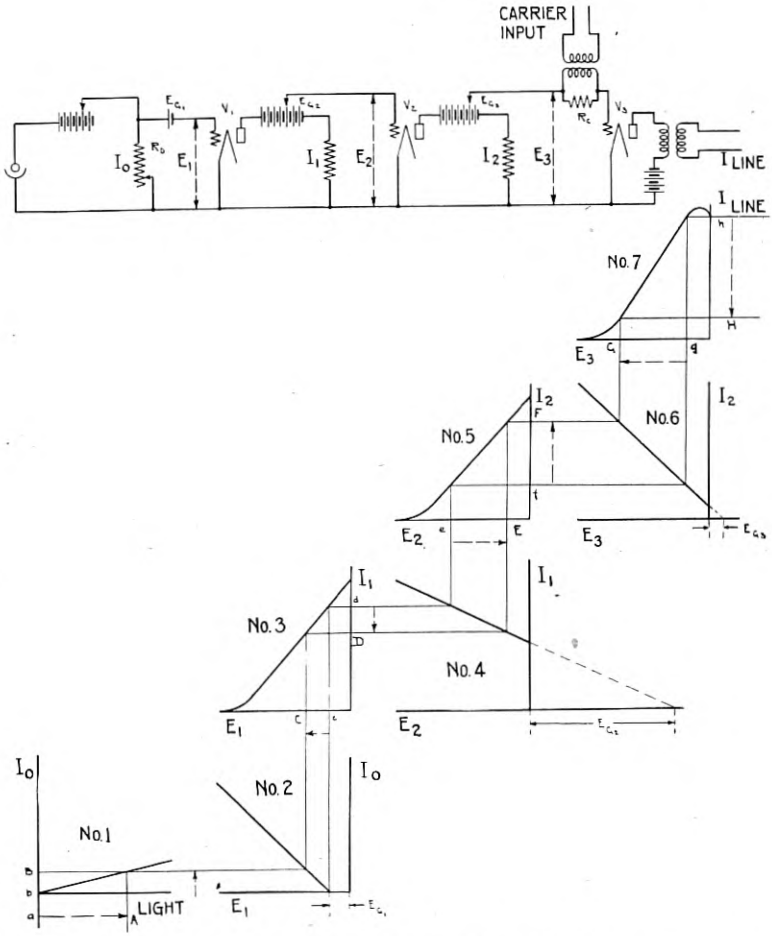


Fig. 11—Circuit schematic of amplifier-modulator with characteristics of each element

photoelectric cell, the current from which passes through a high resistance. The potential tapped off this resistance (of the order of 30 or 40 millivolts) is applied to the grid of the first vacuum tube amplifier. The second vacuum tube amplifier is similarly coupled with the first, and the vacuum tube modulator in turn to it. The relationship between illumination and current in the photoelectric cell is, as shown in diagram No. 1, linear from the lowest to the highest values of illumination. The voltage-current (E versus I) character-

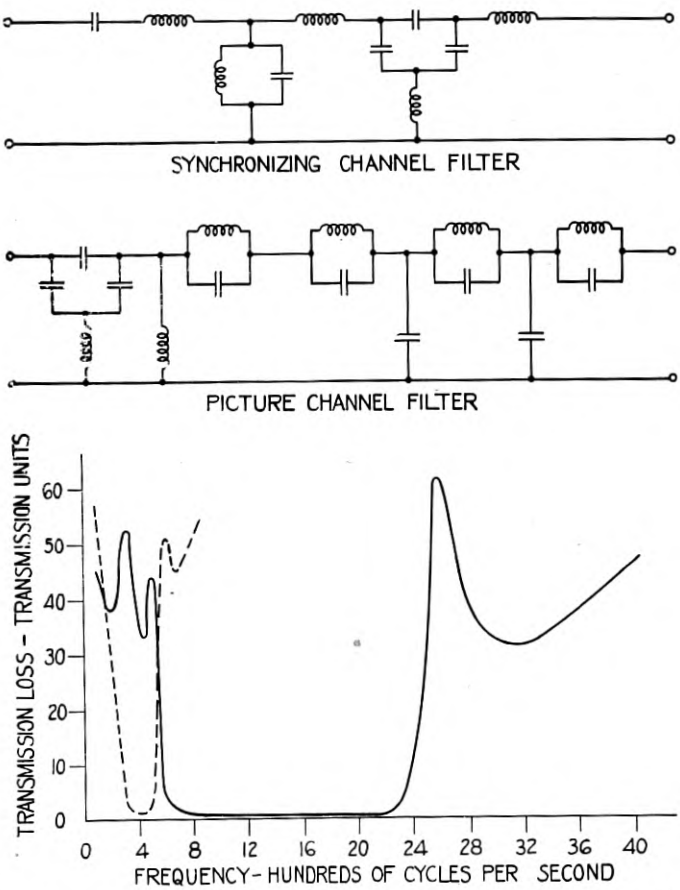


Fig. 12—Circuit schematics (above) and attenuation characteristics (below) of picture (full line) and synchronizing (dashed line) channel filters

istics of the amplifying tubes and the modulating tube circuits are shown in the figure by the diagrams which lie immediately below these tubes. They are not linear over their whole extent. It becomes necessary, therefore, in order to preserve the linear characteristic, which is essential for faithful picture transmission, to locate the range of variation of current in each of the latter tubes on a linear portion of their characteristics. This is accomplished by appropriate biasing voltages (Eg), as shown. As a consequence of this method of utilizing the straight line portions of the tube characteristics, the current received at the far end of the line does not vary between zero and finite value, but between two finite values. This electrical bias is exactly matched in the light valve by a mechanical bias of the jaws of the valve opening.

Fig. 12 shows diagrammatically the form of the band pass filters used for separating the picture and synchronizing channels, together with the transmission characteristics of the filters. The synchronizing channel filter transmits a narrow band in the neighborhood of 400 c. p. s., the picture channel filter a band between 600 and 2,500 c. p. s.

In addition to the main circuits which have been discussed, arrangements are made for starting the two ends simultaneously and for the transmission of signals. These functions are performed by the interruption of the picture current working through appropriate detectors and relays. Testing circuits are also provided for adjusting the various elements without the use of the actual transmission line.

The Transmission Line

In view of the fact already emphasized, that the currents used in picture transmission are caused to be similar both as to frequency and amplitude to those used in speech transmission, it follows that no important changes in the transmission characteristics of the telephone line are called for. With regard to the frequency range of the alternating currents which must be transmitted and also the permissible line attenuation, the transmission of pictures is less exacting on the telephone line than is speech transmission. In certain other respects, however, the requirements for picture transmission are more severe. For speech, the fundamental requirement is the intelligibility of the result, which may be preserved even though the transmission varies somewhat during a conversation. In the case of picture transmission, variations in the transmission loss of the line, or noise appearing even for a brief instant during the several minutes required for transmission are all recorded and presented to view as blemishes in the finished picture. Picture transmission circuits must, therefore, be carefully designed and operated so as to reduce the possibility of such disturbances. In transmitting pictures over telephone lines, it is also necessary to guard against certain other effects, including transient

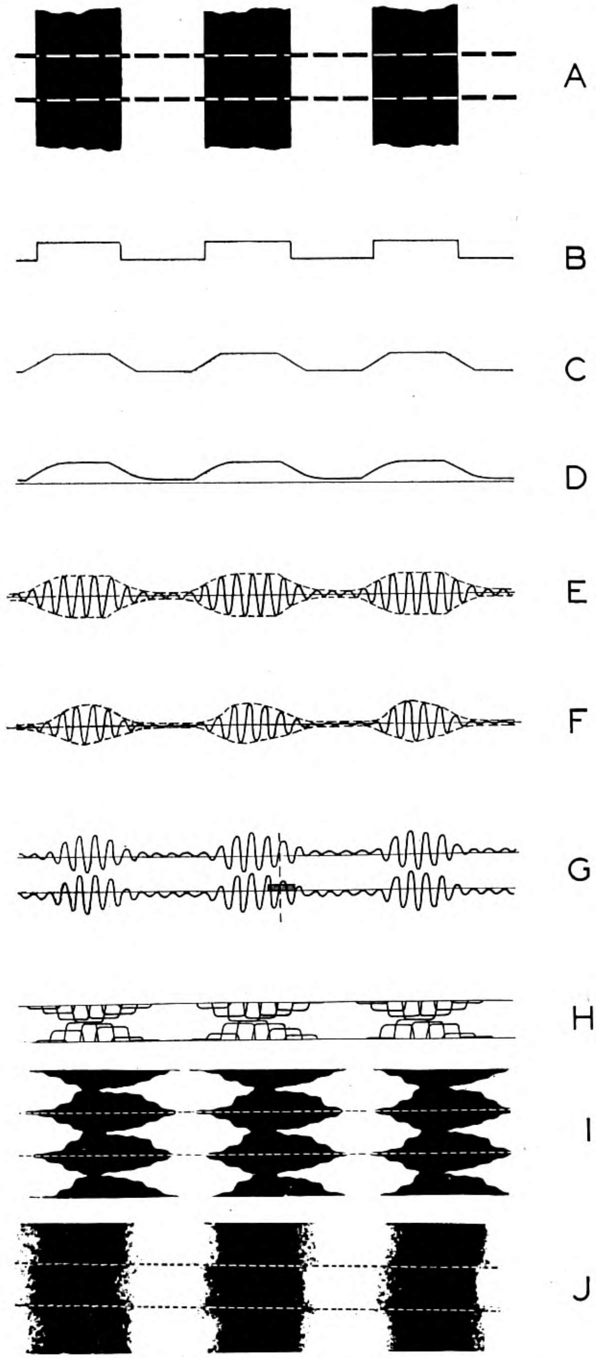


Fig. 13—Diagram illustrating performance of system

effects and "echoes" caused by reflections from impedance irregularities. A high degree of balance between the lines and their balancing networks at repeater points is also required. These conditions can be satisfactorily met on wire telephone lines. Radio communication channels are inherently less stable and less free from interference, and special means to overcome their defects are required in order to secure high-grade pictures.

Characteristics of Received Pictures

All electrically transmitted pictures have, as a result of the processes of scanning at the sending and receiving ends, a certain amount of structure, on the fineness and character of which depends the detail rendering of the result.

The origin and nature of the microscopic structure characteristic of pictures transmitted by the present process is illustrated by the diagrammatic presentation of Fig. 13, which may serve at the same time to give a review of the whole process. We will assume that the original picture consists of a test object of alternating opaque and transparent lines. Such a set of lines is shown at *A*. The lines are assumed to be moving from left to right across the spot of light falling on the film. The width of the spot of light (corresponding to the pitch of the screw) is represented by the pair of dashed lines. If the spot of light were infinitely narrow in the direction of motion of the picture film, the photoelectric current would be represented in magnitude in the manner shown at *B*. Actually the spot must have a finite length, so that the transitions between the maximum and minimum values of current are represented by diagonal lines as shown at *C*. Due to the unavoidable reactances in the amplifying system, there is introduced a certain rounding off of the signal so that the variation of potential impressed on the modulator tube follows somewhat the course shown at *D*. The alternating current introduced by the vacuum tube oscillator is, then, given the characteristics shown at *E*, the envelope being a close copy of *D*. Passing out to the transmission line, the fact that the band of frequencies transmitted by a telephone line is limited in extent results in a certain further rounding off of the envelope of the picture current as shown in *F*. The ribbon of the light valve when traversed by the alternating current from the line performs oscillations to either side of the center of the aperture, consequently opening first one side of the aperture and then the other. The two curves of sketch G represent the excursions of the light valve ribbon, with time, past the

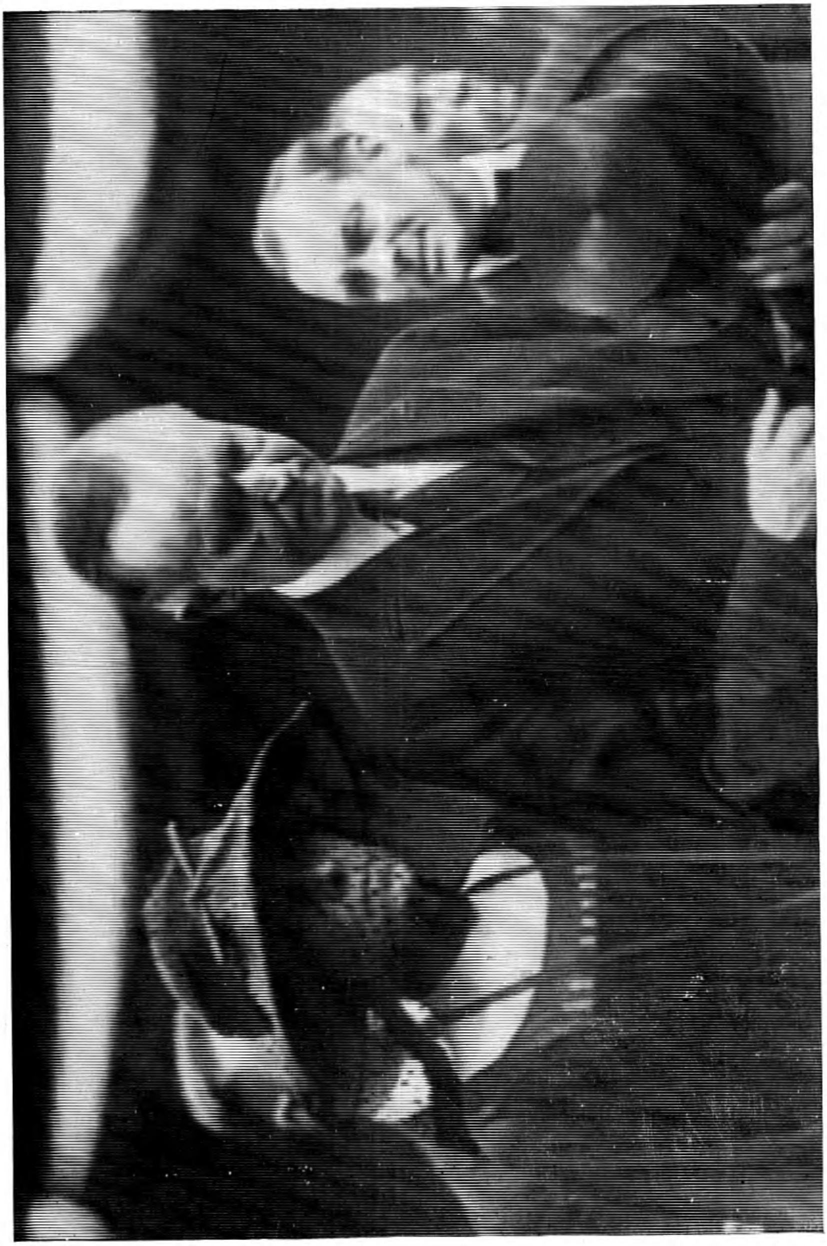


Fig. 14—Example of electrically transmitted news picture—variable width line system—President and Mrs. Coolidge

edges of the aperture, which latter are indicated by parallel straight lines. Owing to the fact that the light valve aperture must have a finite length in the direction of rotation of the cylinder (indicated by the small rectangle in the center of the sketch), there is a certain overlapping of the light pulses on the film. (This is, in fact, necessary for the production of solid blacks.) These are indicated diagrammatically at *H*. In sketch I are shown, from an actual photomicrograph, the variations in the image of the light valve as traced out on the moving photographic film. Here the dashed lines represent the limits of the image as formed by one rotation of the receiving cylinder. It will be noted that the images due to the opening of the light valve in each direction form a double beaded line. These double lines are juxtaposed, so that the right hand image due to one rotation of the cylinder backs up against the left hand image due to the next rotation, thus forming on the film a series of approximately symmetrical lines of variable width. These are exhibited clearly in the enlarged section of a picture, Fig. 5. It will be understood that for purposes of illustration, the grating used as the test object in the preceding discussion has been represented as traversing the spot of light at the sending end at such a high speed that the final picture is close to the limit of the resolving power of the system. Thus the photomicrograph shown in I must be viewed from a considerable distance in order that its difference in structure from the original object *A* will disappear. A practical problem in the design of picture transmission apparatus is to so choose the speed of rotation of the cylinder with reference to the losses in resolving power incident to transmission that definition is substantially the same along and across the constituent picture lines.

There are, in general, two methods by which a transmitted picture may be received. One of these is to form an image of the light valve aperture on the sensitive photographic surface. When this is done, in the manner described in connection with Fig. 13 the picture is made up of lines of constant density and varying width. A picture of this sort is shown in Fig. 14. A merit of this kind of picture (when received in negative form) is that if the structure is of suitable size (60 to 65 lines to the inch) it may be used to print directly on zinc and thus make a typographic printing plate similar to the earlier forms of half tone, whereby the loss of time usually incident to copying a picture for reproduction purposes may be avoided. A disadvantage of this form of picture is that it does not lend itself readily to retouching or to change of size in reproduction.

Another method of picture reception is to let the light from the

Fig. 15—Portion of transmitted picture of variable density line type, enlarged

Fig. 16—Variable density line picture—Cleveland high level bridge

Fig. 17—Variable density line picture—Portrait of Michael Faraday

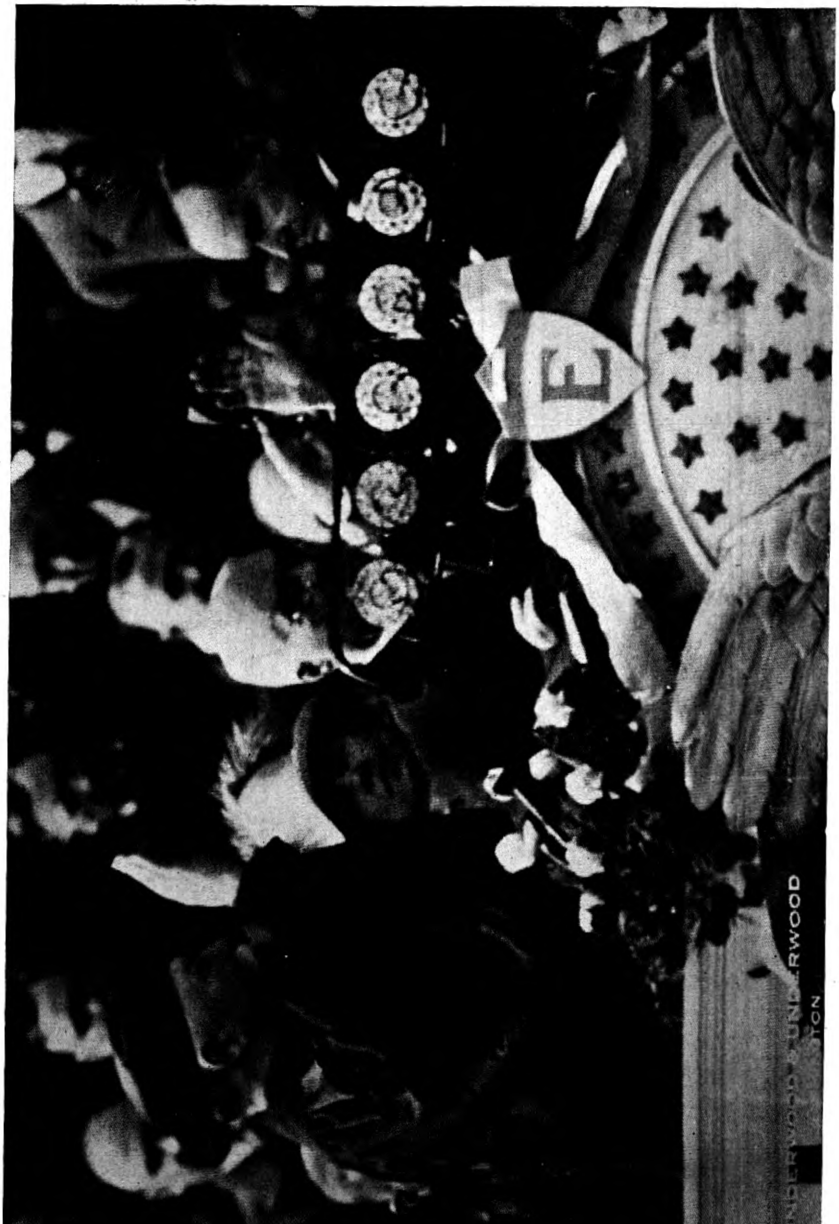


Fig. 18—Variable density line picture—President Coolidge taking the oath of office, March 4, 1925

light valve fall upon the film in a diffused manner through an aperture of fixed length so that lines of constant width (exactly juxtaposed) but of varying density are produced. A photomicrograph of a variable density picture of the opaque line test object previously discussed is shown at *J*, Fig. 13. Prints made from film negatives received in this way, if the structure is chosen fine enough (100 lines to the inch or more) are closely similar in appearance to original photographic prints and may be reproduced through the ordinary half-tone cross-line screen. They may be retouched or subjected to special photographic procedures in any way desired. An enlargement of a portion of a variable density picture is shown in Fig. 15 and examples of complete pictures so received are shown in Figs. 16, 17 and 18.

Electrically transmitted pictures are, in general, suitable for all purposes for which direct photographic prints are used. Such uses include half-tone reproduction for magazines and newspapers, lantern slides, display photographs, etc. Among these uses may be mentioned, as of some interest, the transmission of the three black and white records used for making three-color printing plates. The frontispiece to this article is an example of a three-color photograph transmitted in the form of three black and white records, each corresponding to one of the primary colors, from which printing plates were made at the receiving end.

Some practical details of the procedure followed in the transmission of pictures by the apparatus described may serve to clarify the foregoing description. The picture to be transmitted is usually provided in the form of a negative, which is apt to be on glass and of any one of a number of sizes. From this a positive is made on a celluloid film of dimensions 5″ x 7″, which is then placed in the cylindrical film-holding frame at the sending end. Simultaneously an unexposed film is placed on the receiving end. Adjustments of current values for "light" and "dark" conditions are then made, over the line; after which the two cylinders are simultaneously started by a signal from one end. The time of transmission of a 5″ x 7″ picture is, for a 100 line to the inch picture, about seven minutes. This time is a relatively small part of the total time required from the taking of the picture until it is delivered in the form of a print. Most of this total time is used in the purely photographic operations. When these are reduced to a minimum by using the negative and the sending end positive while still wet, and making the prints in a projection camera without waiting for the received negative to dry, the overall time is of the order of three-quarters of an hour.

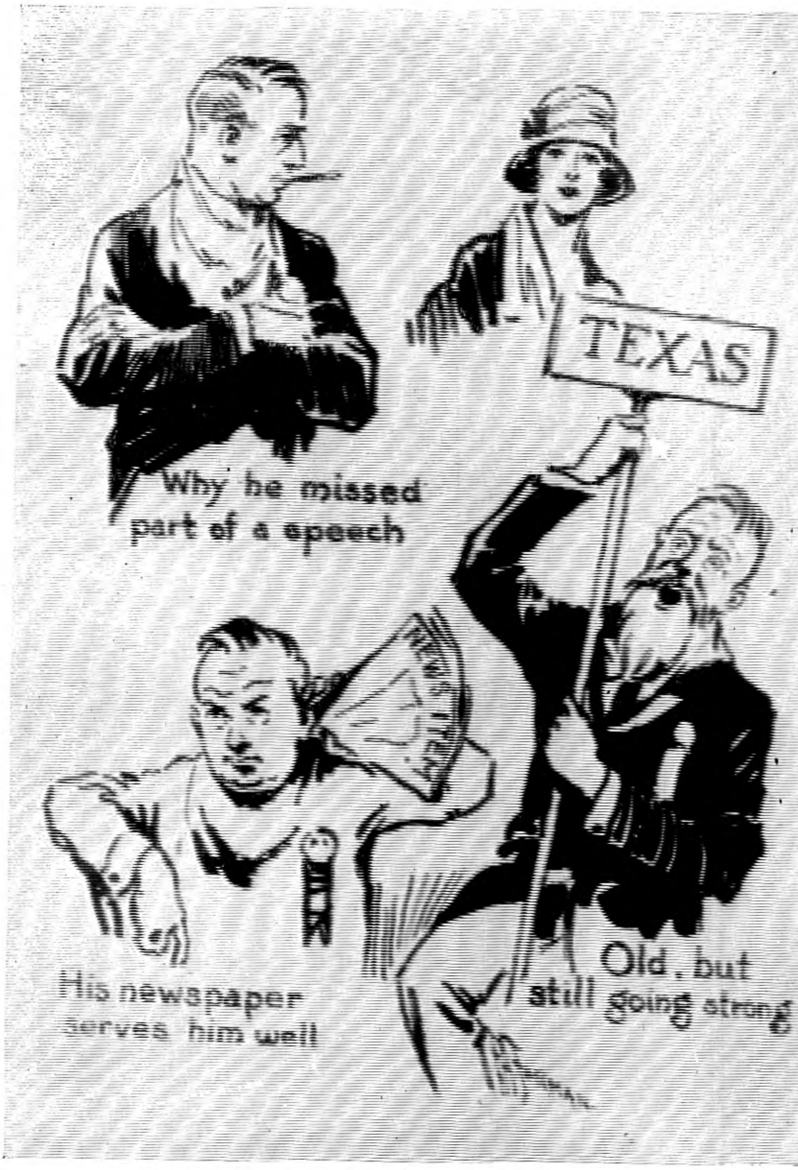


Fig. 19—Electrical transmission of cartoon

Fields of Usefulness

The fields in which electrically transmitted pictures may be of greatest service are those in which it is desired to transmit information which can only be conveyed effectively, or at all, by an appeal to vision. Illustrations of cases where an adequate verbal description is almost impossible, are portraits, as, for instance, of criminals

Fig. 20—Electrically transmitted fingerprint

or missing individuals; drawings, such as details of mechanical parts, weather maps, military maps, or other representations of transient conditions.

The value of electrically transmitted pictures in connection with police work has been recognized from the earliest days of experiments in the transmission of pictures. Besides the transmission of portraits of wanted individuals to distant points, there is now possible the transmission of finger prints. Some of the possibilities of the latter were demonstrated over the New York-Chicago picture sending circuit at the time of the Democratic Convention, July, 1924. The Police Department of New York selected the fingerprint of a criminal whose complete identification data were on file in the Police Department in Chicago. This single fingerprint, together with a code description of the prints of all the fingers, was transmitted to Chicago and identified by the Chicago experts almost instantly. This method of identification will be, it is thought, of value in those cases where difficulty is now experienced in holding a suspect long enough for identification to be completed. Fig. 20 shows a transmitted fingerprint.

The fact that an electrically transmitted picture is a faithful copy of the original, offers a field of usefulness in connection with the

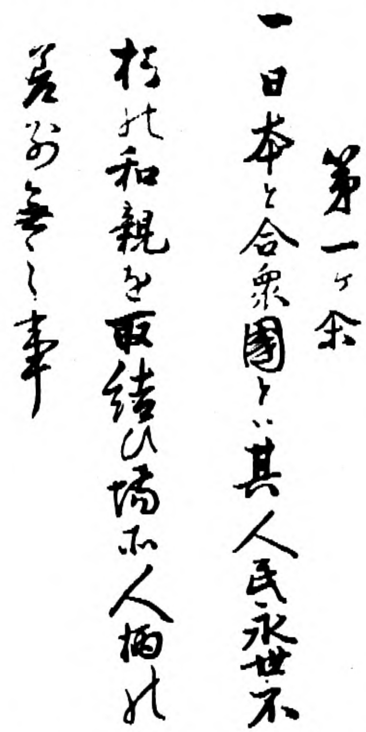


Fig. 21—Transmission of autograph material—First section of Japanese-American Treaty of 1853

transmission of original messages or documents in which the exact form is of significance, such as autographed letters, legal papers, signatures, etc. It would appear that this method might under certain circumstances save many days of valuable legal time and the accumulation of interest on money held in abeyance. For these reasons, it is thought that bankers, accountants, lawyers, and large real estate dealers will find a service of this kind useful. Fig. 22 illustrates the transmission of handwriting.

Messages in foreign languages, employing alphabets of forms not suited for telegraphic coding, are handled to advantage. Thus, Fig. 21 shows the first section of the original Japanese-American treaty in Japanese script, as transmitted from New York to Chicago.

Advertising material, particularly when in the form of special typography and drawings is often difficult and costly to get to dis-

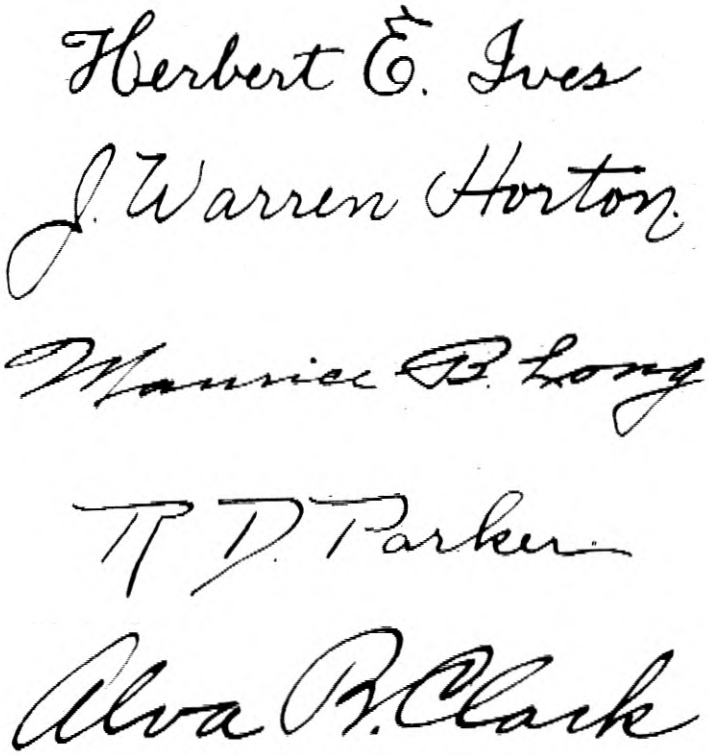


Fig. 22—Transmission of signatures

tant publishers in time for certain issues of periodicals and magazines. A wire service promises to be of considerable value for this purpose.

A very large field for electrically transmitted pictures is, of course, The Press. Their interest in the speedy transportation of pictures has been indicated in the past by the employment of special trains, aeroplanes, and other means for quickly conveying portraits and pictures of special events, to the large news distributing centers. The use of pictures by newspapers seems at present to be growing in favor, and many are now running daily picture pages as regular features.

Some of the possibilities in this direction were demonstrated by the picture news service furnished to newspapers, especially those in New York and Chicago, during the 1924 Republican and Democratic National Conventions at Cleveland and New York. During these conventions several hundred photographs were transmitted between Cleveland and New York and between New York and Chicago, and copies furnished the Press at the receiving points. Photographs made shortly after the opening sessions, usually about noon, were transmitted to New York and Chicago and reproduced in afternoon papers. A demonstration of picture news service on a still larger scale was furnished on March 4th, 1925, when pictures of the inauguration of President Coolidge were transmitted from Washington simultaneously to New York, Chicago and San Francisco, appearing in the afternoon papers in all three cities. Illustrations of typical news pictures are given in Figs. 14 and 18. The transmission of timely cartoons offers another field for service, Fig. 19.

Other news-distributing agencies can also use electrically transmitted pictures to advantage. Among these are the services which make a specialty of displaying large photographs or half-tone reproductions in store windows and other prominent places. Electrically transmitted pictures of interesting events, about which newspapers have published stories, appear suited to this service, and have already been so used by some of these picture service companies. They may also be used as lantern slides for the display of news events of the day by projection either upon screens in front of newspaper offices or in moving picture theaters.

Miscellaneous commercial uses have been suggested. Photographs of samples or merchandise, of building sites, and of buildings for sale may be mentioned. The quick distribution of moving picture "stills" which is now done by aeroplane is one illustration of what may prove to be a considerable group of commercial photographs for which speedy distribution is of value.

Propagation of Electric Waves Over the Earth

By H. W. NICHOLS and J. C. SCHELLENG

SYNOPSIS: The comparatively poor transmission of radio waves of two or three hundred meters indicates some sort of selective effect in the atmosphere. Such an effect is found to result from the existence of free electrons in the atmosphere when the magnetic field of the earth is taken into account. In the earth's magnetic field, which is about one-half gauss, this selective effect will occur at a wave length of approximately 200 meters. Ionized hydrogen molecules or atoms result in resonant effects at frequencies of a few hundred cycles, this being outside of the radio range. The paper, however, takes into account the effects of ionized molecules as well as electrons.

The result of this combination is that the electric vector of a wave traveling parallel to the magnetic field is rotated. Waves traveling perpendicular to the magnetic field undergo double refraction. Critical effects are observed in rotation, bending of the wave and absorption at the resonant frequency. The paper develops the mathematical theory of these phenomena and gives formulas for the various effects to be expected.

THE problem of the propagation over the earth of electromagnetic waves such as are used in radio communication has attracted the attention of a number of investigators who have attacked the problem along somewhat different lines, with the purpose of offering an explanation of how electromagnetic waves can affect instruments at a great distance from the source in spite of the curvature of the earth. No attempt will be made here to describe adequately the various theories, but we remark that the theories of diffraction around a conducting sphere in otherwise empty space did not give satisfactory results and led to the necessity for the invention of a hypothetical conducting layer (Heaviside layer) whose aid is invoked to confine the wave between two concentric spherical shells. In many cases this Heaviside layer was considered to have the properties of a good conductor and it was supposed that a beam of short waves, for example, might be more or less regularly reflected back to the earth. The high conductivity of this layer was supposed to be due to the ionizing action of the sun or of particles invading the earth's atmosphere from outside and producing in the rarefied upper atmosphere a high degree of ionization. The differences in transmission during day and night and the variations which occur at sunrise and sunset were supposed to be due to the different ionizing effects of the sun's rays appropriate to the different times of day. The explanation of the phenomenon of "fading" or comparatively rapid fluctuations in the intensity of received signals could then be built up on the assumption of irregularities in the Heaviside layer producing either interference between waves arriving by different paths or reflection to different points on the earth's surface. The principal difficulty in this explanation is the necessity for rather high conductivity to account for the propagation of waves to great distances without large absorption.

In 1912 there appeared an article by Eccles [1] in which the bending of waves around the surface of the earth was explained on the basis of ions in the upper atmosphere which became more numerous as the vertical height increased and thereby decreased the effective dielectric constant which is a measure of the velocity of propagation of the wave. In this case the velocities at higher levels will be slightly greater than the velocities at lower levels, which will result in a bending downward of the wave normal and a consequent curvature of the wave path to conform to the curvature of the earth. In order to produce this effect without absorption the ions must be relatively free. If they suffer many collisions during the period of a wave, energy will be absorbed from the wave and pass into the thermal agitation of the molecules. Thus absorption of the wave can be computed provided the nature of the mechanism is understood thoroughly.

Sommerfeld and others have worked out the effect of the imperfect conductivity of the ground upon the wave front and such computations lead to a prediction that the electric vector in the wave near the ground will be tilted forward and thus have a horizontal component. This effect of imperfect conductivity is usually given as the cause of the large electromotive force which is induced in the so-called "wave antenna." This effect, however, apparently does not lead to an explanation of the bending of waves around the earth.

There has recently appeared an article by Larmor [2] in which the idea of a density gradient of ions or electrons is developed further to explain the bending of waves around the earth without a large absorption. This paper, as well as that of Eccles, leads to the conclusion that long radio waves will be bent around the earth, and that the effect increases as the square of the wave length, becoming vanishingly small for very short waves.

The large amount of data now available from both qualitative and quantitative observations of radio transmission shows that the phenomena may be more complicated than would be indicated by these theories. It is found that very long waves possess a considerable degree of stability and freedom from fading and that as the wave length decreases the attenuation and the magnitude of fluctuations increases until for a wave length of the order of two or three hundred

[1] Proc. Roy. Soc., June, 1912.

[2] *Phil. Mag.*, Dec., 1924.

meters there is great irregularity in transmission so that reliable communication over land for distances as short as 100 miles is not always possible even with large amounts of power. With decreasing wave length we find also variations in apparent direction of the wave. On the other hand, as the wave length is decreased still further we find, sometimes, rather surprising increases in range and stability. The nature of the fading changes, becoming more rapid, and the absorption in many cases seems to decrease. This peculiarity of wave transmission must be explained in a satisfactory theory. In addition to the apparent selective effect just mentioned, some observations indicate that there are often differences between east and west and north and south transmission at all wave lengths.

The various irregularities in radio transmission, and particularly the apparently erratic and anomalous behavior of electromagnetic waves occurring in the neighborhood of a few hundred meters wave length seem to indicate that as the wave length is decreased from a value of several kilometers to a value of a few meters some kind of selective effect occurs which changes the trend of the physical phenomena. These considerations have suggested to us the possibility of finding some selective mechanism in the earth's surface or in the atmosphere which becomes operative in the neighborhood of 200 meters. A rather superficial examination of the possibility that such a selective mechanism may be found in a possible distribution of charged particles in the atmosphere has resulted in the conclusion that a selective effect of the required kind cannot be produced by such a physical mechanism. There is, however, in the earth's atmosphere—in addition to distributions of ions—a magnetic field due to the earth, which in the presence of ions will have a disturbing effect upon an electromagnetic wave. As is well known, a free ion moving in a magnetic field has exerted upon it, due to the magnetic field, a force at right angles to its velocity and to the magnetic field. If the ion has impressed upon it a simple periodic electric force, it will execute a free oscillation together with a forced oscillation whose projection on a plane is an ellipse which is traversed in one period of the applied force. The component velocities are linear functions of the components of the electric field and at a certain frequency, depending only upon the magnetic field and the ratio $\frac{e}{m}$ of the ion, become very large unless limited by dissipation. This critical frequency is equal to $\frac{He}{2\pi mc}$ if H is measured in electromagnetic units and e in electrostatic units. It is the same as the frequency of free oscillation. For an electron in the earth's magnetic field (assumed to have a value of 1/2 gauss) this resonant frequency is 1.4×10^6 cycles, corresponding to a wave length of 214 meters.[3] We thus have an indication that some at least of the phenomena of transmission at the lower wave lengths may be explained by taking into account the action of the earth's magnetic field upon electrons present in the earth's atmosphere and acted upon by the electric field of the wave. This frequency occurs at approximately the position in the spectrum at which the peculiar effects already mentioned occur. The next resonant frequency which would be encountered would be due to the hydrogen ion which has a ratio, $\frac{e}{m}$, equal to $\frac{1}{1800}$ that of the electron. The resonant frequency of this ion is only 800 cycles and certainly can have no sharply selective effect in the propagation of electromagnetic waves over the earth. We have, therefore, worked out the consequences of the assumption that we have in the upper atmosphere two controlling factors influencing the propagation of electromagnetic waves in the radio range, namely, free electrons and ions together with the earth's magnetic field. The electrons will be dominant in their effects in the neighborhood of the resonant frequency and perhaps above, while the heavy ions will affect the wave at all frequencies and, if much more numerous, may be controlling at frequencies other than the critical one. In working out this theory it is assumed that there are present in the earth's atmosphere free electrons and ions. At high altitudes these are capable, on the average, of vibrating under the influence of the electromagnetic field through several complete oscillations before encountering other ions or neutral atoms. At low altitudes this assumption will not hold, the collisions being so numerous that the importance of the resistance term in the equations of motion becomes much greater. In either case the ions have no restoring forces of dielectric type. The motion of the electron or ion constitutes a convection current which reacts upon the electromagnetic wave and changes the velocity

[3] This frequency does not depend upon the direction of the field, and is practically constant over the earth's surface.

On March 7, after this paper had been written, the February 15 issue of the Proceedings of the Physical Society of London arrived in New York. In this journal there was a discussion on ionization in the atmosphere in which Prof. E. V. Appleton suggested, in an appendix, that the earth's magnetic field acting upon electrons would change the velocity of a wave and produce rotation. A calculation of the critical frequency was given in which, however, only the horizontal component of the earth's field was used, resulting in an incorrect value for the critical frequency, namely less than half the actual value. If the complete equations are written down it is evident at once that the total field is involved in the critical frequency, no matter what may be the direction of propagation.

of propagation of the wave. This is, in fact, the basis for the explanation of the optical properties of transparent and absorbing media and also of media which show magnetic or other rotatory powers. Due to collisions and recombinations, energy will pass continuously from the electromagnetic field and increase the energy of agitation of neutral molecules. Since this process is irreversible it accounts for absorption of energy from the wave.

Assume an electron or ion of charge e and mass m moving with velocity $\boldsymbol{v}$ and acted upon by an electric field $\boldsymbol{E}$ and the earth's magnetic field $\boldsymbol{H}$. The equation of motion of the free ion will be

$$\frac{m}{e}\dot{\boldsymbol{v}}=\boldsymbol{E}+\frac{1}{c}\boldsymbol{v}\times\boldsymbol{H}$$

or

$$a\dot{\boldsymbol{v}}=\boldsymbol{E}+\boldsymbol{v}\times\boldsymbol{h} \tag{1}$$

in which $\boldsymbol{h}$ is written for $\dfrac{\boldsymbol{H}}{c}$ and a for m/e. (When we come to consider absorption it will be necessary to generalize a into $a\left(1-i\dfrac{r}{mn}\right)$ to include a resisting force, rv, proportional to the velocity.)

The total current is given by

$$4\pi\boldsymbol{I}=\dot{\boldsymbol{E}}+\sum 4\pi Ne\,\boldsymbol{v}. \tag{2}$$

In these equations and the following we are using Gaussian units and the summation refers to different kinds of ions.

In order to avoid a complicated mathematical treatment, which, however, is not difficult to carry through if necessary, it will be assumed that the magnetic field $\boldsymbol{H}$ is along the axis of z. When more general results are required, they will be stated. All time variables are assumed periodic with a frequency $\dfrac{n}{2\pi}$, so that $\dfrac{\partial}{\partial t}=in$.

Solving equation (1) for the components of $\boldsymbol{v}$ we find, for each type of ion:

$$v_1=\frac{ina\,X+h\,Y}{h^2-a^2n^2},$$

$$v_2=\frac{-hX+ina\,Y}{h^2-a^2n^2},$$

$$v_3=\frac{Z}{ina},$$

from which it appears that a resonance frequency occurs for

$$n=\frac{h}{a}=n_o.$$

Since e/m for the electron is $-1.77c\times 10^7$, the earth's magnetic field of about 1/2 gauss will produce a resonance frequency at 1.4×10^6 corresponding to a wave length of 214 meters, while all heavier ions have resonance frequencies far outside the spectral region to be considered.

The assumption that the components of the ionic motion are simple harmonic, in spite of the fact that the motion of the ion is rather complicated, is justified as follows. From (1) we find that the velocity of an ion (r), say $\boldsymbol{v}_r$ is made up of the complementary solution, $\boldsymbol{v}_r'$ and the particular solution $\boldsymbol{v}_r''=f(\boldsymbol{E})$. The latter depends upon the impressed force $\boldsymbol{E}$, while the former has constants of integration determined by the position and motion of the ion at the last collision. The complete current is thus

$$\boldsymbol{I}=\frac{1}{4\pi}\dot{\boldsymbol{E}}+\sum e\boldsymbol{v}_r'+Nef(\boldsymbol{E}).$$

The second term, however, averages out over a large number of ions since the initial conditions are random;[4] hence, as far as the effect upon wave propagation is concerned, we may treat all quantities as periodic.

Following the usual procedure for the investigation of the propagation of waves in media of this kind, we shall rewrite equation (2) in terms of the components of the electric field, thus for each type of ion:

$$\begin{aligned}
4\pi I_1&=\left(1+\frac{\sigma N}{n_o^2-n^2}\right)\dot{X}-i\frac{\sigma N\frac{n_o}{n}}{n_o^2-n^2}\dot{Y}=\epsilon_1\dot{X}-i\alpha\dot{Y},\\
4\pi I_2&=i\frac{\sigma N\frac{n_o}{n}}{n_o^2-n^2}\dot{X}+\left(1+\frac{\sigma N}{n_o^2-n^2}\right)\dot{Y}=i\alpha\dot{X}+\epsilon_1\dot{Y}, \qquad (3)\\
4\pi I_3&=\left(1-\frac{\sigma N}{n^2}\right)\dot{Z} \qquad\qquad =\epsilon_2\dot{Z},
\end{aligned}$$

in which $\frac{4\pi e}{a}=\sigma$, or 3.2×10^9 for an electron and $3.2 \,.\, 10^9\ \frac{m}{M}$ for an ion of mass M. In order to avoid complicated formulas, the summations which must be carried in equations (3) to take account

[4] It is here assumed that the mean time between collisions is large compared to $\frac{1}{n}$.

of the effect of ions of different kinds have been omitted, but it is to be understood that the dielectric constants ϵ, α, etc., are built up from the contributions of all types of ions. Thus for an ion of mass M we must put $\sigma\frac{m}{M}$ for σ, $n_o\frac{m}{M}$ for n_o, in equations (3).

The effective dielectric constant, instead of being unity, has thus the structure:

$$(\epsilon)=\begin{pmatrix} \epsilon_1 & -i\alpha & 0 \\ i\alpha & \epsilon_1 & 0 \\ 0 & 0 & \epsilon_2 \end{pmatrix}$$

and we may write equation (2) as

$$4\pi\boldsymbol{I}=(\epsilon)\dot{\boldsymbol{E}}$$

which has the significance of the scalar equations (3). Thus $\boldsymbol{I}$ is a linear vector function of $\boldsymbol{E}$ and the operator (ϵ) is skew symmetric, indicating a rotatory effect about the axis of z.

(The general case in which h has the three components $(h_1\ h_2\ h_3)$ results in a dielectric constant having the structure

$$(\epsilon)=\begin{pmatrix} \epsilon_1 & -\beta_3-i\alpha_3 & -\beta_2+i\alpha_2 \\ -\beta_3+i\alpha_3 & \epsilon_2 & -\beta_1-i\alpha_1 \\ -\beta_2-i\alpha_2 & -\beta_1+i\alpha_1 & \epsilon_3 \end{pmatrix}$$

of which the above is a special case. With this value of (ϵ) the equation (4) below contains the *general* solution of our problem.)

Let $\boldsymbol{H}_1$ be the magnetic force associated with $\boldsymbol{E}$ in the wave so that

$$c \operatorname{curl} \boldsymbol{H}_1=(\epsilon)\,\dot{\boldsymbol{E}}$$

$$c \operatorname{curl} \boldsymbol{E}=-\dot{\boldsymbol{H}}_1.$$

Eliminating $\boldsymbol{H}_1$ from these equations we get

$$-\nabla^2\boldsymbol{E}+\nabla \operatorname{div} \boldsymbol{E}=\frac{n^2}{c^2}(\epsilon)\boldsymbol{E} \tag{4}$$

or in scalar form

$$-\nabla^2X+\frac{\partial}{\partial x} \operatorname{div} \boldsymbol{E}=\frac{n^2}{c^2}(\epsilon_1 X-i\alpha Y),$$

$$-\nabla^2 Y+\frac{\partial}{\partial y}\operatorname{div}\boldsymbol{E}=\frac{n^2}{c^2}(i\alpha X+\epsilon_1 Y), \tag{5}$$

$$-\nabla^2 Z+\frac{\partial}{\partial z}\operatorname{div}\boldsymbol{E}=\frac{n^2}{c^2}(\epsilon_2 Z).$$

These equations for the propagation of light in magnetically active substances have been given by Voigt, Lorentz, Drude and others and form the basis of the explanation of optical phenomena in such substances. As applied to optics, they are worked out, for example, in Drude's "Optics" (English translation), page 433. As applied to this problem, they assume either that the motion of the ions is unimpeded or that the resistance to the motion may be expressed as a constant times the velocity, which, as explained later. may be done in this case. We shall work out some comparatively simple cases and point out the conclusions to be drawn from them.

Consider first a plane polarized ray having its electric vector parallel to the magnetic field and moving in the xy plane; for example parallel to x. In this case the electric vector is a function of x and t only of the form

$$Z=Z_o\,\epsilon^{in\left(t-\frac{\mu x}{c}\right)}$$

in which $\frac{c}{\mu}$ is the velocity of the wave. Substituting in the general equations (5) we find that

$$\mu^2=1-\Sigma\frac{\sigma_i N_i}{n^2}. \tag{6}$$

The velocity of propagation is thus a function of the frequency and of the density N. This particular case corresponds to that treated by Eccles and Larmor in the papers cited. It will be noted that the velocity is greater for long waves than for short waves and that if N is a function of distance from the surface of the earth, the velocity will vary in a vertical direction, causing a curvature of the rays as worked out by the authors mentioned. In this particular case, however, which corresponds completely in practice to conditions obtaining over only a limited area of the earth's surface, the greatest effect is produced on the longer waves. Since electromagnetic waves are in general radiated from vertical antennas so that the electric vector is vertical, this case would correspond to the condition of transmitting across the north or south magnetic poles of the earth.

The second case to be considered is that of propagation along the direction of the magnetic field. In this case X and Y are functions of z and t and the appropriate solutions of the fundamental equations (5) are

$$X' = A \cos n\left(t - \frac{\mu_1 z}{c}\right),$$

$$Y' = -A \sin n\left(t - \frac{\mu_1 z}{c}\right), \quad \mu_1^2 = \epsilon_1 + \alpha,$$

$$X'' = A \cos n\left(t - \frac{\mu_2 z}{c}\right),$$

$$Y'' = A \sin n\left(t - \frac{\mu_2 z}{c}\right), \quad \mu_2^2 = \epsilon_1 - \alpha.$$

which represent two oppositely circularly polarized components traveling with the different velocities $\frac{c}{\mu_1}$ and $\frac{c}{\mu_2}$. The plane of polarization is rotated through an angle of 2π in a distance given by

$$\frac{z_o}{\lambda} = \frac{\epsilon_1}{\alpha}.$$

The third case to be considered is that of propagation at right angles to the magnetic field, say in the direction of x. For this case equations (5) become:

$$X = \frac{i\alpha}{\epsilon_1} Y$$

$$-\frac{c^2}{n^2}\nabla^2 Y = \left(\epsilon_1 - \frac{\alpha^2}{\epsilon_1}\right) Y$$

$$-\frac{c^2}{n^2}\nabla^2 Z = \epsilon_2 Z,$$

of which the solutions are

$$X = \frac{i\alpha}{\epsilon_1} Y_o \epsilon^{in\left(t - \frac{\mu_1 x}{c}\right)}$$

$$Y = Y_o \epsilon^{in\left(t - \frac{\mu_1 x}{c}\right)} \qquad \mu_1^2 = \epsilon_1 - \frac{\alpha^2}{\epsilon_1}.$$

$$Z = Z_o \epsilon^{in\left(t - \frac{\mu_2 x}{c}\right)}.$$

The first of these is merely the (usually small) component of field required to make the total current solenoidal, that is, to balance the convection of electrons. The last two show that the plane polarized ray whose electric vector is parallel to H will travel with the velocity $\frac{c}{\mu_2}$ while the one whose electric vector is at right angles to this direction and to the direction of propagation will travel at a different speed, $\frac{c}{\mu_1}$. There is thus double refraction.

Bending of the rays. If μ is the index of refraction, which is a function of the space variables, the curvature of the ray having this index is $\frac{1}{\mu}\frac{d\mu}{ds}$ where s is taken perpendicular to the direction of the ray. Since μ is practically unity except at the critical frequency, this curvature is $1/2\; d\,\mu^2/ds$. In order that the ray should follow the curvature of the earth it is clear that μ must decrease at higher altitudes; that is, $\frac{d\mu^2}{ds}$ must be negative.

We shall work out the curvatures for the special cases considered. (The first case has been given above and was worked out in the papers cited). For the case of propagation along H, the two circularly polarized beams have indices given by

$$\mu_1^2=\epsilon_1+\alpha=1+\frac{\sigma N}{n^2}\,\frac{1}{\omega-1}, \tag{7}$$

$$\mu_2^2=\epsilon_1-\alpha=1-\frac{\sigma N}{n^2}\,\frac{1}{\omega+1}, \tag{8}$$

$$\left(\omega=\frac{n_o}{n}\right).$$

We are interested in the values of $1/2\,\frac{d\mu^2}{ds}$ in which N and h are functions of distance s and also of the time. These come out to be

$$C_1=\frac{\sigma}{2n_o^2}\left[\frac{\omega^2}{\omega-1}\,\frac{dN}{ds}-\frac{\omega^3}{(\omega-1)^2}\,\frac{N}{h}\,\frac{dh}{ds}\right], \tag{9}$$

$$C_2=\frac{\sigma}{2n_o^2}\left[\frac{-\omega^2}{\omega+1}\,\frac{dN}{ds}+\frac{\omega^3}{(\omega+1)^2}\,\frac{N}{h}\,\frac{dh}{ds}\right]. \tag{10}$$

A striking fact shown by these formulae is that the curvatures of the two rays are in general different. A limited beam entering an ionized medium along a magnetic meridian will be split into two which will traverse different paths. Thus we should expect to find, occasionally, a circularly polarized beam at the receiver due to the fact that the receiving instrument is located at a point toward which one of the beams is diverted after having passed through an upper ionized layer. This is now being investigated experimentally. It is clear that, although the two components do not in general travel over the same path, both may eventually arrive at the same receiver. The first ray, however, may have penetrated much higher in the atmosphere than the other, that is, to a level at which $\frac{dN}{ds}$ has the proper negative value to cause it to return to earth.

For long waves, these curvatures become:

$$C_1 = \frac{\sigma\omega}{2n_o^2}\left[+\frac{dN}{ds} - \frac{N}{h}\frac{dh}{ds}\right], \tag{11}$$

$$C_2 = \frac{\sigma\omega}{2n_o^2}\left[-\frac{dN}{ds} + \frac{N}{h}\frac{dh}{ds}\right]. \tag{12}$$

Hence a limited beam of long waves entering this medium would tend to split into two of opposite polarization and traverse different paths. In the special case for which $\frac{1}{N}\frac{dN}{ds} = \frac{1}{h}\frac{dh}{ds}$ throughout the medium, there will be no such separation of the beam.

For very short waves

$$C_1 = \frac{\sigma}{2n_o^2}\left[-\omega^2\frac{dN}{ds} - \omega^3\frac{N}{h}\frac{dh}{ds}\right], \tag{13}$$

$$C_2 = \frac{\sigma}{2n_o^2}\left[-\omega^2\frac{dN}{ds} + \omega^3\frac{N}{h}\frac{dh}{ds}\right]. \tag{14}$$

Hence if the most effective cause of refraction is the variation in the ionic density both components tend to remain together and to travel with a rotation of the plane of polarization. If variation in the magnetic field is appreciable the two components tend to diverge as in the case of long waves.

For propagation at right angles to H, say along x, we have

$$\mu_1^2 = \epsilon_2 = 1 - \frac{\sigma N}{n^2}, \tag{15}$$

$$\mu_2^2 = \epsilon_1 - \frac{\alpha^2}{\epsilon_1}. \tag{16}$$

The bending of the plane polarized component having the index μ_1 shows no selective effects, being simply

$$C_1 = -\frac{\sigma}{2n^2}\frac{dN}{ds} \tag{17}$$

and is appreciable only for long waves unless N is very large. For the other component we find:

$$C_2 = \frac{\sigma}{2n_o^2}\cdot\frac{\omega^2}{\omega^2-1}\,\frac{1-\dfrac{2\sigma N}{n_o^2}\omega^2-\dfrac{\sigma^2N^2}{n_o^4}\,\dfrac{\omega^2}{\omega^2-1}}{\left(1+\dfrac{\sigma N}{n_o^2}\,\dfrac{\omega^2}{\omega^2-1}\right)^2}\,\frac{dN}{ds} \tag{18}$$

where, in order to simplify the formula, only the term containing $\frac{dN}{ds}$ has been included. This applies to ions of one kind.

For long waves these two curvatures become

$$C_1 = -\frac{\sigma}{2n_o^2}\omega^2\frac{dN}{ds}, \tag{19}$$

$$C_2 = \frac{\sigma}{2n_o^2}\left(1-\frac{2\sigma N}{n_o^2}\omega^2\right)\frac{dN}{ds}. \tag{20}$$

These formulas show that the first curvature is always in the same direction for a given value of $\frac{dN}{ds}$, while the second curvature, which is that of the electric vector perpendicular to the magnetic field, is, for very long waves, in the same direction as C_1 but, as the wave length is decreased or N increased, reverses in sign and becomes opposite to C_1. As an example, if $N=10$, for 6 kilometer waves the curvatures are opposite, so that if the first component tends to bend downward the second will tend to bend upward; while if $N=100$, for the same wave length both curvatures have the same sign and the second is five times as large as the first.

For extremely short waves the two curvatures are equal as they obviously should be, since the magnetic field can then have no effect.

In transmitting from New York to London, for example, waves travel approximately at right angles to the magnetic field, which in this latitude has a dip of about 70°. If we assume a plane polarized ray starting out with its electric vector vertical, the component parallel to the magnetic field will be the larger and will be subject to the curvature C_1 above, while the smaller component will be affected by the magnetic field and will have the curvature C_2. The two components into which the original wave is resolved will travel with different velocities. It is clear that when the distribution of ions in the upper atmosphere is changed by varying sunlight conditions, the resulting effect at a receiver is likely to vary considerably. Some of the possibilities will be discussed later.

Rotation of the plane of polarization. It has been shown that in the second case, namely transmission along the magnetic field, there will be a rotation of the plane of polarization of the wave. This rotation is such that the wave is rotated through a complete turn in a distance given by

$$z_o = \frac{2\pi c}{n_o} \frac{1 + \frac{\sigma N}{n_o^2} \frac{\omega^2}{\omega^2 - 1}}{\frac{\sigma N}{n_o^2} \frac{\omega^2}{\omega^2 - 1}}. \tag{21}$$

It is interesting to note that the distance in which a long wave rotates through $2\,\pi$ approaches the constant value $\frac{2\pi c n_o}{\sigma N}$ as the wave length increases and that for very short waves the rotation of the plane of polarization tends to vanish with the wave length.

Absorption. When an electron strikes a massive neutral atom the average persistence of velocities is negligible and in the steady state of motion of electrons and neutral molecules the element of convection current represented by an impinging electron will be neutralized, so far as the wave is concerned, at every collision. Of the energy which has been put into this element of convection current since the last collision, a part will be spent in accelerating neutral molecules, part will go to increase the average random velocity of the electron and a part will appear as *disordered* electromagnetic radiation. Thus, as far as the wave is concerned, the process of collision with massive neutral molecules is irreversible even if the molecules are elastic, and all the energy picked up by the electron from the wave between collisions is taken from the wave at the next collision. Exactly the same state of affairs would exist if at each collision the electron recombined with a molecule and a new electron were created with zero or random velocity. Thus for massive molecules for which we can neglect the persistence of electron velocities the effect upon the wave is exactly the same whether the collision is elastic or inelastic.

These conclusions are verified by the results of two different computations which we have made of the resistance term, rv, in equation of motion of the electron. Consider in the first place a mixture of electrons and massive neutral molecules, assumed perfectly elastic, in which the persistence of velocities of the electrons after collision is negligible. If an electric field $X\epsilon^{\text{int}}$ operates in the x direction and if the state of motion is a steady one, we can compute the energy w taken from the wave by a single electron at any time after a collision at the time t_1 and before the next collision. Let this time after t_1 be τ. If the mean frequency of collisions is f, the time τ between collisions will be distributed according to the law

$$f\epsilon^{-f\tau}$$

and we shall obtain the mean energy taken from the wave per collision by multiplying w by the above expression, integrating from zero to infinity with respect to τ and then performing an average over all the times t_1. The result of this is that the mean energy loss per collision is simply

$$w=\frac{e^2X^2}{2mn^2}\,\frac{n^2}{f^2+n^2}$$

and consequently the loss per second is f times this. If we equate this to rv^2, which is also the rate at which energy is being dissipated, we find that $r=mf$, which is therefore the resistance term to be inserted in the equation of motion of the electron.

If the convection current is carried partly by heavier ions, it will not be annulled at each collision and all the energy derived from the field will not be lost on impact.

The foregoing computation assumes as obvious that energy is lost from the wave at a rate equal to the number of collisions times the average energy which the electron takes from the wave between collisions. The second method is somewhat more general. The mean velocity at a time t is found for electrons which collided last in an interval at t_1. This is evidently a function of the velocity persisting through the last collision and hence of the average velocity before the impact; so that if the average velocity before collision was v, that after impact would be $\delta\, v$, in which δ is a number less than unity, depending on the relative masses and the nature of the collision. Averaging for all values of t_1 before t and using the same law of distribution assumed above, the mean velocity of the ions since the last collision is obtained. By comparison with the solution obtained for the velocity of forced oscillation in which the resistive force is rv, we find that $r=mf(1-\delta)$. For the special case of electrons, δ may be taken equal to zero, hence $r=mf$. For the case of very heavy ions colliding with light neutral molecules, $r=o$, since $\delta=1$. For equal masses δ would be about one half, hence $r=\frac{1}{2}\, mf$.

Since the resistance factor r is equal to mf, in order to include the effect of attenuation of the wave, we must replace a by

$$a\left(1-i\frac{f}{n}\right).$$

If, as usual, we assume a wave proportional to

$$\epsilon^{\frac{-nk\mu x}{c}} \quad \epsilon^{in\left(t-\frac{\mu x}{c}\right)}$$

the equations (5) show that, in order to calculate the value of the absorption constant k, we must put

$$\mu^2(1-ik)^2=\epsilon$$

in which ϵ is the generalized dielectric constant appropriate to the case considered. We have worked out in this way the absorption for the various cases treated above with the following results.

In the case in which there is either no magnetic field or the magnetic field is parallel to the direction of the electric vector, we find

$$k=\frac{\sigma N}{2n_o^2}\omega^2 \; \frac{f/n}{1+f^2/n^2}.$$

This formula for absorption applies (for electrons) for any value of f or n. Thus near the surface of the earth where the collision frequency f is of the order of 10^9, the fraction $\frac{f}{n}$ will be large even for rather short waves. As we go higher in the atmosphere this ratio decreases for a given wave frequency until at a height for which $\frac{f}{n}=1$ we encounter the maximum absorption *per electron*. Above this level $\frac{f}{n}$ and consequently the absorption per electron decreases. For ions other than electrons the resistance will be somewhat different from mf, depending upon the ratio of the masses, and a corresponding change must be made in the above statement.

In this paper we are considering only the effects which take place at heights above that for maximum absorption so that, generally speaking, $\frac{f}{n}$ will be small or at least less than unity. This approximation will be used in computing the absorption constants which follow.

As an example of the nature of this approximation, at a height of about 100 kilometers, we may expect an atmospheric pressure of 10^{-5} standard and a corresponding collision frequency of the order of 10^5. Thus for very long waves of frequency 40,000 cycles per second we still have $\frac{f}{n} = .4$, while at the critical frequency $\frac{f}{n}$ is only 1/100.

The computation of the collision frequency for electrons is rather involved because of the peculiar nature which such a collision may have and because it probably is not permissible to assume thermal equilibrium with the molecules of the gas. The processes of ionization and recombination will also lead to complications. Probably the most significant information would be the number of electron free paths per second for unit volume.

The question of the behavior of waves in or below the layer of maximum absorption per ion is a somewhat different one and belongs properly in another paper.

For the case of transmission along a magnetic meridian the oppositely circularly polarized rays have the absorption constants:

$$k_1 = \frac{\sigma N}{2n_o^2} \frac{\omega^2 f/n}{(\omega-1)^2+(f/n)^2}, \qquad k_2 = \frac{\sigma N}{2n_o^2} \frac{\omega^2}{(\omega+1)^2} \frac{f}{n}.$$

It will be noted that, at the critical frequency, the first of these waves has the high absorption $\frac{\sigma N}{2n_o^2} \cdot \frac{n}{f}$ and is therefore extinguished in a short distance, while the other wave has a normal absorption constant $\frac{\sigma N}{8n_o^2} \cdot \frac{f}{n}$. Thus for the case of transmission along a meridian at the critical frequency we might expect a receiving station, sufficiently far above the ground, to receive a circularly polarized beam. This would mean that if a loop were used for reception, the intensity of the received signal would be independent of the angle of setting of the loop, provided one diameter of the loop was set parallel to the direction of propagation of the wave. In general, of course, this ideal condition could not be realized due to the disturbing action of the ground and of other conducting or refracting bodies and the most we should expect to receive in practice would be an elliptically polarized beam.

In the third case, namely, that of propagation perpendicular to the direction of the magnetic field, we find that the wave polarized with its electric vector parallel to the magnetic field has the same absorption as before, namely $\frac{2n_o^2}{\sigma N}\,\omega^2\,\frac{f}{n}$ and the other ray whose complex index of refraction is $\epsilon_1 - \frac{\alpha^2}{\epsilon_1}$ has the absorption constant $\frac{1}{2}(k_1+k_2)$ in which k_1 and k_2 are the absorption constants given above for propagation along a magnetic meridian.

At the critical frequency we find, therefore, that the absorption constant is abnormally high and equal to $\frac{\sigma N}{4n_o^2}\cdot\frac{n}{f}$ which is one-half that obtained for the first ray of case 2.

One very striking fact is brought to light by these equations. Thus, referring to the two values of absorption constants for transmission along the magnetic field, we find that for very long waves (for which ω is large) the ionic absorption is very much less with a magnetic field present than without it. This means that in this case and in the next the presence of a magnetic field *assists* in the propagation of an electromagnetic wave by decreasing the absorption. This reduction in absorption may amount to a rather large amount, as may be seen from an inspection of the formula for k_1. For example, if in this case ω is 20, corresponding to 4,000 meter waves, we find that under corresponding conditions the absorption *due to electrons only* is reduced by the magnetic field to 1/400th the value it would have for no magnetic field. Of course, these cases are not directly comparable because the path chosen by the wave would be different in the two cases. It is plausible, however, that the propagation of long waves along the magnetic field may go on with much less attenuation than propagation from East to West over a region in which the magnetic field is nearly vertical, in which case the effect of the magnetic field is largely absent. This conclusion, however, cannot be made in general since a number of other causes are influential in determining the propagation, for example, the bending of the rays, so that it is not certain that transmission over a region in which the magnetic field is vertical is always more difficult than in the other cases.

The reason for the decreased absorption of long waves when the magnetic field can operate (that is, in all cases in which the electric vector is not parallel to the field) is that the velocities acquired by the free electrons are much less for small values of n when the magnetic field is present.

Fading. By this is meant a variation with time of the strength of a received signal at a given point. It is clear that a wave starting originally with constant amplitude and frequency can be received as one of variable amplitude only if certain characteristics of the medium are variable with the time. So far as the atmosphere is concerned, these characteristics may be the distribution of electrons and heavier ions and the intensity and direction of the earth's magnetic field. If these are functions of the time, the velocities, bending, absorption and rotation of the plane of polarization will all be variable, the amplitude of variation depending upon the variations of N, $\frac{dN}{ds}$, H, $\frac{dH}{ds}$, as well as the frequency of the wave, the effects being in many cases magnified greatly in the neighborhood of the critical frequency. These effects are obviously sufficiently numerous to account for fading of almost any character and suggest a number of experiments to determine the most effective causes. The question of rotation of the plane of polarization, fading and distortion is now being examined experimentally.

From the formulas it is clear that the velocity, curvature and absorption of an electromagnetic wave as well as the rotation of its plane of polarization can all be affected by a time variation in the intensity and direction of the earth's field. An examination of the probable time and space variations of each, however, lead us to the conclusion that these are not of primary importance in determining large amplitude fading except, perhaps, during magnetic storms. One result of the last two years of consistent testing between New York and London at about 60,000 cycles has shown that severe magnetic storms are always accompanied by corresponding variations in the strength of received signals. Thus, although the earth's magnetic field can well exercise a large influence upon the course and attenuation of radio waves, it does not seem likely that its time variation is ordinarily a large contributing cause to fading.

This leaves as the probable principal cause of time variations the number and distribution of ions in the earth's atmosphere. It is impossible in this paper, which is devoted primarily to a development of a theory of transmission involving the earth's magnetic field, to consider adequately all the possibilities resulting from changes in ionic distributions, but some general remarks may be made. Imagine a wave traveling from the source to the receiver. At a short distance from the source the wave front will be more or less regular but as it progresses, due to the irregularities in ionic distribution, the wave front will develop crinkles which become exaggerated as the wave goes on. These crinkles in the wave front will be due to irregularities in the medium and can be obtained by a Huyghen's construction at any point. If we consider the wave a short distance before it reaches the receiver, we will find regions in which the wave front is concave to the receiver and regions of opposite curvature. Thus at certain portions of the wave front energy will be concentrated toward a point farther on and at other parts will be scattered. The location of these convex or concave portions of the wave in the neighborhood of a given receiving point will be very sensitive to changes in ionic distribution along all the paths of the elementary rays contributing to the effect at the receiver. Hence, if we knew the location and movement of all the ions between the transmitter and the receiver, it would be possible, theoretically, to predict the resultant effect at the latter point.

To explain fading it is essential that there be a time variation in this distribution. It is clear that effects of this kind should be more marked at short waves than at long waves since a region of the medium comparable in dimensions to a wave length must suffer some change in order to produce an effect upon the received signal. If, for example, there were space irregularities in the medium comparable to the wave length, a kind of diffraction effect would be produced at the receiver which would be very sensitive to slight changes in grating space.

A possible cause of irregularity may be found in the passage across the atmosphere of long waves of condensation and rarefaction, each of which results in a change in the density and gradient of the ions, even though the average density remains constant throughout a large volume. If, as seems plausible, the upper atmosphere is traversed by many such atmospheric waves of great wave length, the resulting effect at a given receiving point would be fluctuations in signal strength due to a more or less rapid change in the configuration of the wave front near the receiver.

For radio waves whose length is of the order of a few hundred meters, fading experimentally observed occurs at a rate of the order of one per minute (of course, it is not implied by this statement that there is any regular periodicity to the fading). The pressure wave referred to would travel in the upper atmosphere with a velocity of the order of 300 meters per second at lower levels or 1,000 meters in the hydrogen atmosphere, so that the wave length of these "sound" waves would be of the order of 50 of the radio wave lengths. The irregularities of the medium would thus be of sufficient dimensions with respect to the electromagnetic waves so that one of the characteristics referred to above might be developed. In this way we might explain variations in intensity of the wave at the receiver recurring at intervals of a minute or so.

These effects, of course, might be produced even without a magnetic field but the results of this paper indicate that conditions in the wave front will be complicated still further by a rotation of the electric vector and by the existence of bending and double refraction due to the magnetic field, these effects being exaggerated in the neighborhood of the critical frequency. Due to the magnetic field we have also the possibility of summation effects between components of the wave which were split off by the action of the field and consequently had traveled by different paths at different speeds. It is obviously impossible to make any general statement concerning the nature of the effects which will be produced by this complicated array of causes but future experimental work will, we hope, allow us to estimate the relative importance of the various elements.

Open Tank Creosoting Plants for Treating Chestnut Poles

By T. C. SMITH

Introduction

FOR a number of years chestnut timber, because of its many desirable characteristics, has served a broad field of usefulness in telephone line construction work, not only in its native territory, the eastern and southeastern part of the United States, but also in neighboring states. In fact, as an average, about 200,000 chestnut poles are set annually in the Bell System plant as replacements and in new lines.

In areas which are gradually being extended from the northern part of the chestnut growing territory into the southern sections, blight is rapidly making serious inroads into this class of pole timber. North of the Potomac River practically all chestnut territories have been visited by the blight and it has in a major sense crossed into areas south and southwest of this river, where it is developing from scattered spots. While many poles are yet secured in the blighted areas, they must be cut within a very few years after becoming affected, in order to save them from the decay which destroys blighted poles after they are killed.

A chestnut pole lasts satisfactorily above the ground line but decays at and within a few inches below the ground, thus weakening it at a critical location. In order to protect the poles from decay at this location, the open tank creosote treatment seems to be the most satisfactory, where the facilities for applying the treatment are available. In general this treatment consists of standing the poles in an open tank and treating them in a creosote bath which covers them from the butt ends to a point about one foot above what will be the ground line when the poles are set. The method of applying the treatment will be explained in more detail further along in the paper.

Due to the scattered locations of the chestnut timber and also to the fact that in many places this timber is rapidly being depleted by the blight, it has required considerable study to establish locations for open tank treating plants which would be convenient for applying the treatments and would also have a sufficient available pole supply to permit the operation of the plants long enough to warrant the necessary investment in them. However, suitable locations have been established and plants have been constructed which will, when operating to their planned capacities, treat about 139,000 chestnut poles per year, and these plants may easily be enlarged to treat additional quantities as the demand for treated poles develops.

These plants have been designed by our engineers and are being operated for applying preservative treatments to poles used by the Bell System.

Locating the Treating Plants

It might be interesting to bring out the governing considerations in locating the chestnut open tank treating plants, as compared with commercial plants for treating cedar poles, which are operating in the north central and northwest portions of the Uhited States. Due to the geographical locations in which the cedar poles grow, in relation to the centers of distribution en route to the locations where they will be used, treating plants of large capacities can be supplied for many years with poles which pass them in the normal course of transporting the poles from the timber to their destinations. Commercial pole treating companies seem to have had no difficulty in establishing locations for handling 100,000 or more cedar poles per year through a single plant; whereas the scattered locations of the chestnut poles, as outlined above, make it more economical to build the chestnut treating plants in units varying between 10,000 and 36,000 poles per year capacity.

Several factors were considered in determining the proper locations for the seven Bell System treating plants which have been built. It was often possible to select a location which was admirably adapted to the purpose when considered from two or three viewpoints but which was found undesirable when considered from all of the necessary angles. The principal points considered were:

1. Quantity of poles of the desired sizes available locally which could be delivered to a proposed plant by wagons, motor vehicles, etc.

2. Quantity of poles which could be conveniently routed past the plant during the rail shipments from the timber to their destinations.

3. Quality of the available timber.

4. The length of time during which a plant of the desired size could be supplied with timber for treatment. This estimated figure would, of course, determine the length of life of the proposed plant.

5. Railroad facilities and freight distances from the proposed plant to points where the poles would be used.

6. Availability of labor for operating the plant.

7. Locating a suitable site for the plant.

Experience of the Western Electric Company's Purchasing Department and the local Associated Telephone Company representatives, together with information from Government reports, provided the

Fig. 1—Land upon which Sylva Plant was Built

answers to the first five items. Studies upon the ground were made to settle the remaining two items after a preliminary survey of the situation had indicated what locations seemed to warrant consideration.

The unevenness of the land as shown by Fig. 1, which is typical of the many available locations studied, made it difficult to secure a comparatively level tract of the proper area and dimensions adjoining a railroad siding or at a location where a siding could conveniently be established. In fact it soon became evident in making the preliminary studies, that it would be necessary to design the various treating plants to fit the best of the available tracts.

As a result of these studies, seven plants were established and placed in operation in five states as outlined below:

Location	Date when Plant Was Placed in Operation	Annual Pole Capacity Now	Total Annual Pole Capacity When Additions Now Planned Are Completed
Shipman, Va...........	Oct. 1922	10,000	15,000
Danbury, Conn.........	Dec. 1922	10,000	10,000
Natural Bridge, Va......	Apr. 1923	10,000	18,000
Willimantic, Conn......	Aug. 1923	10,000	10,000
Sylva, N. C............	May 1924	18,000	25,000
Nashville, Tenn........	July 1924	18,000	25,000
Ceredo, W. Va..........	Sept. 1924	23,000	36,000
Totals.............		99,000	139,000

It will be noted from the above table that several of the plants are not yet working to their capacities as now planned. In designing the plants, the plans were made to provide for the total annual capacities shown above. However, when they were built the initial capacities were made somewhat lower as indicated by the table, by omitting in some cases tanks and in other cases pole handling equipment which could readily be added in conformity with the plans, later when the additional capacities would be required.

Yard Sizes

It might not seem necessary to occupy a very great area in the operation of a pole treating plant. However, experience with some of the earlier plants indicated that a reasonably large yard was very desirable because of the number of poles necessarily carried in piles on skids in the yard both in the untreated stock and in the treated stock. In so far as practicable the poles in the various treating plants are arranged in such a manner that each length and class is piled separately. This greatly facilitates handling the poles, but requires considerable space. Ordinarily about 80 pole piles are necessary in a yard.

From four to ten acres of land has been used for each of the various pole treating yards. Fig. 2, which includes about half of a comparatively small capacity yard, shows the necessity for plenty of room for the pole piles.

Yard Layouts

Since the pole treating yard layouts are necessarily built around the railroad sidings which handle the poles in and out of the yards and transfer them from one location to another inside the yards, it is desirable to build the yards long and narrow.

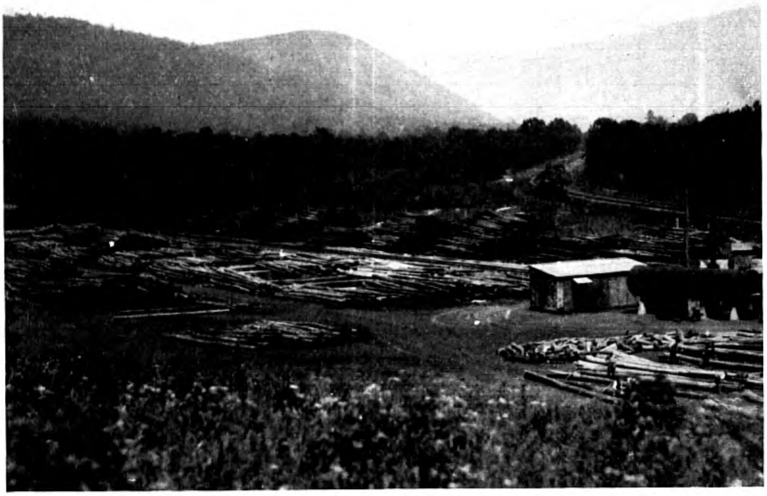


Fig. 2—Portion of Pole Yard at One of the Smaller Plants. (Tool House and Creosote Storage Tank at Right)

Of course, the sharper the railroad curves can be made in laying out a siding from the railroad into the pole treating yard, the easier it is to accommodate the siding to cramped yard conditions or to spread out the tracks over a short, wide yard. However, due to the use of heavy locomotives on the main lines and the desirability of having switch curves suitable for the locomotives ordinarily used, it has been necessary to use 12 degree railroad curves in planning most of the yard entrances, and in no case has a curve been used which is sharper than 18 degrees.

It will be noted from Fig. 3 that the pole treating apparatus is so located that the work of handling poles to and from the treating tanks will not interfere in any way with loading outgoing cars of treated poles from the skids. It will also be noted that the poles which are received from the river are treated during the natural course of their passage to the "treated" skids.

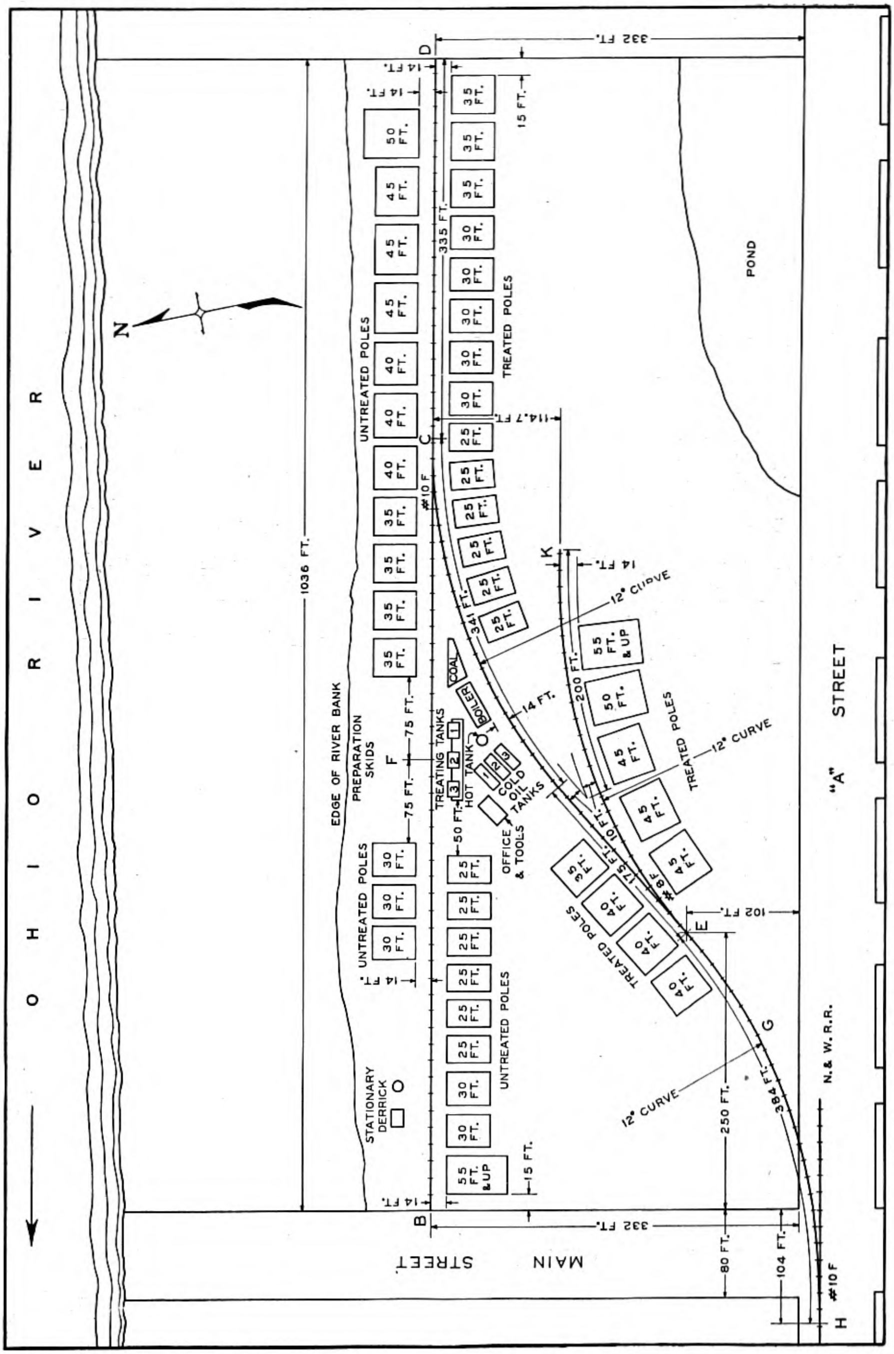


Fig. 3—Plan of Yard Layout for Ceredo. (Each Skid Shown is Separated into Two Pole Piles)

Car loads of poles which are received by rail may be backed into the track leading to the pole treating plant for treatment or may be unloaded upon the "untreated" skids if desired. In any event, there should be a minimum of confusion in the pole moving operations.

Fig. 4 shows the skids at one end of the Sylva yard before poles had been piled upon them. It illustrates the desirability of having a

Fig. 4—Skid Layout at One End of Sylva Yard

long, narrow yard and also shows that the switch track is the backbone of the pole yard.

It will also be noted from Fig. 4 that in the Sylva yard the ends of the skids are brought up close to the track. This is because the pole handling in the Sylva yard is done by means of a locomotive crane which runs on the track and works from the ends of the cars.

In the Natural Bridge yard, which is shown in Fig. 5, a tractor crane is used for pole handling. This unit has crawlers and wheels which operate on the narrow roadways at either side of the spur tracks. The tractor crane runs up to the side of a car to unload it. By operating at the sides of the cars a much shorter boom is required by the tractor crane than for the locomotive crane working at the ends of the cars handling the same lengths of poles.

Delivery of Poles to Plants

Various methods are used for delivering poles to the treating plants, from the locations where they are cut. In addition to the use of automobile trucks with their trailers, and to the use of horse-drawn wagons which may be seen along the road in Fig. 4, poles are delivered by railroad cars, river rafts and ox-teams.

In the timber the poles are ordinarily loaded on cars for shipment to the treating plants by means of a logging loader shown in Fig. 6.

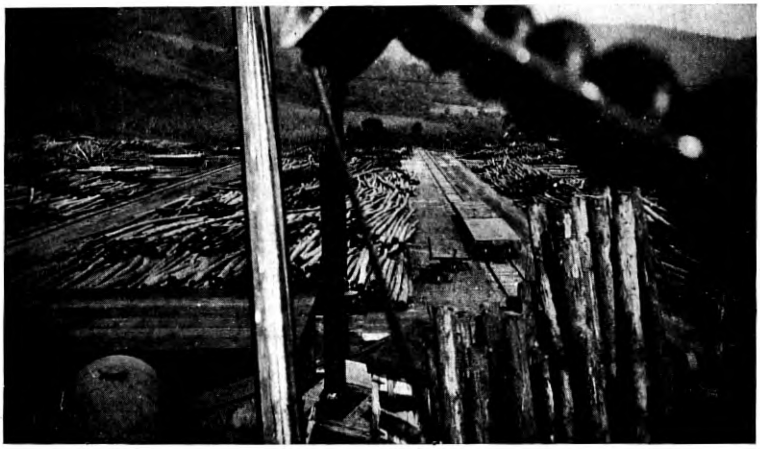


Fig. 5—Yard Layout at One End of Natural Bridge Yard, Viewed from Mast of Derrick

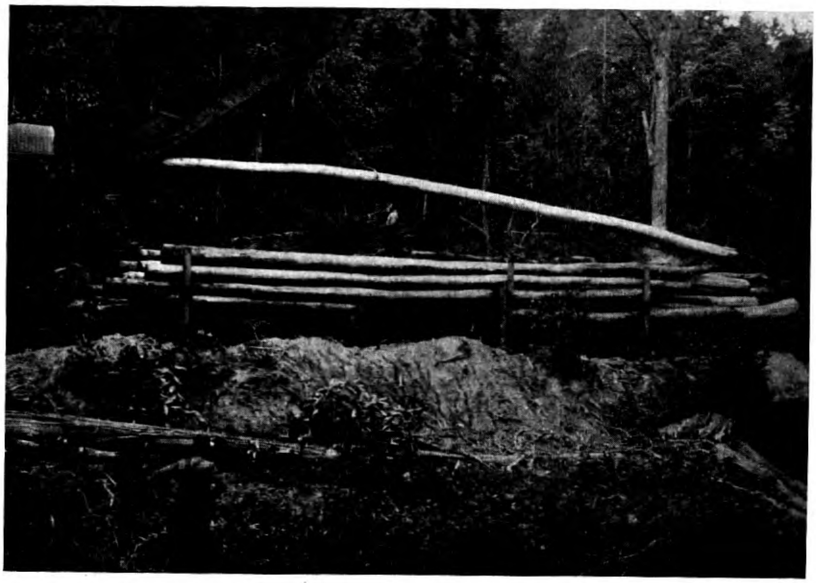


Fig. 6—Placing Poles on Logging Car by Means of Logging Loader

Although it has a short boom, it is able to handle very long poles because of the method in which it lifts them. One end of the pole, either top or butt, is rested against the middle point of the boom and the pole lifted by the winch line which may be attached only one-third or one-fourth of the distance from the loader end to the

Fig. 7—Geared Locomotive in Use on Logging Road Which Supplies Poles to Treating Plant

free end of the pole. In lifting long poles by this method, they spring considerably, and brash timber usually breaks under this treatment. Thus in handling poles by this method, they are given a test before they leave the timber.

The winch line is attached to the pole by means of hooks which resemble ice tongs. From long experience in handling these tongs, the pole men are able to throw them several feet and catch a pole at any point they desire, to pull it from the pole pile. This operation is very fast. In fact, under favorable conditions, 35 foot chestnut poles have been loaded on a car at the rate of two per minute.

The pole piles along the logging road are usually disorderly, resembling a lot of giant tooth-picks which might have been carelessly dropped in a heap.

Steep grades on the logging roads make it very desirable to use locomotives which have a maximum amount of traction. For this reason, a geared type locomotive is used which permits a big reduction between the engine and drive wheels, and also transmits the driving torque to all wheels of the engine and coal tender which is shown, and also to the wheels of the water tender which is not shown in Fig. 7.

From one to ten car loads of poles in a group arrive at the treating plants. A car load varies between 40 and 65 poles depending upon

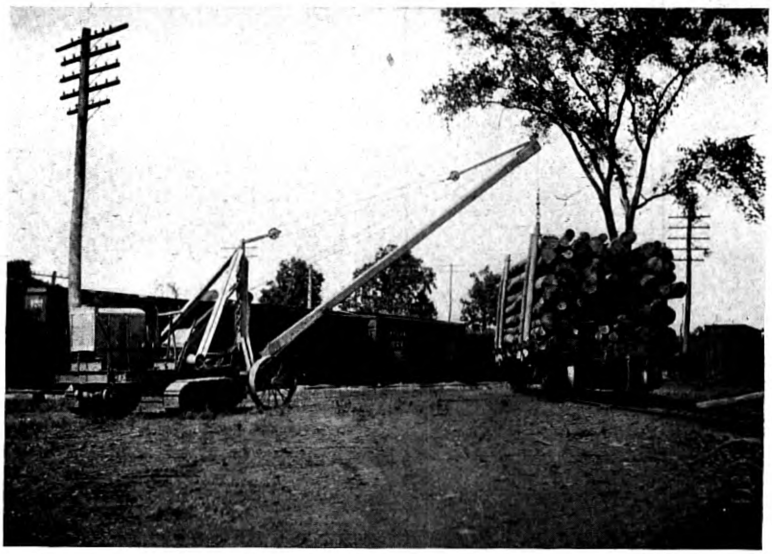


Fig. 8—Car Load of Poles Arriving at the Danbury Treating Plant

the sizes of the poles. They may be unloaded by a locomotive crane or a tractor crane or by the method shown in Fig. 9.

At the Shipman Yard the poles are unloaded by cutting the stakes and permitting the poles to roll down an embankment into piles from which they are drawn to the treating plant by means of a steel rope from a tractor winch.

Utilization of the cheapest method of delivering poles to the treating plants is possible at Ceredo and Nashville where the plants are located on the river banks. These poles are securely tied in rafts of about 100 poles each and either floated down the rivers or handled by stern wheel, river steamboats.

Fig. 9—Unloading Poles at the Shipman Yard

Fig. 10—Four Rafts of Poles at Ceredo Plant

It may be of interest to note that the photograph shown in Fig. 10 was taken from the West Virginia bank of the river, while the Ohio bank is seen across the river and the Kentucky hills are visible beyond the bridge.

Particularly in the Carolinas, ox-teams are used to draw pole loads down from the mountains.

Fig. 11—Pole Delivery by Ox-Teams

Fig. 12—Derrick for Handling Poles from River Rafts to Piles or Pole Cars in the Yard

HANDLING POLES IN THE YARD

Where the derrick is used for lifting poles out of the river it is necessary to set it at a distance from the water's edge which, of course, approaches and recedes depending upon the height of the river. Because of this distance, the poles are dragged as well as lifted up the sloping side of the bank.

Fig. 13—Handling Poles by Man Power

Fig. 14—Tractor Crane Handling Poles from Rail Dollies in Danbury Yard

It has been found that wherever it is possible to eliminate the handling of poles by man-power, a considerable economy can be realized. Less men are required for crane or derrick operation, and the cranes and derricks do the work much more rapidly.

In order to move the poles about the yard it is not necessary to retain a freight car to carry them, since small rail dollies have been provided for this purpose. The two dollies shown in Fig. 14 are separate and can be located under the poles at any distance apart depending upon the lengths of the poles.

The tractor crane which is used for pole handling in the smaller plants is operated by a heavy duty gasoline engine and it is able

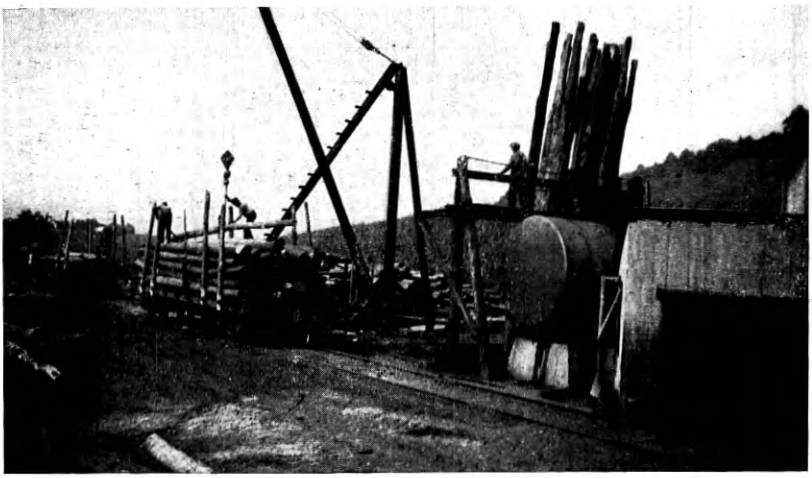


Fig. 15—Stiff Leg Derrick Removing Poles from Treating Tank and Loading Them on Flat Car

to handle a 4,000 lb. load at a 15 foot radius through an arc of about 270 degrees. It has a 30 foot boom. Since a very large percentage of the chestnut poles handled, weigh less than one ton each, this tractor crane has sufficient capacity for the service.

In the smaller plants where it has been found desirable to increase the pole treating capacities above what could be handled by means of the tractor cranes, stiff leg derricks have been installed. These derricks are of 6-tons capacity, having 45-foot booms. They are operated by steam from the treating plant boiler, which feeds the 8 H.P. hoisting engines. In these installations the swingers are operated by the hoisting engines.

Where the treating plant is of large enough capacity to warrant the investment in a locomotive crane, this type of unit has proven to be the most satisfactory in operation. The cranes which are suitable for this type of work have a 50 foot boom and are rated at 17½ tons capacity. Actually they can safely handle a 3-ton load at 50 feet radius from the king pin of the crane, perpendicular to the

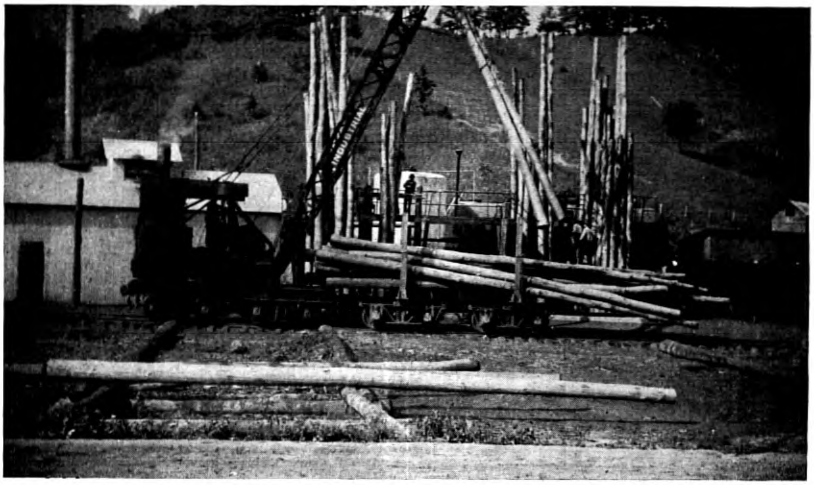


Fig. 16—Unloading Poles from the Treating Tanks to the Dollies, with Locomotive Crane

track, without tipping the car body of the crane. Of course, with the boom in a position above the track the maximum safe load is considerably greater.

The method of handling poles most commonly used is illustrated in Fig. 17 where the poles are lifted in a balanced condition, swung to one side of the track and piled parallel to it.

Another method which is applicable, particularly to handling a 40-foot and longer pole, consists of butting the pole end against the boom of the locomotive crane and swinging it to a pile which lies perpendicular to the track. This method of handling poles is similar to that shown in use with the logging outfit in Fig. 6.

When the poles are piled either parallel or perpendicular to the track as shown by Figs. 17 and 18, respectively, there should be frequent breaks in the piles in order to permit the air to circulate around the poles and keep them dry, and to reduce the fire hazard.

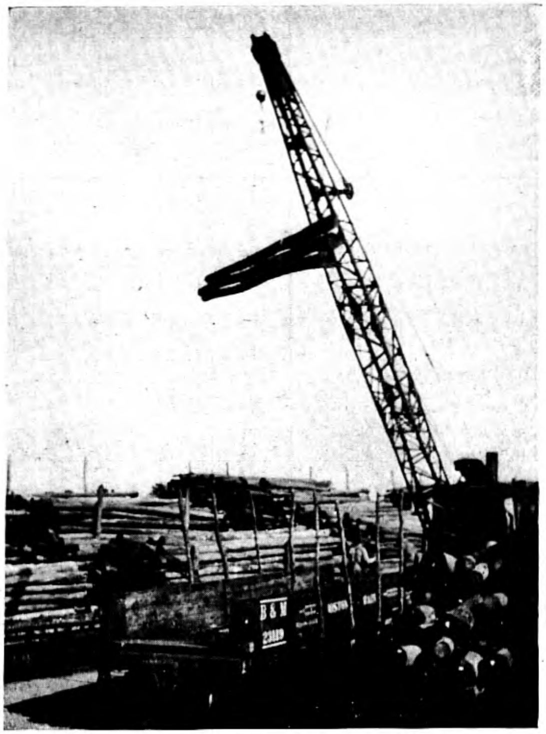


Fig. 17—Handling Poles by Balanced Method with Locomotive Crane

Fig. 18—Handling Pole with End Butted Against Boom of Locomotive Crane

Preparing Poles for Treatment

Although efforts were originally made to clean and prepare the poles on the cars at the time they were received at the plant, in order to be able to unload them from the cars directly into the treating tanks, it was found to be more satisfactory to first unload them upon skids where they would be more accessible for the removal

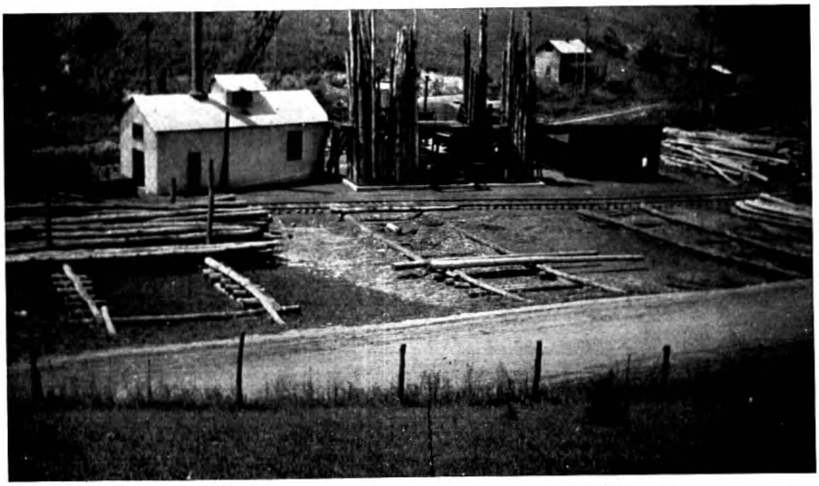


Fig. 19—Preparation Skids Opposite Treating Tanks at Sylva Plant

of all bark and foreign matter from the area to be treated and where any defective poles could be culled out before treatment.

The preparation skids are ordinarily not used for storage purposes. When a load of poles is placed upon them it can be spread in such a manner that every pole will be accessible.

In Fig. 20 the load of poles from the dollies has just been laid on the preparation skids where they will be cleaned for treatment in the far tank which is shown empty. Due to the desirability of having a continuous supply of poles for treatment, also of having the poles seasoned for several months before treatment, it is not practicable in a very large percentage of cases to ship the poles direct from the timber to the yard and unload them on the preparation skids for immediate treatment. For this reason it is necessary first to pile them in the untreated section of the pole yard and later to bring them to the preparation skids on dollies as illustrated in Fig. 20.

Treatment

The following is a very brief outline of the method pursued in treating the poles and also of the results obtained.

In so far as practicable the poles are seasoned 6 months or more before being treated. The method of treatment consists of immersing the butts to a level of about 1 foot above what will be the ground line of the poles, for not less than 7 hours in creosote at a temperature between 212° and 230° Fahrenheit. At the end of the hot treatment, the hot oil is quickly removed from the tank and cold oil at a temperature of from 100° to 110° Fahrenheit is permitted to flow quickly into the treating tank to the level previously reached by the hot oil. The cold oil treatment lasts for at least 4 hours.

Fig. 20—Preparation Skids Opposite Treating Tanks in Nashville Yard

Heat is absorbed by the pole butts in the hot oil bath until the moisture contained in the sapwood is either expanded into steam or entirely driven out. During the short interval while the oil is being changed, the surfaces to be treated remain covered by oil from the hot treatment. The oil change is made so quickly that the pole butts cool very little before it is completed. Then, as soon as the cold oil is admitted, these surfaces are covered by the creosote which remains until the pole butts become cool. In the sapwood, from which the moisture has been driven by the hot treatment, the cooling process condenses the steam, thus forming a partial vacuum in the

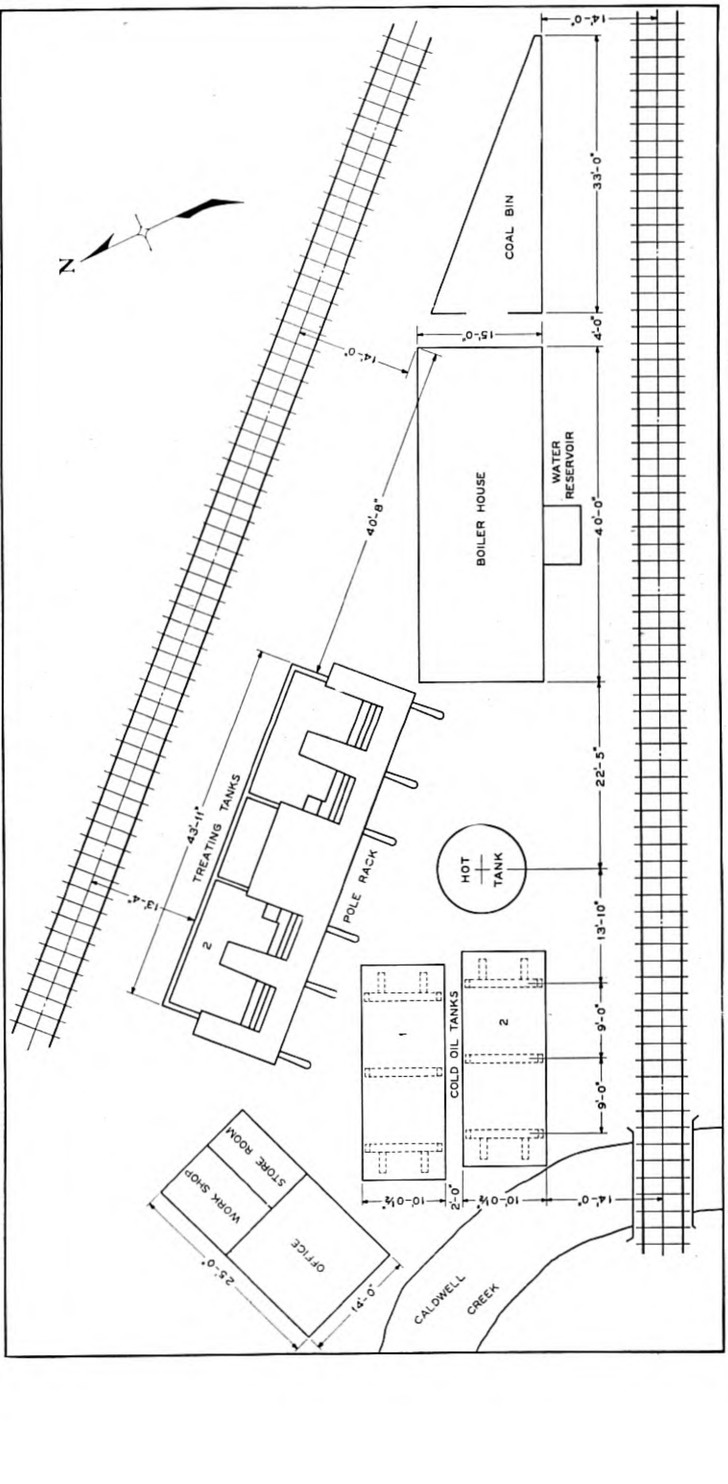


Fig. 21—Plan of Sylva Plant Layout

wood. This causes the oil, in which these surfaces are immersed, to be forced into the wood by atmospheric pressure.

During the treatment, the creosote is absorbed by the pole to such an extent that as an average, about 95 per cent. of the sapwood in the treated section of the pole is saturated. This requires from 2 to 4 gallons of oil per pole, depending upon the size and condition of the pole being treated.

Assembly Layout

The same general features of design were followed in all the pole treating plant layouts in so far as practicable. However, the number

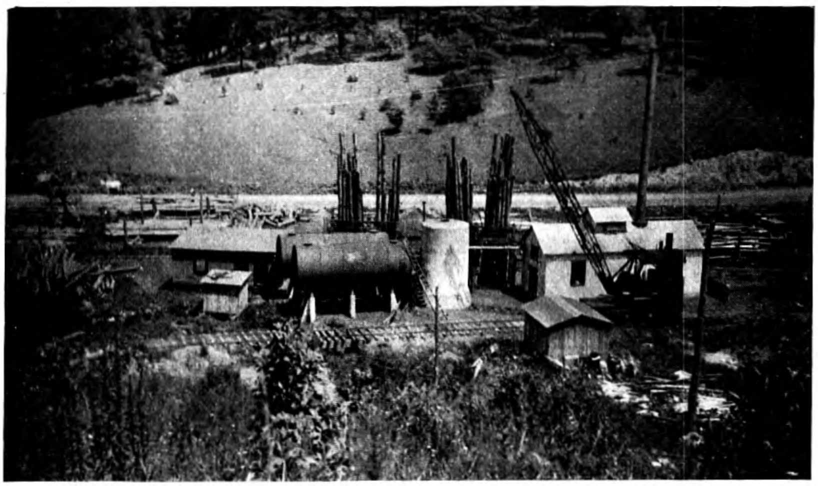


Fig. 22—View of Treating Equipment at Sylva Plant

of the different units used was varied to provide the plant capacities required.

In designing the plants it was found desirable to separate the poles into two or three treating tanks in order that the treating gang could be continuously employed in either preparing or handling poles from or to one of the tanks while the treatment would be in progress in other tanks. By dividing the tanks it was also possible to use a smaller quantity of hot creosote, since the hot oil could be used in one tank and when that treatment was finished, pumped to another tank containing fresh poles ready for treatment. Cutting down the hot oil capacity, of course, reduced the amount of radiation in the heating tank and also the amount of radiation in use at any particular time in the treating tanks, thus resulting in considerably less steam boiler capacity than would be necessary with a very large single treating tank unit.

Handling poles at smaller tanks is much easier because less boom action of the derrick is required and the men at the tanks can reach all poles more easily for attaching and removing the derrick winch line.

It was found that a vertical cylindrical tank served better than a horizontal one for the storage of hot oil, while the horizontal cylindrical

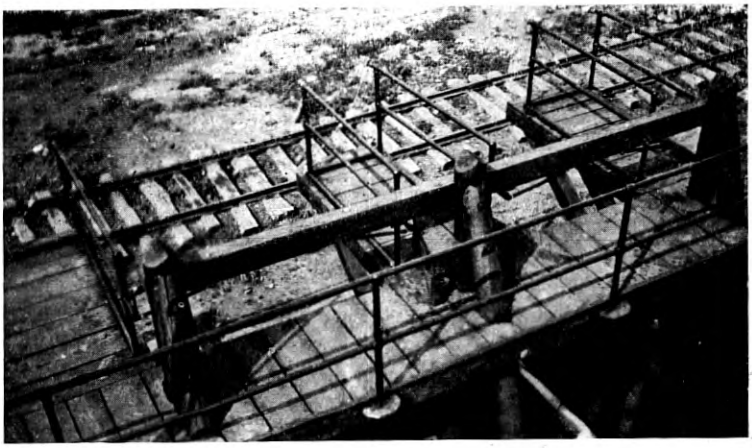


Fig. 23—Plan View of Pole Rack

tanks were preferable for cold oil storage. The radiation from a vertical hot tank is considerably reduced by the jacket of hot air rising along its side.

Particularly during the summer months care must be taken to keep down the temperature of the cold oil. It has been found that the long cylindrical steel tanks when lying horizontally radiate heat from the oil to the atmosphere satisfactorily and thus keep the oil cool.

Care has been taken in the design, to locate the various units so that all hot oil leads would be as short as possible in order to minimize radiation. Wherever possible, both the hot and cold oil are handled by gravity. The steam boiler is located as near as practicable to the heavy banks of steam radiators.

In all cases, careful study has been given to facilitating the handling of poles, since a considerable part of the cost of the pole treating process is due to pole handling.

Pole Racks

For supporting the poles standing in the treating tanks, it is necessary to have a very strong rack surrounding each of the tanks. Fig. 20 shows a view at one end and the front side of the two-tank rack in the Nashville plant. The poles shown, stand 8½ feet below the ground level. They are supported at the ends and middle of the rack by timbers under the rack platform at a height of 12 feet above

Fig. 24—Excavation for Treating Tanks

the ground. At the back, the poles are supported by a timber which is 16 feet above the ground. This arrangement permits the treatment of any size of pole up to and including 65 feet in length.

It will be noted in Fig. 23, which shows the rack above one tank, that the poles in each tank are divided at the middle by the platform of the pole rack. This feature of the rack has proved to be very desirable in that it permits the platform man to reach any pole in the rack during the loading and unloading process, so that there is no delay and no hazard in attaching the winch line sling to, or detaching it from the poles. The taper of the poles is such that ample space is provided for holding the sections of the poles at the platform level even though the area of the opening at this level is somewhat smaller than the area of the bottom of the treating tank.

Suitable railings have been provided around all parts of the platform to protect the platform man. They are substantial enough to protect the operator and yet flexible enough to compensate for the irregular sections of poles which may lie against them.

Tanks

As was mentioned above, in so far as practicable the tanks for the various plants are made in multiples of standard units. The treating

Fig. 25—Concrete Foundation and Protecting Walls for Treating Tanks

tanks for the smaller plants are 11 feet long and 5 feet 6 inches wide with 6 inches in each end of the tanks taken up by the vertical radiators. These tanks are of proper size to treat ½ carload of poles each.

The larger plants are provided with treating tanks, each of which will easily handle one carload of poles. These tanks are 15 feet long, 8 feet wide and 9 feet 8 inches deep in the clear.

Some idea of the sizes and arrangement of the treating tanks can be had from the excavation for them shown in Fig. 24. Each of the raised levels shown, will support the bottom of a tank while the pits between will contain the steam and oil piping, oil handling machinery, etc. This is a three-tank pit with space for two tanks shown.

In order to provide dry pits for the equipment below the treating tank bottoms and also to facilitate removal of a tank from the ground in case it might need repair, it has been found desirable to build concrete foundations and walls around the treating tanks.

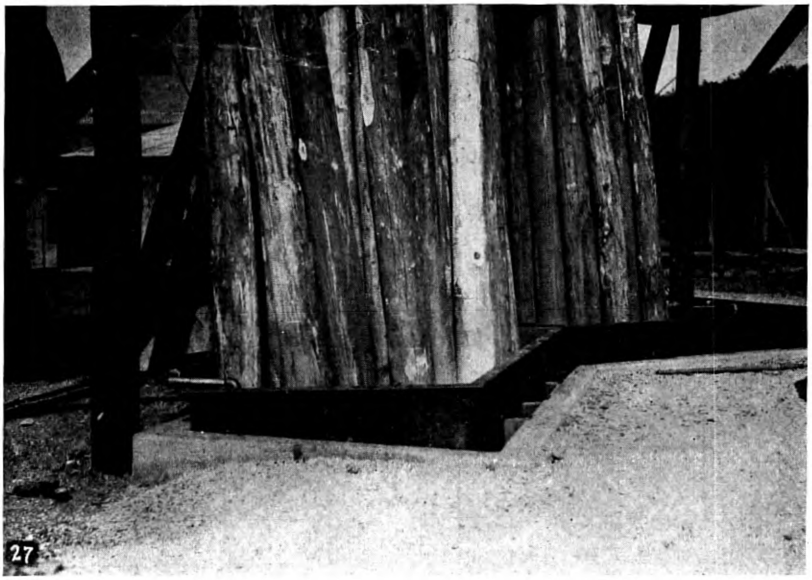


Fig. 26—Treating Tanks in Place

A few inches of space is left between the concrete retaining walls and the sides of the treating tanks. This space serves two purposes: it permits placing or removing the tanks with ease and it also provides air spaces around the sides of the tanks, which tend to insulate them from the ground. As has been mentioned, it is necessary to change the temperature of the oil in the tanks quickly from about 220° to about 105° Fahrenheit. There is very little lag in making the temperature change due to heat retained by the tank walls. However, if the ground around the tanks were wet and in contact with them, considerable lag would be experienced in making the temperature change of the oil because of heat which would be retained by the ground.

The poles in the tanks as shown by Fig. 26 rest in a position inclined slightly back toward the racks so that they remain in this position without being tied. Inclining the tank bottoms toward the rear facilitates the drainage of oil from them.

The bottom of the tank is practically perpendicular to the poles as they stand on it, which minimizes the tendency for the butts to slip on the tank bottom. In order to further prevent any danger from this happening, the bottom of each tank is covered by extra heavy Irving grids similar to those used at subway ventilating openings. These grids are supported by a suitable I-beam framework in

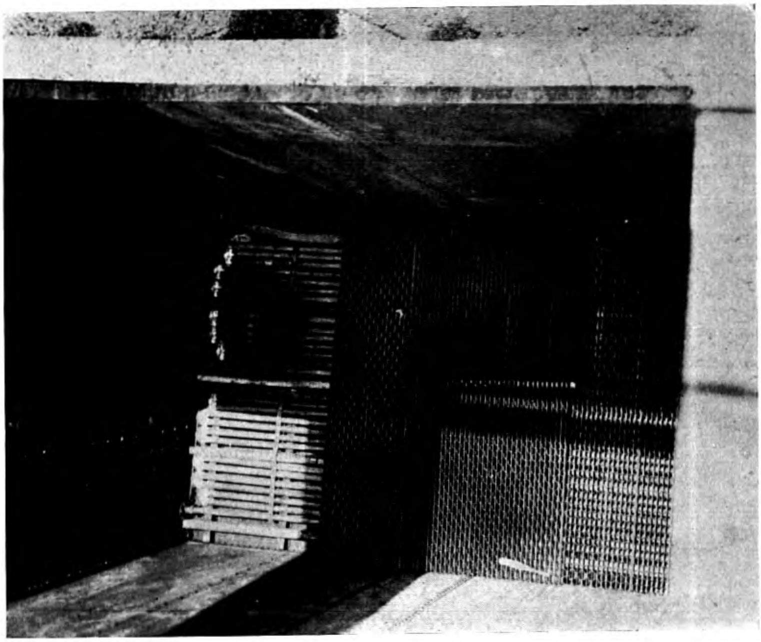


Fig. 27—Bottom of Treating Tank Showing Horizontal Radiators and Grids Covering Them

which the steel pipe radiators are placed. The grids do not interfere with the circulation of the hot oil and form a good protection for the radiators.

Each of the horizontal cold oil tanks has a capacity of about 14,000 gallons. Tanks of this size will easily take a tank-car load of creosote each, leaving some reserve capacity for residual oil which may be in the tanks at the time the additional cars of oil are received. The tank cars ordinarily carry from 8,000 to 12,000 gallons of oil.

The hot oil tanks vary in capacity between 3,000 and 13,000 gallons each, depending upon the sizes of the plants. One hot oil tank suffices for each installation. In order to conserve the heat, these tanks are covered by a $1\frac{1}{2}$ inch coat of magnesia block heat insulating material, the outside of which is covered by $\frac{1}{4}$ inch of asbestos cement and $\frac{1}{4}$ inch of half and half asbestos and Portland cement.

Boilers, Radiators, Pressure Regulators and Other Steam Equipment

For these installations, a self-contained type of steam boiler was used because of its comparatively high efficiency in the sizes required and also because of the ease of installation. The boilers used vary

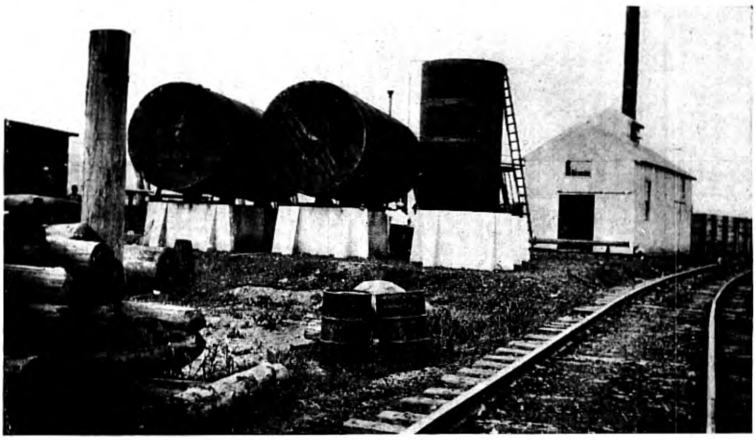


Fig. 28—Horizontal Cold Oil Tanks and Vertical Hot Oil Tank

from 30 to 80 horsepower capacity depending upon the sizes of the plants. These boilers are of the return tubular type with the fire boxes and smoke boxes lined with keyed-in fire brick.

The boilers are operated at a pressure of about 100 lbs. which is a suitable pressure for the steam turbine and for the steam hoisting engines in the plants where these are used. This boiler steam pressure is too high for the cast iron radiators which are used to heat oil in the hot and cold tanks and, for the smaller plants, in the treating tanks. Steam for these radiators should be supplied at a pressure of about 40 pounds. In order to meet this requirement a pressure reducer is used to convert the steam from the boiler pressure, whatever it may be, to a pressure of about 40 pounds, before it enters the radiators.

The water condensed from the various radiators is returned to the boiler in order to conserve its heat. Small automatic steam traps pass the water condensed in the radiators as fast as it is made, but do not permit the steam to pass. On the water side of these small steam traps, the piping from the various radiators is brought together and led to a point above the steam boiler where it is connected to a large tilting trap. The traps automatically raise the water to a

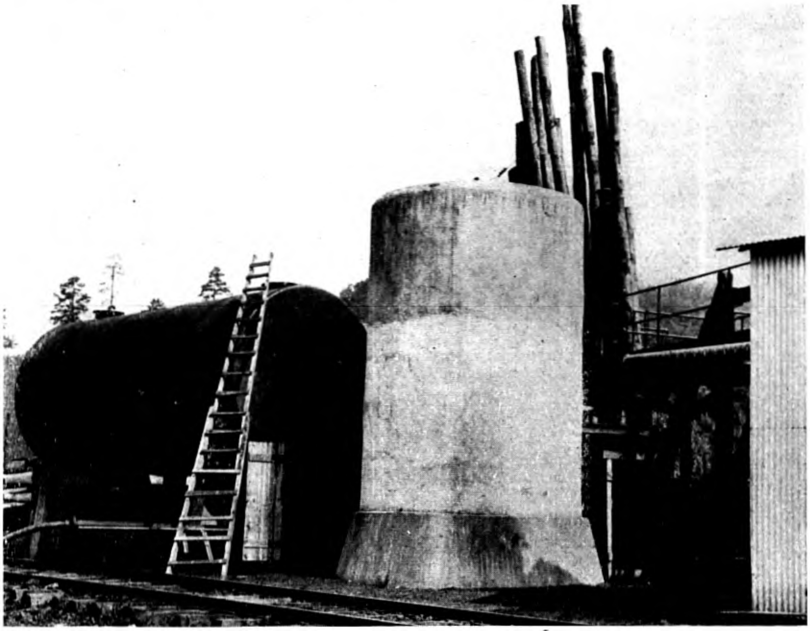


Fig. 29—Vertical Hot Oil Tank with Insulated Covering

receiver above the boiler and the tilting trap injects it into the boiler as fast as it is delivered to the water pipe lines by the small traps.

It is very desirable in the operation of the steam turbines that they be supplied with dry steam in order that slugs of water cannot enter the turbine chambers at high velocities and injure the vanes. A large water trap is located above the treating tank pit at each plant to insure dry steam for the turbine which is mounted in the pit directly below it.

Temperature Control

A continuous record is kept of the temperature of the oil in the treating tanks by means of recording thermometers mounted in the boiler room and connected by flexible thermometer tubes to the bulbs which are immersed in oil along the inside of the tanks after the poles are in place. In the cold and hot tanks the temperature does not change rapidly, so their temperatures can be read by means of stationary indicating thermometers mounted on the sides of the tanks and having bulbs which project into the insides of the tanks through suitable fittings. The oil temperatures, of course, are controlled by the steam valves to the radiators in the various tanks.

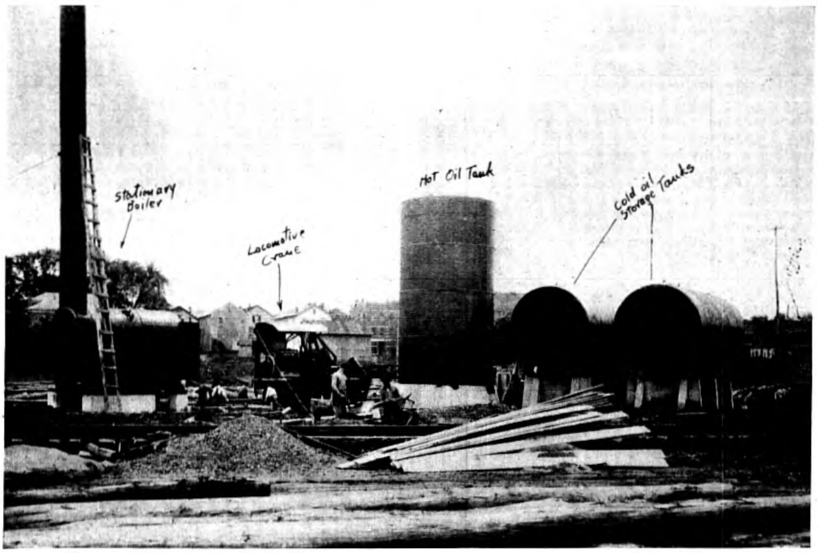


Fig. 30—Steam Boiler During Installation

Oil Handling

The heart of the oil handling apparatus, of course, is the centrifugal pump which has been mentioned and which is direct connected to the 20 H.P. steam turbine. In some of the smaller plants the centrifugal pumps are operated by 5 H.P. gasoline engines.

Both cold oil and hot oil are fed from the storage tanks to the treating tanks by gravity. The centrifugal pump is used for returning the oil from the treating tanks to the proper storage tank, for moving it from one storage tank to another or for delivering oil from the tank cars to the storage tanks.

Since the creosote which is used in pole treating may solidify at any temperature below 100° Fahrenheit, even in comparatively warm weather it is sometimes necessary to provide a steam connection to the radiators inside the tank car in order to make the oil fluid enough to flow through the flexible hose and pipes to the centrifugal pump. The solidifying of the creosote at comparatively high temperatures also requires a small bank of radiators in each cold tank.

The steam pipe runs, between the steam boiler and the various tanks, and the oil pipe lines between the various tanks and the pump,

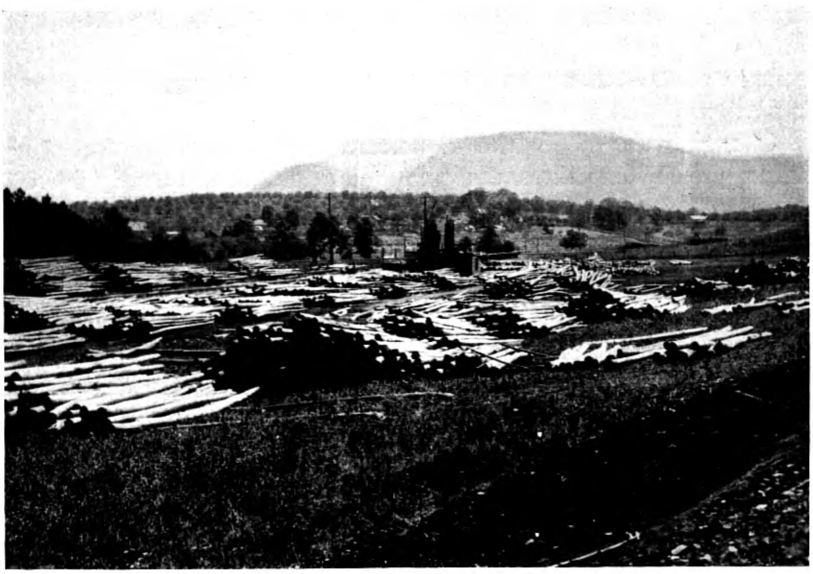


Fig. 31—General View of Natural Bridge Plant in Operation

are grouped so that both the steam lines and oil lines can be enclosed in boxes. The heat radiated from the steam lines warms the air in the boxes to such an extent that the oil remains liquid.

The valve controls for the oil and steam lines which are led through the boxes, are grouped so that several can be reached by opening the door of each of the boxes.

In the smaller plants which have the one-half-car pole capacity of treating tanks, the centrifugal pump handles the oil at a rate of about 200 gallons per minute. In the larger plants, however, where the treating tanks have one-car capacity of poles, the oil is handled through the centrifugal pump at the rate of about 600 gallons per minute. As mentioned in the above section describing the treatment, the high rate of oil movement is necessary in order to accomplish the change from hot to cold oil in the treating tanks in such a short time that the heated pole butts will not be permitted to cool when not immersed in oil. The oil change ordinarily is made in from 7 to 12 minutes from the time the pump starts to remove the hot oil until the cold oil is up to the proper level.

Experience indicates that no material loss in penetration of the creosote into the poles is experienced by having the treated section uncovered for this short length of time. Practically the same penetration is obtained as would be secured by keeping the poles in hot oil for the same length of time and then permitting the hot oil to remain around them until its temperature had gradually fallen by radiation to that specified for the cold oil bath.

Changing the oil instead of permitting it to cool in the treating tanks greatly expedites the treatments and consequently increases the plant capacity, which, of course, results in a corresponding economy in the cost of treating the poles.

Conclusion

In this paper an endeavor has been made to cover in a general way, the principal engineering and operating features involved in building creosoting plants designed specially for applying open tank treatment to chestnut poles. It has, of course, been necessary to omit practically all of the details of construction, which were followed in building the various plants.

These treating plants are valuable assets to the Bell System in providing concentration points where preservative treatment can be economically applied to the chestnut poles, thus becoming an important factor in the general program for the conservation of natural resources, by making possible the utilization of this valuable and rapidly diminishing type of timber over a considerably longer period.

Selective Circuits and Static Interference*

By JOHN R. CARSON

SYNOPSIS: The present paper has its inception in the need of a correct understanding of the behavior of selective circuits when subjected to irregular and random interference, and of devising a practically useful figure of merit for comparing circuits designed to reduce the effects of this type of interference. The problem is essentially a statistical one and the results must be expressed in terms of mean values. The mathematical theory is developed from the idea of the spectrum of the interference and the response of the selective circuit is expressed in terms of the mean square current and mean power absorbed. The application of the formulas deduced to the case of static interference is discussed and it is shown that deductions of practical value are possible in spite of meagre information regarding the precise nature and origin of static interference.

The outstanding deductions of practical value may be summarized as follows:

1. Even with absolutely ideal selective circuits, an irreducible minimum of interference will be absorbed, and this minimum increases linearly with the frequency range necessary for signaling.

2. The wave-filter, when properly designed, approximates quite closely to the ideal selective circuit, and little, if any, improvement over its present form may be expected as regards static interference.

3. As regards static or random interference, it is quite useless to employ extremely high selectivity. The gain, as compared with circuits of only moderate selectivity, is very small, and is inevitably accompanied by disadvantages such as sluggishness of response with consequent slowing down of the possible speed of signaling.

4. A formula is developed, which, together with relatively simple experimental data, provides for the accurate determination of the spectrum of static interference.

5. An application of the theory and formulas of the paper to representative circuit arrangements and schemes designed to reduce static interference, shows that they are incapable of reducing, in any substantial degree, the mean interference, as compared with what can be done with simple filters and tuned circuits. The underlying reason lies in the nature of the interference itself.

I

THE selective circuit is an extremely important element of every radio receiving set, and on its efficient design and operation depends the economical use of the available frequency range. The theory and design of selective circuits, particularly of their most conspicuous and important type, the electric wave filter, have been highly developed, and it is now possible to communicate simultaneously without undue interference on neighboring channels with a quite small frequency separation. On the other hand too much has been expected of the selective circuit in the way of eliminating types of interference which inherently do not admit of elimination by any form of selective circuit. I refer to the large amount of inventive thought devoted to devising ingenious and complicated circuit ar-

* Presented at the Annual Convention of the A. I. E. E., Edgewater Beach, Chicago, Ill., June 23–27, 1924.

rangements designed to eliminate *static interference.* Work on this problem has been for the most part futile, on account of the lack of a clear analysis of the problem and a failure to perceive inherent limitations on its solutions by means of selective circuits.

The object of this paper is twofold: (1) To develop the mathematical theory of the behavior of selective circuits when subjected to random, irregular disturbances, hereinafter defined and designated as *random interference.* This will include a formula which is proposed as a measure of the *figure of merit of selective circuits with respect to random interference.* (2) On the basis of this theory to examine the problem of *static interference* with particular reference to the question of its elimination by means of selective circuits. The mathematical theory shows, as might be expected, that the complete solution of this problem requires experimental data regarding the frequency distribution of static interference which is now lacking. On the other hand, it throws a great deal of light on the whole problem and supplies a formula which furnishes the theoretical basis for an actual determination of the spectrum of static. Furthermore, on the basis of a certain mild and physically reasonable assumption, it makes possible general deductions of practical value which are certainly qualitatively correct and are believed to involve no quantitatively serious error. These conclusions, it may be stated, are in general agreement with the large, though unsystematized, body of information regarding the behavior of selective circuits to static interference, and with the meagre data available regarding the wave form of elementary static disturbances.

The outstanding conclusions of practical value of the present study may be summarized as follows:

(1) Even with absolutely ideal selective circuits, an irreducible minimum of interference will be absorbed, and this minimum increases linearly with the frequency range necessary for signaling.

(2) The wave-filter, when properly designed, approximates quite closely to the ideal selective circuit, and little, if any, improvement over its present form may be expected as regards static interference.

(3) As regards static or random interference, it is quite useless to employ extremely high selectivity. The gain, as compared with circuits of only moderate selectivity, is very small, and is inevitably accompanied by disadvantages such as sluggishness of response with consequent slowing down of the possible speed of signaling.

(4) By aid of a simple, easily computed formula, it should be possible to determine experimentally the frequency spectrum of static.

(5) Formulas given below for comparing the relative efficiencies of selective circuits on the basis of signal-to-interference energy ratio are believed to have considerable practical value in estimating the relative utility of selective circuits as regards static interference.

II

Discrimination between signal and interference by means of selective circuits depends on taking advantage of differences in their wave forms, and hence on differences in their *frequency spectra.* It is therefore the function of the selective circuit to respond effectively to the range of frequencies essential to the signal while discriminating against all other frequencies.

Interference in radio and wire communication may be broadly classified as *systematic* and *random,* although no absolutely hard and fast distinctions are possible. *Systematic interference* includes those disturbances which are predominantly steady-state or those whose energy is almost all contained in a relatively narrow band of the frequency range. For example, interference from individual radiotelephone and slow-speed radio telegraph stations is to be classified as systematic. *Random interference,* which is discussed in detail later, may be provisionally defined as the aggregate of a large number of elementary disturbances which originate in a large number of unrelated sources, vary in an irregular, arbitrary manner, and are characterized statistically by no sharply predominate frequency. An intermediate type of interference, which may be termed either *quasi-systematic* or *quasi-random,* depending on the point of view, is the aggregate of a large number of individual disturbances, all of the same wave form, but having an irregular or random time distribution.

In the present paper we shall be largely concerned with random interference, as defined above, because it is believed that it represents more or less closely the general character of *static* interference. This question may be left for the present, however, with the remark that the subsequent analysis shows that, as regards important practical applications and deductions, a knowledge of the exact nature and frequency distribution of static interference is not necessary.

Now when dealing with random disturbance, as defined above, no information whatsoever is furnished as regards instantaneous values. In its essence, therefore, the problem is a statistical one and the conclusions must be expressed in terms of mean values. In the present paper formulas will be derived for the *mean energy* and *mean square current* absorbed by selective circuits from random interference, and their applications to the static problem and the protection afforded by selective networks against static will be discussed.

The analysis takes its start with certain general formulas given by the writer in a recent paper[1], which may be stated as follows:

Suppose that a selective network is subjected to an impressed force $\phi\ (t)$. We shall suppose that this force exists only in the time interval, or epoch, $o \leqq t \leqq T$, during which it is everywhere finite and has only a finite number of discontinuities and a finite number of maxima and minima. It is then representable by the Fourier Integral

$$\phi(t) = 1/\pi \int_0^\infty |f(\omega)| \cdot \cos\left[\omega t + \theta(\omega)\right] d\omega \tag{1}$$

where

$$|f(\omega)|^2 = \left[\int_0^\infty \phi(t) \cos \omega t dt\right]^2 + \left[\int_0^\infty \phi(t) \sin \omega t dt\right]^2. \tag{2}$$

Now let this force $\phi\ (t)$ be applied to the network in the *driving* branch and let the resulting current in the *receiving* branch be denoted by $I\ (t)$. Let $Z\ (i\ \omega)$ denote the steady-state *transfer* impedance of the network at frequency $\omega/2\pi$: that is the ratio of e.m.f. in *driving* branch to current in *receiving* branch. Further let $z\ (i\ \omega)$ and $\cos \alpha\ (\omega)$ denote the corresponding impedance and power factor of the receiving branch. It may then be shown that

$$\int_0^\infty [I(t)]^2 dt = 1/\pi \int_0^\infty \frac{|f\ (\omega)|^2}{|Z(i\omega)|^2} d\omega \tag{3}$$

and that the total energy W absorbed by the receiving branch is given by

$$W = 1/\pi \int_0^\infty \frac{|f(\omega)|^2}{|Z(i\omega)|^2} |z(i\omega)| \cos \alpha(\omega) \cdot d\omega. \tag{4}$$

To apply the formulas given above to the problem of random interference, consider a time interval, or epoch, say from $t=o$ to $t=T$, during which the network is subjected to a disturbance made up of a large number of unrelated elementary disturbances or forces, $\phi_1\ (t)$, $\phi_2\ (t) \ldots \phi_n\ (t)$.

If we write

$$\Phi(t) = \phi_1(t) + \phi_2(t) + \ldots + \phi_n(t),$$

then by (1), $\Phi(t)$ can be represented as

$$\Phi(t) = 1/\pi \int_0^\infty |F(\omega)| \cdot \cos\left[\omega t + \theta(\omega)\right] d\omega$$

[1] Transient Oscillations in Electric Wave Filters, Carson and Zobel, *Bell System Technical Journal*, July, 1923.

and

$$\int_0^{\infty} [I(t)]^2 \, dt = 1/\pi \int_0^{\infty} \frac{| F(\omega) |^2}{| Z(i\omega) |^2} \, d\omega. \tag{3}$$

We now introduce the function $R\ (\omega)$, which will be termed the *energy spectrum* of the random interference, and which is analytically defined by the equation

$$R(\omega) = \frac{1}{T} | F(\omega) |^2 \tag{5}$$

Dividing both sides of (3) and (4) by T we get

$$\bar{I}^2 = 1/\pi \int_0^{\infty} \frac{R(\omega)}{Z | (i\omega) |^2} \, d\omega, \tag{6}$$

$$\bar{P} = 1/\pi \int_0^{\infty} \frac{R(\omega)}{|Z(i\omega)|^2} \, | z(i\omega)| \cdot \cos \alpha\ (\omega)\ .\ d\omega. \tag{7}$$

$\bar{I}^2$, $\bar{P}$ and $R\ (\omega)$ become independent of the T provided the epoch is made sufficiently great. *$\bar{I}^2$ is the mean square current and $\bar{P}$ the mean power absorbed by the receiving branch from the random interference.*

In the applications of the foregoing formulas to the problem under discussion, the mean square current $\bar{I}^2$ of the formula (6) will be taken as the relative measure of interference instead of the mean power $\bar{P}$ of formula (7). The reason for this is the superior simplicity, both as regards interpretation and computation, of formula (6). The adoption of $\bar{I}^2$ as the criterion of interference may be justified as follows:

(1) In a great many important cases, including in particular experimental arrangements for the measurement of the static energy spectrum, the receiving device is substantially a pure resistance. In such cases multiplication of $\bar{I}^2$ by a constant gives the actual mean power $\bar{P}$.

(2) It is often convenient and desirable in comparing selective networks to have a standard termination and receiving device. A three-element vacuum tube with a pure resistance output impedance suggests itself, and for this arrangement formulas (6) and (7) are equal within a constant.

(3) We are usually concerned with relative amounts of energy absorbed from static as compared with that absorbed from signal. Variation of the receiver impedance from a pure constant resistance would only in the extreme cases affect this ratio to any great extent. In other words, the ratio calculated from formula (6) would not differ greatly from the ratio calculated from (7).

(4) While the interference actually apperceived either visually or by ear will certainly depend upon and increase with the energy absorbed from static, it is not at all certain that it increases linearly therewith. Consequently, it is believed that the additional refinement of formula (7) as compared with formula (6) is not justified by our present knowledge and that the representation of the receiving device as a pure constant resistance is sufficiently accurate for present purposes. It will be understood, however, that throughout the following argument and formulas, $\overline{P}$ of formula (7) may be substituted for $\overline{I}^2$ of (6), when the additional refinement seems justified. The theory is in no sense limited to the idea of a pure constant resistance receiver, although the simplicity of the formulas and their ease of computation is considerably enhanced thereby.

The problem of random interference, as formulated by equations (6) and (7) was briefly discussed by the writer in "Transient Oscillations in Electric Wave Filters" [1] and a number of general conclusions arrived at. That discussion will be briefly summarized, after which a more detailed analysis of the problem will be given.

Referring to formula (6), since both numerator and denominator of the integrand are everywhere $\geqq o$, it follows from the mean value theorem that a value ω of ω exists such that

$$\overline{I}^2 = \frac{R(\overline{\omega})}{\pi}\int_0^\infty \frac{d\omega}{|Z(i\omega)|^2}. \tag{8}$$

The approximate location of $\overline{\omega}$ on the frequency scale is based on the following considerations:

(a) In the case of efficient selective circuits designed to select a continuous finite range of frequencies in the interval $\omega_1 \leqq \omega \leqq \omega_2$, the important contributions to the integral (6) are confined to a finite continuous range of frequencies which includes, but is not greatly in excess of, the range which the circuit is designed to select. This fact is a consequence of the impedance characteristics of selective circuits, and the following properties of the spectrum $R(\omega)$ of random interference, which are discussed in detail subsequently.

(b) $R(\omega)$ is a continuous finite function of ω which converges to zero at infinity and is everywhere positive. It possesses no sharp maxima or minima, and its variation with respect to ω, where it exists, is relatively slow.

On the basis of these considerations it will be assumed that $\overline{\omega}$ lies within the band $\omega_1 \leqq \omega \leqq \omega_2$ and that without serious error it may be taken as the mid-frequency ω_m of the band which may be defined either as $(\omega_1+\omega_2)/2$ or as $\sqrt{\omega_1\omega_2}$. Consequently

$$\bar{I}^2=\frac{R(\omega_m)}{\pi}\int_0^\infty \frac{d\omega}{|Z(i\omega)|^2}. \tag{9}$$

From (9) it follows that the mean square current $\bar{I}^2$, due to random interference, is made up of two factors: one $R\ (\omega_m)$ which is proportional to the energy level of the interference spectrum at mid-frequency $\omega_m/2\ \pi$: and, second, the integral

$$\rho=1/\pi\int_0^\infty \frac{d\omega}{|Z\ (i\omega)|^2} \tag{10}$$

which is independent of the character and intensity of the interference. Thus

$$\bar{I}^2=\rho R(\omega_m). \tag{11}$$

Formula (11) is of considerable practical importance, because by its aid the spectral energy level $R\ (\omega)$ can be determined, once $\bar{I}^2$ is experimentally measured and the frequency characteristics of the receiving network specified or measured. It is approximate, as discussed above, but can be made as accurate as desired by employing a sufficiently sharply selective network.

The formula for the *figure of merit of a selective circuit with respect to random interference* is constructed as follows:

Let the signaling energy be supposed to be spread continuously and uniformly over the frequency interval corresponding to $\omega_1\leqq\omega\leqq\omega_2$. Then the mean square signal current is given by

$$\frac{E^2}{\pi}\int_{\omega_1}^{\omega_2}\frac{d\ \omega}{|Z(i\omega)|^2}$$

or, rather, on the basis of the same transmitted energy to

$$\frac{E^2}{\pi(\omega_2-\omega_1)}\int_{\omega_1}^{\omega_2}\frac{d\omega}{|Z(i\omega)|^2}=E^2\frac{\sigma}{\omega_2-\omega_1}. \tag{12}$$

The ratio of the mean square currents, due to signal and to interference, is

$$\frac{E^2}{R(\omega_m)}\cdot\frac{1}{\omega_2-\omega_1}\frac{\sigma}{\rho}. \tag{13}$$

The first factor $\frac{E^2}{R(\omega_m)}$ depends only on the signal and interference energy levels, and does not involve the properties of the network. The second factor depends only on the network and measures the efficiency with which it excludes energy outside the signaling range. It will therefore be termed *the figure of merit of the selective circuit* and denoted by S, thus

$$S=\frac{1}{\omega_2-\omega_1}\frac{\sigma}{\rho}=\frac{1}{\omega_2-\omega_1}\int_{\omega_1}^{\omega_2}\frac{d\omega}{Z(i\omega)^2}\div\int_0^{\infty}\frac{d\omega}{|Z(i\omega)|^2}. \tag{14}$$

Stated in words, *the figure of merit of a selective circuit with respect to random interference is equal to the ratio of the mean square signal and interference currents in the receiver, divided by the corresponding ratio in an ideal band filter which transmits without loss all currents in a "unit" band* ($\omega_2-\omega_1=1$) *and absolutely extinguishes currents outside this band.*

III

Before taking up practical applications of the foregoing formulas further consideration will be given to the hypothesis, fundamental to the argument, that over the frequency range which includes the important contributions to the integral $\int_0^{\infty}\frac{d\omega}{|Z(i\omega)|^2}$ the spectrum $R(\omega)$ has negligible fluctuations so that the integral

$$\int_0^{\infty}\frac{R(\omega)}{|Z(i\omega)|^2}d\omega$$

may, without appreciable error, be replaced by

$$R(\omega_m)\int_0^{\infty}\frac{d\omega}{|Z(i\omega)|^2}$$

where $\omega_m/2\pi$ is the "mid-frequency" of the selective circuit.

The original argument in support of this hypothesis was to the effect that, since the interference is made up of a large number of unrelated elementary disturbances distributed at random in time, any sharp maxima or minima in the spectrum of the individual diturbances would be smoothed out in the spectrum of the aggregate disturbance. This argument is still believed to be quite sound: the importance of the question, however, certainly calls for the more detailed analysis which follows:

Let
$$\Phi(t)=\sum_1^N\phi_r(t-t_r) \tag{15}$$

where t_r denotes the time of incidence of the r^{th} disturbance $\phi_r(t)$. The elementary disturbances $\phi_1, \phi_2 \ldots \phi_N$ are all perfectly arbitrary, so that $\Phi(t)$ as defined by (15) is the most general type of disturbance possible. The only assumption made as yet is that the instants of incidence $t_1 \ldots t_N$ are distributed at random over the epoch $o \leqq t \leqq T$; an assumption which is clearly in accordance with the facts in the case of static interference. If we write

$$C_r(\omega)=\int_0^\infty \phi_r(t)\cos\omega t\,dt,$$

$$S_r(\omega)=\int_0^\infty \phi_r(t)\sin\omega t\,dt, \qquad (16)$$

it follows from (2) and (15), after some easy rearrangements that

$$|F(\omega)|^2=\sum_{r=1}^{N}\sum_{s=1}^{N}\cos\omega(t_r-t_s)[C_r(\omega)C_s(\omega)+S_r(\omega)S_s(\omega)]=$$
$$\sum C_r^2(\omega)+S_r^2(\omega) \qquad (17)$$
$$+\sum\sum\cos\omega(t_r-t_s)\,[C_r(\omega)C_s(\omega)+S_r(\omega)\,S_s(\omega)],\ r\neq s.$$

The first summation is simply $\sum|f_r(\omega)|^2$. The double summation involves the factor $\cos\omega\,(t_r-t_s)$. Now by virtue of the assumption of random time distribution of the elementary disturbances, it follows that t_r and t_s, which are independent, may each lie anywhere in the epoch $o \leqq t \leqq T$ with all values equally likely. The mean value of $|F(\omega)|^2$ is therefore gotten by averaging[2] with respect to t_r and t_s over all possible values, whence

$$|F(\omega)|^2=\sum|f_r(\omega)|^2+2/T^2\,\frac{1-\cos\omega T}{\omega^2}$$
$$\times\sum\sum[C_r(\omega)C_s(\omega)+S_r(\omega)S_s(\omega)] \qquad (18)$$

and

$$\bar{I}^2=\frac{1}{\pi T}\sum\int_0^\infty\frac{|f_r(\omega)|^2}{|Z(i\omega)|^2}d\omega+\frac{2}{\pi T^2}\sum\sum\int_0^\infty\frac{1-\cos\omega T}{\omega^2 T}[C_r(\omega)C_s(\omega)$$
$$+S_r(\omega)S_s(\omega)]\frac{d\omega}{|Z(i\omega)|^2}.$$

[2] The averaging process with respect to the parameters t_r and t_s employed above logically applies to the average result in a very large number of epochs during which the system is exposed to the same set of disturbances with different but random time distributions. Otherwise stated, the averaging process gives the mean value corresponding to all possible equally likely times of incidence of the elementary disturbances. The assumption is, therefore, that if the epoch is made sufficiently large, the actual effect of the unrelated elementary disturbances will in the long run be the same as the average effect of all possible and equally likely distributions of the elementary disturbances.

Now in the double summation if the epoch T is made sufficiently great, the factor $\frac{(1-\cos \omega T)}{\omega^2 T}$ vanishes everywhere except in the neighborhood of $\omega = o$. Consequently, the double summation can be written as

$$\frac{2}{\pi T^2}\int_0^\infty \frac{1-\cos\omega T}{\omega^2 T^2}\,d\omega T \cdot \sum\sum \frac{C_r(o)C_s(o)}{|Z(o)|^2} = \frac{1}{T^2}\sum\sum\frac{C_r(o)C_s(o)}{|Z(o)|^2}.$$

Finally if we write $N/T = n =$ average number of disturbances per unit time, and make use of formula (2), we get

$$\bar{I}^2 = \frac{n}{N}\sum 1/\pi \int_0^\infty \frac{|f_r(\omega)|^2}{|Z(i\omega)|^2}\,d\omega + \frac{n^2}{N^2}\cdot\frac{1}{|Z(o)|^2}\cdot\sum\sum\int_0^\infty \phi_r(t)dt \cdot \int_0^\infty \phi_s(t)dt, \tag{19}$$

which can also be written as

$$\bar{I}^2 = \frac{n}{N}\sum\int_0^\infty i_r^2 dt + \frac{n^2}{N^2}\sum\sum\int_0^\infty i_r dt\cdot\int_0^\infty i_s dt\cdot \tag{20}$$

when $i_r = i_r\ (t)$ is the current due to the r^{th} disturbance $\phi_r\ (t)$.

Now the double summation vanishes when, due to the presence of a condense or transformer, the circuit does not transmit direct current to the receiving branch. Furthermore, if the disturbances are oscillatory or alternate in sign at random, it will be negligibly small compared with the single summation. Consequently, it is of negligible significance in the practical applications contemplated, and will be omitted except in special cases. Therefore, disregarding the double summation, the foregoing analysis may be summarized as follows:

$$R(\omega) = \frac{n}{N}\sum|f_r(\omega)|^2 = n\cdot r(\omega), \tag{21}$$

$$\bar{I}^2 = \frac{n}{N}\sum 1/\pi\int_0^\infty \frac{|f_r(\omega)|^2}{|Z(i\omega)|^2}\,d\omega \tag{22}$$

$$= \frac{n}{N}\sum\int_0^\infty i_r^2 dt = n\int_0^\infty \bar{i}^2 dt, \tag{23}$$

$$\bar{P} = \frac{n}{N}\int_0^\infty \frac{r(\omega)}{|Z(i\omega)|^2}|z(i\omega)|\cdot\cos\alpha\ (\omega)\cdot d\omega \tag{24}$$

$$= \frac{n}{N}\sum w_r = n\cdot\bar{w}. \tag{25}$$

In these formulas n denotes the average number of elementary disturbances per unit time, w_m the energy absorbed from the r^{th} disturbance $\phi_r(t)$, and $\bar{P}$ the mean power absorbed from the aggregate disturbance. $r(\omega)$ is defined by formula (20) and is the mean spectrum of the aggregate disturbance, thus

$$r(\omega)=1/N\sum|f_r(\omega)|^2=R(\omega)/N. \tag{26}$$

We are now in a position to discuss more precisely the approximations, fundamental to formulas (9)–(14),

$$\int_0^\infty \frac{R(\omega)}{|Z(i\omega)|^2}d\omega=R(\omega_m)\int_0^\infty \frac{d\omega}{|Z(i\omega)^2|}. \tag{27}$$

The approximation involved in this formula consists in identifying $\omega_m/2\pi$ with the "mid-frequency" of the selective circuit, and is based on the hypothesis that over the range of frequencies, which includes the important contribution to the integral (22), the fluctuation of $R(\omega)$ may be ignored.

Now it is evident from formulas (21)–(22) that the theoretically complete solution of the problem requires that $R(\omega)$ be specified over the entire frequency range from $\omega=o$ to $\omega=\infty$. Obviously, the required information cannot be deduced without making some additional hypothesis regarding the character of the interference or the mechanism in which it originates. On the other hand, the mere assumption that the individual elementary disturbances $\phi_1 \ldots \phi_N$ differ among themselves substantially in wave form and duration, or that the maxima of the corresponding spectra $|f_r(\omega)|$ are distributed over a considerable frequency range, is sufficient to establish the conclusion that the individual fluctuations are smoothed out in the aggregate and that consequently $r(\omega)$ and hence $R(\omega)$ would have negligible fluctuations, or curvature with respect to ω, over any limited range of frequencies comparable to a signaling range.

It is admitted, of course, that the foregoing statements are purely qualitative, as they must be in the absence of any precise information regarding the wave forms of the elementary disturbances constituting random interference. On the other hand, the fact that static is encountered at all frequencies without any sharp changes in its intensity as the frequency is varied, and that the assumption of a systematic wave form for the elementary disturbances would be physically unreasonable, constitute strong inferential support of the hypothesis underlying equation (27). Watt and Appleton (*Proc. Roy. Soc.*, April 3, 1923) supply the only experimental data regarding the wave forms of the elementary disturbances which they found to be classifiable under general types with rather widely variable amplitudes and durations. Rough calculations of $r\ (\omega)$, based on their results, are in support of the hypothesis made in this paper, at least in the radio frequency range. In addition, the writer has made calculations based on a number of reasonable assumptions regarding variations of wave form among the individual disturbances, all of which resulted in a spectrum $R\ (\omega)$ of negligible fluctuations over a frequency range necessary to justify equation (27) for efficient selective circuits. However the problem is not theoretically solvable by pure mathematical analysis, so that the rigorous verification of the theory of selectivity developed in this paper must be based on experimental evidence. On the other hand, it is submitted that the hypothesis introduced regarding static interference is not such as to vitiate the conclusions, qualitatively considered, or in general to introduce serious quantitative errors. Furthermore, even if it were admitted for the sake of argument that the figure of merit S was not an accurate measure of the ratio of mean square signal to interference current, nevertheless, it is a true measure of the excellence of the circuit in excluding interference energy outside the necessary frequency range.

IV

The practical applications of the foregoing analysis depend upon the formulas

$$\bar{I}^2 = \frac{R(\omega_m)}{\pi}\int_0^\infty \frac{d\omega}{|Z(i\omega)|^2} = \rho \cdot R\ (\omega_m) \tag{11}$$

and

$$S = \frac{1}{\omega_2 - \omega_1} \cdot \int_{\omega_1}^{\omega_2} \frac{d\omega}{|Z(i\omega)|^2} \div \int_0^\infty \frac{d\omega}{|Z(i\omega)|^2} = \frac{1}{\omega_2 - \omega_1} \cdot \frac{\sigma}{\rho} \tag{14}$$

which contain all the information which it is possible to deduce in the case of purely random interference. They are based on the principle that the effect of the interference on the signaling system is measured by the mean square interference current in the receiving branch, and that the efficiency of the selective circuit is measured by the ratio of the mean square signal and interference currents. As stated above, in the case of random interference results must be expressed in terms of mean values, and it is clear that either the mean square current or the mean energy is a fundamental and logical criterion.

Referring to formula (11), the following important proposition is deducible.

If the signaling system requires the transmissions of a band of frequencies corresponding to the interval $\omega_2-\omega_1$, *and if the selective circuit is efficiently designed to this end, then the mean square interference current is proportional to the frequency band width* $\frac{(\omega_2-\omega_1)}{2\pi}$.

This follows from the fact that, in the case of efficiently designed band-filters, designed to select the frequency range $\frac{(\omega_2-\omega_1)}{2\pi}$ and exclude other frequencies, the integral $\int_0^\infty \frac{d\omega}{|Z(i\omega)|^2}$ is proportional to $\omega_2-\omega_1$ to a high degree of approximation.

The practical consequences of these propositions are important and immediate. It follows that as the signaling speed is increased, the amount of interference inevitably increases practically linearly and that this increase is inherent. Again it shows the advantage of single vs. double side-band transmission in carrier telephony, as pointed out by the writer in a recent paper.[3] It should be noted that the increased interference with increased signaling band width is not due to any failure of the selective circuit to exclude energy outside the signaling range, but to the inherent necessity of absorbing the interference energy lying inside this range. The only way in which the interference can be reduced, assuming an efficiently designed band filter and a prescribed frequency range $\frac{(\omega_2-\omega_1)}{2\pi}$, is to select a carrier frequency, at which the energy spectrum $R(\omega)$ of the interference is low.

Formula (11) *provides the theoretical basis for an actual determination of the static spectrum.* Measurement of $\bar{I}^2$ over a sufficiently long interval, together with the measured or calculated data for evaluating the integral $\int_0^\infty \frac{d\omega}{|Z(i\omega)|^2}$, determines $R(\omega_m)$ and this determination can be made as accurate as desired by employing a sufficiently sharply tuned circuit or a sufficiently narrow band filter. It is suggested that the experimental data could be gotten without great difficulty, and that the resulting information regarding the statistical frequency distribution of static would be of large practical value.

The selective figure of merit S as defined by (14) is made up of two factors, $\frac{1}{(\omega_2-\omega_1)}$ which is inversely proportional to the required signaling frequency range; and the ratio of the integrals σ/ρ. This

[3] Signal-to-Static-Interference Ratio in Radio Telephony, *Proc. I. R. E. E.*, June, 1923.

ratio is unity for an ideally designed selective circuit, and can actually be made to approximate closely to unity with correctly designed band-filters. Formula (14) is believed to have very considerable value in comparing various circuits designed to eliminate interference, and is easily computed graphically when the frequency characteristics of the selective circuit are specified.

The general propositions deducible from it may be briefly listed and discussed as follows:

With a signaling frequency range $\frac{(\omega_2-\omega_1)}{2\pi}$ *specified, the upper limiting value of S with a theoretically ideal selective circuit is* $\frac{1}{(\omega_2-\omega_1)}$, *and the excellence of the actual circuit is measured by the closeness with which its figure of merit approaches this limiting value.*

Formula (14) for the figure of merit S has been applied to the study of the optimum design of selective circuits and to an analysis of a large number of arrangements designed to eliminate or reduce static interference. The outstanding conclusions from this study may be briefly reviewed and summarized as follows:

The form of the integrals σ and ρ, taking into account the signaling requirements, shows that the optimum selective circuit, as measured by S, is one which has a constant transfer impedance over the signaling frequency range $\frac{(\omega_2-\omega_1)}{2\pi}$, and attenuates as sharply as possible currents of all frequencies outside this range. Now this is precisely the ideal to which the band filter, when properly designed and terminated, closely approximates, and leads to the inference that *the wave filter is the best possible form of selective circuit, as regards random interference.* Its superiority from the steady-state viewpoint has, of course, long been known.

An investigation of the effect of securing extremely high selectivity by means of filters of a large number of sections was made, and led to the following conclusion:

In the case of an efficiently designed band-filter, terminated in the proper resistance to substantially eliminate reflection losses, the figure of merit is given to a good approximation by the equation

$$S=\frac{1}{\omega_2-\omega_1}\,\frac{1}{1+1/16\,n^2}$$

where n is the number of filter sections and $\frac{(\omega_2-\omega_1)}{2\pi}$ the transmission band. It follows that *the selective figure of merit increases inappreciably with an increase in the number of filter sections beyond* 2, *and that the* *band filter of a few sections can be designed to have a figure of merit closely approximating the ideal limiting value,* $\frac{1}{(\omega_2-\omega_1)}$,

This proposition is merely a special case of the general principle that, as regards static interference, it is useless to employ extremely high selectivity. The gain obtainable, as compared with only a moderate amount of selectivity is slight and is inherently accompanied by an increased sluggishness of the circuit. That is to say, as the selectivity is increased, the time required for the signals to build up is increased, with a reduction in quality and possible signaling speed.

Another circuit of practical interest, which has been proposed as a solution of the "static" problem in radio-communication consists of a series of sharply tuned oscillation circuits, unilaterally coupled through amplifiers.[4] This circuit is designed to receive only a single frequency to which all the individual oscillation circuits are tuned. The figure of merit of this circuit is approximately

$$S=L/R\frac{2^{2n-2}\,(n-1)!^2}{(2n-2)!}$$

where n denotes the number of sections or stages, and L and R are the inductance and resistance of the individual oscillation circuits. The outstanding fact in this formula is the slow rate of increase of S with the number of stages. For example, if the number of stages is increased from 1 to 5, the figure of merit increases only by the factor 3.66, while for a further increase in n the gain is very slow.[5] This gain, furthermore, is accompanied by a serious increase in the "sluggishness" of the circuit: That is, in the partciular example cited, by an increase of 5 to 1 in the time required for signals to build up to their steady state.

The analysis of a number of representative schemes, such as the introduction of resistance to damp out disturbances, balancing schemes designed to neutralize static without affecting the signal, detuning to change the natural oscillation frequency of the circuit, demodulation through several frequency stages, etc., has shown that they are one and all without value in increasing the ratio of mean square signal to interference current. In the light of the general theory, the reason for this is clear and the limitation imposed on the solution of the static problem by means of selective circuits is seen to be inherent in the nature of the interference itself.

[4] See U. S. Patent No. 1173079 to Alexanderson.

[5] When the number of stages n is fairly large, the selective figure of merit becomes proportional to $\sqrt{n}$ and the building-up time to n.

Some Contemporary Advances in Physics—VII
Waves and Quanta

By KARL K. DARROW

THE invaluable agent of our best knowledge of the environing world, and yet itself unknown except by inference; the intermediary between matter and the finest of our senses, and yet itself not material; intangible, and yet able to press, to strike blows, and to recoil; impalpable, and yet the vehicle of the energies that flow to the earth from the sun—light in all times has been a recognized and conspicuous feature of the physical world, a perpetual reminder that the material, the tangible, the palpable substances are not the only real ones. Yet its apparent importance, to our forerunners who knew only the rays to which the eye responds and suspected no others, was as nothing beside its real importance, which was realized very gradually during the nineteenth century, as new families of rays were discovered one after the other with new detecting instruments and with new sources. Radiation is not absent from the places where there is no eye-stimulating light; radiation is omnipresent; there is no region of space enclosed or boundless, vacuous or occupied by matter, which is not pervaded by rays; there is no substance which is not perpetually absorbing rays and giving others out, in a continual interchange of energy, which either is an equilibrium of equal and opposite exchanges, or is striving towards such an equilibrium. Radiation is one of the great general entities of the physical world; if we could still use the word "element," not to mean one of the eighty or ninety kinds of material atoms, but in a deeper sense and somewhat as the ancients used it, we might describe radiation and matter, or possibly radiation and electricity, as coequal elements. Also the problem of the nature and structure of radiation is of no lesser importance than the problem of the structure and nature of matter; and in fact neither can be treated separately; they are so inextricably intertwined that whoever sets out to expound the present condition of one soon finds himself outlining the other. One cannot write a discourse on the nature of radiation alone nor on the structure of the atom alone, one can but vary the relative emphasis laid upon these two subjects, or rather upon these two aspects of a single subject; and in this article I shall restate many things about the atom which were stated in former articles, but the emphasis will be laid upon light.

Speaking very generally and rather vaguely, light has been much more tractable to the theorists than most of the other objects of enquiry in physics or chemistry. Over a rather long period of years, it was indeed generally regarded as perfectly intelligible. The famous battle between the corpuscular theory adopted by Newton, and the wave-theory founded by Descartes and Huyghens, died out in the earlier years of the nineteenth century with the gradual extinction of the former. The history of optics in the nineteenth century, from Fresnel and Young to Michelson and Rayleigh, is the tale of a brilliant series of beautiful and striking demonstrations of the wave-theory, of experiments which were founded upon the wave-theory as their basis and would have failed if the basis had not been firm, of instruments which were designed and competent to make difficult and delicate measurements of all sorts—from the thickness of a sheet of molecules to the diameter of a star—and would have been useless had the theory been fallacious. The details of the bending of light around the sides of a slit or the edge of a screen, the intricate pattern of light and shade formed where subdivisions of a beam of light are reunited after separation, the complexities of refraction through a curved surface, are represented by the theory with all verifiable accuracy; and so are the incredibly complicated phenomena attending the progress of light through crystals, phenomena which have slipped out of common knowledge because few are willing to undertake the labour of mastering the theory. The wave-theory of light stands with Newton's inverse-square law of gravitation, in respect of the many extraordinarily precise tests which it has undergone with triumph; I know of no other which can rival either of them in this regard.

By the term "wave-theory of light" I have meant, in the foregoing paragraph, the conception that light is a wave-motion, an undulation, a periodic form advancing through space without distorting its shape; I have not meant to imply any particular answer to the question, *what is it of which light is a wave-motion?* It may seem surprising that one can make and defend the conception, without having answered the question beforehand; but as a matter of fact there are certain properties common to all undulations, and these are the properties which have been verified in the experiments on light. There are also certain properties which are not shared by such waves as those of sound, in which the vibration is confined to a single direction (that normal to the wavefront) and may not vary otherwise than in amplitude and phase, but are shared by transverse or distortional waves in elastic solids, in which the vibration may lie in any of an infinity of directions (any direction tangent to the wavefront). Light possesses these properties, and therefore the wave-motion which is radiation may not be compared with the wave-motion which is sound; but a wide range of comparisons still remains open.

Of course, very many have proposed images and models for "the thing of which the vibrations are light", and many have believed with an unshakable faith in the reality of their models. The fact that light-waves may be compared, detail by detail, with transverse vibrations in an elastic solid, led some to fill universal space with a solid elastic medium to which they gave the sonorous name of "luminiferous aether". It is not many years since men of science used to amaze the laity with the remarkable conception of a solid substance, millions of times more rigid than steel and billions of times rarer than air, through which men and planets serenely pass as if it were not there. Even now one finds this doctrine occasionally set forth.[1]

In that image of the elastic solid, the propagation of light was conceived to occur because, when one particle of the solid is drawn aside from its normal place, it pulls the next one aside, that one the next one to it, and so on indefinitely. Meanwhile, each particle which is drawn aside exerts a restoring force upon the particle of which the displacement preceded and caused its own. Set one of the particles into vibration, and the others enter consecutively into vibration. Maintain the first particle in regular oscillation, and each of the others oscillates regularly, with a phase which changes from one to the next; a wave-train travels across the medium. One particle influences the next, because of the attraction between them. But in the great and magnificent theory of light which Maxwell erected upon the base of Faraday's experiments, the propagation was explained in an altogether different manner. Vary the magnetic field across a loop of wire in a periodic manner, and you obtain a periodic electric force around the loop, as is known to everyone who has dabbled in electricity. Vary the electric field periodically, and you obtain a periodic magnetic field—this a fact not by any means so well known as the other, one which it was Maxwell's distinction to have anticipated, and which was verified after the event. In a traveling train of light-waves the electric field and the magnetic field stimulate one another alternately and reciprocally, and for this reason the wave-train travels. Since the periodic electric field may point in any one of the infinity of directions in the plane of the wave-front, the wave-motion possesses all the freedom and variability of

[1] Apparently the image of the elastic solid was never quite perfected; one recalls the question as to whether its vibrations were in or normal to the plane of polarization of the light, which required one answer in order to agree with the phenomena of reflection, and another in order to agree with those of double refraction. Probably a *modus vivendi* could have been arranged if the whole idea had not been superseded.

form which are required to account for the observed properties of light.

Maxwell's theory immediately achieved the stunning success of presenting a value for the speed of the imagined electromagnetic waves, determined exclusively from measurements upon the magnetic fields of electric currents, and agreeing precisely with the observed speed of light. Two supposedly distinct provinces of physics, each of which had been organized on its own particular basis of experience and in its own particular manner, were suddenly united by a stroke of synthesis to which few if any parallels can be found in the history of thought. And this is by no means the only achievement of the electromagnetic theory of light; there will shortly be occasion to mention some of the others.

Now that there was so much evidence that light travels as a wave-motion, and that its speed and other properties are those of electromagnetic waves, it became urgently desirable to inquire into the nature of the *sources* of light. Granted that light *en route* outwards from a luminous particle of matter is of the nature of a combination of wave-trains, what is taking place in the luminous particle? To this question all our experience and all our habits of thought suggest one sole obvious answer—that in the luminous particle there is a vibrating something, a *vibrator*, or more likely an enormous number of vibrators—one to each atom, possibly—and the oscillations of these vibrators are the sources of the waves of light, as the oscillations of a violin-string or a tuning-fork are the sources of waves of sound. This analogy drawn from acoustics, this picture of the vibrating violin-string and the vibrating tuning-fork, has been powerful—indeed, it begins to seem, too powerful—in guiding the formation of our ideas on light. It is profitable to reflect that the evolution of thought in acoustics must have traveled in the opposite sense from the evolution of thought in optics. Whoever it was who was the first to conceive that sound is a wave-motion in air, must certainly have arrived at the idea by noticing that sounding bodies vibrate. One feels the trembling of the tuning-fork or the bell, one sees the violin-string apparently spread out into a band by the amplitude of its motion; it is not difficult to build apparatus which, like a slowed-down cinema film, makes the vibrations separately visible, or, like the stroboscope, produces an equivalent and not misleading illusion. This was not possible in optics, and never will be. In acoustics, one may sometimes accept the vibrations of the sounding body as an independently-given fact of experience, and reason forward to the wave-motion spreading outwards into the environing air; in optics, this entrance to the path is closed, one must reason in the inverse sense from the wave-motion to the qualities of the shining body. Inevitably, it was assumed that when the path should at last be successfully retraced, the shining body would be found in the semblance of a vibrator.

For a few years at the end of the nineteenth century and the beginning of the twentieth, it seemed that the desired vibrator had been found. Apparently it was the electron, the little corpuscle of negative electricity, of which the charge and the mass were rather roughly estimated in the late nineties, although Millikan's definite measurements were not to come for a decade yet. Maxwell had not conceived of particles of electricity, his conception of the "electric fluid" was indeed so sublimated and highly formal that it gave point to the celebrated jest (I think a French one) about the man who read the whole of his "Electricity and Magnetism" and understood it all except that he was never able to find out what an electrified body was. H. A. Lorentz incorporated the electron into Maxwell's theory. Conceiving it as a spherule of negative electricity, and assuming that in an atom one or more of these spherules are held in equilibrium-positions, to which restoring-forces varying proportionally to displacement draw them back when they are displaced, Lorentz showed that these "bound" electrons are remarkably well adapted to serve as sources and as absorbents for electromagnetic radiation. Displaced from its position of equilibrium by some transitory impulse, and then left to itself, the bound electron would execute damped oscillations in one dimension or in two, emitting radiation of the desired kind at a calculable rate. Or, if a beam of radiation streamed over an atom containing a bound electron, there would be a "resonance" like an acoustic resonance—the bound electron would vibrate in tune with the radiation, absorbing energy from the beam and scattering it in all directions, or quite conceivably delivering it over in some way or other to its atom or the environing atoms. There were numerical agreements between this theory and experience, some of them very striking.[2] Apparently the one thing still needful was to produce a plausible theory of these binding-forces which control the response of the "bound" electron to disturbances of all kinds. Once these were properly described, the waves of light would be supplied with

[2] Notably, the trend of the dispersion-curves for certain transparent substances, recently extended by Bergen Davis and his collaborators to the range of X-ray frequencies; the normal Zeeman effect; Wien's observations on the exponential dying-down of the luminosity of a canal-ray beam, interpreted as the exponential decline in the vibration-amplitudes of the bound electrons in the flying atoms; the dependence of X-ray scattering on the number of electrons in the atom.

their vibrators, the electromagnetic theory would receive a most valuable supplement. And, much as a competent theory of the binding-forces was to be desired, a continuing failure to produce one would not impugn the electromagnetic theory, which in itself was a coherent system, self-sustaining and self-sufficient.

This was the state of affairs in the late nineties. The wave-conception of light had existed for more than two centuries, and it was seventy-five years since any noticeable opposition had been raised against it. The electromagnetic theory of light had existed for about thirty years, and now that the electron had been discovered to serve as a source for the waves which in their propagation through space had already been so abundantly explained, there was no effective opposition to it. Not all the facts of emission and absorption had been accounted for, but there was no reason to believe that any particular one of them was unaccountable. Authoritative people thought that the epoch of great discoveries in physics was ended. It was only beginning.

In the year 1900, Max Planck published the result of a long series of researches on the character of the radiation inside a completely-enclosed or nearly-enclosed cavity, surrounded by walls maintained at an even temperature. Every point within such a cavity is traversed by rays of a wide range of wave lengths, moving in all directions. By the "character" of the radiation, I mean the absolute intensities of the rays of all the various frequencies, traversing such a point. The character of the radiation, in this sense, is perfectly determinate; experiment shows that it depends only on the temperature of the walls of the cavity, not on its material. According to the electromagnetic theory of radiation, as completed by the adoption of the electron, the walls of the cavity are densely crowded with bound electrons; nor are these electrons all bound in the same manner, so that they would all have the same natural frequency of oscillation—they are bound in all sorts of different ways with all magnitudes of restoring-forces, so that every natural frequency of oscillation over a wide range is abundantly represented among them. Now the conclusion of Planck's long study was this:

If the bound electrons in the walls of the cavity (i.e., *in any solid body*) *did really radiate while and as they oscillate, in the fashion prescribed by the electromagnetic theory, then the character of the radiation in the cavity would be totally different from that which is observed.*[3]

[3] The belief that the character of radiation within a cavity could not be explained without doing some violence to the "classical mechanics" had already been gaining ground for some years, by reason of extremely recondite speculations of a statistical nature. It is very difficult to gauge the exact force and bearing of such considerations.

However, *if the bound electrons do not radiate energy while they oscillate, but accumulate it and save it up and finally discharge it in a single outburst when it attains some one of a certain series of values $h\nu$, $2h\nu$, $3h\nu$,* etc. (h stands for a constant factor, ν for the frequency of vibration of the electrons and the emitted radiation)—*then the character of the radiation will agree with that which is observed*, provided a suitable value be chosen for the constant h.

The value required [4] for h in C.G.S. units (erg seconds) is $6.53.10^{-27}$.

Here, then, was a phenomenon which the electromagnetic theory seemed to be fundamentally incapable of explaining. For this notion of a bound electron, which oscillates and does not meanwhile radiate, is not merely foreign to the classical theory, but very dangerous to it; one does not see how to introduce it, and displace the opposed notion, without bringing down large portions of the structure (including the numerical agreements which I cited in a foregoing footnote). However, Planck had arrived at this conclusion by an intricate process of statistical and thermodynamical reasoning. Statistical reasoning is notoriously the most laborious and perplexing in all physics, and many will agree that thermodynamical reasoning is not much less so. Planck's inference made an immense impression on the most capable thinkers of the time; but in spite of the early adherence of such men as Einstein and Poincaré, I suspect that even to this day it might practically be confined to the pages of the more profound treatises on the philosophical aspects of physics, if certain experimenters had not been guided to seek and to discover phenomena so simple that none could fail to apprehend them, so extraordinary that none could fail to be amazed.

Honour for this guidance belongs chiefly to Einstein. Where Planck in 1900 had said simply that bound electrons emit and absorb energy in fixed finite quantities, and shortly afterwards had softened his novel idea as far as possible by making it apply only to the act of emission, Einstein in 1905 rushed boldly in and presented the idea that these fixed finite quantities of radiant energy retain their identity throughout their wanderings through space from the moment of emission to the moment of absorption. This idea he offered as a "heuristic" one—the word, if I grasp its connotation exactly, is an apologetic sort of a word, used to describe a theory which achieves successes though its author feels at heart that it really is too absurd to

[4] I take the numerical values of the constant h scattered through this article from Gerlach. The weighted mean of the experimental values, with due regard to the relative reliability of the various methods, is taken as 6.55 or $6.56.10^{-27}$. None of the individual values cited in these pages is definitely known to differ from this average by more than the experimental error.

be presentable. The implication is, that the experimenters should proceed to verify the predictions based upon the idea, quite as if it were acceptable, while remembering always that it is absurd. If the successes continue to mount up, the absurdity may be confidently expected to fade gradually out of the public mind. Such was the destiny of this heuristic idea.

I will now describe some of these wonderfully simple phenomena—wonderfully simple indeed, for they stand out in full simplicity in domains where the classical electromagnetic theory would almost or quite certainly impose a serious complexity. If Planck's inference from the character of the radiation within a cavity had been deferred for another fifteen years, one or more of these phenomena would assuredly have been discovered independently. What would have happened in that case, what course the evolution of theoretical physics would have followed, it is interesting to conjecture.

The *photoelectric effect* is the outflowing of electrons from a metal, occurring when and because the metal is illuminated. It was discovered by Hertz in 1889, but several years elapsed before it was known to be an efflux of electrons, and several more before the electrons were proved to come forth with speeds which vary from one electron to another, upwards as far as a certain definite maximum value, and never beyond it.

Here is a rather delicate point of interpretation, which it is well to examine with some care; for all the controversies as to continuity versus discontinuity in Nature turn upon it, in the last analysis. What is meant, or what reasonable thing can be meant, when one says that the speeds of all the electrons of a certain group are confined within a certain range, extending up to a certain limiting topmost value? If one could detect each and every electron separately, and separately measure its speed, the meaning would be perfectly clear. For that matter, the statement would degenerate into a truism. The fact is otherwise. The instruments used in work such as this perceive electrons only in great multitudes. Suppose that one intercepts a stream of electrons with a metal plate connected by a wire to an electrometer. If a barrier is placed before the electrons in the form of a retarding potential-drop, which is raised higher and higher, the moment eventually comes when the current into the electrometer declines. This happens because the slower electrons are stopped and driven back before they reach the plate, the faster ones surmount the barrier. As the potential-drop is further magnified, the reading of the electrometer decreases steadily, and at last becomes inappreciable. Beyond a certain critical value of the retarding voltage, the electrometer reports no influx of electrons. Does this really mean that there are *no* electrons with more than just the speed necessary to overpass a retarding voltage of just that critical value? Or does it merely mean that the electrons flying with more than that critical speed are plentiful, but not quite plentiful enough to make an impression on the electrometer? Is there any topmost speed at

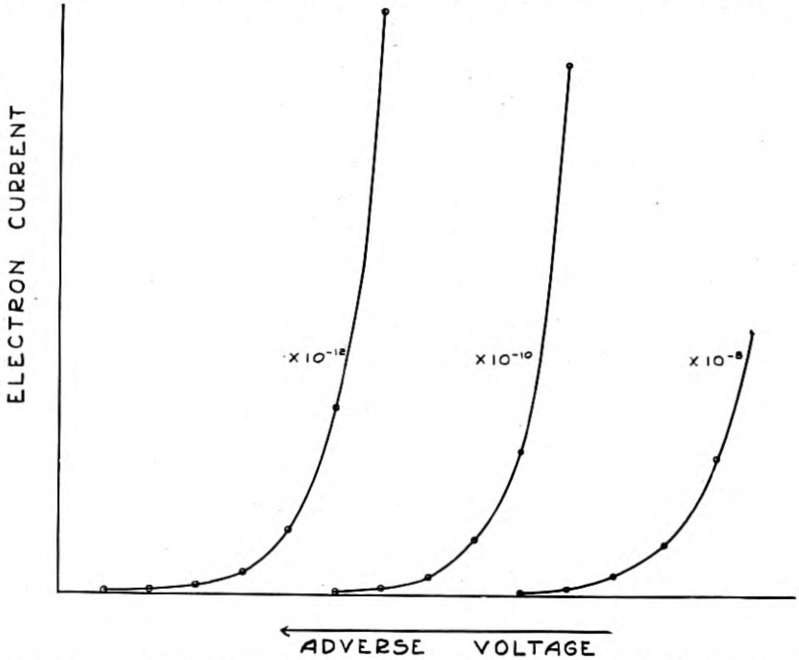


Fig. 1—Curves showing thermionic electron-current versus opposing voltage, demonstrating a distribution-in-speed extending over an unlimited range of speeds. Multiply the ordinates of the middle curve by 100, those of the right-hand curve by 10,000, to bring them to the same scale and make them merge into a single curve. (L. H. Germer)

all, or should we find, if we could replace the current-measuring device with other and progressively better ones *ad infinitum*, that the apparent maximum speed soared indefinitely upwards?

Absolute decisions cannot be rendered in a question of this kind; but it is possible, under the best of circumstances, to pile up indicatory evidence to such an extent that only an unusually strong will-to-disbelieve would refuse to be swayed by it. The judgment depends on the shape of the curve which is obtained by plotting the electrometer-reading *vs.* the retarding potential—in other words, the fraction y of the electrons of which the energy of motion surpasses the amount x, determined from the retarding-voltage by the relation $x = eV$. Look for example at the curves of Fig. 1, which refer to the electron-

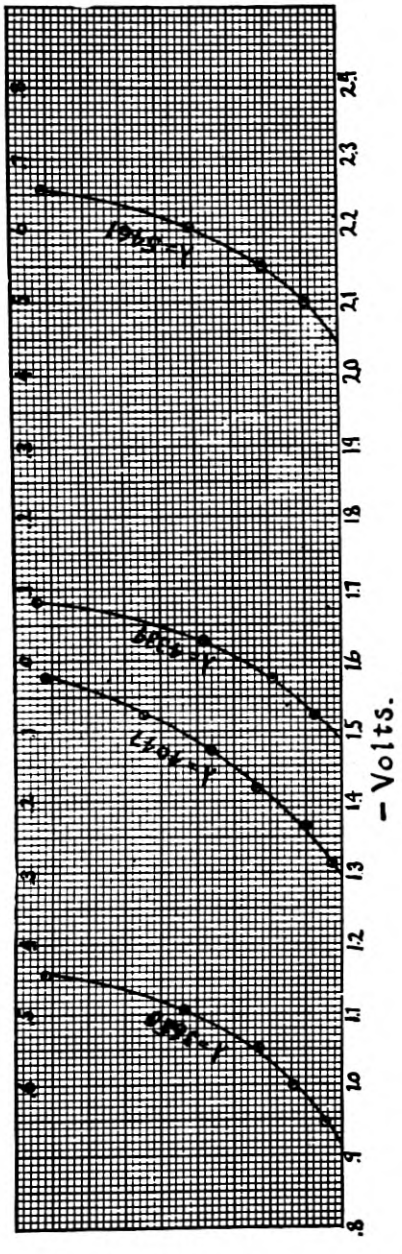


Fig. 2—Curves showing photoelectric electron-current versus opposing voltage, demonstrating a distribution-in-speed extending over a range limited at the top. (R. A. Millikan, *Physical Review*)

stream flowing spontaneously out of an incandescent wire; they are three segments of one single curve, plotted on different scales as the numerals show. This curve bends so gradually around towards tangency with the axis of abscissae, that one can hardly avoid the inference that it is really approaching that axis as if to an asymptote, and that if the electrometer at any point ceases to declare a current, it is because the electrometer is too insensitive to respond to the smaller currents, and not because there are no faster electrons. Look instead at the curves of Fig. 2, which refer to the electrons emerging from an illuminated surface of sodium. These curves slant so sharply towards the axis of abscissae, they bend so slightly in the portions of their courses where the data of experiment determine them, that the linear extrapolation over the little interval into the axis commends itself as natural and inevitable. Because the curves for the thermionic electrons approach the axis so gently, it is agreed that their velocities are distributed continuously over an unlimited range; because the curves for the photoelectrons cut into it so acutely, it is felt that their velocities are confined below a definite maximum value.

This therefore is the photoelectric effect: waves of light inundate the surface of a metal, and electrons pour out with various velocities, some nearly attaining and none exceeding a particular topmost value. I will designate this maximum speed, or rather the corresponding maximum kinetic energy, by E_{max}. Analyzing the process in the classical manner, one must imagine the waves entering into the metal and setting the indwelling electrons into forced oscillations; the oscillations grow steadily wider; the speed with which the electron dashes through its middle position grows larger and larger, and at last it is torn from its moorings and forces its way through the surface of the metal. Some of the energy it absorbed during the oscillations is spent (converted into potential energy) during the escape; the rest is the kinetic energy with which it flies away. Even if the electron were free within the metal and could oscillate in response to the waves, unrestrained by any restoring force, it would still have to spend some of its acquired energy in passing out through the boundary of the metal (the laws of thermionic emission furnish evidence enough for this). It is natural to infer that E_{max} is the energy absorbed by an electron originally free, minus this amount (let me call it P) which it must sacrifice in crossing the frontier; the electrons which emerge with energies lower than E_{max} may be supposed to have made the same sacrifice at the frontier and others in addition, whether in tearing themselves away from an additional restraint or in colliding with atoms during their emigration. This is not the only conceivable interpretation, but it seems unprofitable to enter into the others. It is therefore $E_{\max}$ which appears to merit the most attention.

Now the mere fact that there is a maximum velocity of the escaped electrons, that there is an $E_{\max}$, is not in itself of a nature to suggest that the classical theory is inadequate. It is the peculiar dependence of this quantity on the two most important controllable qualities of the light—on its intensity and on its frequency—which awakens the first faint suspicions that something has at last been discovered, which the classical theory is ill adapted to explain. One would predict with a good deal of confidence that the greater the intensity of the light, the greater the energy acquired by the electron in each cycle of its forced oscillation would be, the greater the energy with which it would finally break away, the greater the residuum of energy which at the end would be left to it. But $E_{\max}$ is found to be independent of the intensity of the light. This is strange; it is as though the waves beating upon a beach were doubled in their height and the powerful new waves disturbed four times as many pebbles as before, but did not displace a single one of them any farther nor agitate it any more violently than the original gentle waves did to the pebbles that they washed about. As for the dependence of $E_{\max}$ on the frequency of the light, it would be necessary to make additional assumptions to calculate it from the classical theory; in any case it would probably not be very simple. But the actual relation between $E_{\max}$ and ν is the simplest of all relations, short of an absolute proportionality; this is it:

$$E_{\max}=h\nu-P \tag{1}$$

Fig. 3 shows the relation for sodium, observed by Millikan.

The maximum energy of the photoelectrons increases linearly with the frequency of the light. P is a constant which varies from one metal to another. In the terms of the simple foregoing interpretation, P is the energy which an electron must spend (more precisely, the energy which it must invest or convert into potential energy) when it passes through the frontier of the metal on its way outward. Comparing the values of P for several metals with the contact potentials which they display relatively to one another, one finds powerful evidence confirming this theory. Having discussed this particular aspect of the question in the fifth article of this series, I will not enter further into it at this point.

The constant h is the same for all the metals which have been used in such experiments. The best determinations have been made upon two or three of the alkali metals, for these are the only metals

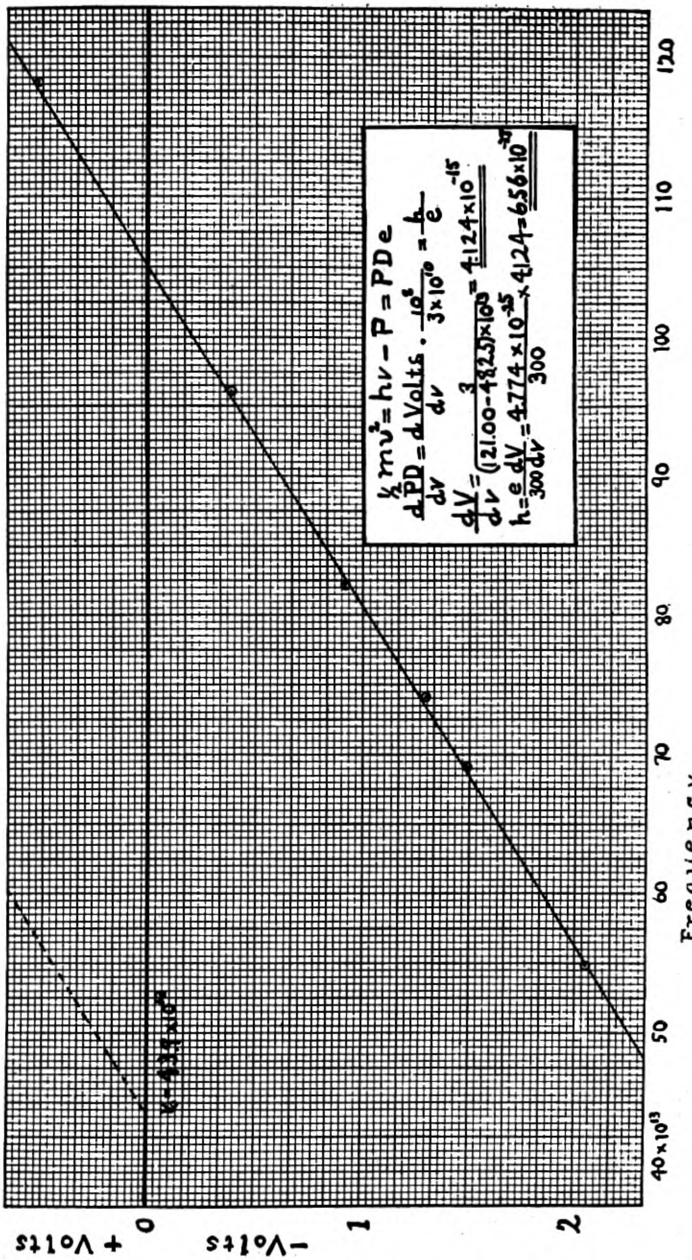


Fig. 3—Curve showing the linear relation between the maximum energy of photo-electrons and the frequency of the light which excites them. (Millikan, *Physical Review*)

which release electrons when illuminated with light of wide convenient ranges of frequency and color. Most metals must be irradiated with ultraviolet light, and the experiments become very difficult if they must be performed with light of frequencies far from the visible spectrum. The values which Millikan obtained for sodium and for lithium agree within the experimental error with one another and with the mean value

$$h = 6.57.10^{-27} \tag{2}$$

The maximum energy of the electrons released by light of the frequency ν is therefore equal to a quantity $h\nu$ which is the same, whatever metal be illuminated by the light—a quantity which is characteristic of the light, not of the metal—minus a quantity P which, there is every reason to believe, is the quota of energy surrendered by each electron in passing out across the boundary-surface of the metal. It is *as if* each of the released electrons had received a quantity $h\nu$ of energy from the light. I will go one step further, and lay down this as a rule, with another cautiously-inserted *as if* to guard against too suddenly daring an innovation:

Photoelectric emission occurs as if *the energy in the light were concentrated in packets, or units, or corpuscles of amount $h\nu$, and one whole unit were delivered over to each electron.*

This is a perfectly legitimate phrasing of equation (1), but I doubt whether anyone would ever have employed it, even with the guarded and apologetic *as if*, but for the fact that the value of h given in (2) agreed admirably well with the value of that constant factor involved in Planck's theory, the constant to which he had given this very symbol and a somewhat similar role. Deferring for a few pages one other extremely relevant feature of the photoelectric effect (its "instantaneity") I will proceed to examine these other situations.

An effect which might well be, though it is not, called the *inverse photoelectric effect*, occurs when electrons strike violently against metal surfaces. Since radiation striking a metal may elicit electrons, it is not surprising that electrons bombarding a metal should excite radiation. Electrons moving as slowly as those which ultraviolet or blue light excites from sodium do not have this power; or possibly they do, but the radiation they excite is generally too feeble to be detected. Electrons moving with speeds corresponding to kinetic energies of hundreds of equivalent volts,[5] and especially electrons

[5] One equivalent volt of energy = the energy acquired by an electron in passing across a potential-rise of one volt $= e/300$ ergs $= 1.591.10^{-12}$ ergs. This unit is usually called simply a "volt of energy", or "volt", a bad usage but ineradicable. Also "speed" is used interchangeably with "energy" in speaking of electrons, and one finds (and, what is worse, cannot avoid) such deplorable phrases as "a speed of 4.9 volts" ! ! !

with energies amounting to tens of thousands of equivalent volts, do possess it. This is in fact the process of excitation of X-rays, which are radiated from a metal target exposed to an intense bombardment of fast electrons. The protagonists of the electromagnetic theory had an explanation ready for this effect, as soon as it was discovered. A fast electron, colliding with a metal plate, is brought to rest by a slowing-down process, which might be gradual or abrupt, uniform or *saccadé*, but in any case must be continuous. Slowing-down entails radiation; the radiation is not oscillatory, for the electron is not oscillating, but it is radiation none the less; it is an outward-spreading single pulsation or *pulse*, comparable to the narrow spherical shell of condensed air which diverges outward through the atmosphere from an electric spark and has been photographed so often, or to a transient in an electrical circuit.

One may object that the pulse is just a pulse and nothing more, while the X-rays are wave-trains, for otherwise the X-ray spectroscope (which is a diffraction apparatus) would not function. The objection is answered by pointing out the quite indubitable fact that any pulse, whatever its shape (by "shape" I mean the shape of the curve representing the electric field strength, or whatever other variable one chooses to take, as a function of time at a point traversed by the wave) can be accurately reproduced by superposing an infinity of wave-trains, of all frequencies and divers properly-adjusted amplitudes, which efface one another's periodic variations, and in fact efface one another altogether at all moments except during the time-interval while the pulse is passing over—during this interval they coalesce into the pulse. Thence, the argument leads to the contention that the actual pulse is made up of just such wave-trains, and the sapient diffracting crystal recognizes them all and diffracts each of them duly along its proper path. The problem is not new, nor the answer; white light has long been diagnosed as consisting of just such pulses, and the method of analyzing transient impulses in electrical circuits into their equivalent sums of wave-trains has been strikingly successful.

The application of the method to this case of X-ray excitation enjoyed one qualitative success. The spherical pulse diverging from the place where an electron was brought to rest should not be of equal thickness at all the points of its surface; it should be broader and flatter on the side towards the direction whence the electron came, thinner and sharper on the side towards the direction in which the electron was going when it was arrested. Analyzing the pulse, it is found that at the point where it is broad and low, the most intense of its equivalent wave-trains are on the whole of a lower frequency than the most intense of the wave-trains which constitute it where it is narrow and high. By examining and resolving the X-rays radiated from a target, at various inclinations to the direction of the bombarding electrons, this was verified—verified in part, not altogether. The X-rays radiated nearly towards the source of the electron-stream include a

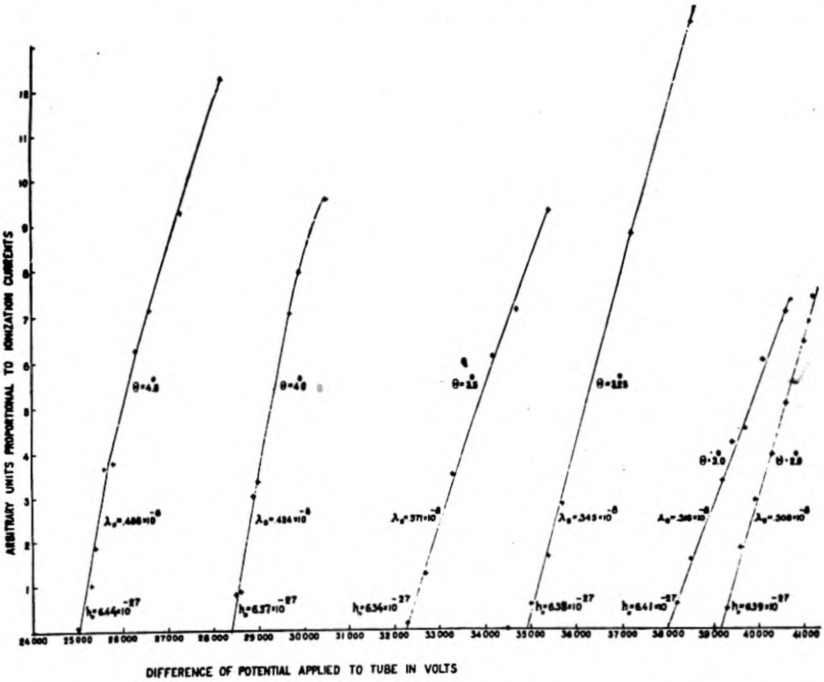


Fig. 4—Curves ("isochromatics") each representing the intensity of X-radiation of a very narrow range of frequencies, plotted versus the energy of the bombarding electrons. (Duane & Hunt, *Physical Review*)

lesser proportion of high-frequency wave-trains, they are *softer* as the phrase is, than the X-rays radiated nearly along the prolongation of the electron-stream. In the spectrum of each of these beams of X-rays, there is a wave length where the density of radiant energy attains a maximum, and this wave length is longer in the former beam than in the latter one. So much is implied in the classical theory.

But it is nowhere implied in the classical theory that the spectrum of an X-ray beam, produced when electrons of a constant energy rain down upon a metal, should extend upwards only to a certain maximum frequency, and then and there come to a sudden end; yet apparently it does. There is a *high-frequency limit* to each X-ray spectrum, and wave-trains of frequencies exceeding that limit are not detected; whereas the spectrum of the hypothetical pulses ought to include wave-trains of every frequency low or high, the amplitudes indeed declining to infinitely low values as one goes along the spectrum to infinitely high frequencies, but certainly declining smoothly and gradually. To demonstrate this high-frequency limit is a delicate experimental problem, quite like that other problem of demonstrating a sharply definite topmost value for the energies of photo-electrons. That question whether the curves of photoelectric current *vs.* retarding voltage, the curves of Fig. 2, cut straightly and sharply enough into the axis of abscissae to prove that there are no photo-electrons with velocities higher than the one corresponding to x_o, returns again in a slightly altered form.

The most reliable of the methods actually used to demonstrate the high-frequency limit depends on the fact that the high limiting frequency (which I will call ν_{max}) varies with the energy of the bombarding electrons, increasing as their velocity increases. Therefore, if the radiant energy belonging to rays of a certain fixed wave length or a certain fixed narrow range of wave lengths is separated out from the X-ray beam by a spectroscope, and measured for various velocities of the impinging electrons, passing from very high velocities step by step to very low ones; it will decrease from its first high value to zero at some intermediate velocity, and thereafter remain zero. But according to the classical theory also, it must decrease from its first high value to an imperceptibly low one; the descent however will be gradual and smooth. Thus the only question which can be settled by experiment is the question whether the descent from measurable intensities to immeasurably small ones resembles the gentle quasi-asymptotic decline of the curve of Fig. 1 or the precipitate slope of the curve of Fig. 2. The data assembled by Duane and Hunt are shown in Fig. 4 plotted in the manner I have described; there is little occasion for doubt as to which sort of curve these resemble most.[6]

Each of the curves in Fig. 4 represents that portion of the total intensity of an X-ray beam, which belongs to rays of wave lengths near the marked value of the frequency ν. This frequency is the high

[6] Three simple curves of the intensity-distribution in the X-ray spectrum are shown in Figure 5. The abscissa is neither frequency or wavelength, but a variable which varies continuously with either (it is actually *arc sin* of a quantity proportional to wavelength) so that the acute angle between each curve and the axis of abscissae, at the point where they meet, corresponds to and has much the same meaning as the acute angles in Figure 2—not so conspicuously.

limiting frequency ν_{max} for that value of the energy E of the bombarding electrons, which corresponds to the point on the axis of abscissae where the curve (extrapolated) intersects it. The relation between ν_{max} and E is the simplest of all relations:

$$E = \text{constant} \cdot \nu_{max} = h\ \nu_{max} \tag{3}$$

The constant h is the same for all the metals on which the experiment has been performed—a few of the least fusible ones, for metals of a low melting-point would be melted before E could be lifted far enough to give an adequate range for determining the relation between it and

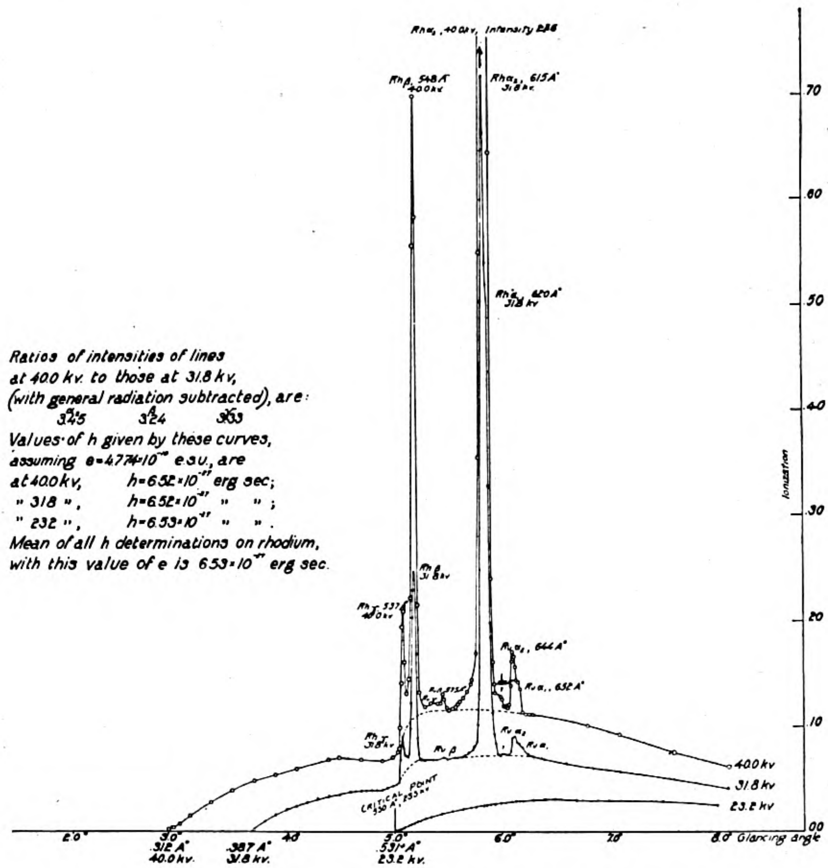


Fig. 5—The continuous X-ray spectrum for three values of the energy of the bombarding electrons, intensity being plotted versus a quantity varying uniformly with frequency. Ignore the peaks. (D. L. Webster, *Physical Review*.) See footnote 6

ν_{max}. The value[7] given for it by Gerlach, after a critical study of all the determinations, is

$$h = 6.53.10^{-27} \tag{4}$$

The highest frequency of radiation which electrons moving with the energy E are able to excite, when they are brought to rest by colliding with a metal target, is therefore equal to E divided by a constant independent of the kind of metal. So far as this high limiting frequency is concerned, it is perfectly legitimate to express equation (3) in these words,

Excitation of radiation by electrons stopped in their flight by collision with a metal occurs as if *the energy in the radiation were concentrated in units of amount $h\nu$, and one such unit were created out of the total energy which each electron surrenders when it is stopped.*

As for the radiation of frequencies inferior to the high limiting frequency, it is very easily explained by asserting that most of the electrons come to rest not in one operation, but in several successive ones, dividing their energy up among several units of frequencies inferior to ν_{max} or E/h; or possibly they lose energy in various sorts of impacts or various other ways before making the first impact of the sort which transforms their energy into energy of X-rays. Nothing about it contradicts the italicized rule. Still it is not likely that anyone would have formulated equation (3) in such language, if the value of the constant h which appears in it were not identical with the value which we have already once encountered in analyzing the photoelectric effect, and with the value at which Planck earlier arrived.

I think it is too early in this discourse to fuse these italicized Rules for the release of electrons by radiation and the excitation of radiation by electrons into a single Rule; but by contemplating the two Rules side by side one arrives without much labor at an inference which could be tested even though we had no way of measuring the frequency of a radiation, and in fact was verified before any such way existed. For if electrons of energy E can excite radiation of frequency E/h, and radiation of frequency E/h striking a piece of metal can elicit electrons of energy $h\ (E/h) - P$; then, if a target is bombarded with electrons, and another metal target is exposed to the radiation which emanates from the first one, the fastest of the electrons which escape from the second target will move with the same velocity and

[7] Gerlach regards this as the most accurate of all the methods for determining h, an opinion in which probably not all would concur. It has been maintained that the high-frequency limit, like the wavelength of maximum intensity in the X-ray spectrum, depends on the inclination of the X-ray beam to the exciting electron-stream. I do not know whether the experiments adduced in support of this claim have been adequately confuted.

the same energy as the electrons which strike the first one (minus the quantity P which, however, is immeasurably small and perfectly negligible in comparison with the energy of the electrons which excite ordinary X-rays). This fact emerged from a series of experiments which were performed by various people in the first decade of this century, the results of which were generally phrased somewhat in this way, "the energy of the secondary electrons depends only on the energy of the primary electrons, not on the nature of the material which the primary electrons strike or on that from which the secondary electrons issue, nor on the distance over which the X-rays travel." Upon these results Sir William Bragg based his corpuscular theory of X-rays; for (he argued) the most sensible interpretation of the facts is surely this, that some of the electrons striking the first target rebound with their full energy, and rebound again with their full energy from the second target, each of them carrying with it from the first to the second target a positive particle which neutralizes its charge over that part of its course, and so defeats all the methods devised to recognize a flying electron. Not many years later, Sir William cooperated in the slaying of his own theory, by developing the best of all methods for proving that X-rays are undulatory and measuring their wave-lengths; but it was only the imagery of the theory that perished, for its essence, the idea that the energy of the first electron travels as a unit or is carried as a parcel to the place where the second electron picks it up, had to be resurrected. All the mystery of the contrast between wave-theory and quantum-theory is implicit in this phenomenon, for which Sir William found an inimitable simile: "It is as if one dropped a plank into the sea from a height of 100 feet, and found that the spreading ripple was able, after travelling 1,000 miles and becoming infinitesimal in comparison with its original amount, to act upon a wooden ship in such a way that a plank of that ship flew out of its place to a height of 100 feet."

Among the radiations excited from a metal by electrons of a single energy E, there are many of which the frequencies differ from the interpreted frequency E/h, being lower. Among the electrons expelled from a metal by radiation of a single frequency ν, there are many of which the energies differ from the interpreted energy-value $h\nu$, being lower. These were accounted for by supposing that the electrons are troubled by repeated encounters with closely-crowded atoms. If then a metal vapor or a gas were bombarded with electrons or exposed to radiation, would all the excited radiation have a single frequency conforming to equation (3), would all the released electrons have a single energy conforming to equation (1)? One could not affirm this *a priori*, for a solid metal is not a collection of free atoms close together as a gas is an assemblage of free atoms far apart, but rather a structure of atoms which interfere with one another and are distorted, and there are many electrons in a solid of which the bonds and the constraints are very different from those by which the electrons of free atoms are controlled and vice versa. When a plate of sodium or a pool of mercury is exposed to a rain of electrons, not exceeding say 10 equivalent volts in energy, nothing apparent happens.[8] When the vapor of either metal is similarly exposed, the atoms respond in a manner from which they are inhibited, when they are bound together in the tight latticework of a solid or the promiscuous crowding of a liquid; and light is emitted.

The phenomena are clearest when the bombarded vapor is that of a volatile metal, such as mercury, sodium, or magnesium. The atoms in such vapors are not usually bound together two by two or in greater clusters, as they are in such gases as oxygen or hydrogen, of which the response to electron-impacts or to radiation is not quite understood to this day; and the first radiations which they emit are not in the almost inaccessible far ultra-violet, like those of the monatomic noble gases, but in the near ultra-violet or even in the visible spectrum. Dealing with such a vapor, I will say mercury for definiteness, one observes that so long as the energy of the bombarding electrons remains below a certain value, no perceptible light is emitted; but beyond, there is a certain range of energies, such that electrons possessing them are able to arouse one single frequency of radiation from the atoms. Ordinarily, as when a vapor is kept continuously excited by a self-sustaining electric discharge throughout it, the atoms emit a great multitude of different frequencies of radiation, forming a rich and complicated spectrum of many lines. But if the energy of the bombarding electrons is carefully adjusted to some value within the specified range, only one line of this spectrum makes its appearance; under the best of circumstances this single line may be exceedingly bright, so that the absence of its companions—some of which, in an ordinary arc-spectrum, are not much inferior to it in brightness—is decidedly striking. The one line which constitutes this *single-line spectrum* is the first line of the principal series in the complete arc-spectrum of the element; its wave length is (to take a few examples) 2536A for mercury, 5890 for sodium (for which it is a doublet), 4571 for magnesium.

[8] According to a very recent paper by C. H. Thomas, radiations from iron excited by electrons with as low an energy as some two or three equivalent volts have been detected.

Does this single line appear suddenly at a precise value of the energy of the impinging electrons? This question suggests itself, when one has already studied the excitation of X-rays from solids by electrons and the excitation of electrons from solids by light. Here again we meet that tiresome but ineluctable problem, as to what constitutes a *sudden* appearance, and how we should recognize it if it really occurred. The only consistent way to meet it (consistent, that is, with the ways already employed in the prior cases) would be to measure the intensity of the line for various values of the energy of the electrons, plot the curve, and decide whether or not it cuts the axis of abscissae at a sharp angle. This is in principle the same method as is used in determining whether a given X-ray frequency appears suddenly at a given value of the energy of the electrons bombarding a solid; the curves of Fig. 4 were so obtained. Attempting to apply this same method to such a radiation as 2,536 of mercury, one has the solitary advantage that the frequency of the light is sharp and definite (it is not necessary to cut an arbitrary band of radiations out of a continuous spectrum) and two great counteracting disadvantages: the intensity of the light cannot be measured accurately (one has to guess it from the effect upon a photographic plate) and the impinging electrons never all have the same energy. Owing probably to these two difficulties, there is no published curve (that I know of) which cuts down across the axis of abscissae with such a decisive trend as the curves of Figs. 2 and 4. Still it is generally accepted that the advent of the single line is really sudden. The common argument is, that one can detect it on a photographic film exposed for a few hours when the energy of the bombarding electrons is (say) 5 equivalent volts, and not at all on a plate exposed for hundreds of hours when the bombarding voltage is (say) 4.5 volts. In this manner the energy of the electrons just sufficient to excite 2536 of mercury has been located at 4.9 equivalent volts. Dividing this critical energy (expressed in ergs) by the frequency of the radiation, we get

$$(4.9e/300) \;/\; (c/.00002536) = 6.59 \cdot 10^{-27} \tag{5}$$

It agrees with the values of the constant which I designated by h in the two prior cases, and the data obtained with other kinds of atoms are not discordant. Gerlach arrives at $6.56 \cdot 10^{-27}$ as the mean of all values from experiments of this type upon many vapours. The evidence is not quite so strong as in the prior cases, but fortunately it is supplemented and strengthened by testimony of a new kind.

When electrons strike solids and excite X-rays, it is impossible to follow their own later history, or the adventures of a beam of radiation after it sinks into a metal. We have inferred that the electrons which collide with a piece of tungsten and disappear into it transfer their energy to X-rays, but the inference lacked the final support which would have been afforded by a demonstration of these very electrons, still personally present after the collision but deprived of their energy. Now when electrons are fired against mercury atoms, this demonstration is possible, and the results are very gratifying. I have already several times had occasion to remark, in this series of articles, that when an electron strikes a free atom of mercury, the result of the encounter is very different, according as its energy of motion was initially less than some 4.9 equivalent volts, or greater. In the former case, it rebounds as from an elastic wall, having lost only a very minute fraction of its energy, and this fraction spent in communicating motion to the atom; but in the latter case, it may and often does lose 4.9 equivalent volts of its energy *en bloc*, in a single piece as it were, retaining only the excess of its original energy over and above this amount. Thus if electrons of an energy of 4.8 equivalent volts are shot into a thin stratum of mercury vapor, nothing but electrons of that energy arrives at the far side; but if electrons of an only slightly greater energy, say 5.0 equivalent volts, are fired into the stratum, those which arrive at the far side will be a mixture of electrons of that energy, and very slow ones. The very slow ones can be detected by appropriate means, and the particular value of the energy of the bombarding electrons, at which some of them are for the first time transformed into these very slow ones, can be determined. Once more we meet that question as to whether the transformation does make its first appearance *suddenly*, but in this case the indications that it does are rather precise and easy to read. Furthermore it is possible to measure the energy of the slow electrons, and one finds that it is equal to the initial energy of the electrons, minus the amount 4.9 equivalent volts. (These measurements are not so exact as is desirable, and it is to be hoped that somebody will take up the task of perfecting them.)

We, therefore, see both aspects of the transaction which occurs when an electron whereof the energy is 4.9 equivalent volts, or greater, strikes a mercury atom. It loses 4.9 equivalent volts of energy, and we measure the loss; the atom sends forth radiation of a certain frequency, and no other; the atom does not send forth even this frequency of radiation, if none of the electrons fired against it has at least so much energy. We have already compared the energy transferred with the frequency radiated, and as in the case of X-rays excited from a solid target by very fast electrons, it is legitimate to say for these radiations which form the single-line spectra of metallic atoms, that

Excitation of the ray forming a single-line spectrum, by the collision of an electron against an atom, occurs as if *the energy in the radiation were concentrated in units of amount $h\nu$, and one such unit were created out of the total energy which the electron surrenders.*

There are yet several phenomena which I might treat by the same inductive method, arriving after each exposition at a Rule which would resemble one or the other of those which I have thus far written in italics; but it is no longer expedient, I think, to pass in each instance through the same elaborate inductive detour. These three phenomena which I have discussed already combine into an impressive and rather formidable obstacle to the classical manner of thinking. Here is a mercury atom, which receives a definite quantity of energy U from an electron, and distributes it in radiation of a definite frequency U/h. Here again is a multitude of atoms locked together into a solid, and when an electron conveys its energy U to the solid, it redistributes that energy in radiation of a definite frequency U/h. (It is true that many other radiations issue from the solid, but they are all explicable if one assumes that the electron may deliver over its energy in stages, and there is no radiation of the sort which would controvert the theory by virtue of its frequency exceeding U/h.) And when that radiation of frequency U/h in its turn strikes a metal, it is liable and able to release an electron from within the metal, conferring upon it an energy which is apparently equal to U. Apparently there is some correlation between an energy U and a frequency U/h, between a frequency ν and an energy $h\nu$. Apparently a block of energy of the amount U tends to pass into a radiation of the frequency U/h; apparently a radiation of the frequency ν tends to deliver up energy in blocks of the amount $h\nu$. The three italicized Rules coalesce into this one:

Photoelectric emission, and the excitation of X-rays from solids by electrons, and the excitation of single-line spectra from free atoms, occur as if *radiant energy of the frequency ν were concentrated into packets, or units, or corpuscles, of energy amounting to $h\nu$, and each packet were created in a single process and were absorbed in a single process.*

If the neutralizing *as if* were omitted, this would be the corpuscular theory *rediviva*. It is good policy to leave the *as if* in place for awhile yet. But conservatism such as this need not and should not deter anyone from using the idea as basis for every prediction that can be founded upon it, and testing every one of the predictions that can be tested by any possible way. Just so were the three phenomena cited in these Rules discovered. All of them involve either the emission or the absorption of radiation, and so do all the others which I could have quoted in addition, if this account had been written three years ago. Reserving to the end the one new phenomenon that transcends this limitation, I must explain the relation between this problem and the contemporary Theory of Atomic Structure.

The classical notion of a source of radiation is a vibrating electron. The classical conception of an atom competent to emit radiations of many frequencies is this: a family or a system of electrons, each electron remaining in an equilibrium-position so long as the system is not disturbed, one or more of the electrons vibrating when the system is jarred or distorted. A system with these properties would have to contain other things than electrons, otherwise it would fly apart; it would have to contain other things than particles of positive and particles of negative electricity intermixed, otherwise it would collapse together. One would have to postulate some sort of a framework, some imaginary analogue to a skeleton of springs and rods and pivots, to hold the electrons together in an ensemble able to vibrate and not liable to coalesce or to explode. This would not be satisfying, for in making atom-models one wants to avoid the elaborate machinery and in particular the non-electrical components; it would be much more agreeable to build an atom out of positive and negative electricity associated with mass, omitting all masses or structures not electrified. Nevertheless, if anyone had succeeded in devising a framework having the same set of natural frequencies as (say) the hydrogen atom exhibits in its spectrum—if anyone expert in dynamics or acoustics had been able to demonstrate that some peculiar shape of drumhead or bell, if anyone versed in electricity had been able to show that some particular arrangement of condensers and induction-coils has such a series of natural vibrations as some one kind of atom displays—then, it is quite safe to say, that framework or that membrane or that circuit would today be either the accepted atom-model, or at least one of the chief candidates for acceptance. Nobody ever succeeded in doing this; it is the consensus of opinion today that the task is an impracticable one.[9]

[9] It is difficult to put this statement into a more precise form. Rayleigh was of the opinion that the hydrogen spectrum could not be regarded as the ensemble of natural frequencies of a mechanical system, because it is the general rule for such systems that the *second* power of the frequency conforms to simple algebraic formulae, while in the hydrogen spectrum it is the *first* power for which the algebraic expression is simple. He admitted, however, that it was possible to find "exceptional" mechanical systems for which the first power of the frequency is given by a simple formula; which goes far to vitiate the conclusion. Another aspect of the formula (6) for

This set of natural frequencies which baffled all the efforts to explain it, the set constituting the two simplest of all spectra (the spectrum of atomic hydrogen and the spectrum of ionized helium), is given by the formula

$$\nu = R\left(\frac{1}{m^2} - \frac{1}{n^2}\right) \tag{6}$$

the different lines being obtained by assigning different integral values to the parameters m and n; lines corresponding to values of m ranging from 1 to 5 inclusive, and to values of n ranging from 2 to 40 inclusive, have already been observed, and there is no reason to doubt that lines corresponding to much higher values of m and n actually are emitted, but are too faint to be detected with our apparatus. The constant R has one value for hydrogen, another almost exactly four times as great for ionized helium.

Here, then, is the problem in its simplest presentation: How can a model for a hydrogen atom be constructed, which shall emit rays of the frequencies given by the formula (6), only these and no others? The obvious answer "By constructing a mechanical framework having precisely these natural frequencies" is practically excluded; it seems infeasible. Something radically different must be done. The achievement of Niels Bohr consisted in doing a radically different thing, with such a degree of success that the extraordinary divergence of his ideas from all foregoing ones was all but universally condoned. I do not know how Bohr first approached his theory; but it will do no harm to pretend that the manner was this.

Look once more at the formula for the frequencies of the hydrogen spectrum. It expresses each frequency as a difference between two terms, and the algebraic form of each term is of an extreme sim-

---

the hydrogen spectrum is this, that it specifies infinitely many frequencies within finite intervals enclosing certain critical values, such as R, $4R$, $9R$, and so forth. Poincaré is said to have proved that the natural frequencies of an elastic medium with a rigid boundary cannot display this feature, so long as the displacements are governed by the familiar equation $d^2\phi/dt^2 = k^2\nabla^2\phi$. For a membrane this equation is tantamount to the statement that the restoring-force acting upon an element of the membrane is proportional to the curvature of the membrane at that element. Ritz was able to show that the natural frequencies of a square membrane would conform to the formula (6), *if* the restoring-force upon each element of the membrane, instead of being proportional to the curvature of the membrane at that element, depended in an exceedingly involved and artificial manner upon the curvature of the membrane elsewhere. He apologized abundantly for the extraordinary character of the properties with which he had been obliged to endow this membrane, in order to arrive at the desired formula; but his procedure might have proved unsuspectedly fruitful, if Bohr's interpretation had not supplanted it.

plicity. Multiply now each member of the formula by h, that same constant h which we have encountered three times in the course of this article; and reverse the signs of the terms.[10] The formula becomes

$$h\nu = (-h\ R/n^2) - (-h\ R/m^2) \tag{7}$$

In the left-hand member there stands $h\nu$. The reader will have become more or less accustomed to the notion that, under certain conditions and circumstances of Nature, radiant energy of the frequency ν apparently goes about in packets or corpuscles of the amount $h\nu$; now and then, here and there, energy is absorbed from such radiation in such amounts, or energy is converted into such radiation in such amounts. *Suppose that this also happens when a hydrogen atom radiates*, whatever the cause which sets it to radiating. Then the left-hand member of the equation (5) represents the energy which the hydrogen atom radiates; so also does the right-hand member; but the right-hand member is obviously the difference between two terms; *these terms are respectively the energy of the atom before it begins to radiate, and the energy of the atom after it ceases from radiating.*

The problem of the hydrogen atom has now experienced a fundamental change. The proposal to make a mechanical framework, having the natural vibration-frequencies expressed by (6), has been laid aside. The new problem, or the new formulation of the old problem, is this: how can a model for a hydrogen atom be constructed, which shall be able to abide only in certain peculiar and distinctive states or shapes or configurations, in which various states the energy of the atom shall have the various values $-hR$, $-hR/4$, $-hR/9$, $-hR/16$, and so forth?

Bohr's own model has become one of the best-known and most-taught conceptions of the whole science of physics, in the twelve years of its public existence. He based it upon the conception, then rapidly gaining ground and now generally accepted, that the hydrogen atom is a microcosmic sun-and-planet system, a single electron revolving around a much more massive nucleus bearing an electric charge equal in magnitude and opposite in sign to its own. This is really a most unpromising conception, very ill adapted to the modification we need to make. We want an atom which shall be able to assume only those definite values of energy which were listed above: $-hR$, $-hR/4$, $-hR/9$ and the rest. Now the energy of this sun-and-planet atom depends on the orbit which the electron is describing.

[10] For the explanation of this rather confusing reversal, see my third article (page 278; or page 11 of the reprint).

If the energy may assume only those definite values, the electron may describe only certain definite orbits. But there is no obvious reason why the electron should not describe any of an infinity of other orbits, circular or elliptical. To consider only the circular orbits: if the atom may have no other values of energy than $-hR$, and $-hR/4$, and $-hR/9$, and the rest of the series, then it may not revolve in any other circular orbits than those of which the radii are $e^2/2hR$, and $e^2/2(hR/4)$, and $e^2/2(hR/9)$, and so forth; but why just these? What prevents it from revolving in a circular orbit of radius $e^2/2(hR/2)$, or any other value not in the series? And for that matter how can it revolve in a closed orbit at all, since according to the fundamental notions of the electromagnetic theory it must be radiating its energy as it revolves, and so must sink into the nucleus in a gradually narrowing spiral?

Bohr did not resolve these difficulties, and no one has ever resolved them except by ignoring them. The customary procedure is to select some common feature of these permitted orbits, and declare that it is this feature which makes these orbits permissible, and forbids the electron to follow any other. For example, there is the fact that the angular momentum of the electron in any one of the permitted circular orbits is an integer multiple of the constant quantity $h/2\pi$, h being the same constant as we have met hitherto, which is hardly an accidental coincidence. If one could only think of some plausible reason why an electron should want to revolve only in an orbit where it can have some integer multiple of $h/2\pi$ for its angular momentum, and should radiate no energy at all while so revolving, and should refuse to revolve in an orbit where it must have a fractional multiple of $h/2\pi$, the model would certainly be much fortified. Failing this it is necessary to put this assertion about the angular momentum as a downright assumption, in the hope that its value will be so great and its range of usefulness so widespread that it will commend itself as an ultimate basic principle such as no one thinks of questioning. So far this hope has not been thoroughly realized. On the one hand, Sommerfeld and W. Wilson did succeed in generalizing it into a somewhat wider form, and using it in this wider form they explained the fine structure of the lines of hydrogen and ionized helium, and Epstein explained the effect of an electric field upon these lines. These are truly astonishing successes, and no one, I think, can work through the details of these applications to the final triumphant comparisons of theory with experiment, and not experience an impression amounting almost or quite to conviction. Yet on the other hand this generalization does not account for the frequencies forming the spectra of other elements.[11] There is the spectrum of neutral helium, for example, and the spectrum of sodium, and the spectrum of mercury; in each of these there are series of lines, of which the frequencies are clearly best expressed each as the difference between a pair of terms, and these terms should be the energies of the atom before and after radiating. But we have not the shadow of an idea what the corresponding configurations of the atom are; it may be that the outermost electron has certain permissible orbits, but we do not know what these orbits are like nor what common feature they possess.

Is it then justifiable to write down a Rule such as this: *the frequencies of the rays which free atoms emit are such as to confirm the idea that radiant energy of the frequency ν is emitted in packets or corpuscles of the amount $h\nu$*? Very few men of science, I imagine, would hesitate to approve this. However one may fluctuate in his feelings about Bohr's model of the atom, there always remains that peculiar relation among the frequencies emitted by the hydrogen atom, which is so nearly copied by analogous relations in the spectra of other elements. When one has once looked at the general formula

$$h\nu = \left(-\frac{hR}{n^2}\right) - \left(-\frac{hR}{m^2}\right) \tag{7}$$

and has once interpreted the first term on the right as the energy of an atom before radiating, the second term on the right as the energy of the atom after radiating, and the quantity $h\nu$ as the amount of the packet of energy radiated, it is very difficult to admit that this way of thinking will ever be superseded; particularly when one remembers the auxiliary facts, such as that fact about the electrons transferring just 4.9 equivalent volts to the mercury atoms which they strike, no more and no less. Analyzing the mercury spectrum in the same way as the hydrogen spectrum was analyzed, we find the frequencies expressible as differences between terms; interpreting the terms as energy-values, we find that between the normal state of the mercury atom and the next adjacent state, there is a difference in energy of 4.9 equivalent volts, and between this and the next adjacent state there is a further difference of 1.8 volts. This then is the reason why an electron with less than 4.9 equivalent volts of

[11] The mathematical experts who have laboured over the theory of the helium atom (two electrons and a nucleus of charge $+2e$) seem to have convinced themselves that the features which distinguish the permitted orbits of the electrons in this atom, whatever they may be, are definitely not the same features as distinguish the permitted orbits of the electron in the hydrogen atom. This cannot be said with certainty for any other atom.

energy can communicate no energy at all to a mercury atom; and an electron with 5 or 6 equivalent volts of energy can transfer only 4.9 of them. It is conceivable that other conditions may be found to govern the orbits of the electrons, so that the atoms shall have only the prescribed energy-values and no others; it is even conceivable that the conception of electron-orbits may be discarded; but the interpretation of the terms in the formula (7) as energies will, in all human probability, be permanent.

The foregoing Rule is thus very strongly based; but let us nevertheless rephrase it in a somewhat milder form as follows: *The idea that radiant energy of frequency ν is emitted in packets of the amount $h\nu$, and the contemporary theory of atomic structure, between them give a attractive and appealing account of spectra in general, and a convincingly exact explanation of two spectra in particular.*

But what has happened meanwhile to the Vibrator, to the oscillating electron, to the postulated electrified particle of which the vibrations caused light-waves to spread out from around it like sound waves from a bell? It has disappeared from the picture; or rather, since the attempt to account for the frequencies of a spectrum as the natural frequencies of an elastic framework was abandoned, no one has tried to re-insert it. But there are some who will never be quite happy with any new conception, until the vibrator is established as a part of it.

Ionization, the total removal of an electron from an atom, affords another chance to see whether radiant energy behaves as though it could be absorbed only in complete packets of amount $h\nu$. That it requires a certain definite amount of energy to deprive an atom of its loosest electron, an amount characteristic of the atom, may now be regarded as an experimental result quite beyond question, and not requiring the support of any special theory. Thus, a free-flying electron may remove the loosest electron from a free mercury atom which it strikes, if its energy amounts to 10.4 equivalent volts, not less; or the loosest electron from a helium atom if its energy amounts to at least 24.6 equivalent volts. If radiant energy of frequency ν goes about in parcels of magnitude $h\nu$, the frequency of a parcel which amounts just exactly to 10.4 equivalent volts is $\nu_o = 2.53.10^{16}$, corresponding to a wave length of 1188A. Light of inferior frequency should be unable to ionize a mercury atom; light of just that frequency should just be able to ionize it; light of a higher frequency ν should be able to ionize the atom, and in addition confer upon the released electron an additional amount of kinetic energy equal to $h\ (\nu - \nu_o)$. The same could be said, with appropriate numerical changes, for every other kind of atom. Of all the phenomena which might serve to illuminate this difficult question of the relations between radiation and atoms, this is the one which has been least studied. The experimental material is scanty and dubious. There is no reason to suppose that light of a lower frequency than the one I have called ν_o is able to ionize; but it is not clear whether perceptible ionization commences just at the frequency ν_o, although it has been observed at frequencies not far beyond. The energy of the released electrons has not been measured.

The removal of deep-lying electrons, the electrons lying close to the nuclei of massive atoms, is much better known; and the data confirm in the fullest manner the idea that radiant energy of the frequency ν is absorbed in units amounting to $h\nu$. When a beam of X-rays of a sufficiently high frequency is directed against a group of massive atoms, various streams of electrons emanate from the atoms, and the electrons of each stream have a certain characteristic speed. The kinetic energy of each electron of any particular stream is equal to $h\nu$, minus the amount of energy which must be spent in extracting the electron from its position in the atom; for this amount of energy is independently known, being the energy which a free-flying electron must possess in order to drive the bound electron out of the atom, which is measurable and has been separately measured. Here again I touch upon a subject which has been treated in an earlier article of this series—the second—and to prevent this article from stretching out to an intolerable length, I refrain from further repetition of what was written there. The analogy of this with the photoelectric effect will escape no reader. Here as there, we observe electrons released with an energy which is admittedly not $h\nu$, but $h\nu$ minus a constant; the idea that this constant represents energy which the electrons have already spent in escaping, in one case through the surface of the metal and in the other case from their positions within atoms, is fortified by independent measurements of these energies which give values agreeing with these constants.

We have considered various items of evidence tending to show that radiant energy is born, so to speak, in units of the amount $h\nu$, and dies in units of the amount $h\nu$. Whether energy remains subdivided into these units during its incarnation as radiation remains unsettled; to settle this question absolutely, one would have to devise some way of testing the energy in a beam of radiation, otherwise than by absorbing it in matter; and such a way has not yet been discovered. There is, however, another quality which radiant energy possesses.

Conceive a stream of radiation in the form of an extremely long train of plane waves, flowing against a blackened plate facing normally against the direction in which they advance, which utterly absorbs them. This wave-train shall have an intensity I; by which it is meant, that an amount of energy I appears, in the form of heat, in unit area of the blackened plate in unit time. Furthermore, the radiation is found to exert a pressure p against the blackened plate; by which it is meant, that unit area of the plate (or the framework upholding it) acquires in unit time an amount of momentum p. According to the classical electromagnetic theory, verified by experience, p is equal to I/c. Unit area of the plate acquires, in unit time, energy to the amount I and momentum to the amount I/c.

Where is this energy, and where is this momentum, an instant before they appear in the plate? One might say that they did not exist, that they had vanished at the moment when the radiation left its source, not to reappear until it arrived at the plate; but such an answer would be contrary to the spirit of the electromagnetic theory, and we have long been accustomed to think of the energy as existing in the radiation, from the moment of its departure from the source to the moment of its arrival at the receiver; the term "radiant energy" implies this. Momentum has the same right to be conceived as existing in the radiation, during all the period of its passage from source to receiver. In the system of equations of the classical electromagnetic theory, the expression for the stream of energy through the electromagnetic field stands side by side with the expression for the stream of momentum flowing through the field. If the second expression is not so familiar as the first, and the phrase "radiant momentum" has not entered into the language of physics together with "radiant energy," the reason can only be that the pressure which light exerts upon a substance is very much less conspicuous than the heat which it communicates, and seems correspondingly less important,—which is no valid reason at all. Radiant energy and radiant momentum deserve the same standing; it is admitted that the energy I is the energy which is brought by the radiation in unit time to unit area of the plate which blocks the wave-train, and with it the radiation brings momentum I/c in unit time to unit area of the plate. The density of radiant energy in the wave-train is obviously I/c, the density of radiant momentum is I/c^2.

Now let that tentative idea, that radiant energy of the frequency ν is emitted and absorbed in packets of the amount $h\nu$, be completed by the idea that these packets travel as entities from the place of their birth to the place of their death. Let me now introduce the word "quantum" to replace the alternative words *packet*, or *unit*, or *corpuscle*; I have held to these alternative words quite long enough, I think, to bring out all of their connotations. Then the energy I is brought to unit area of the plate, in unit time, by $I/h\nu$ of the quanta; which also bring momentum amounting to I/c. Shall we not divide up the momentum equally among the quanta as the energy is divided, and say that *each is endowed with the inherent energy* $h\nu$ *and with the inherent momentum* $h\nu/c$?

The idea is a fascinating one, but not so easy to put to the trial as one might at first imagine. None of the phenomena I have described in the foregoing pages affords any means of testing it. In studying the photoelectric effect, we concluded that each of the electrons released from an illuminated sodium plate had received the entire energy of a packet of radiation; but this does not imply that each of them had received the momentum associated with that energy; the momentum passed to the plate, to the framework supporting it, eventually to the earth. The same statement holds true for the release of electrons from the deep levels of heavy atoms, such as de Broglie and Ellis observed. Even if the same experiments should be performed on free atoms, as for example on mercury vapor, no clear information could be expected; for the momentum of the absorbed radiation may divide itself between the released electron and the residuum of the atom, and this last is so massive that the speed it would thus acquire is too low to be noticed. Only one way seems to be open; this is, to bring about an encounter between a quantum of radiation and a free electron, so that whatever momentum and whatever energy are transferred to the electron must remain with it, and cannot be passed along to more massive objects where the momentum, so far as the possibility of observing it goes, is lost. *A priori* one could not be certain that even this way is open; radiation might ignore electrons which are not tightly bound to atoms.

Arthur H. Compton, then of Washington University, is the physicist whose experiments were the first that clearly and strikingly disclosed such encounters between quanta of radiation and sensibly free electrons. Others had observed the effect which reveals them, but his were the first measurements accurate enough for inference. Unaware at the moment of the meaning of his data, he realized it almost immediately afterward, and so established the fact and the explanation both—a twofold achievement of a very unusual magnitude, whence the phenomenon received the name of "Compton effect" by a universal acceptance, and deservedly.

What Compton observed was not the presence of electrons possessed of momentum acquired from radiation—these electrons were however to be discovered later, as I shall presently mention—but the presence of radiation of a new sort, come into being by virtue of the encounters between the original radiation and free electrons. We have not encountered anything of this sort heretofore. When a quantum of radiant energy releases an electron from an atom, it dies completely and confers its entire energy upon the electron. The disposal of its momentum gives no trouble, for as I have mentioned the atom takes care of that. When the electron is initially free, and there is no atom to swallow up the momentum of the radiation, it cannot be ignored in this simple fashion. For if the quantum did utterly disappear in an encounter with a free electron, the velocity which the electron acquired would have to be such that its kinetic energy and its momentum were separately equal to the energy and momentum of the quantum; but these distinct two conditions would generally be impossible for the electron to fulfil. Hence in general, a quantum possessed of momentum cannot disappear by the process of transferring its energy to a free electron, whatever may be the case with an electron bound to a massive atom. This reflection might easily have led to the conclusion that radiation and free electrons can have nothing to do one with the other.

What actually happens is this: the energy and the momentum of the quantum are partly conferred upon the electron, the residues of each go to form a new quantum, of lesser energy and of lesser and differently-directed momentum, hence lower in frequency and deflected obliquely from the direction in which the original quantum was moving. The encounter occurs much like an impact between two elastic balls; what prevents the analogy from being perfect is, that when a moving elastic ball strikes a stationary one, it loses some of its speed but remains the same ball, whereas the quantum retains its speed but changes over into a new and smaller size. It is as though a billiard-ball lost some of its weight when it touched another but rolled off sidewise with its original speed. I do not know what this innovation would do to the technique of billiards, but it would at all events not make technique impossible; the result of an impact would still be calculable, though the calculations would lead to a new result. The rules of this microcosmic billiard-game in which the struck balls are electrons and the striking balls are quanta of radiant energy are definite enough to control the consequences. The rules are these:

Conservation of energy requires that the energy of the impinging quantum, $h\nu$, be equal to the sum of the energy of the resulting quantum, $h\nu'$, and the kinetic energy K of the recoiling electron. For this last quantity the expression prescribed by the special relativity-theory [12] is used, viz.

$$K=mc^2\left(\frac{1}{\sqrt{1-\beta^2}}-1\right)$$

in which m stands for the mass of the electron and $c\beta=v$ for its speed. The equation of conservation of energy is then

$$h\nu=h\nu'+mc^2\left(\frac{1}{\sqrt{1-\beta^2}}-1\right). \qquad (8a)$$

Conservation of momentum requires that the momentum of the impinging quantum be equal to the sum of the momenta of the resulting quantum and the recoiling electron. Momentum being a vector quantity, this rule requires three scalar equations to express it, which three may be reduced to two if we choose the x-axis to coincide with the direction in which the impinging quantum travels, and the y-axis to lie in the plane common to the paths of the recoiling electron and the resulting quantum. Designate by ϕ the angle between the paths of the impinging quantum and the recoiling electron; by θ the angle between the paths of the two quanta. The magnitude of the momentum-vector is, by the special relativity-theory, $mv/\sqrt{1-\beta^2}$. Conservation of momentum then requires:

$$\begin{aligned} h\nu/c&=(h\nu'/c)\cos\theta+\frac{mv}{\sqrt{1-\beta^2}}\cos\phi,\\ O&=(h\nu'/c)\sin\theta+\frac{mv}{\sqrt{1-\beta^2}}\sin\phi. \end{aligned} \qquad (8b)$$

Eliminating ϕ and v between these three equations, we arrive at this relation between ν and ν', the frequencies of the impinging quantum and the recoiling quantum—or, as I shall hereafter say, between the frequencies of the primary X-ray and the scattered X-ray—and the angle θ between the directions of the primary X-ray and the scattered X-ray:

$$\frac{\nu'}{\nu}=\frac{1}{1+\dfrac{h\nu}{mc^2}(1-\cos\theta)}. \qquad (9)$$

[12] If the reader prefers to use the familiar expressions $\frac{1}{2}mv^2$ for the kinetic energy and mv for the magnitude of the momentum of the electron, he will arrive at a formula for ν^1 which, while apparently dissimilar to (9) and not so elegant, is approximately identical with it when v is not too large—or, which comes practically to the same thing, when $h\nu$ is small in comparison with mc^2; a condition which is realized for all X-rays now being produced.

The relation between λ' and λ, the wavelengths of the primary beam and of the scattered beam, is still simpler, being

$$\lambda' - \lambda = \frac{h}{mc}(1 - \cos \theta). \tag{10}$$

The intrusion of this angle θ into the final equation may seem to contradict my earlier statement that the results of the impact are calculable; for it is true that there are not equations enough to eliminate θ, and yet I have offered no additional means of calculating it. In fact it cannot be calculated with the data at our command. All that we are able to say is that *if* the resulting quantum goes off in the direction θ, then its frequency is given by (9). What determines θ in any particular case? Reverting to the image of the billiard-balls, it is easy to see that the direction in which the rebounding ball rolls away depends on whether it gave a central blow, or a glancing blow, or something in between, to the initially stationary ball. If we knew just which sort of a blow was going to be given, we could calculate θ; otherwise we can only apply our conditions of conservation of energy and conservation of momentum to ascertain just how much of its energy the rebounding ball retains when θ has some particular value, and then produce—or, if we cannot produce at will, await—a collision which results in that value, and make our comparison of experiment with theory. So it is in this case of the rebounding quantum. When a beam of primary electrons is scattered by encountering a piece of matter, some quanta rebound in each direction, and all the values of θ are represented. We cannot know what determines the particular value of θ in any case; but we can at least select any direction we desire, measure the frequency of the quanta which have rebounded in that direction, and compare it with the formula. Fig. 6 is a diagram illustrating these relations.[13]

The comparison, which has now been made repeatedly by Compton, repeatedly by P. A. Ross, and once or oftener by each of several other physicists—notably de Broglie in Paris—is highly gratifying. The value of the frequency-difference between the primary X-rays and the scattered X-rays, that is to say, between the impinging quanta and the rebounding quanta, is in excellent accord with the formula, whether the measurements be made on the quanta recoiling at 45°, at 90° or at 135°, or at intermediate values of the angle θ. The method consists in receiving the beam of scattered X-rays into an X-ray spectroscope, whereby it is deflected against an ionization-chamber or a photographic plate at a particular point, of which the location is the measure of the wave-length. An image can be made on the same plate at the point where the beam would have struck it, if it had retained the frequency of the primary beam. The two images then stand sharply and widely apart. Indeed it is not necessary to make a special image to mark the place on the plate where a scattered beam of unmodified wave-length would fall, for there

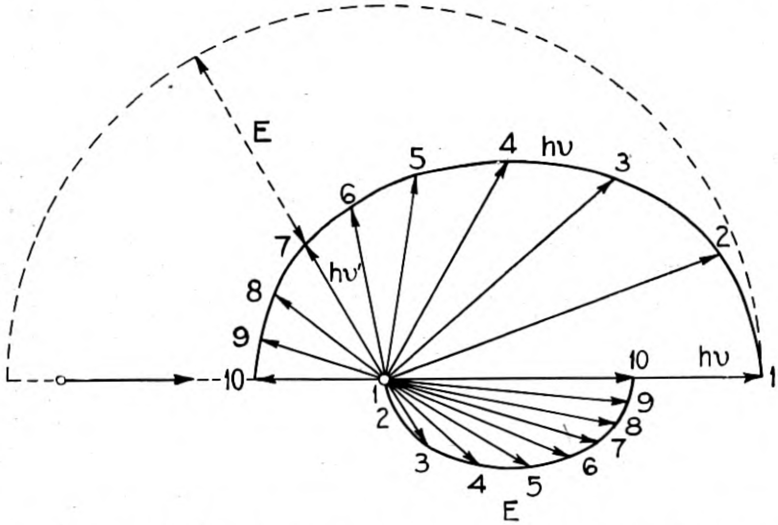


Fig. 6—Diagram showing the energy-relations ensuing upon an impact between a quantum and a free electron. (After Debye.) See footnote 13

nearly always is such a beam and such an image. A plausible explanation is easy to find; one has only to assume that the quanta composing this beam have rebounded from electrons so rigidly bound into atoms that they did not budge when the impinging quanta struck them, and these were reflected as from an immovable wall.[14]

[13] The diagram in Fig. 6 is designed to illustrate the relations between the energy of the primary quantum (radius of the dotted semicircle), the energy of the rebounding quantum (radius of the upper continuous curve), and the energy of the recoiling electron (radius of the lower continuous curve). Thus the two arrows marked with a 5 are proportional respectively to the energies of the secondary quantum and of the recoiling electron, when the encounter has taken place in such a fashion that the angle θ is equal to the angle between the arrow 10 and the upper arrow 5. In the same case, the angle between arrow 10 and lower arrow 5 is equal to ϕ of the equations (9).

[14] As a matter of fact we have no independent means of knowing that the recoiling electrons are initially free, or that the scattered beam with the modified frequency originates from collisions of primary quanta with initially free electrons; we know only that the frequency of the scattered quanta is such as would be expected if little or no energy is spent in freeing the electrons, and little or no momentum is transferred otherwise than to the electrons—which, of course, is not quite the same

In the photographs which I reproduce,[15] the imprints of these two beams stand side by side. In the first of them, Fig. 7, the spectrum of the primary rays is specially depicted on the upper half of the plate; one sees the α, β, and γ lines of the K-series of molybdenum, three lines (the first a doublet) of which the wavelengths are respec-

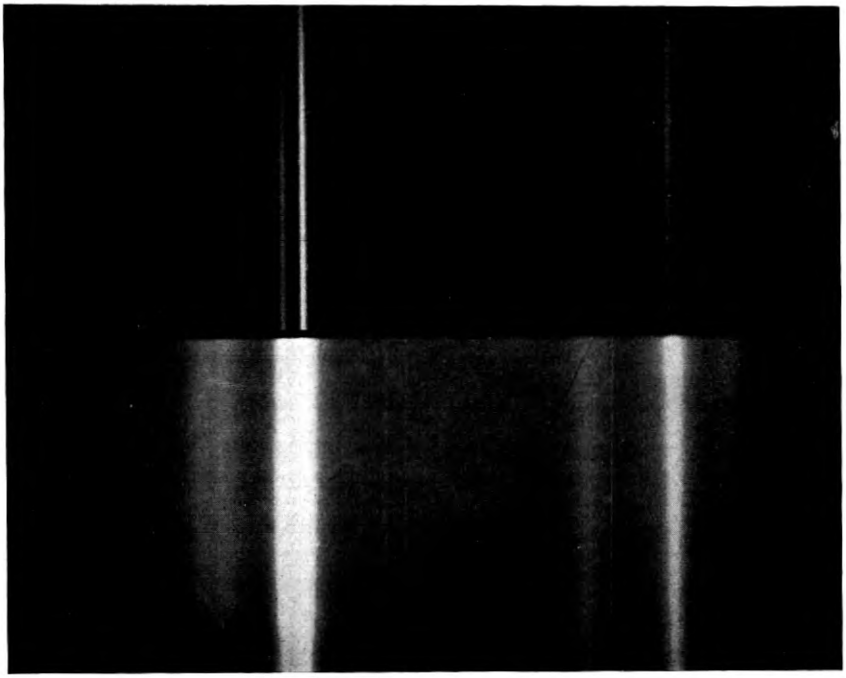


Fig. 7—Above, the K-spectrum of molybdenum (α-doublet, β-line, γ-line from left to right); below, the spectrum of this same radiation after scattering at 90° from aluminum (each line doubled). (P. A. Ross)

tively .710 − .714A, .633A, .618A. Below, the spectrum of the secondary rays scattered at the angle θ is spread out; to each of the primary rays there corresponds a scattered ray of the same wavelength, and beside it another ray of which the wavelength exceeds that of its companion by the required amount.

thing. The Compton effect has been demonstrated only where there are electrons associated with atoms. It may be that the rebound occurs only from an electron which is connected to an atom by some peculiar liaison, weak so far as the energy required to break it is concerned, but able to control the response of the electron to an impact. Something of this sort may have to be assumed to explain why the effect is apparently not greater for conductive substances than for insulating ones and is certainly feebler for massive atoms with numerous loosely-bound electrons than for light atoms with few.

[15] I am indebted to Professor Ross for these photographs.

Another series of photographs, in Fig. 8, shows the two scattered rays produced when a beam of the $K\alpha$-radiation of molybdenum falls upon various scattering substances: carbon (the sixth element of the periodic table), aluminium (the thirteenth), copper (the twenty-ninth), and silver (the forty-seventh). The relative intensity of the two rays—that is to say, the proportion between the number of quanta which rebound as from free electrons, and the number of quanta which recoil as from immobile obstacles—varies in a curious manner

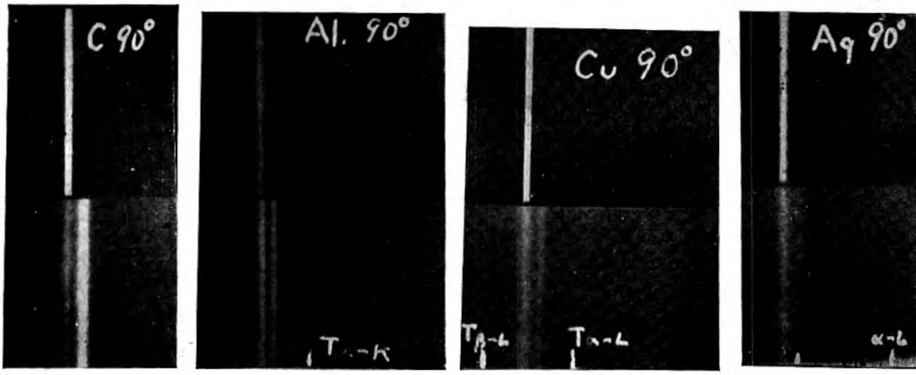


Fig. 8—Above, the $K\ \alpha$-line of molybdenum; below, the same radiation after scattering at 90° from carbon, aluminium, copper and silver. (P. A. Ross)

from one of these elements to another. Most of the quanta scattered by lithium undergo the alteration in wavelength which we have calculated; nearly all of the quanta scattered by lead emerge with the same frequency as the incident quanta. Apparently, the heavier the atoms of a substance are, the less conspicuous does Compton's effect become. Further, the relative intensity of the two rays assumes different values for one and the same substance, depending on the direction of scattering. This is illustrated in Fig. 9, the curves of which may be interpreted as graphical representations of photographs like those of the foregoing Figure, the ordinate standing for the density of the image on the photographic plate. (Actually, the ordinate stands for a quantity which is much more nearly proportional to the true intensity of the rays—that is, the amount of ionization which they produce in a dense gas.) These curves show, in the first place, that the separation between the two scattered rays has the proper theoretical values at the angle 45°, at 90°, and at 135°; in the second place, among the quanta scattered at 45°, those that retain the primary wavelength are more abundant than the altered quanta, while among the quanta scattered at 135° the modified ones have the predominance. Why the relative commonness of these two kinds of scattering, of these two modes of interaction between quanta

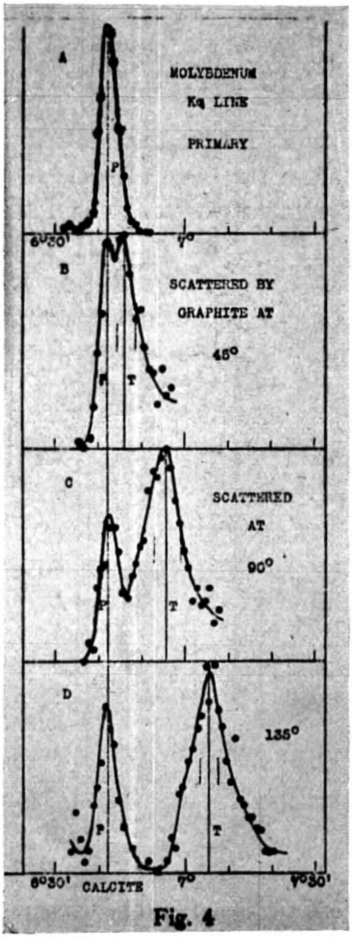


Fig. 9—The modified and unmodified scattered rays, at various inclinations, recorded by the ionization-chamber method. The vertical line T represents the position calculated from (9) for the modified ray. (A. H. Compton, *Physical Review*)

and matter, should depend on the substance and on the angle θ is a deeper question than any we have considered.

The recoiling electrons also have been detected; and Figs. 10 and 11, which are photographs of the trails left by flying electrons as they

Fig. 10—Trails of recoiling electrons, mingled with long sinuous trails of electrons ejected from atoms by totally-absorbed quanta. (C. T. R. Wilson, *Proceedings of the Royal Society*)

proceed through air supersaturated with water vapor, shows evidence for these.[16] The long sinuous trails are those of fast electrons, which were liberated from their atoms by high-frequency quanta proceeding across the gas; each of these electrons possesses the entire energy of a vanished quantum (minus such part of it as was sacrificed when

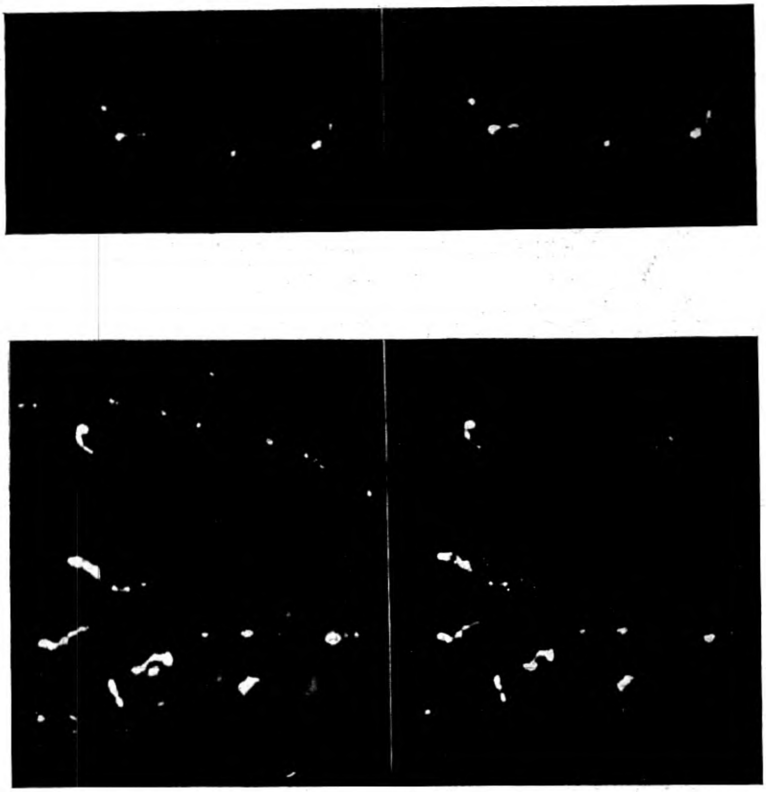


Fig. 11—Trails of recoiling electrons (C. T. R. Wilson, *Proceedings of the Royal Society*)

the electron emerged from its atom). The small slightly-elongated comma-like "blobs", the "fish tracks" as C. T. R. Wilson called them, are the trails of very slow electrons—these are the electrons from which quanta rebounded, transferring in the rebound a little of their energy and a little of their momentum. These appear only when the frequency of the X-ray quanta exceeds a certain minimum amount —a circumstance which, combined with others, shows that the com-

[16] I am indebted to Professor C. T. R. Wilson and to the Secretary of the Royal Society for permission to reproduce these photographs.

monness of the Compton effect depends not merely on the nature of the atoms and on the angle at which the scattering is observed, but also upon the frequency of the radiation. High-frequency quanta are liable to rebound in the manner prescribed by Compton's assumptions, but low-frequency quanta are not. Light of the visible spectrum suffers no change in wavelength when it is scattered.

Must we now concede that radiant energy travels about through space in the form of atom-like units, of corpuscles, of *quanta* every one of which, for a radiation of a specific frequency ν, possesses always the same energy $h\nu$ and always the same momentum $h\nu/c$? How indeed can we longer avoid admitting it? The phenomena which I have cited do certainly seem to close the case beyond any possibility of reopening it. Yet they might be interpreted in another way—a way which will probably seem extremely elaborate and artificial to the reader, a way which will seem like a mere excuse to avoid a simple and satisfying explanation; and yet this would not be sufficient to condemn it utterly. We might lay the whole blame and burden for all these "quantum" phenomena upon the atom. We might say that there is some mysterious mechanism inside every atom, which constrains it never to emit radiation of a frequency ν unless it has a quantity of energy $h\nu$ all packed up and ready to deliver, and never to absorb radiation of a frequency ν unless it has a special storeroom ready to receive just exactly the quantity of energy $h\nu$. This indeed is not a bad formulation of Bohr's theory of the atom. It would be necessary to go much further, and to say that not only every atom, but likewise every assemblage of atoms forming a liquid or a solid body, contains such a mechanism of its own; for the phenomena which I have called the "photoelectric effect" and the "inverse photoelectric effect" are qualities not of individual atoms, but of pieces of solid metal.[17] And it would be necessary to go much further yet, and make mechanisms to account for the transfer of momentum from radiation to electrons.

Yet even this would not be sufficient; for the most surprising and inexplicable fact of all is still to be presented. Here is the crux of the great dilemma. Imagine radiation of the frequency ν emerging from an atom, for a length of time determined by the condition that

[17] It was formerly contended that this explanation, while applicable to the behavior of free atoms which respond only to certain discrete frequencies, would not avail for a solid substance like sodium which delivers up electrons with energy $h\nu$, whatever the frequency ν may be. This contention, however, is probably not forcible, as it can be supposed that the solid has a very great number of natural frequencies very close together. This in fact was the inference from Epstein's theory of the photoelectric effect.

the total energy radiated shall be $h\nu$ exactly. According to the wave-theory, it emerges as a spherical wave-train, of which the wave-fronts are a series of expanding spheres, widening in all directions away from the atom at their common centre. Place another atom of the same kind some little distance away. Apparently it can absorb no radiant energy at all, unless it absorbs the whole amount $h\nu$ radiated from the first atom. But how can it do this, seeing that only a very small portion of each wavefront touched it or came anywhere near it, and much of the radiant energy went off from the first atom in a diametrically opposite direction? How can it reach and suck up all the energy from the entire wavefront, so little of which it actually intercepts? And the difficulty with the momentum is even greater.

But, of course, this experiment is unrealizable. In any laboratory experiment, there are always great multitudes of radiating atoms close together, and the atoms exposed to the radiation are bathed in myriads of wave-trains proceeding from myriads of sources. Does then the atom which absorbs the amount $h\nu$ of energy take it in little bits, one from this wavetrain and another from that, until the proper capital is laid up? But if so, it surely would require some appreciable time to gather up the separate amounts. According to the classical electromagnetic theory, a bound electron placed in a wavetrain of wavelength λ will gather up energy from an area of each wavefront, of the order of magnitude of the quantity λ^2. Hence we should not expect that the exposed atom would finish the task of assembling the amount of energy $h\nu$ from the various wavetrains which pass by it, until the lapse of a time-interval sufficient for so much energy to flow against a circle of the area λ^2, set up facing the rays at the point where the atom stands. Set up a mercury arc, or better still, an X-ray tube, and measure the intensity of the radiation from it at various distances. You will easily find a position sufficiently near to it for convenience, and yet sufficiently far from it, so that if a circular target of this area were set in that position, the radiant energy falling upon it would not mount up in one minute—nor in one day—nor in one year, to the amount $h\nu$. Yet cover the source of rays with a shutter, and then put a piece of matter in that position, and then lift the shutter; and you will not have to wait a year, nor a day, nor a minute, for the first electron which emerges from the matter with a whole quantum of energy; it will come out so quickly that no experimenter has, as yet, demonstrated a delay. What possible assumptions about the structure of the *atom* can account for this?

More and more the evidence is piled up to compel us to concede that radiation travels around the world in corpuscles of energy $h\nu$ and momentum $h\nu/c$, which never expand, or at all events always remain small enough to be swallowed up in one gulp by an atom, or to strike an electron with one single concentrated blow.

But it is unfair to close the case without pleading once more the cause of the undulatory theory—the more so because, in the usual fashion, I have understated the old and presumptively familiar arguments in its favor, and given all the advantages to the arguments of the opposition, which still have the force and charm of novelty. Furthermore, I may have produced the impression that the conception of the quantum actually unites the corpuscular theory with the wave-theory, mitigating discord instead of creating it. Why are we not really voicing a perfectly competent wave-theory of light, when we imagine wave-trains limited both in length and in breadth, so narrow that they can dive into an atom, but so long that they contain $h\nu$ of energy altogether? *filamentary* wave-trains, so to speak, like the tracing of a sine-wave in chalk upon a blackboard, or the familiar picture of a sea-serpent?

Well, the difficulty is that the phenomena of interference and of diffraction, which are the basis of the wave-theory, imply that the wave-trains are broad, that they have a considerable cross-sectional area; these phenomena should not occur, if the wave-trains were filaments no thicker than an atom, or even so wide that their cross-sectional area amounted to λ^2. Let me cite one or two of these phenomena, in tardy justice to the undulatory theory, as a sort of a makeweight to all the "quantum" phenomena I have described. Imagine an opaque screen with a slit in it; light flows against the screen from behind, some passes through the slit. The slit may be supposed to be half a millimetre wide, or even wider. If light consists of quanta only as thick as an atom, or even as thick as the wave-length of the light, they will shoot through the slit like raindrops or sand-grains through a wide open skylight. If they are all moving in parallel directions before they reach the slit, they will continue so to move after they pass through it—for how shall they know that the slit has any boundaries, since they are so small and the slit is so large? The beam of light which has passed through the slit will always retain the same cross-section as the slit. But we know that in truth the beam widens after it goes through the slit, and it develops a peculiar distribution of intensity which is accurately the same as we should expect, if the wavefront is *wider* than the slit—so much wider, that the slit cuts a piece out of it, which piece spreads outwards independently in its own fashion.[18]. Therefore the quantum must be wider than the widest slit which displays clear diffraction-phenomena—and this makes it at least a millimetre wide! But this is not the limit! Cut another slit in the screen, parallel to the first one, a distance *d* away from it. Where the widening diffracted light-beams from the two slits interpenetrate one another, they will produce interference-patterns of light and shade, accurately the same as we should expect if the wavefront is wider than the distance *d*. The quantum must therefore be wider than the greatest distance between two slits, the light-beams passing through which are able to interfere with one another. The slits may be put quite far apart, and the light-beams brought together by systems of prisms and mirrors. This is the principle of Michelson's famous method of determining the diameters of stars. He obtained interference fringes when the two beams of light were taken from portions of the wavefront *twenty feet apart!*[19]

Therefore the quantum is twenty feet wide! This is the object from which an atom one ten-millionth of a millimetre wide can suck up all its energy! this is what enters as a unit into collision with an electron ten thousandfold smaller yet!

The evidence is now before the reader; not the entire evidence for either of the two conceptions of radiation, but, I think, a fair sampling for both. If either view has been inequitably treated, it is the undulatory theory which has been underrated; for, as I have said already but cannot say too often, the evidence that light partakes of the nature of a wave-motion is tremendously extensive and tremendously compelling; it seems the less powerful only because it is so thoroughly familiar, and through much repetition has lost the force of novelty. Still, it is not necessary to hold all the relevant facts continually in mind. If one could reconcile a single typical fact of the one sort, such as the interference between beams of light brought together from parallel courses far apart, with a single outstanding fact of the other sort, such as the instantaneous emergence of electrons with great energy from atoms upon which a feeble beam of light has only just been directed—if one could unify two such phenomena as these, all of the others would probably fuse spontaneously into a harmonious system. But in thinking about these things, there is one more all-important

[18] One might, of course, inquire, why should a *piece* of the wavefront of a quantum, cut out of it by the edges of a slit, expand after passing through the slit when the quantum itself apparently rushes through space without expanding?

[19] It might be argued that these quanta from stars have come an enormously long way, and possibly have had a better chance to expand than the quanta passing across a laboratory room from an X-ray tube or a mercury arc to a metal plate. However, since the photoelectric cell is used to measure the brightness of a star, they evidently produce the same sort of photoelectric effect as newborn quanta.

fact that must never be forgotten: the quantum-theory involves the wave-theory in its root and basis, for *the quantum of a given radiation is defined in terms of the frequency of that radiation, and the frequency* is determined from the wavelength, and the wavelength *is determined by applying the wave-theory* to measurements on interference and diffraction patterns. Was there ever an instance in which two such apparently contradictory theories were woven so intimately the one with the other!

The fusion of the theories is not likely to result from new experimental evidence. Indeed there are already indications that further experiments will merely accentuate the strangeness, much as happened with the numerous experiments devised and performed three or four decades ago in the hope of settling whether the earth does or does not move relatively to the aether. More probably what is required is a modification, indeed a revolutionary extension in the art of thinking—such a revolution as took place among a few mathematicians when non-Euclidean geometry was established by the side of Euclidean, as is taking place today among the disciples of Einstein who are striving to unlearn the habitual distinctions between time and space—such a revolution, to go centuries back into the past, as occurred in the minds of men generally when they learned to realize that the earth is round, and yet at every place upon it the sky is above and the ground is below. Our descendants may think pityingly of us as we of our ancestors, who could not comprehend how a man can stand upright at the Antipodes.

Wave Propagation Over Parallel Tubular Conductors: The Alternating Current Resistance

By SALLIE PERO MEAD

SYNOPSIS: On the basis of Maxwell's laws and the conditions of continuity of electric and magnetic forces at the surfaces of the conductor, the fundamental equations are established for the axial electric force and the tangential magnetic force in a non-magnetic tubular conductor with parallel return. The alternating current resistance per unit length is then derived as the mean dissipation per unit length divided by the mean square current. The general formula is expressed as the product of the alternating current resistance of the conductor with concentric return and a factor, termed the "proximity effect correction factor," which formulates the effect of the proximity of the parallel return conductor. The auxiliary functions which appear in the general formula are each given by the product of the corresponding function for the case of a solid wire and a factor involving the variable inner boundary of the conductor.

In general, the resistance may be calculated from this formula, using tables of Bessel functions. The most important practical cases, however, usually involve only the limiting forms of the Bessel functions. Special formulae of this kind are given for the case of relatively large conductors, with high impressed frequencies, and for thin tubes. A set of curves illustrates the application of the formulae.

I. INTRODUCTION

WHERE circular conductors of relatively large diameter are under consideration, the effect on the alternating current resistance of the tubular as distinguished from the solid cylindrical form becomes of practical importance. Mr. Herbert B. Dwight has worked on a special case of this problem and developed a formula for the ratio of alternating to direct current resistance in a circuit composed of two parallel tubes when the tubes are thin.[1] As infinite sums of infinite series are involved, however, his result is not well adapted to computation.

Mr. John R. Carson has given a complete solution for the alternating current resistance of two parallel solid wires in his paper "Wave Propagation Over Parallel Wires: The Proximity Effect," *Phil. Mag.*, April, 1921. The analysis of that paper may readily be extended to the more general case of propagation over two tubular conductors by a parallel method of development. This is done in the present paper. As the underlying theory is identical in the two problems, familiarity with the former paper will be assumed and the analysis will merely be sketched after the fundamental equations are established.

[1] "Proximity Effect in Wires and Thin Tubes," Trans. *A. I. E. E.*, Vol. XLII (1923), p. 850.

In this paper formulae for the alternating current resistance have been worked out in detail with particular reference to the case of relatively large conductors at high frequencies and to relatively thin tubes. In general the auxiliary functions involved are expressed as the product of the corresponding functions for solid wires by a correction factor which formulates the greater generality due to the variable inner boundary of the conductors. As far as possible the symbols are the same as in the solid wire case but refer now to the system of tubular conductors. Primes are added where the letters denote the corresponding functions for the solid wire case. This will hardly lead to confusion with the primes used in connection with the Bessel functions to denote differentiation.

The general solution is developed in section II. The alternating current resistance of one of the tubular conductors is expressed as the product of the alternating current resistance of the conductor with concentric return and a factor which formulates the effect of the proximity of the parallel return conductor. Section III is a summary of the general formula, special asymptotic forms and forms for thin conductors.

II. Mathematical Analysis and Derivation of Formulae

We require the expression for the axial electric force, E_z, in the conductors. Since the tubular conductor does not extend to $r=0$, the electric force must be expressed by the more general Fourier-Bessel expansion,

$$E_z = \sum_{n=0}^{\infty} A_n \left[J_n(\rho) + \lambda_n K_n(\rho)\right] \cos n\theta,$$

where

$$\begin{aligned}\rho &= ir\sqrt{4\pi\lambda\mu i\omega}\\ &= \xi = xi\sqrt{i} \text{ when } r=a\\ &= \zeta = yi\sqrt{i} \text{ when } r=\alpha,\end{aligned}$$

a and α being the outer and inner radii, respectively, of the conductors. The additional set of constants $\lambda_0, \lambda_1 \ldots \lambda_n$ is to be determined by the conditions of continuity at the inner boundary of the conductor. It is necessary to satisfy the boundary conditions at the surface of one conductor only, since the symmetry of the system insures that they will then be satisfied at the surface of the other also.

In the dielectric space inside the tube where $r<\alpha$, the axial electric force may be written

$$E_z=\sum_{n=O}^{\infty} C_n J_n(\rho)\cos n\theta, \tag{1}$$

or replacing the Bessel functions by their values for vanishingly small arguments,

$$E_z=\sum_{n=O}^{\infty} D_n r^n \cos n\theta \tag{2}$$

where $D_o, D_1 \ldots D_n$ are constants determined by the boundary conditions. Applying Maxwell's law relating the normal and tangential magnetic forces H_r and H_θ to the axial electric force, gives

$$\mu i\omega H_\theta=\frac{\rho}{r}\sum_{n=O}^{\infty} A_n\,[J_n'(\rho)+\lambda_n K'_n(\rho)]\cos n\theta, \tag{3}$$

$$\mu i\omega H_r=\frac{1}{r}\sum_{n=O}^{\infty} A_n\,[J_n(\rho)+\lambda_n K_n(\rho)]\sin n\theta, \tag{4}$$

for the space inside the conductor, and

$$i\omega H_\theta=\sum_{n=O}^{\infty} nD_n r^{n-1}\cos n\theta, \tag{5}$$

$$i\omega H_r=\sum_{n=O}^{\infty} nD_n r^{n-1}\sin n\theta, \tag{6}$$

for the inner dielectric ($\mu=1$). Equating the two expressions for the tangential magnetic force H_θ and for the normal magnetic induction μH_r term by term at the surface $r=\alpha$,

$$[\zeta J_n'(\zeta)-\mu n J_n(\zeta)]+\lambda_n[\zeta K_n'(\zeta)-\mu n K_n(\zeta)]=0. \tag{7}$$

Whence, for the practically important case of non-magnetic conductors in which $\mu=1$, we have

$$\lambda_n=-\frac{J_{n+1}(\zeta)}{K_{n+1}(\zeta)} \tag{8}$$

and

$$E_z=\sum_{n=O}^{\infty} A_n\left[J_n(\rho)-\frac{J_{n+1}(\zeta)}{K_{n+1}(\zeta)}K_n(\rho)\right]\cos n\theta. \tag{9}$$

In the subsequent analysis $J_n\ (\xi)$ of the solution for the solid wire case is replaced by

$$J_n(\xi)-\frac{J_{n+1}(\zeta)}{K_{n+1}(\zeta)}K_n(\xi)=M_n(\xi), \tag{10}$$

and $J_n'\ (\xi)$ is replaced by

$$J_n'(\xi)-\frac{J_{n+1}(\zeta)}{K_{n+1}(\zeta)}K_n'(\xi)=M_n'(\xi). \tag{11}$$

Otherwise the formulation of the alternating current resistance of the conductor proceeds exactly as in the solid wire case. For the electric force at the surface $r=a$ in the conductor, we write

$$E_z=A_o[M_o(\xi)+h_1M_1(\xi)\cos\theta+h_2M_2(\xi)\cos 2\theta+\ .\ .\ .] \tag{12}$$

and determine the fundamental coefficient A_o in terms of the current in the conductor. The resistance R of the tubular conductor per unit length is defined as the mean dissipation per unit length divided by the mean square current where the mean dissipation is calculated by Poynting's theorem. Accordingly, we get

$$R=\text{Real}\frac{2\mu i\omega}{\xi}\left\{\frac{M_o\ (\xi)}{M_o'(\xi)}+\frac{1}{2}\sum_{n=1}^{\infty}|h_n|^2\frac{M_n\ (\xi)}{M_o'(\xi)}\text{conj.}\frac{M_n'(\xi)}{M_o'\ (\xi)}\right\}. \tag{13}$$

To determine the harmonic coefficients $h_1\ .\ .\ .\ h_n$ or $A_1\ .\ .\ .\ A_n$, the total tangential magnetic force and the total normal magnetic induction at the outer surface of a conductor are expressed in terms of the coordinates of that conductor alone, and the conditions of continuity at the surface are applied. This leads to the set of equations

$$q_n=(-1)^n2\rho_nk^n-\frac{(-1)^n}{(n-1)!}\rho_nk^n\sum\nolimits_n(q) \qquad n=1,2,3\ .\ .\ .\ \infty \tag{14}$$

where

$$\sum\nolimits_n(q)=\frac{n!}{1!}kq_1-\frac{(n+1)!}{2!}k^2q_2+\ .\ .\ .\ ,$$

$$\sigma_n=(\xi M_n'\ (\xi)-n\mu M_n\ (\xi))/\xi\ M'(\xi),$$

$$\rho_n=(\xi M_n'\ (\xi)-n\mu M_n(\xi))/(M_n'(\xi)+n\mu M_n(\xi)),$$

$$q_n=\sigma_nh_n,$$

$$\frac{a}{c}=k.$$

When the permeability is unity, the solution, to the same order of approximation as in the solid wire case, is

$$|h_n|^2 = \frac{u_1^2+v_1^2}{u_{n-1}^2+v_{n-1}^2}\left|\frac{1+\lambda_o K_1(\xi)/J_1(\xi)}{1+\lambda_n K_{n-1}(\xi)/J_{n-1}(\xi)}\right|^2 p_n^2(1+2ngk^2/s^{n-1}) \qquad (16)$$

where

$$g = \frac{\sqrt{2}}{x}\,\frac{p[u_1(u_o+v_o)-v_1(u_o-v_o)]-q[u_1(u_o-v_o)+v_1(u_o+v_o)]}{u_o^2+v_o^2}, \qquad (17)$$

$$p+iq = \frac{1+\lambda_1 K_1(\xi)/J_1(\xi)}{1+\lambda_1 K_o(\xi)/J_o(\xi)}, \qquad (18)$$

$$J_n(\xi) = u_n + iv_n,$$

$$p_n = (-1)^n 2k^n s^n,\ n = 1, 2 \ldots \infty,$$

$$s = 2\,\frac{1-\sqrt{1-(2k)^2}}{(2k)^2}.$$

Since the resistance R_o of an isolated tubular conductor is given by

$$R_o = \text{Real}\,\frac{2\mu i p}{\xi}\,\frac{M_o(\xi)}{M_o'(\xi)} \qquad (19)$$

equation (13) becomes equation (I) of the formulae in the next section. This is the general solution for the case of non-magnetic conductors.

In general R may be calculated from this formula and tables of Bessel functions. The ber, bei, ker and kei functions [2] and the recurrence formulae are sufficient to evaluate the Bessel functions but the process is long. In the most important practical cases, the conductors are rather large and the applied frequencies fairly high. When this is true as well as when the tubes are very thin the formulae usually involve only the limiting forms of the Bessel functions. These special results are given in the next section.

III. Alternating Current Resistance Formulae for Non-Magnetic Conductors

The symbols used are:

a = outer radius of conductor in centimeters,

α = inner radius of conductor in centimeters,

c = interaxial separation between conductors in centimeters,

$k = a/c$

λ = conductivity of conductor in electromagnetic c.g.s. units,

[2] A convenient table of these functions for arguments from 0 to 10 at intervals of 0.1 is incorporated in Mr. Dwight's paper "A Precise Method of Calculation of Skin Effect in Isolated Tubes," *J. A. I. E. E.*, Aug., 1923.

μ = permeability of conductor in electromagnetic c.g.s. units,

$\omega = 2\pi$ times frequency in cycles per second,

$i = \sqrt{-1}$

$x = a\sqrt{4\pi\lambda\omega}$

$y = \alpha\sqrt{4\pi\lambda\omega}$

$\xi = xi\sqrt{i}$

$\zeta = yi\sqrt{i}$

$\lambda_n = -J_{n+1}(\zeta)/K_{n+1}(\zeta)$

$J_n(\xi) = u_n + iv_n$

= Bessel function of first kind of order n and argument $xi\sqrt{i}$,

$J_n'(\xi) = \dfrac{dJ_n(\xi)}{d\xi}$

$u_n' + iv_n' = \dfrac{dJ_n(\xi)}{dx}$

$K\ (\xi)$ = Bessel function of second kind of order n and argument $xi\sqrt{i}$,

$K_n'(\xi) = \dfrac{dK_n(\xi)}{d\xi}$

R = resistance per unit length of tubular conductor with parallel return,

R_o = resistance per unit length of tubular conductor with concentric return in electromagnetic c.g.s. units,

C = proximity effect correction factor,

$$R = C\,R_o. \tag{I}$$

The auxiliary functions involved are:

$$^{3}\ R_o = R_o' m\left(1 - \frac{n}{m}\,\frac{u_o u_o' + v_o v_o'}{u_o v_o' - u_o' v_o}\right) \tag{20}$$

where

$$R_o' = \frac{1}{a}\sqrt{\frac{\omega}{\pi\lambda}}\,\frac{u_o v_o' - u_o' v_o}{u_1^2 + v_1^2} \tag{21}$$

= resistance of solid wire with concentric return,

$$m + in = \frac{1 + \lambda_o K_o(\xi)/J_o(\xi)}{1 + \lambda_o K_o'(\xi)/J_o'(\xi)}, \tag{22}$$

$$g = g'p\left\{1 - \frac{q\,[u_1(u_o - v_o) + v_1(u_o + v_o)]}{p\,[u_1(u_o + v_o) - v_1(u_o - v_o)]}\right\}, \tag{23}$$

[3] The ratio R_o/R_o' oscillates about unity which it approaches more and more closely as the frequency increases. It is due to the fact that the phase of the current in the inner portion of the solid conductor may be such as to oppose the current in the outer portion, that the resistance of the solid conductor may be greater than that of the tube even though the heating effect in the latter is the greater.

where

$$g' = \frac{\sqrt{2}}{x} \frac{u_1(u_o+v_o) - u_1(u_o - v_o)}{u_o^2 + v_o^2}, \tag{24}$$

$$p+iq = \frac{1+\lambda_1 K_1(\xi)/J_1(\xi)}{1+\lambda_1 K_o(\xi)/J_o(\xi)}, \tag{25}$$

$$w_n = w_n' \frac{a_n}{|1+\lambda_n K_{n-1}(\xi)/J_{n-1}(\xi)|^2}\left(1 - \frac{b_n}{a_n}\frac{u_n u_n' + v_n v_n'}{u_n v_n' - u_n' v_n}\right), \tag{26}$$

where

$$w_n' = \frac{u_n v_n' - u_n' v_n}{u_{n-1}^2 + v_{n-1}^2}, \tag{27}$$

$$a_n + ib_n = \left(1+\lambda_n \frac{K_n(\xi)}{J_n(\xi)}\right) \text{conj.} \left(1+\lambda_n \frac{K_n'(\xi)}{J_n'(\xi)}\right), \tag{28}$$

$$s = 2\frac{1-\sqrt{1-(2k)^2}}{(2k)^2}. \tag{29}$$

The formula for the correction factor C is then

$$C = 1 + \frac{2}{aR_o}\sqrt{\frac{\omega}{\pi\lambda}}(S_1 + 2gk^2S_2) \tag{II}$$

where

$$S_1 = \sum_{n=1}^{\infty} w_n k^{2n} s^{2n}, \tag{30}$$

$$S_2 = \sum_{n=1}^{\infty} n w_n k^{2n} s^{n+1}. \tag{31}$$

For large values of the argument

$$R_o = R_o'\left[m - n\left(1 - \frac{1}{\sqrt{2}x}\right)\right] \tag{32}$$

and the correction factor is

$$C = 1 + 2\frac{\sqrt{2} - 1/x}{m - n(1 - 1/\sqrt{2}x)}\left(S_1 - \frac{2\sqrt{2}}{x}\left[p + q\left(1 - \frac{1}{\sqrt{2}x}\right)\right]k^2 S_2\right) \tag{III}$$

When x and y are both large quantities, the auxiliary functions are as follows, provided terms of the second order in $1/x$ and $1/y$ are negligible, n in d and h below being equal to the number of terms in which S_1 and S_2 converge to a required order of approximation.

With the notation

$$\cos = \cos\sqrt{2}(x-y),$$

$$\sin = \sin\sqrt{2}(x-y),$$

$$\exp = \exp\left[-\sqrt{2}(x-y)\right],$$

$$R_o = R_o' \frac{1 + [(1+a)\sin - (1-a)\cos]\exp - a\exp^2}{1 - [(1-b)\sin + (1+b)\cos]\exp + b\exp^2} \tag{33}$$

where

$$a = 1 - \frac{1}{2\sqrt{2}x} - \frac{3}{2\sqrt{2}y},$$

$$b = 1 + \frac{3}{2\sqrt{2}x} - \frac{3}{2\sqrt{2}y},$$

$$\frac{1}{aR_o'}\sqrt{\frac{\omega}{\pi\lambda}} = \sqrt{2} - \frac{1}{x}, \tag{34}$$

$$g = g' \frac{1 + [(1-c)\cos - (1+c)\sin]\exp - c\exp^2}{1 - [(1+c)\cos + (1-c)\sin]\exp + c\exp^2}, \tag{35}$$

where

$$c = 1 - \frac{1}{2\sqrt{2}x} - \frac{15}{2\sqrt{2}y},$$

$$g' = -\sqrt{2}/x, \tag{36}$$

$$w_n = w_n' \frac{1 - [(1-d)\cos - (1+d)\sin]\exp - d\exp^2}{1 - [(1+h)\cos + (1-h)\sin]\exp + h\exp^2}, \tag{37}$$

where

$$d = 1 + \frac{4n^2 - 1}{2\sqrt{2}x} - \frac{4(n+1)^2 - 1}{2\sqrt{2}y},$$

$$h = 1 + \frac{4(n-1)^2 - 1}{2\sqrt{2}x} - \frac{4(n+1)^2 - 1}{2\sqrt{2}y},$$

$$w_n' = \frac{1}{\sqrt{2}} - \frac{2n-1}{2x}. \tag{38}$$

At frequencies sufficiently high to afford practically skin conduction, the following formulae indicate the way in which the resistance of the tubular conductor approaches its limit, the resistance of the solid wire.

$$R_o = R_o' \frac{1 + 2\sin\exp}{1 - 2\cos\exp}, \tag{39}$$

$$\frac{1}{aR_o'}\sqrt{\frac{\omega}{\pi\lambda}} = \sqrt{2} - \frac{1}{x},$$

$$C = C_m(1 - A/x), \tag{IV}$$

$$C_m = \frac{1 + k^2s^2}{1 - k^2s^2}, \tag{40}$$

$$A = 2\sqrt{2}\,\frac{k^2s^2}{1 - k^4s^4}\left\{1 + 2k^2\frac{(1-k^2s^2)^2}{(1-k^2s)^2}\,\frac{1 - 2\sin\exp}{1 - 2\cos\exp}\right\}. \tag{41}$$

When the conductors are very thin tubes, i.e., thin as compared to the radius, $(a-\alpha)/a$ is necessarily small and, in general, $x-y$ is small. Of course, when the frequency is high enough, $x-y$ becomes large in any case. When this is true with respect to thin tubes, however, x and y will usually be large enough to make the asymptotic formulae applicable; but, if $x-y$ is small, the approximations

$$J_n(\zeta)=J_n(\xi)-(\xi-\zeta)J_n'(\xi)+\frac{(\xi-\zeta)^2}{2!}J_n''(\xi),$$

$$K_n(\zeta)=K_n(\xi)-(\xi-\zeta)K_n'(\xi),+\frac{(\xi-\zeta)^2}{2!}K_n''(\xi),$$

reduce the correction factor to

$$C=1+2\,\beta^2 f\left\{\sum_{n=1}^{\infty}k^{2n}s^{2n}\frac{d_n}{D_n}-2k^2\frac{\dfrac{4d_1^2}{x^4}-\beta^3c_1}{D_1}\sum_{n=1}^{\infty}k^{2n}s^{n+1}n\frac{d_n}{D_n}\right\} \qquad \text{(V)}$$

where $\beta=\dfrac{a-\alpha}{a}$,

$$f=\frac{(1+\beta/2)^2}{1+\beta+\beta^2}=\frac{c_o^2}{d_o},$$

$$D_n=\beta^2c_n^2+\frac{4n^2}{x^4}d_n^2,$$

$$c_n=1+\frac{2n+1}{2}\beta,$$

$$d_n=1+(n+1)\beta+\frac{(n+1)\,(n+2)}{2}\beta^2.$$

and the resistance with concentric return to

$$R_o=\frac{1}{2\pi\lambda a(a-\alpha)}\,\frac{1+\beta+\beta^2}{1+\beta/2}. \qquad (42)$$

$1/2\pi\lambda a(a-\alpha)$ is, of course, the direct current resistance of a very thin conductor.

If $(a-\alpha)/a$ is very small and negligible compared with $2n/x^2$, where n is the number of terms in which the series of (V) converge to a required order of approximation,

$$C=1+\frac{x^4}{2}\left(\frac{a-\alpha}{a}\right)^2\left\{\begin{array}{l}\left(1-\dfrac{a-\alpha}{a}\right)\left\{\displaystyle\sum_{n=1}^{\infty}\frac{k^{2n}s^{2n}}{n^2}+2k^2s\log(1-k^2s)\right\}\\ \quad+\dfrac{a-\alpha}{a}\left\{\log(1-k^2s^2)+2\dfrac{k^4s^2}{1-k^2s}\right\}\end{array}\right\} \qquad \text{(VI)}$$

As a check on formulae (V) and (VI), the limiting cases may be arrived at directly as follows. If the conductors are thin tubes, the harmonic coefficients are given by

$$h_n = (-1)^{n+1} 2k^n \frac{\xi - \zeta}{\frac{2n}{\xi} - (\xi - \zeta)\left(1 - \frac{2n(n+1)}{\xi^2}\right)}$$

$$- (-1)^{n+1} \frac{\xi - \zeta}{\frac{2n}{\xi} - (\xi - \zeta)\left(1 - \frac{2n(n+1)}{\xi^2}\right)} k^n \left[nkh_1 - \frac{n(n+1)}{2!} k^2 h_2 + \dots \right]. \quad (43)$$

When ξ is very large

$$h_n = (-1)^n 2k^n \left[1 - \frac{1}{2} \left\{ nkh_1 - \frac{n(n+1)}{2!} k^2 h_2 + \dots \right\} \right]$$

$$= (-1)^n 2k^n s^n, \quad (44)$$

and

$$\frac{M_n}{M_o} = \frac{M_n'}{M_o'} = 1 \quad (45)$$

so that

$$C = \text{Real} \left[1 + \frac{1}{2} \sum_{n=1}^{\infty} |h_n|^2 \frac{M_n}{M_o} \text{conj.} \frac{M_n'}{M_o'} \right]$$

$$= \frac{1 + k^2 s^2}{1 - k^2 s^2}, \quad (46)$$

the same result as for the corresponding limiting case of a solid conductor.

On the other hand, if ξ is not large and $\xi - \zeta$ is very small,

$$h_n = (-1)^{n+1} \frac{k^n}{n} \xi(\xi - \zeta), \quad (47)$$

$$\frac{M_n}{M_o} = 1, \quad (48)$$

$$\frac{M_n'}{M_o'} = -\frac{in}{x(x-y)}, \quad (49)$$

so that

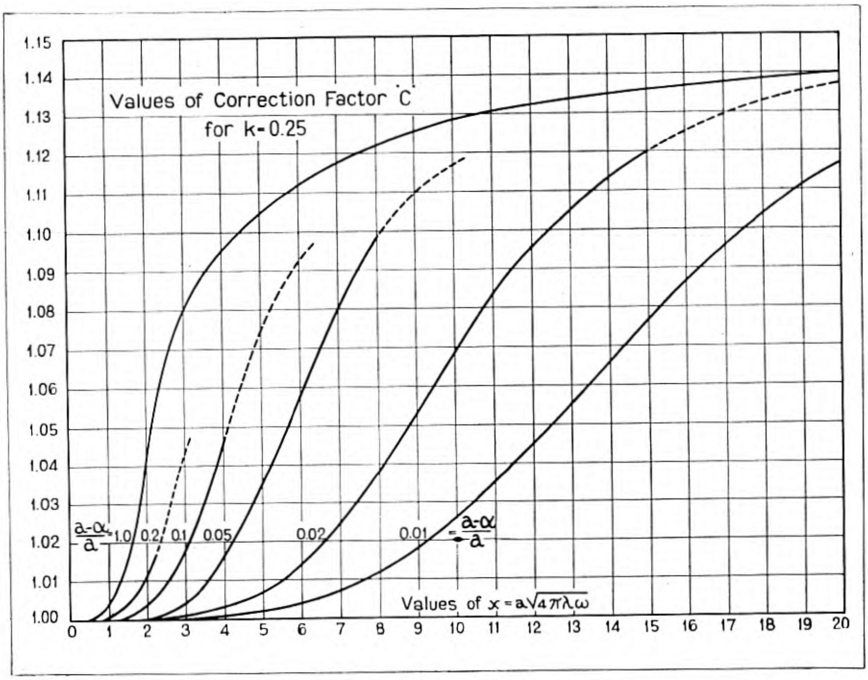
$$C = 1, \quad (50)$$

and

$$R = R_o = R_{d.c.}, \quad (51)$$

where $R_{d.c.}$ is the direct current resistance of the thin tubular conductor. Eqs. (46) and (50) agree with the corresponding limits of formulae V and VI respectively.

The curves of the accompanying figure do not pretend to represent the proximity effect correction factor with precision. They are, however, accurate for thin tubes, and indicate the order of magnitude of the factor for various values of the thickness of the tubular conductor and show the nature of its variation with respect to the applied



frequency. They are computed from formula (V) which is valid for quite high frequencies when the tubes are thin. When the thickness of the tubes is greater, however, the range of validity with respect to frequency is smaller, the dotted portions indicating a doubtful degree of precision. It was previously pointed out in connection with formula (IV) and is immediately deducible from physical considerations, that all of the curves eventually coincide with the curve for the solid wire which approaches the value 1.155 asymptotically.

As a simple application, suppose the resistance is required of a tubular conductor with an outer radius of 0.4125 cm. (that of No. O gauge A.W.G. copper wire) whose resistivity is 1696.5 electromagnetic units per cm., where there is an equal parallel return so situated that $k=0.25$ and a frequency of 5,000 cycles per second is applied to the circuit. Then $m=\sqrt{4\pi\lambda\omega}=15.26$ and $x=ma=15.26\times 0.4125=6.30$. When the ratio of the thickness of the conductor to the radius is greater than about 0.01 the proximity effect correction factor C is appreciable. If the ratio is 0.05, reading C from the curves, gives $C=1.064$. From formula (42), $R_o=5.24$ ohms per mi. which makes the resistance $R=5.53$ ohms per mi.

Abstracts of Bell System Technical Papers Not Appearing in this Journal

Voice-Frequency Carrier Telegraph System for Cables.[1] B. P. HAMILTON, H. NYQUIST, M. B. LONG and W. P. PHELPS. Carrier telegraph systems using frequencies above the voice range have been in use for a number of years on open-wire lines. These systems, however, are not suitable for long toll cable operation because cable circuits greatly attenuate currents of high frequencies. The system described in this paper uses frequencies in the voice range and is specially adapted for operation on long four-wire cable circuits, ten or more telegraph circuits being obtainable from one four-wire circuit. The same carrier frequencies are used in both directions and are spaced 170 cycles apart. The carrier currents are supplied at each terminal station by means of a single multi-frequency generator.

Metallic Polar-Duplex Telegraph System for Long Small-Gage Cables.[2] JOHN H. BELL, R. B. SHANCK, and D. E. BRANSON. In connection with carrying out the toll-cable program of the Bell System, a metallic-circuit polar-duplex telegraph system was developed. The metallic-return type of circuit lends itself readily to the cable conditions, its freedom from interference allowing the use of low potentials and currents so that the telegraph may be superposed on telephone circuits. The new system represents an unusual refinement in direct current telegraph circuits, the operating current being of the same order of magnitude as that of the telephone circuits on which the telegraph is superposed.

The following are some of the outstanding features of the present system. Sensitive relays with closely balanced windings are employed in the metallic circuit, and "vibrating circuits" are provided for minimizing distortion of signals. Repeaters are usually spaced about 100 miles apart. Thirty-four-volt line batteries are used and the line current is four or five milli-amperes on representative circuits. Superposition is accomplished by the compositing method which depends upon frequency discrimination, the telegraph occupying the frequency range below that of the telephone. New local-circuit arrangements have been designed, employing polar relays for repetition of the signals; these arrangements are suitable for use in making up circuits in combination with carrier-current and ground-return polar-duplex telegraph sections. New forms of mounting are em-

[1] Journal A. I. E. E., Vol. 44, p. 213, 1925.

[2] Presented at the mid-winter convention of the A. I. E. E., Feb., 1925.

ployed in which a repeater is either built as a compact unit or is made up of several units which are mounted on I-beams, and subsequently interconnected. In the latter case the usual arrangements for sending and receiving from the repeater are omitted, and a separate "monitoring" unit provided for connection to any one of a group of repeaters.

The metallic system is suitable for providing circuits up to 1,000 miles or more in length, the grade of service being better than that usually obtained from ground-return circuits on open-wire lines for such distances. About 55,000 miles of this type of telegraph circuit are in service at present.

Polarized Telegraph Relays.[3] J. R. FRY and L. A. GARDINER. This paper discusses two forms of polarized telegraph relay which have been developed by the Bell System for metallic telegraph circuits and for carrier current telegraph circuits. Both relays are of the same general construction except that one is more sensitive and carries an auxiliary accelerating winding. The more sensitive relay is required to operate on reversals of line current of one milliampere, and at the same time retain its adjustment over long periods and faithfully and accurately repeat signals. It is interesting to note that under average conditions the ratio of power controlled by the contact circuit to that required by the line windings is about 5,000 to one. The parts entering into the magnetic circuit of this relay except for a permanent magnet, are made of the new magnetic alloy (permalloy) recently developed in the Bell Telephone Laboratories. Permalloy lends itself to use in this relay because of its high permeability and very small residual effects. The design of the relay armature and the support for the moving contacts is such that contact chatter is practically eliminated. Photo-micrograms showing practically no destructive action are given of the contacts of a relay which was in continuous service for 8½ months, during which time each contact made and broke its circuit approximately 45,000,000 times.

Supervisory Systems for Remote Control.[4] J. C. FIELD. With the great growth in power distributor systems and especially with the advent of the automatic substations with no attendant there has arisen need for a supervisory system to indicate to the central load dispatcher the position or operating condition of each important power unit in the outlying stations and also to give him means to operate promptly these power units when desired.

[3] Journal A. I. E. E., Vol. 43, p. 223, 1925.

[4] Electrical Communications, Vol. 3, pp.127–133, 1924.

By the turning of a key the dispatcher can open or close any switch or circuit breaker, start or stop any of the machines and receive back almost instantly a visual and continuous signal of a red or green lamp. The present systems provide in effect a key and two lamps, one red, one green, for each unit supervised mounted in easy access of the dispatcher.

Two main systems known as the distributor supervisory and the selector supervisory have been developed to meet the varying conditions of service.

The distributor system is recommended when there is a large number of units to be supervised in a given station. It consists essentially of two motor-driven distributors, one in each station, running in synchronism. Brushes on each distributor pass over corresponding segments of two sets of 50 segments at the same instant. Thus by means of only four connecting wires between the stations the control and continuous indication of 50 power units is possible.

The selector system is recommended when there is only a few switches to be supervised in a single station or in several stations located some distance apart. It consists essentially of hand operated keys to send predetermined codes of impulses to operate selectively step by step selectors at the distant stations. After the selector has operated the power unit, an auxiliary contact on this unit operates a motor-driven key to send coded impulses to operate a selector at the dispatcher's station to indicate the condition of the unit by lighting a red or green lamp. Several stations can be supervised over the same three-line wires.

The dispatcher, by looking at the lamps on his control board, can thus tell at all times the electrical and mechanical conditions at all points in the system and has means to change the operating conditions at any substation according to the demand for power.

Note on Dr. Louis Cohen's Paper on Alternating Current Cable Telegraphy.[5] L. A. MacColl. This is a criticism of two papers which were published in the Journal of the Franklin Institute by Dr. Louis Cohen. It is shown that Cohen's development of the theory of cable telegraphy contains many defects and errors, and in particular that his criticisms of H. W. Malcolm's book, "The Theory of the Submarine Telegraph and Telephone Cable," are without foundation.

Telephone Circuit Unbalances, Determination of Magnitude and Location.[6] L. P. Ferris and R. G. McCurdy. This paper dis-

[5] Journal of the Franklin Institute, Vol. 199, p. 99, 1925.
[6] Journal A. I. E. E., Vol. 43, p. 1133, 1924.

cusses the effects of unbalances of telephone circuits on noise and crosstalk, and describes methods for detecting the presence of these unbalances and locating them when detected. The maintenance of telephone circuits in a high state of efficiency with respect to balance is important since unbalances contribute to crosstalk between telephone circuits and to noise when such circuits are involved in inductive exposures. Different types of unbalances are included and their effects under different conditions of energization of the unbalanced circuit and neighboring conductors are discussed. Methods are described for determining:

(1) The general condition of circuits with respect to balance by crosstalk measurements from their terminals.

(2) The approximate location of unbalances along a line by measurements over a range of frequencies with a bridge at one end of the line.

(3) The final location of unbalances by field measurements with an unbalance detector which may be operated by a lineman and which usually does not require interruption of telephone service, except momentarily.

Toll circuit office unbalances are briefly discussed and a special bridge for detecting and measuring the unbalances of composite sets is described. A mathematical treatment of the bridge method for locating unbalances and a discussion of the necessity of terminating the circuits involved in the tests in their characteristic line impedances are given in an appendix. The methods and apparatus described are widely used in the Bell System and afford operating telephone companies means for maintaining their circuits in the condition of minimum practicable unbalance.

The Theory of Probability and Some Applications to Engineering Problems.[7] E. C. MOLINA. The purpose of this paper is to suggest a wider recognition by engineers of a body of principles which, in its mathematical form, is a powerful instrument for the solution of practical problems. Certain fundamental principles of the theory of probabilities are stated and applied to three problems from the field of telephone engineering.

Note on the Least Mechanical Equivalent of Light.[8] HERBERT E. IVES. In this paper the value for the brightness of the black body at the melting point of platinum recently obtained by the writer is

[7] Journal A. I. E. E., Vol. 44, p. 122, 1925.

[8] Journal of the Optical Society of American and Rev. of Scientific Instruments, Vol. 10, No. 3, March, 1925, p. 289.

used to find a value for the least mechanical equivalent of light using the latest values for the black body constants and the melting point of platinum. The spectral luminous efficiency curve obtained by Tyndall and Gibson is employed. It is found that over the entire range of probable values of the black body constants, the values for the least mechanical equivalent of light may be plotted as a straight line in terms of $\frac{C_2}{T}$ so that the present computations may be expressed in a simple equation in which any desired values of the black body constants may be inserted. Using the latest values the least mechanical equivalent of light is found to be .00161 watts per lumen. This is practically identical with the value obtained by using the author's earlier experimental determination using the monochromatic green mercury light, when combined with the Gibson and Tyndall luminous efficiency curve.

Photoelectric Properties of Thin Films of Alkali Metals.[9] HERBERT E. IVES. The thin films of alkali metals which deposit spontaneously on clean metal surfaces in highly exhausted inclosures are studied. The alkali metals, sodium, potassium, rubidium, and caesium, in the thin film form all exhibit, to a striking degree, the selective photoelectric effect first discovered in sodium-potassium alloy. Experiments on varying the thickness of the deposited film show that the selective effect only occurs at a certain stage of the film's development; for very thin films the selective effect is absent, and it disappears again for thick layers of the pure alkali metal. The wavelength maxima of emission previously ascribed to the selective effect in the pure alkali metals on the basis of observations with rough or colloidal surfaces are absent in these thin films.

The Normal and Selective Photoelectric Effects in the Alkali Metals and Their Alloys.[10] HERBERT E. IVES and A. L. JOHNSRUD. The photoelectric currents from specular surfaces of molten sodium, potassium, rubidium, and caesium, and their alloys are studied at various angles of incidence for the two principal planes of polarization. The selective photoelectric effect is clearly exhibited only in the case of the liquid alloy of sodium and potassium. Wave-length distribution curves show maxima of emission, which are usually, but not always, most pronounced for light polarized with the electric vector parallel to the plane of incidence. The wave-length maxima previously assigned to the several elements are not confirmed; the

[9] Astrophysical Journal, Vol. LX, No. 4, November, 1924.

[10] Astrophysical Journal, Vol. LX, No. 4, November, 1924.

maxima vary in position for the same element with the condition and mode of preparation of the surface.

Theory of the Schroteffekt.[11] T. C. Fry. The current from a vacuum tube is composed of discrete particles of electricity which emerge according to no regular law but in an accidental, statistical fashion. The current therefore fluctuates with time. If the fluctuations are amplified sufficiently they may be heard in a telephone receiver as "noise"—a type of noise which is due to the mechanism of electron emission itself and not to outside interference. This noise is called the "Schroteffekt."

The effect is of certain importance from the telephone standpoint, for it appears that signals, the intensity of which is lower than that of the accidental current fluctuations, can never be rendered intelligible by vacuum tube amplification since the noise due to the statistical fluctuations of space current would be amplified to the same extent and would mask the signals. Fortunately, however, the effect is much less pronounced under operating conditions than it is under the conditions which are most favorable for laboratory study. This is due to the fact that the presence of space charge under operating conditions smooths out the electron stream to a very material extent, and thus reduces the tube noise. The limitation imposed upon amplification is therefore not serious.

The present paper deals with what we have termed "laboratory conditions" as distinct from "operating conditions." Its principal result, arrived at by theoretical consideration, is: That if the electrons are emitted independently of one another the intensity of the noise in the measuring instrument is

$$S=\nu\overline{w}_1,$$

where ν is the number of electrons emitted per unit time and $\overline{w}_1$ is the average over all electrons of the energy that each would have caused to be dissipated in the measuring device if not other had ever been emitted.

When this formula is applied to the type of simply tuned circuit that was considered by earlier writers, it leads to substantially the the same results as they had obtained. It is more general than these earlier results, however, and rests on less questionable methods of derivation. It is, in fact, more general than the problem of the Schroteffekt itself and applies equally well to the absorption of energy from any type of accidental disturbance which satisfies the condition that the individual electromotive impulses occur inde-

[11] Journal of Franklin Institute, Vol. 199, p. 203, 1925.

pendently of one another. Static in radio telephony and certain types of crosstalk probably satisfy these conditions.

The Transmission Unit.[12] R. V. L. Hartley. The Bell System has recently adopted a new transmission unit, abbreviated *TU*, for expressing those quantities which heretofore have been expressed in miles of standard cable, or in Europe in terms of the βl unit. It is shown that units of this type measure the logarithm of a ratio, and that the present art requires that this ratio be that of two amounts of power. Any of the proposed units may be so defined. Their essential difference is in the ratio chosen to correspond to one unit. The ratio chosen for the *TU*, $10^{0.1}$, makes it nearly the same in size as the 800-cycle mile, which has advantages. It also facilitates the use of common logarithms in preference to natural logarithms for which the ratio e of the βl unit is adapted. A distortionless reference system calibrated in *TU* is discussed, and conversion tables for the various units are given.

The Thermionic Work Function of Oxide Coated Platinum.[13] C. Davisson and L. H. Germer. Measurements of the thermionic work function of pure platinum coated with oxides of barium and strontium have been made simultaneously by two methods for the same segment of a uniformly heated filament. The theory of the measurements and the experimental arrangements are the same as used in an earlier experiment on the thermionic work function of pure tungsten.[14] Filament temperatures accurate to $\pm 5°$, were found from the resistance of the filament at 0° C. in conjunction with the temperature coefficients of resistance. (1) In the Calorimetric method the equivalent voltage of the work function was computed from the sudden voltage change resulting from switching off the space current, due to the cooling effect of the emission. The determination was much more difficult that in the case of the tungsten filament, and measurements were made at the signle temperature, 1064° K. At this temperature the work function ϕ was found to be equal to $1.79 \pm .03$ volts. (2) In the temperature variation method it was found that, after the temperature had been changed suddenly from one value to another, the emission changed approximately exponentially from an initial value to a final steady value. The half value period of this change varied from a few seconds at high temperature to over a quarter of an hour at low temperature. Interpreting this

[12] Electrical Communications, July, 1924. London Electrician, January 16 and 23, 1925.

[13] Physical Review, Vol. 24, p. 666, 1924.

[14] Davisson and Germer, Phys. Rev., 20, 300 (1922).

phenomenon as due to a progressive and reversible change of the character of the filament with temperature, the initial emissions after temperature changes from 1064° K, were used to determine the b constant of Richardson's equation corresponding to the equilibrium character of the filament at 1064° K, and similar measurements were made for the b constant corresponding to the character of the filament at 911° K. The two determinations lead, through the relationship $\phi = bk/e$, to 1.79 volts and 1.60 volts for the corresponding values of ϕ. For 1064° K, then, the two methods give values for ϕ in agreement. The measurements are, however, not sufficiently accurate to give any indication whether or not an electron within the metal possesses the thermal energy $3kT/2$. The various corrections made and possible errors are thoroughly discussed. It is pointed out that if the transition from the equilibrium state at one temperature to that at another had occurred so rapidly as to avoid observation, a disagreement of 25 per cent. between the values of ϕ given by the two methods would have been obtained which might have been misinterpreted.

Contributors to this Issue

Herbert E. Ives, B.S., University of Pennsylvania, 1905; Ph.D., Johns Hopkins, 1908; assistant and assistant physicist, Bureau of Standards, 1908–09; physicist, Nela Research Laboratory, Cleveland, 1909–12; physicist, United Gas Improvement Company, Philadelphia, 1912–18; U. S. Army Air Service, 1918–19; research engineer, Western Electric Company (Bell Telephone Laboratories), 1919 to date. Dr. Ives' work has had to do principally with the production, measurement and utilization of light.

J. W. Horton, B.S., Massachusetts Institute of Technology, 1914; instructor in physics, 1914–16; Engineering Department of the Western Electric Company, 1916—. Mr. Horton has been closely connected with the development of apparatus for carrier current communication.

Ralzemond D. Parker, B.S., University of Michigan, 1905; M.S., 1906; instructor in Electrical Engineering, University of Michigan, 1906–09; assistant professor, 1909–13; Engineering Department, American Telephone and Telegraph Company, 1913–19; Department of Development and Research, 1919—. Mr. Parker's work has related particularly to telegraphy, included the development of printing telegraph apparatus, carrier, and metallic circuit systems for fine wire cables.

A. B. Clark, B.E.E., University of Michigan, 1911; American Telephone and Telegraph Company, Engineering Department, 1911–19; Department of Development and Research, 1919—. Mr. Clark's work has been connected with toll telephone and telegraph systems.

H. W. Nichols, B.S., 1908, E.E., 1911, Armour Institute of Technology; M.S., 1909, Ph.D., 1918, University of Chicago; Assistant Professor of Electrical Engineering, Armour Institute of Technology, 1909–14; Engineering Department, Western Electric Company (Bell Telephone Laboratories), 1914—. Since 1916 Mr. Nichols has been in charge of the laboratories research in radio communication.

J. C. Schelleng, A.B., 1915; instructor in physics, Cornell University, 1915–18; Engineering Department, Western Electric Company (Bell Telephone Laboratories), 1919—. Since 1918, Mr. Schelleng has been engaged in research in radio communication.

T. C. SMITH, B.S., Purdue University, 1910; Plant Engineering, New York Telephone Company, 1910–14; engineering construction of high tension lines and municipal electric light plants, 1915; Outside Plant Engineering, New York Telephone Company, 1916–19; Automotive Engineering, New York Telephone Company, 1919–21; Automotive and Construction Apparatus Engineering, American Telephone and Telegraph Company, 1921—.

JOHN R. CARSON, B.S., Princeton, 1907; E.E., 1909; M.S., 1912; Research Department, Westinghouse Electric and Manufacturing Company, 1910–12; instructor of physics and electrical engineering, Princeton, 1912–14; American Telephone and Telegraph Company, Engineering Department, 1914–15; Patent Department, 1916–17; Engineering Department, 1918; Department of Development and Research, 1919—. Mr. Carson's work has been along theoretical lines and he has published several papers on theory of electric circuits and electric wave propagation.

KARL K. DARROW, S.B., University of Chicago, 1911; University of Paris, 1911–12; University of Berlin, 1912; Ph.D., in physics and mathematics, University of Chicago, 1917; Engineering Department, Western Electric Company, 1917–24; Bell Telephone Laboratories, Inc., 1925—. Mr. Darrow has been engaged largely in preparing studies and analyses of published research in various fields of physics.

SALLIE PERO MEAD, A.B., Barnard College, 1913; M.A., Columbia University, 1914; American Telephone and Telegraph Company, Engineering Department, 1915–19; Department of Development and Research, 1919—. Mrs. Mead's work has been of a mathematical character relating to telephone transmission.